



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

RNA Folding and Hydrolysis Terms Explain ATP Independence of RNA Interference in Human Systems

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Ali, Nicole, and Vinothan N. Manoharan. 2009. RNA folding and hydrolysis terms explain ATP-independence of RNA interference in human systems. <i>Oligonucleotides</i> 19(4): 341-346.
Published Version	doi:10.1089/oli.2009.0186
Accessed	January 27, 2018 4:30:57 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:8301599
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

RNA folding and hydrolysis terms explain ATP-independence of RNA interference in human systems

Nicole Ali^{1†}, Vinothan N. Manoharan^{1,2*}

Abstract

Although RNA interference (RNAi) has emerged as an important tool for studying the effects of gene knockdown, it is still difficult to predict the success of RNAi effectors in human systems. By examining the basic thermodynamic equations for RNA interactions in RNAi, we demonstrate how the free energies of RNA folding and phosphoester bond hydrolysis can drive RNAi without ATP. Our calculations of RNAi efficiency are close to actual values obtained from *in vitro* experimental data from two previous studies, for both silencing complex formation (2.50 vs. 2.40 for relative efficiency of RISC formation) and mRNA cleavage (0.50 vs 0.56 for proportion cleaved). Our calculations are also in agreement with previous observations that duplex unwinding and target site folding are major energy barriers to RNAi.

¹ Department of Physics, Harvard University, Cambridge MA USA

[†] Present address: Yale School of Medicine, New Haven, CT

² School of Engineering and Applied Sciences, Harvard University, Cambridge MA USA

* To whom correspondence should be addressed: vnm@seas.harvard.edu

Introduction

The regulation of gene expression via the RNA interference (RNAi) process requires a double-stranded RNA molecule with a sequence that is complementary to the targeted transcript (Fire, 1998). MicroRNAs (miRNAs) are short endogenous regulatory RNAs that act through RNAi. miRNAs are incorporated into the miRNA loading complex (miRISC), forming the RNA induced silencing complex (RISC), which binds to complementary mRNA and cleaves at the target site or induces silencing by other mechanisms including translational repression and RNA destabilization. The basic mechanism modeled in this paper is summarized in Figure 1. We focus on the RISC-based cleavage mechanism of RNAi, and note that while other mechanisms of RNAi may be more prevalent in cellular systems, RISC formation is necessary for most modes of RNAi and the mechanism described in this paper is important for design of synthetic RNAi effectors (Shabalina et al, 2006).

Although many synthetic constructs have been made to mimic miRNA function (Kim and Rossi, 2007), the design of such constructs is typically based on a complex set of empirical rules rather than more fundamental RNA thermodynamic data. Fundamental models of RNAi based on thermodynamic calculations could lead to more accurate and efficient rational design strategies for synthetic miRNAs. Here we demonstrate a necessary first step toward that goal: we outline the basic thermodynamics of RNAi, and we show that a model based on RNA thermodynamics can accurately reproduce data for RNAi efficiency in *in vitro* human systems.

An interesting characteristic of human RNAi is that, in contrast with RNAi in *Drosophila*, both RISC formation and multiple rounds of mRNA cleavage do not seem to require energy from ATP hydrolysis (Maniataki and Mourelatos, 2005; Gregory, 2005; Provost, 2002; Zhang, 2002). Some hypotheses are that coupling with the phosphoester bond-breaking steps and binding with Ago2 help drive duplex-unwinding (Maniataki and Mourelatos, 2005; Gregory et al., 2005). Here we use existing data to verify these hypotheses.

While previous studies (Shabalina et al, 2006; Taxman et al, 2006; Ui-Tei et al, 2006) have examined a wide range of thermodynamic “features” in order to find correlations with RNAi effector efficiency, little work has been done to model RNAi function from first principles. *In vitro* systems provide closed, thermodynamically controlled systems, and the apparent lack of ATP dependence makes absolute thermodynamic modeling possible. We examine data from two *in vitro*-based papers in which detailed quantitative information is available on concentrations of protein, RNAi effectors, target mRNA, and levels of knockdown (Maniataki and Mourelatos, 2005; Gregory et al., 2005). Using an equilibrium statistical thermodynamic model based on

nearest-neighbor interactions, we accurately reproduce patterns in ATP-independent knockdown levels measured previously *in vitro*. Our calculations support the hypothesis that phosphoester bond-breaking steps are sufficient to drive RNAi through the RISC cleavage mechanism.

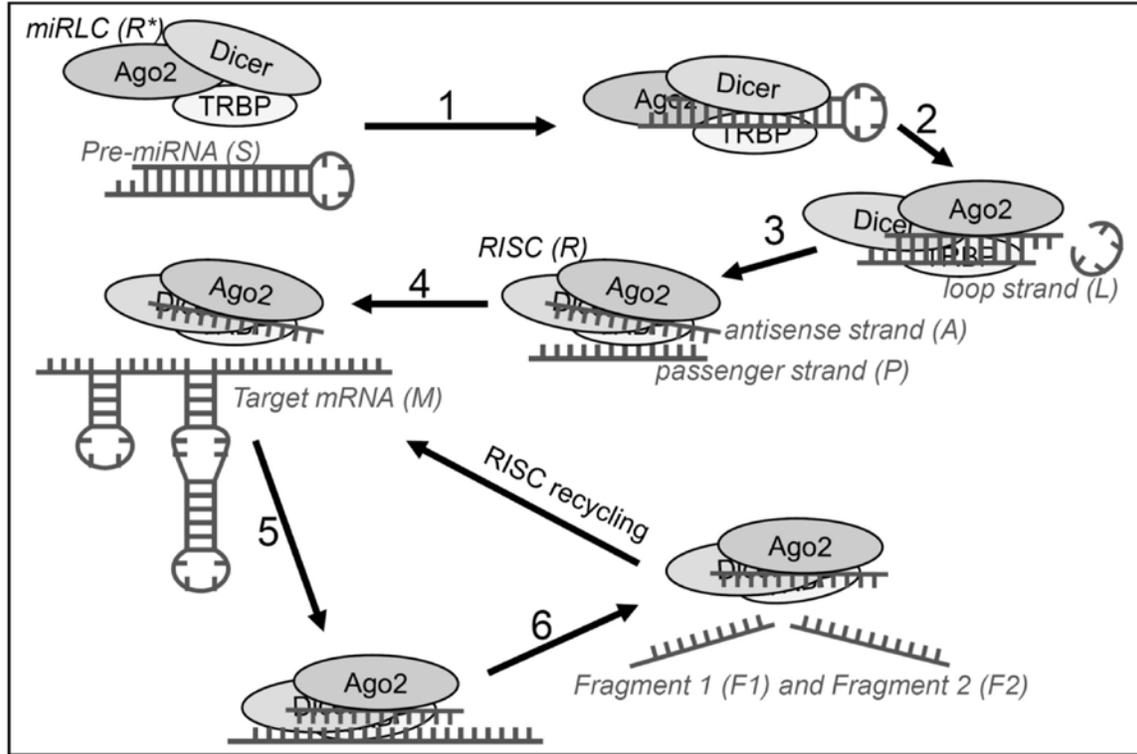


Figure 1. The RNA interference mechanism modeled in the Methods and Calculations section. Steps 1 to 3 correspond to the methods section *Calculation of in vitro RISC formation levels and comparison to data from Gregory et al (2005)*, with the net reaction $S + R^* \rightarrow R + L + P$. Steps 4 to 6 correspond to the methods section *Calculation of in vitro mRNA cleavage levels and comparison to data from Maniataki and Mourelatos (2005)*, with the net reaction $M \xrightarrow{R} F1 + F2$. The energies involved in the labeled steps are: (1) ΔG_{bind} , (2) $\Delta G_{hydrolysis}$, (3) ΔG_S and $\Delta G_{L,P}$, (4) & (5) ΔG_M , (6) $\Delta G_{F1,2}$ and $\Delta G_{hydrolysis}$. Note that in steps 4 to 6, the RISC binding and detaching terms cancel and are not included in the calculations.

Methods and Calculations

Energy Calculations

In a closed, non-cellular *in vitro* system without ATP, the behavior of the system depends on the distribution of energy states of the participating molecules. To briefly review, the probability that a system is in any particular microstate s is given by

$$P(s) = \frac{1}{Z} \exp\left(\frac{E(s)N}{RT}\right),$$

where E is the energy of the state, N is the number of molecules, and Z is the partition function,

$$Z = \sum_s \exp\left(\frac{E(s)N}{RT}\right).$$

The partition function for the system is a central quantity, and it is used to derive a number of important energies as described below. Changes in ensemble energies of the system determine the final outcomes of the associated reactions.

The stabilities of RNA structures can be determined using the nearest-neighbor model, in which the stability of a particular base pair depends on the adjacent base pair (Markham and Zuker, 2005). The DINAMelt webserver (Markham and Zuker, 2005) was used for RNA folding and hybridization calculations by the nearest-neighbor model at 37°C. In the calculations we assume that there is no energy input from ATP, consistent with the experimental studies.

DINAMelt takes the ensemble of unfolded and unhybridized strands as its reference state (Dmitrov and Zuker, 2004), so the energy change required to fold and hybridize a mixture of RNAs can be taken as the ensemble energy given by DINAMelt for that mixture at 37°C at the concentration entered. To summarize Dmitrov and Zuker, given an initial amount N_A^0 of molecules A , with canonical partition function $Z(N_A)$, the chemical potentials for A and self-hybridized AA are

$$\mu_A = -RT \ln[Z(N_A)] + RT \ln\left[\frac{N_A}{N_A^0}\right]$$

and

$$\mu_{AA} = -RT \ln[Z(N_{AA})] + RT \ln\left[\frac{N_{AA}}{(N_A^0)^2}\right].$$

Taking the free energy of unfolded states to be zero, the free energy

$F = \mu_A N_A + \mu_{AA} N_{AA}$ of the ensemble is the energy change from the ensemble of molecules A in completely unfolded states to the final folded ensemble. Likewise, the energy

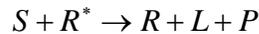
required to unfold and unhybridize the mixture is given by the negative of the (folded) ensemble energy given by DINAMelt. Calculations were done at 37°C ($R_{gas}T = 0.616268$ kcal/mol) and at the RNA concentrations as indicated from Maniataki and Mourelatos (2005) and Gregory et al (2005). DINAMelt does not account for variations in ion concentrations.

Hydrolysis energies were taken from previous experiments (Alberty, 2007), but we estimate that discrepancies up to 50% may arise since the systems in the experiments were at 25°C and possibly different ionic conditions. From Table 7 of Alberty (2007), the energy of hydrolysis of CMP, UMP, and TMP is 12.96 kJ/mol, or approximately 3.1 kcal/mol per bond. Two phosphoester bonds are hydrolyzed in the processing of pre-miRNA, and thus we take this energy of hydrolysis to be 6.2 kcal/mol.

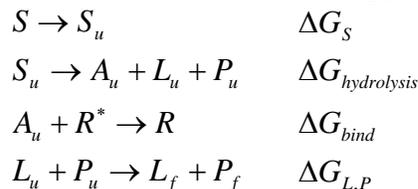
The first step of RISC formation is the binding of the pre-miRNA with Argonaute 2 (Ago2) in the miRLC (Kim and Rossi, 2007). Yan et al (2003) measured a dissociation constant $K_d \approx 1\mu M$ for binding of the Argonaute PAZ domain with short single-stranded RNA. In the experiments examined in this paper, the concentrations of RISC and guide RNA were all near or much lower than K_d , so the free energy change for binding $\Delta G_{bind} \geq 0$. Thus, for ensuing calculations, we took $\Delta G_{bind} \approx 0$ as a first order approximation.

Calculation of in vitro RISC formation levels and comparison to data from Gregory et al (2005)

To compare RISC formation using pre-miRNA vs. dsRNA, we compared energies using the let-7 pre-miRNA and dsRNA sequences from Gregory et al (2005). All calculations were done at 0.5 nM RNA in accordance with the experiments. Let R^* = inactivated miRLC, R = activated RISC, S = complete pre-miRNA, L = loop fragment, and P = passenger strand. The net reaction for RISC formation is given by:



Let the subscript f stand for “folded” and u stand for “unfolded”, and let A stand for the antisense or guide strand of the pre-miRNA. For the purposes of calculating the free energy of formation, the total reaction can be thought of as the sum of four simpler reactions, each with an associated free energy change ΔG :



ΔG_S and $\Delta G_{L,P}$ can be calculated using DINAMelt. For the let-7 pre-miRNA, which undergoes the reaction above, inputting the sequence yields a ΔG_S of 11.0472 kcal/mol, and inputting the cleaved fragments gives $\Delta G_{L,P}$ of -2.56656 kcal/mol. For the let-7 pre-miRNA, $\Delta G_{hydrolysis}$ is the energy change due to cleavage of 2 phosphoester bonds to release three new fragments, and can be approximated as -6.2 kcal/mol (Alberty, 2007). As mentioned in the previous section, we take $\Delta G_{bind} \approx 0$ as a first order approximation. The total energy change for RISC formation is given by $\Delta G_{RISC} = \Delta G_S + \Delta G_{hydrolysis} + \Delta G_{bind} + \Delta G_{L,P}$. The corresponding equilibrium constant is given as:

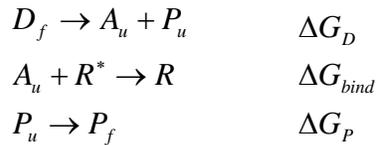
$$K_{RISC} = \exp(-\Delta G_{RISC}/R_{gas}T) \quad (1)$$

This equilibrium constant comes out to 0.024706 for the let-7 pre-miRNA. Using the initial concentration of pre-miRNA $[S_0]$ as the reference concentration for RNA folding and subsequent reactions, if r is the fraction of S converted to RISC according to experimental data, the equilibrium can be written as:

$$K_{RISC} = \frac{\frac{[R]}{[S_0]} \frac{[L]}{[S_0]} \frac{[P]}{[S_0]}}{\frac{[R^*]}{[S_0]} \frac{[S]}{[S_0]}} = \frac{\frac{r[S_0]}{[S_0]} \frac{r[S_0]}{[S_0]} \frac{r[S_0]}{[S_0]}}{\frac{(1-r)[S_0]}{[S_0]} \frac{(1-r)[S_0]}{[S_0]}} = \frac{r^3}{(1-r)^2} \quad (2)$$

Combining equations (1) and (2), it is possible to solve for r given ΔG_{RISC} or vice versa. Solving for r gives the value 0.242 for the let-7 pre-miRNA.

A similar reaction occurs if double-stranded RNA (dsRNA, no loop) is used. Letting D stand for the dsRNA, the reaction can be written as:



For the let-7 dsRNA, ΔG_D is 3.67417 kcal/mol (energy to separate the two strands), there is no hydrolysis energy (no loop to hydrolyze), and ΔG_P is the unfolding energy of the passenger strand, which is -0.9837 kcal/mol. Again, we are taking $\Delta G_{bind} \approx 0$ as a first order approximation. The RISC formation energy is given by $\Delta G_{RISC,D} = \Delta G_D + \Delta G_{bind} + \Delta G_P$, which sums to 2.6905 for the let-7 dsRNA, and the corresponding equilibrium constant is given as:

$$K_{RISC,D} = \exp(-\Delta G_{RISC,D}/R_{gas}T) \quad (3)$$

For the let-7 dsRNA, the equilibrium constant comes out to 0.0127. If r_D is the fraction of D converted to R , the equilibrium constant can also be written as:

$$K_{RISC,D} = \frac{r_D^2}{(1-r_D)^2} \quad (4)$$

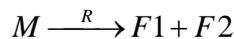
Combining (3) and (4), it is possible to solve for r_D given $\Delta G_{RISC,D}$. For the let-7 dsRNA, this method gives a value of 0.101 for r_D .

Assuming the reaction rate is roughly proportional to enzyme concentration (Liu, 2004), r/r_D approximates the difference in RNAi efficiency for pre-miRNA versus dsRNA. In this case, $r/r_D = 0.242/0.101$ which is approximately 2.40. r/r_D can also be found by measuring actual RNAi efficiencies in vitro. We used data from Gregory et al Figure 4, which had experiments performed at 0.5 nM RNA concentrations. In Figure 4B, the % conversion to cleaved product for RISC loaded let-7 dsRNA without ATP is given as 6, and in Figure 4C, the % conversion for let-7 pre-mRNA without ATP is 15. Taking this to be equal to r/r_D , we obtain a value of $15/6 = 2.50$, close to our calculated prediction of 2.40.

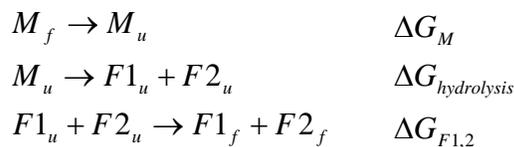
By inputting the let-7 mRNA target sequence and the sequences of the cleaved product into DINAMelt, we found the mRNA target site unfolding energy, and mRNA fragment strands ensemble energy, which are also included in Table 1.

Calculation of in vitro mRNA cleavage levels and comparison to data from Maniataki and Mourelatos (2005)

We next attempt to predict absolute mRNA knockdown levels using the GL3 target and GL3 pre-miRNA sequences from Maniataki and Mourelatos (2005) and compare to their experimental data. Calculations are at 10 nM concentrations in accordance with the experiments. Let R stand for the active RISC complex, let M stand for the target mRNA, and let $F1$ and $F2$ stand for the two fragments resulting from cleavage of M . Then the cleavage reaction is given by:



As before, this reaction can be thought of as the sum of some simpler reactions:



GAGUACUUUGAGC) and GL3 target
(UGGACAUCACUUACGCUGAGUACUUCGAAAUG).

Results and Discussion

Previous work in the literature has focused on trends in energy dependence for RNA interference, drawing from large bodies of *in vivo* data (Shabalina et al, 2006; Taxman et al, 2006; Ui-Tei et al, 2006). In contrast, we attempt to predict absolute levels of mRNA knockdown or RNAi effector efficiency based on first principles. We provide a proof of principle by comparing thermodynamic calculations to a limited data pool.

In Table 1 we show calculated values from our model for the free energy changes in various steps of the RNAi mechanism. Details for the derivation of each value are given in the Methods. Important energy contributions include folding energies of the pre-miRNA and the mRNA target site (Table 1), in line with previous observations that pre-miRNA unwinding and mRNA target site unfolding provide the greatest energy barriers to RNAi (Shabalina et al, 2006).

With these values we first examine the difference in efficiency between pre-miRNA and dsRNA against the same target site, then we calculate the absolute fractions of mRNA that would be cleaved *in vitro* by a given RNAi effector. Overall, numerical results from the model agree well with the experimentally measured levels of RISC formation and mRNA cleavage *in vitro*, with no adjustable parameters.

Calculated RISC formation levels and comparison to data from Gregory et al (2005)

If r is the fraction of pre-miRNA converted to RISC and r_D is the fraction of dsRNA converted to RISC, the value r/r_D should correspond roughly to the ratios of mRNA cleaved for reactions using pre-miRNA versus dsRNA. For RNAi effectors against let-7 mRNA targets used in Gregory et al (2005), we calculate a theoretical r/r_D of 2.4, close to the experimental value of 2.50.

The hypothesis in the original experimental work was that energy for the duplex-unwinding step of RISC is provided by the cleavage of two phosphoester bonds to release the loop fragment of the pre-miRNA. The agreement between the results from our model and the experimental data supports this hypothesis. While it is well established that stem-loop structures are better RNAi effectors than dsRNA, our calculations suggest that the primary reason for the difference in efficiency is that pre-miRNA has a loop that can be cleaved whereas dsRNA does not. The extra 6.2 kcal/mol liberated by the cleavage reaction is a significant fraction of the total ΔG .

Experiment	Energy Terms for RISC formation			Energy terms for mRNA cleavage		
	pre-miRNA or dsRNA unfolding ΔG_S ΔG_D	pre-miRNA or dsRNA hydrolysis $2 \times \Delta G_{hydrolysis}$	pre-miRNA or dsRNA loop/passenger strand ensemble energy $\Delta G_{L,P}$ ΔG_P	mRNA target site unfolding ΔG_M	mRNA hydrolysis $\Delta G_{hydrolysis}$	mRNA fragment strands ensemble energy ΔG_F
Gregory et al (2005), pre-let-7-miRNA	11.05	-6.20	-2.57	0.12	-3.10	-1.05
Gregory et al (2005), ds-let-7-miRNA	3.67	0.00	-0.98	0.12	-3.10	-1.05
Maniataki and Mourelatos (2005), PML-GL3	31.81	-6.20	-5.93	4.68	-3.10	-1.38

Table 1. Calculated energy terms, in kcal/mol, for various steps in the RNAi mechanism. Reference concentrations were set at 0.5 nM for the Gregory paper and 10 nM for the Maniataki paper, in concordance with their respective experimental concentrations. Details on how each value was obtained are included in the methods section.

Calculated in vitro mRNA cleavage levels and comparison to data from Maniataki and Mourelatos (2005)

Maniataki and Mourelatos (2005) had previously performed experiments with purified human minimal RISC (Dicer, Ago2, and the RNA binding protein TRBP), monitoring cleavage of GL3 target mRNA (10 nM) for 90 minutes. Assuming that the cleavage reaction had reached equilibrium by that time, our calculation for the fraction c of cleaved mRNA should match the observed fraction. Calculating folding energies with DINAMelt and solving for c as outlined in the methods, we calculate a fraction c of 0.56, close to the measured value, 0.50. The calculation suggests that energy changes due to RNA folding and phosphoester bond hydrolysis are sufficient to explain the levels of mRNA cleavage by minimal RISC *in vitro*.

Thus the model shows quantitative support for the notion that RNA bond cleavage energies drive both RISC formation and mRNA cleavage. But we also note that energy changes due to binding to RISC proteins were not included in our model, and were in fact not necessary to explain the *in vitro* RNAi results. A key implication is that direct modeling from basic physical principles and RNA thermodynamic data may be sufficient

to accurately describe human RNAi systems; that is, the protein binding energies do not seem to make a substantial contribution to the total free energy, at least for the sequences studied in the *in vitro* experiments.

However, this may not be the case for dynamic living systems. Future experiments measuring binding constants of human Ago2 with a variety of single-stranded RNAs could help make better estimates of the relevant energy changes in living human cells. We note also that the simple equilibrium thermodynamic model presented in this paper does not take into account many of the complexities of the cellular system, where longer sequences are involved, additional proteins may participate, discarded RNA fragments are quickly cleaved by cellular machinery, additional mechanisms of RNAi are prevalent, and rates of mRNA or miRNA production are of utmost importance.

Nonetheless, our calculations show that nearest neighbor thermodynamic model can provide useful quantitative insight into RNA interference. More thorough characterization of RISC formation and activity in live human cells should help develop more accurate kinetic models of RNA interference.

References

- ALBERTY, A. (2007). Thermodynamic properties of enzyme-catalyzed reactions involving cytosine, uracil, thymine, and their nucleosides and nucleotides. *Biophysical Chemistry* 127, 91-96.
- DIMITROV, R.A. and ZUKER, M. (2004) Prediction of hybridization and melting for double-stranded nucleic acids. *Bipohys. J.* 87, 215-226.
- FIRE, A., XU, S., MONTGOMERY, M.K., KOSTAS, S.A., DRIVER, S.E., AND MELLO, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 391(6669), 806-811.
- GREGORY, R., CHENDRIMADA, T., COOCH, N., and SHIEKHATTAR, R. (2005). Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. *Cell* 123, 631-640.
- KIM, D.H., and ROSSI, J.J. (2007). Strategies for silencing human disease using RNA interference. *Nature Rev. Genetics* 8, 173-184.
- LIU, J., CARMELL, M., RIVAS, F., MARSDEN, C., THOMSON, J., SONG, J.J., HAMMOND, S., JOSHUA-TOR, L., and HANNON, G. (2004). Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441.
- MARKHAM, N.R., and ZUKER, M. (2005). DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* 33, W577-W581.
- MANIATAKI, E., AND MOURELATOS, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes & Dev.* 19, 2979-2990.
- PROVOST, P., DISHART, D., DOUCET, J, FRENDEWEY, D., SAMUELSSON, B., and RADMARK, O. (2002). Ribonuclease activity and RNA binding of recombinant human Dicer. *The EMBO Journal* 21, 5864-5874.
- SHABALINA, S., SPRIDONOV, A., and OGURTXOV, A. (2006). Computational models with thermodynamic and composition features improve siRNA design. *BMC Bioinformatics* 7:65.

TAXMAN, D., LIVINGSTON, L., ZHANG, J., CONTI, B., IOCCA, H., WILLIAMS, K., LICH, J., TING, J., and REED, W. Criteria for effective design, construction, and gene knockdown by shRNA vectors. (2006). *BMC Biotech.* 6:7.

UI-TEI, K., NAITO, Y., and SAIGO, K. Essential Notes Regarding the Design of Functional siRNAs for Efficient Mammalian RNAi. (2006). *J. of Biomed. and Biotech.* 2006, Article ID 65052, 8 pages.

YAN, K., YAN, S., FAROOQ, A., HAN, A., ZENG, L., and ZHOU, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. *Nature* 426:27, 469-274.

ZHANG, H., KOLB, F., BRONDANI, V., BILLY, E., and WITOLD, F. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *The EMBO Journal* 21, 5875-5885.