Thermodynamics of RNA-RNA Binding

Ulrike Mückstein[1], Hakim Tafer[1], Jörg Hackermüller[2], Stephan Bernhard[1], Peter F. Stadler[3,1,4,], and Ivo L. Hofacker[1,0]

[1]Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria
[2]Novartis Institutes for Biomedical Research Vienna, Informatics and Knowledge Management at NIBR, Insilico Sciences, Brunnerstrasse 59, A-1235 Vienna, Austria
[3]Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstrasse 16-18, D-04107 Leipzig, Germany
[4]The Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, New Mexico

Abstract. We present an extension of the standard partition function approach to RNA secondary structures that computes the probabilities $P_u[i,j]$ that a sequence interval $[i,j]$ is unpaired. Comparison with experimental data shows that $P_u[i,j]$ can be applied as a significant determinant of local target site accessibility for RNA interference (RNAi). Furthermore these quantities can be used to rigorously determine binding free energies of short oligomers to large mRNA targets. The resource consumption is comparable to a single partition function computation for the large target molecule. We can show that RNAi efficiency correlates well with the binding probabilities of the siRNAs to their respective mRNA target.
Keywords: RNA secondary structure, RNA hybridization, dynamic programming,

1. Introduction

Secondary structure prediction for a single RNA molecule is a classical problem of computational biology, which has received increasing attention in recent years due to mounting evidence emphasizing the importance of RNA structure in a wide variety of biological processes [19, 27, 31, 28]. Despite its limitations, free energy minimization [34, 42, 40] is at present the most accurate and most generally applicable approach of RNA structure prediction, at least in the absence of a large set of homologous sequences. It is based upon a large number of measurements performed on small RNAs and the assumption that stacking base pairs and loop entropies contribute additively to the free energy of an RNA secondary structure [21, 20]. In this framework, a secondary structure is interpreted as the collection of all the three-dimensional structures that share a common pattern of base pairs, hence we speak of a free energy of an individual secondary structure.

Under the assumption that RNA secondary structure are pseudo-knot free, i.e., that base pairs do not cross[1], there are efficient exact dynamic programming algorithms that solve not only the folding problem [39] but also provide access to the full thermodynamics of the model via its partition function [22]. Two widely used software packages implementing these algorithms are available, `mfold` [40, 42] and the `Vienna RNA Package` [17, 16].

More recently, the secondary structure approach has been applied to the problem of interacting RNA molecules. Algorithmically, the "co-folding" of two RNAs can be dealt with in the same way as folding a single molecules by concatenating the two sequences and using different energy

[0]Address for correspondence: Ivo L. Hofacker
....
email: `ivo@tbi.univie.ac.at`

[1]Two base pairs (i,j) and (k,l) are crossing if $i < k < j < l$.

parameters for the loop that contains the cut-point between the two sequences. A corresponding `RNAcofold` program is described in [17], the `pairfold` program [1] also computes suboptimal structures in spirit of `RNAsubopt` [36]. A restricted variant of this approach is implemented in the program `RNAhybrid` [29], see also [41, 4]: here secondary structures within both monomers are neglected so that only intermolecular base pairs are taken into account. The program `bindigo` uses a variation of the Smith-Waterman sequence alignment algorithm for the same purpose [15].

The restriction of the folding algorithm to pseudo-knot-free structures, however, excludes a large set of structures that should not be excluded when studying the hybridization of a short oligonucleotide to a large mRNA. In particular, there is no biophysically plausible reason to exclude elaborate secondary structures in the target molecule (as in the case of `RNAhybrid`). On the other hand, binding of the oligo is in practice not restricted to the exterior loop of the target RNA, as is implicitly assumed in the `RNAcofold` approach.

Here we extend previous RNA/RNA cofold algorithms by taking into account that the oligo can bind also to unpaired sequences in hairpin, interior, or multi-branch loops. These cases could in principle be handled using a generic approach to pseudo-knotted RNA structures [7, 8] at the expense of much more costly computations. Instead we conceptually decompose RNA/RNA binding into two stages: (1) we calculate the partition function for secondary structures of the target RNAs subject to the constraint that a certain sequence interval (the binding site) remains unpaired. (2) We then compute the interaction energies given that the binding site is unpaired in the target. The total interaction probability is than obtained as the sum over all possible binding sites. The advantage is that the memory and CPU requirements are drastically reduced: For a target RNA of length n and an oligo of length $w < n$ we need only only $\mathcal{O}(n^2)$ memory and $\mathcal{O}(n^3 w)$ time (compared to $\mathcal{O}(n^2)$ memory and $\mathcal{O}(n^3)$ time for folding the target alone).

We apply this approach to published data from RNAi experiments [31]: we demonstrate that siRNA/mRNA binding can be quantitatively predicted our procedure. The predicted binding energies correlate well with expression data, showing that the RNAi effect depends quantitatively on siRNA/mRNA binding. In addition to assessing the interactions at known binding sites, our approach also provides an effective way of identifying alternative binding sites, since the computational effort for scanning target mRNA is small compared to the initial partition function calculation.

2. Energy-Directed RNA Folding

All dynamic programming algorithms for RNA folding can be viewed as more sophisticated variants of the maximum circular matching problem [26]. The basic idea is that each base pairs in a secondary structure divides the structure an in interior and an exterior part that can be treated separately as a consequence of the additivity of the energy model. The problem of finding, say, the optimal structure of a subsequence $[i, j]$ can thus decomposed into the subproblems on the subsequence $[i + 1, j]$ (provided i remains unpaired) and on pairs of intervals $[i + 1, k - 1]$ and $[k + 1, j]$ (provided i forms a base pairs with some position $k \in [i, j]$). In the case of the more realistic "loop-based" energy models the approach is the same. In addition, however, one now has to distinguish between the possible types of loops that are enclosed by the pair (i, k) because hairpin loops, interior loops, and multiloops all come with different energetic contributions.

Algorithms that are designed to enumerate all structure (with a below-threshold energy) [36], that compute averages over all structures [22], or that sample from a (weighted [6] or unweighted [33]) ensemble of secondary structures, need to make sure that the decomposition of the structures into substructures in unique, so that each secondary is counted once and only once in the dynamics programming algorithm.

The basis of our algorithm is a modified version of the recursions for the equilibrium partition function introduced by McCaskill [22] as implemented in the Vienna RNA package [17].

3. Probability of an Unpaired Region

In the following let $F(S)$ denote the free energy of a secondary structure S, and write β for the inverse of the temperature times Boltzmann's constant. The equilibrium partition function is defined as $Z = \sum_S \exp(-\beta F(S))$. The partition function is the gateway to the thermodynamics of RNA folding. Quantities such as ensemble free energy, specific heat, and melting temperature can be readily computed from Z and its temperature dependence.

Since the frequency of particular structure S in equilibrium is given by $P(S) = \exp(-\beta F(S))/Z$, partition functions also provide the starting point for computing the frequency of a given structural motif. In particular we are interested in the probability $P_u[i,j]$ that the sequence sequence interval $[i,j]$ is unpaired. Denoting the set of secondary structures in in which $[i,j]$ remains unpaired by $\mathcal{S}^u_{[i,j]}$ we have

$$P_u[i,j] = \frac{1}{Z} \sum_{S \in \mathcal{S}^u_{[i,j]}} e^{-\beta F(S)} \tag{1}$$

Clearly, the set $\mathcal{S}^u_{[i,j]}$ will be exponentially large in general. The program `Sfold` [6, 5] adds a stochastic backtracking procedure to McCaskill's partition function calculation [22] to generate a properly weighted sample of structures. One then simply counts the fraction of structures with the desired structural feature. This approach becomes infeasible, however, when $P_u[i,j]$ becomes smaller than the inverse of the sample size. Nevertheless, even very small probabilities $P_u[i,j]$ can be of importance in the context of interacting RNAs, as we shall see below.

We therefore present here an exact algorithm. In the special case of an interval of length 1, i.e., a single unpaired base, however, $P_u[i,i]$ can be computed by dynamics programming. Indeed, $P_u[i,i] = 1 - \sum_j P_{ij}$, where P_{ij} is the base pairing probability of pair (i,j), which is obtained directly from McCaskill's partition function algorithm [22]. It is natural, therefore, to look for a generalization of the dynamic programming approach to longer unpaired stretches[2].

We first observe that the unpaired interval $[i,j]$ is either part of the "exterior loop", (i.e., it is not enclosed by a basepair), or it is enclosed by a base pair (p,q) such that (p,q) is the closing pair of the loop that contains the unpaired interval $[i,j]$. We can therefore express $P_u[i,j]$ in terms of restricted partition functions for these two cases:

$$P_u[i,j] = \underbrace{\frac{Z[1,i-1] \times 1 \times Z[j+1,N]}{Z}}_{exterior} + \sum_{\substack{p,q \\ p<i<j<q}} \underbrace{P_{pq} \times \frac{Z_{pq}[i,j]}{Z^b[p,q]}}_{enclosed} \tag{2}$$

The first term is the ratio of the partition function of all structure in which $[i,j]$ separates arbitrary 5' from arbitrary 3' structures. In the second term, $Z_{pq}[i,j]$ is the partition function over all structures on the subsequence $[p,q]$ subject to the restriction that $[i,j]$ is unpaired and (p,q) forms base pair, while $Z^b[p,q]$ counts all structures on $[p,q]$ that form the pair (p,q). Multiplying the ratio of these two partition functions by the probability P_{pq} that (p,q) is indeed paired yields the desired fraction of structure in which $[i,j]$ is left unpaired.

The tricky part of the algorithm is the computation of the restricted partition functions $Z_{pq}[i,j]$. The recursion is built upon enumerating the possible types of loops that have (p,q) as their closing

[2]Note that we cannot simply use $\prod_{k=i}^{j} P_u[k,k]$ since these probabilities are not even approximately independent.

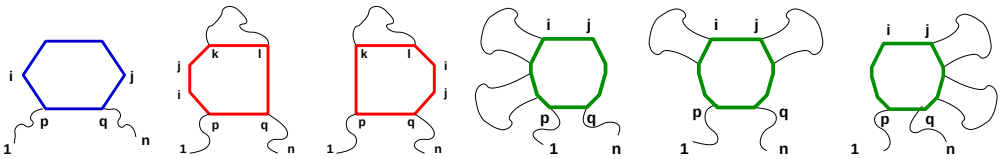

FIGURE 1. A base pair p, q can close various loop types. According to the loop type different contributions have to be considered. a) A hairpin loop is depicted in blue. b) In case of an interior loop, which is shown in red, two indepent contributions to $Q_{pq}[i,j]$ are possible: The unstructured region $[i,j]$ can be located on either side of the stacked pairs (p,q) and (k,l). c) If region $[i,j]$ is contained with in multiloop we have to account for three different conformations, indicated in the green structures, a more detailed description is given in the text.

pair and contain $[i,j]$, see Fig. 1. From this decomposition one derives:

$$\begin{aligned} Z_{pq}[i,j] = & \underbrace{\exp(-\beta H(p,q))}_{(a)} \\ & + \sum_{\substack{p<i\le j<k \text{ or} \\ l<i\le j<q}} \underbrace{Z^b[k,l]\exp(-\beta I(p,q,k,l))}_{(b)} \\ & + \sum_{p<i\le j<q} \underbrace{Z^{m2}[p+1,i-1]\exp(-\beta u(q-i))}_{(c)} \\ & + \sum_{p<i\le j<q} \underbrace{Z^m[p+1,i-1]Z^m[j+1,q-1]\exp(-\beta u(j-i+1))}_{(d)} \\ & + \sum_{p<i\le j<q} \underbrace{Z^{m2}[j+1,q-1]\exp(-\beta u(j-p))}_{(e)} \end{aligned} \tag{3}$$

The computation of the multiloop contributions (c-e) requires two additional types of restricted partitions functions: $Z^m[p,q]$ is the partition function of all conformation on the interval $[p,q]$ that are part of a multiloop and contain at least one component, i.e., at least one substructure that is enclosed by a base pair within $[p,q]$. These quantities are computed and tabulated already in the course of McCaskill's algorithm. There, the computation of Z^m requires an auxiliary array Z^{m1} which counts structures in multiloops that have *exactly* one component, the closing pair of which starts at the first position of the interval. For the one-sided multiloop cases (c) and (e) above we need in addition the partition functions of multiloop configurations that have *at least* two components. These are readily obtained using

$$Z^{m2}[p,q] = \sum_{p<u<q} Z^m[p,u]Z^{m1}[u+1,q]\,. \tag{4}$$

It is not hard to verify that this recursion corresponds to a unique decomposition of the "M2" configurations into a 3' part that contains exactly one component and a 5' part with at least one component.

It is clear from the above recursions that, in comparison to McCaskill's partition function algorithm, we need to store only one additional matrix, Z^{m2}. The CPU requirements, however, increases to $\mathcal{O}(n^4)$ (assuming the usual restriction of the length of interior loops). In practice, however, the probabilities for very long unpaired intervals are negligible, so that $P_u[i,j]$ is of

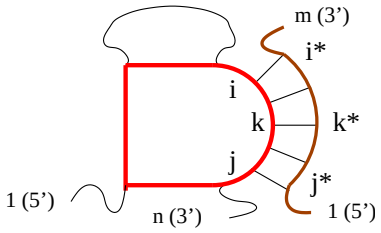

FIGURE 2. Calculation of the probability of an interaction between a short RNA and its target

interest only for limited interval length $|j - i + 1| \leq w$. Taking this constraint into account shows that the CPU requirements are actually only $\mathcal{O}(n^2 \cdot w)$.

4. Interaction Probabilities

The values of $P_u[i, j]$ can be of interest in their own right: Hackermüller, Meisner, and collaborators [**?**, 23] showed that the binding of the HuR protein to its mRNA target depends quantitatively on the probability that the HuR binding site has an unpaired conformation. The values of $P_u[i, j]$ also play a crucial role in modeling RNA-RNA interactions: The energetics of RNA-RNA interactions can be viewed as a step-wise process, $\Delta G = \Delta G_u + \Delta G_h$, in which the free energy of binding consists of the contribution ΔG_u that is necessary to expose the binding site in the appropriate conformation, and contribution ΔG_h that describes the energy gain due to hybridization at the binding site. For an unpaired binding motif in the interval $[i, j]$, we have of course $\Delta G_u = (-1/\beta)(\ln P_u[i, j] - \ln Z)$. Since the energy gain from the hybridization can be substantial, it becomes necessary to deal also with very small values of $P_u[i, j]$. The sampling approach thus becomes infeasible.

The computation of the hybridization part is performed using algorithms similar to `RNAcofold` or `RNAhybrid`. In this contribution we are mostly interested in the binding of miRNAs and siRNAs to a target mRNA, so that complex internal structures in the short RNA can be neglected.

Thus we assume that the oligo binds, possibly with mismatches, to an unpaired stretch of the target molecule. The partition function over all interior loops between a region $[i^*, j^*]$ in the small RNA and a segment $[i, j]$ in the target RNA is obtained recursively by summing up all possible interior interactions in region $[i, k; i^*, k^*]$ for each interior loop closed by base pairs (k, k^*) and (j, j^*) times the Boltzmann weight corresponding to the energy contribution of this interior loop $\exp(-\beta I(k, k^*; j, j^*))$, see Figure 2.

$$Z^*[i, j, i^*, j^*] = P_u[i, j] \sum_{\substack{i<k<j \\ i^*>k^*>j^*}} Z^*[i, k, i^*, k^*] e^{-\beta I(k,k^*;j,j^*)}. \tag{5}$$

The memory requirement for this step is $\mathcal{O}(n \cdot w^3)$, the required CPU time scales as $\mathcal{O}(n \cdot w^5)$, which at least for long target RNAs is dominated by the first step, i.e., the computation of the $P_u[i, j]$.

5. Results

In order to demonstrate that our algorithm produces biologically reasonable results, we compared predicted binding probabilities with data from RNA interference experiments. Small interfering RNAs (siRNAs) are short (21-23nt) RNA duplices with symmetric 2-3 nt overhangs [10, 24, 25]. They are used to silence gene expression in a sequence-specific manner in a process known as RNA interference (RNAi). Recently, there has been mounting evidence that the biological activity of siRNAs is influenced by local structural characteristics of the target mRNA [25, 19, 2, 37]: a target sequence must be available for hybridization in order to achieve efficient translational

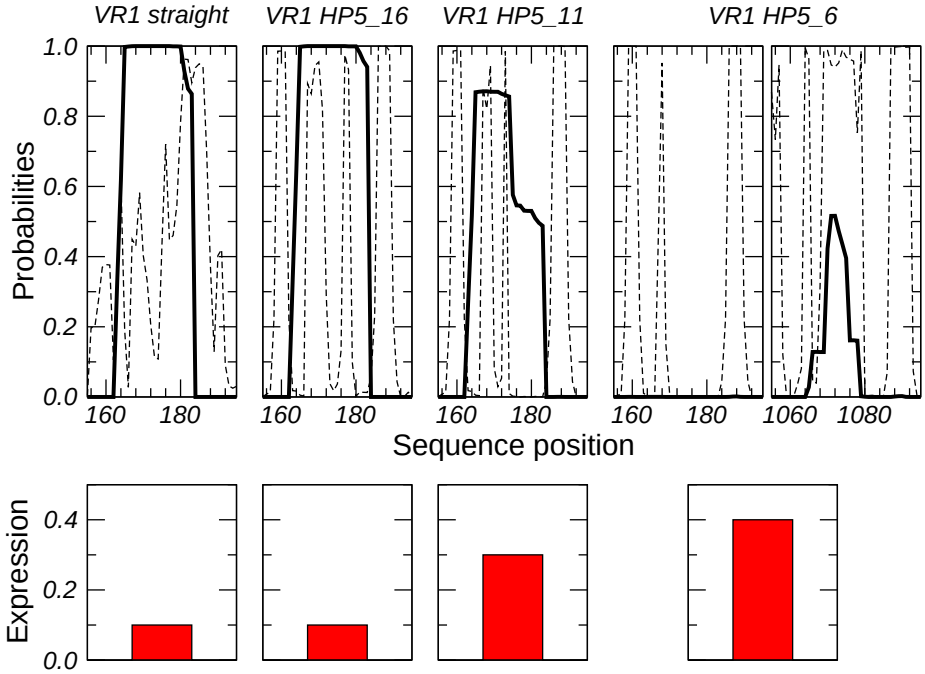

FIGURE 3. Predicted fraction of unpaired structure (dashed line) and fraction of bound siRNA (thick line) near the known binding site of RsiRNA1. Below the protein expression levels in experimental data [31] are indicated. The isolated 21mer target sequence, that has the same activity as the wild type mRNA, and 3 mutants are shown. In the case of the HP5_6 mutant the primary binding site is inaccessible. About 1kb downstream, however, there is an alternative binding position that is partially accessible in this sequence and could be responsible for the observed silencing effect.

repression, hence single stranded regions in RNA secondary structures are important. Conversely, nucleotides sequestered in a stable secondary structure can be inaccessible, resulting in decreased repression of mRNA translation [27, 31].

An obstacle for effective application of siRNAs is the fact that the extent of gene inactivation by different siRNAs varies considerably. Several groups have proposed basically empirical rules for designing functional siRNAs, see e.g. [11, 30], but the efficiency of siRNAs generated using these rules is highly variable. Recent contributions [14, 28] suggest that two of the many parameters studied so far are highly significant: The difference in the thermodynamic stability of the ends of siRNAs, which influences the choice of the strand included into RISC, the protein-siRNA complex that mediates target RNA cleavage [18, 32] and the local secondary structure of the siRNA target site, which determines the accessibility for an interaction between siRNA and mRNA target [14, 31, 27, 25, 19, 2, 37].

Schubert et al. [31] systematically analyzed the contributions of these parameters to siRNA activity. They designed a serie of constructs, which all contained the same target site for the same siRNA. However this binding site was sequestered in local secondary structure elements of different stability and extension. This experimental set up allowed a sound analysis of secondary structure effects by minimizing influences of other possible parameters for RNAi efficiency. Schubert et al. [31] observed a significant obstruction of gene silencing for the same siRNA caused by structural features of the substrate RNA. A clear correlation was found between the number of exposed nucleotides and the efficiency of gene silencing: When all nucleotides were incorporated

in a stable hairpin, silencing was reduced drastically, while exposure of 16 nucleotides resulted in efficient inhibition of expression virtually indistinguishable from the wild type.

We applied our methods to study the target sites provided by Schubert et al. [31]. Our results are in perfect agreement with the Schubert's results. Figure 3, which was calculated using our method, clearly demonstrated the correlation between target site accessibility and siRNA efficiency. The target site of the VR1 straight construct has a high probability of being unstructured, which means it is highly accessible. Accessibilty increases from the 5' to the 3' end of the target site, as proposed by Schubert, reaching a probability of being unstructured by more than 0.9 in the 3' region of the target site. The stepwise reduction of the target accessibilty is clearly visible using our programm. In case of construct VR HP5_6 no interaction at the proposed target site is possible any more. We can also offer an explanation for the still quite high efficiency of translational repression by this construct: Although the siRNA cannot bind to the proposed target site, an interaction is possible at an alternative binding site at position 1066 to 1078. Since siRNAs can also function as miRNAs [9, 38], the siRNA might act in a miRNA like fashion binding to this alternative target site. The incomplete complementarity of the siRNA to the alternative target site should be no obstacle to functionality, since it was shown that miRNA target site can be active even if the length of a continuous helix at the target site is as short as 4 - 5 basepairs [3].Using our new accessibility prediction tool we are not only able to predict potential binding site but we might be able to explain differences in si/miRNA efficiency caused by secondary structure effects. We are positive that our method can be useful for further elucidation of the mechanism of RNAi.

6. Concluding Remarks

We have demonstrated here that variants of McCaskill's partition function algorithm can be implemented efficiently to compute the probability that a given sequence interval $[i, j]$ is unpaired. Since the computation is rigorous it can also be used for small probabilities, i.e., in cases where large free energy changes are necessary to expose a binding site. Since these free energy changes are compensated by sometimes substantial hybridization energies, as in the case siRNA/mRNA binding, even very small probabilities have to be included. The approach presented here therefore overcomes inherent limitations in sampling approaches such as `Sfold` [6, 5].

Conceptually, it is not hard to extend this approach to other structural features, see also [12]. In practice, however, general purpose implementations are least tedious. Such practical limitations can be circumvented, however, in the framework of Algebraic Dynamic Programming, as exemplified in `RNAshapes` [13], which allows computations with RNA structures subject to constraints on a coarse grained level.

The accuracy of the predictions obtained from our approach are limited by two factors: (1) All probabilities above are conditional probabilities given that the molecules interact at all. Comparison with the partition function of the isolated systems and standard statistical thermodynamics, however, could be used to explicitly compute the concentration dependence of RNA-RNA binding, see e.g. [4]. (2) A more general problem is the energetics of RNA-RNA interactions in loops: the binding of the oligo to a loop will of course alter the energy contribution of the loop itself. In the model above we have implicitly assumed that this energy change is a constant. Additional measurement along the lines of the investigation of kissing-interactions [35] are required to improve the energy parameters for interacting RNAs.

Acknowledgments. This work was supported in part by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Project No. P15893, and by the German *DFG* Bioinformatics Initiative BIZ-6/1-2.

References

[1] M. Andronescu, Z. Zhang, and A. Condon. Secondary structure prediction of interacting RNA molecules. *J. Mol. Biol.*, 345(5):987–1001, 2005.

[2] E. A. Bohula, A. J. Salisbury, M. Sohail, M. P. Playford, J. Riedemann, E. M. Southern, and V. M. Macaulay. The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. *J. Biol. Chem.*, 278(18):15991–15997, 2003.

[3] J. Brennecke, A. Stark, R. Russell, and S. Cohen. Principles of microrna-target recognition. *PLoS Biol.*, 3(3):e85, 2005.

[4] R. A. Dimitrov and M. Zuker. Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys. J.*, 87(1):215–226, 2004.

[5] Y. Ding, C. Y. Chan, and C. E. Lawrence. Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Research*, 32(Web Server issue):W135–141, 2004.

[6] Y. Ding and C. E. Lawrence. A statistical sampling algorithm for rna secondary structure prediction. *Nucleic Acids Res.*, 31:7280–7301, 2003.

[7] R. Dirks and N. Pierce. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J. Comput. Chem.*, 24(13):1664–1677, 2003.

[8] R. Dirks and N. Pierce. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *J. Comput. Chem.*, 25(10):1295–1304, 2004.

[9] J. Doench, C. Petersen, and P. Sharp. sirnas can function as mirnas. *Genes Dev.*, 17(4):438–442, 2003.

[10] D. Dykxhoorn, C. Novina, and P. Sharp. Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.*, 4(6):457–467, 2003.

[11] S. Elbashir, H. J., K. Weber, and T. Tuschl. Analysis of gene function in somatic mammalian cells using small interfering rnas. *Methods*, 26(2):199–213, 2002.

[12] C. Flamm, I. L. Hofacker, and . P. F. Stadler. Computational chemistry with RNA secondary structures. *Kemija u industriji*, 53:315–322, 2004. (Proceedings CECM-2 Varaždin 2003).

[13] R. Giegerich, B. Voss, and M. Rehmsmeier. Abstract shapes of RNA. *Nucleic Acids Res.*, 32:4843–4851, 2004.

[14] B. S. Heale, H. S. Soifer, C. Bowers, and J. J. Rossi. sirna target site secondary structure predictions using local stable substructures. *Nucleic Acids Res.*, 33(3):e30, 2005.

[15] N. Hodas and D. Aalberts. Efficient computation of optimal oligo-RNA binding. *Nucleic Acids Research*, 32(22):6636–6642, 2004.

[16] I. Hofacker. Vienna RNA secondary structure server. *Nucleic Acids Res.*, 31(13):3429–3431, 2003.

[17] I. L. Hofacker, W. Fontana, P. F. Stadler, S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.*, 125:167–188, 1994.

[18] A. Khvorova, A. Reynolds, and S. D. Jayasena. Functional sirnas and mirnas exhibit strand bias. *Cell*, 115(2):209–16, 2003.

[19] R. Kretschmer-Kazemi Far and G. Sczakiel. The activity of siRNA in mammalian cells is related to structural target accessibility: a comparison with antisense oligonucleotides. *Nucleic Acids Res.*, 31(15):4417–4424, 2003.

[20] D. H. Mathews. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA*, 10(8):1178–1190, 2004.

[21] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, 288(5):911–940, 1999.

[22] J. S. McCaskill. The equilibrium partition function and base pair binding probabilities for RNA secondary structures. *Biopolymers*, 29:1105–1119, 1990.

[23] N.-C. Meisner, J. Hackermüller, V. Uhl, A. Aszódi, M. Jaritz, and M. Auer. mRNA openers and closers: A methodology to modulate AU-rich element controlled mRNA stability by a molecular switch in mRNA conformation. *Chembiochem.*, 5:1432–1447, 2004.

[24] G. Meister and T. Tuschl. Mechanisms of gene silencing by double-stranded RNA. *Nature*, 431(7006):343–349, 2004.

[25] V. Mittal. Improving the efficiency of RNA interference in mammals. *Nat. Rev. Genet.*, 5(5):355–365, 2004.

[26] R. Nussinov, G. Piecznik, J. R. Griggs, and D. J. Kleitman. Algorithms for loop matching. *SIAM J. Appl. Math.*, 35(1):68–82, 1978.

[27] M. Overhoff, M. Alken, R. K. Far, M. Lemaitre, B. Lebleu, G. Sczakiel, and I. Robbins. Local RNA target structure influences siRNA efficacy: A systematic global analysis. *J. Mol. Biol.*, 348(4):871–881, 2005.

[28] J. S. Parker, S. M. Roe, and D. Barford. Structural insights into mrna recognition from a PIWI domain-siRNA guide complex. *Nature*, 434:663–666, 2005.

[29] M. Rehmsmeier, P. Steffen, M. Hochsmann, and R. Giegerich. Fast and effective prediction of microRNA/target duplexes. *RNA.*, 10(10):1507–17, 2004.

[30] . Reynolds A, D. Leake, Q. Boese, S. S., W. Marshall, and A. Khvorova. Rational sirna design for rna interference. *Nat. Biotechnol.*, 22(3):326–30, 2004.

[31] S. Schubert, A. Grunweller, V. Erdmann, and J. Kurreck. Local rna target structure influences siRNA efficacy: Systematic analysis of intentionally designed binding regions. *J. Mol. Biol.*, 348(4):883–93, 2005.

[32] D. Schwarz, G. Hutvagner, T. Du, Z. Xu, N. Aronin, and P. Zamore. Asymmetry in the assembly of the rnai enzyme complex. *Cell.*, 115(2):99–208, 2003.

[33] M. Tacker, P. F. Stadler, E. G. Bornberg-Bauer, I. L. Hofacker, and P. Schuster. Algorithm independent properties of RNA structure prediction. *Eur. Biophy. J.*, 25:115–130, 1996.

[34] D. Turner, N. Sugimoto, and S. Freier. RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.*, 17:167–92, 1988.

[35] A. Weixlbaumer, A. Werner, C. Flamm, E. Westhof, and R. Schroeder. Determination of thermodynamic parameters for HIV DIS type loop-loop kissing complexes. *Nucleic Acids Res.*, 32:5126–5133, 2004.

[36] S. Wuchty, W. Fontana, I. L. Hofacker, and P. Schuster. Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers*, 49:145–165, 1999.

[37] K. Yoshinari, M. Miyagishi, and T. K. Effects on RNAi of the tight structure, sequence and position of the targeted region. *Nucleic Acids Res.*, 32(2):691–9, 2004.

[38] Y. Zeng, R. Yi, and B. Cullen. Micrornas and small interfering rnas can inhibit mrna expression by similar mechanisms. *Proc. Natl. Acad. Sci. USA.*, 100(17):9779–9784, 2003.

[39] M. Zuker. On finding all suboptimal foldings of an RNA molecule. *Science*, 7(244):48–52, 1989.

[40] M. Zuker. Calculating nucleic acid secondary structure. *Curr. Opin. Struct. Biol.*, 10(3):303–10, 2000.

[41] M. Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, 31(13):3406–15, 2003.

[42] M. Zuker and S. P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.*, 9(1):133–148, 1981.