Thermodynamics of RNA-RNA Binding

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

[1]Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria
[2]Novartis Institutes for Biomedical Research Vienna, Discovery Technologies, Innovative Screening Technologies, 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 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 [KKFS03, OAF+05, SGEK05, PRB05]. Despite its limitations, free energy minimization [TSF88, ZS81, Zuk00] 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 [MSZT99, Mat04]. 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 structures are pseudo-knot free, i.e., that base pairs do not cross[1], there are efficient exact dynamic programming algorithms that solve

*Address for correspondence: Ivo L. Hofacker, ivo@tbi.univie.ac.at
[1]Two base pairs (i,j) and (k,l) are crossing if $i < k < j < l$.

not only the folding problem [Zuk89] but also provide access to the full thermodynamics of the model via its partition function [McC90]. Two widely used software packages implementing these algorithms are available, `mfold` [Zuk00, ZS81] and the `Vienna RNA Package` [HFS$^+$94, Hof03].

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 molecule by concatenating the two sequences and using different energy parameters for the loop that contains the cut-point between the two sequences. A corresponding `RNAcofold` program is described in [HFS$^+$94], the `pairfold` program [AZC05] also computes suboptimal structures in the spirit of `RNAsubopt` [WFHS99]. A restricted variant of this approach is implemented in the program `RNAhybrid` [RSHG04] as well as `RNAduplex` from the Vienna RNA package, see also [Zuk03, DZ04]: 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 [HA04].

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 [DP03, DP04] 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 at a possible binding site is then obtained as the sum over all possible types of binding. The advantage is that the memory and CPU requirements are drastically reduced: For a target RNA of length n and an oligo of length $m < n$ we need only $\mathcal{O}(n^2)$ memory and $\mathcal{O}(n^3 m)$ 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 [SGEK05]: we demonstrate that siRNA/mRNA binding can be quantitatively predicted by our procedure. The predicted binding energies correlate well with expression data, showing that the effect of RNAi 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 [NPGK78]. The basic idea is that each base pair in a secondary structure divides the structure in an 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 be 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 more realistic "loop-based" energy models the same approach is used. 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 structures (with a below-threshold energy) [WFHS99], that compute averages over all structures [McC90], or that sample from a (weighted [DL03] or unweighted [TSBB$^+$96]) ensemble of secondary structures, need to make sure that the decomposition of the structures into substructures is unique, so that each secondary structure is counted once and only once in the dynamic programming algorithm.

The basis of our algorithm is a modified version of the recursions for the equilibrium partition function introduced by McCaskill [McC90] as implemented in the Vienna RNA package [HFS$^+$94].

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 a 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 interval $[i, j]$ is unpaired. Denoting the set of secondary structures 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` [DL03, DCL04] adds a stochastic backtracking procedure to McCaskill's partition function calculation [McC90] 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, $P_u[i,i]$ can be computed by dynamic programming. Indeed, $P_u[i,i] = 1 - \sum_{j \neq i} P_{ij}$, where P_{ij} is the base pairing probability of pair (i,j), which is obtained directly from McCaskill's partition function algorithm [McC90]. 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:

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

The first term accounts for the ratio between the partition functions of all sub-structures on the 5' and 3' side of the interval $[i,j]$ and the total partition function. 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 a 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 structures in which $[i,j]$ is left unpaired.

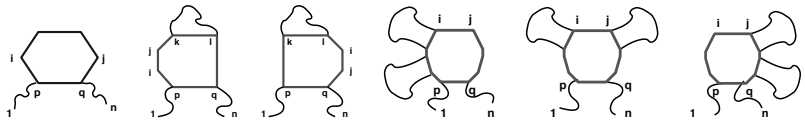

Figure 1: A base pair p,q can close various loop types. According to the loop type different contributions have to be considered. a) A hairpin loop is depicted in blue. b) In case of an interior loop, which is shown in red, two indepent contributions to $Q_{pq}[i,j]$ are possible: The unstructured region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

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

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

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

where $H(p,q)$ and $I(p,q;k,l)$ are functions that compute the loop energies of hairpin and interior loops given their enclosing base pairs; c is an energy parameter for multiloops describing the penalty for increasing the loop size by one. 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 conformations on the interval $[p,q]$ that are part of a multiloop and contain at least one component, i.e., that contain at least one substructure that is enclosed by a base pair. 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) in Fig. 1 we additionally need 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 increase 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 interest only for limited interval length $|j-i+1| \le w$. Taking this constraint into account shows that the CPU requirements are actually only $\mathcal{O}(n^3 \cdot w)$.

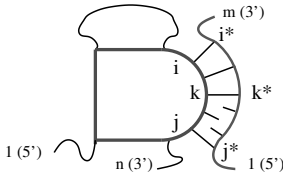

Figure 2: Calculation of the probability of an interaction between a short RNA and its target.

4 Interaction Probabilities

The values of $P_u[i,j]$ can be of interest in their own right: Hackermüller, Meisner, and collaborators [HMA$^+$05, MHU$^+$04] 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. While not much is known about the energetics of RNA-protein interactions, the case of RNA-RNA interactions can be modeled in more detail: The energetics of RNA-RNA interactions is 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. This additivity assumes that the energy of the original loop is unchanged by the binding of the oligo. For an unpaired binding motif in the interval $[i,j]$, we have of course $\Delta G_u = (-1/\beta)(\ln Z_u[i,j] - \ln Z) = (-1/\beta)\ln P_u[i,j]$. 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 similar to `RNAduplex` or `RNAhybrid`: We assume that the binding region may contain mismatches and bulge loops. Thus the partition function over all interactions between a region $[i^*, j^*]$ in the small RNA and a segment $[i,j]$ in the target RNA is obtained recursively by summing over all possible interior loops closed by base pairs (k,k^*) and (j,j^*), see Figure 2.

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

Since we are mostly interested in the binding of miRNAs and siRNAs to a target mRNA, we will neglect internal structures in the short RNA, and include unfolding of the mRNA target site. Thus only Z^I and $P_u[i,j]$ are needed to compute $Z^*[i,j]$, the partition function over all structures where the short RNA binds to region $[i,j]$, and for the computation of the corresponding binding probability, $P^*[i,j]$.

$$Z^*[i,j] = P_u[i,j] \sum_{i^*>j^*} Z^I[i,j,i^*,j^*]; \quad P^*[i,j] = Z^*[i,j] / \sum_{k<l} Z^*[k,l] \tag{6}$$

From $P^*[i,j]$ we can readily compute the probability P^*_k that a position k lies some-

where within the binding site. Note that these are conditional probabilities given that the two molecules bind at all. Furthermore $Z^*[i,j]$ can be used to calculate $\Delta G[ij] = (-1/\beta) \ln Z^*[i,j]$ the free energy of binding, where the binding site is in region $[i,j]$. For visual inspection $\Delta G[ij]$ can be reduced to the optimal free energy of binding at a given position i, $\Delta G_i = min_{k \le i \le l}\{\Delta G[kl]\}$. The memory requirement for these steps 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 [DNS03, MT04, Mit04]. 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 [Mit04, KKFS03, BSS$^+$03, YMT04, OAF$^+$05, SGEK05]: a target sequence must be accessible for hybridization in order to achieve efficient translational repression. An obstacle for effective application of siRNAs is the fact that the extent of gene inactivation by different siRNAs varies considerably. Several groups have proposed basically empirical rules for designing functional siRNAs, see e.g. [EJWT02, RALB$^+$04], but the efficiency of siRNAs generated using these rules is highly variable. Recent contributions [PRB05, SGEK05] suggest two significant parameters: The stability difference between 5' and 3' end of the siRNA, that determines which strand is included into the RISC complex [KRJ03, SHD$^+$03] and the local secondary structure of the target site [SGEK05, OAF$^+$05, Mit04, KKFS03, BSS$^+$03, YMT04].

Schubert et al. [SGEK05] systematically analyzed the contribution of mRNA structure to siRNA activity. They designed a series of constructs, all containing the same target site for the same siRNA. These binding sites, however, were sequestered in local secondary structure elements of different stability and extension. They 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. [SGEK05]. Our predictions, shown in Fig. 3, are in perfect agreement with the experimental results. The target site of the "VR1straight" construct has a high probability of being unstructured, consequently ΔG_i, the optimal free energy of binding, is highly favorable and the siRNA will bind almost exclusively to the intended target site. The stepwise reduction of the target accessibility is directly correlated to a weaker optimal free energy of binding and decreasing silencing efficiency. In case of construct VR HP5_6 the optimal free energy of binding at an alternative binding site at positions 1066 to 1078 nearly equals that at the pro-

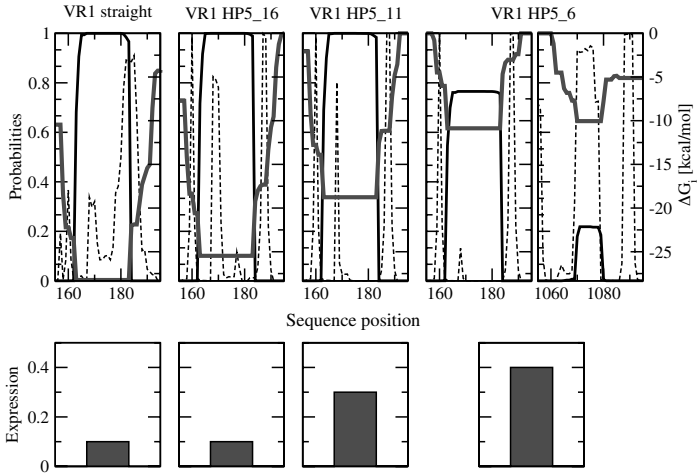

Figure 3: Probability of being unpaired $P_u[i,i]$ (dashed line), probability of binding to siRNA at position i, P_i^*, (thick black line) and ΔG_i, the optimal free energy of binding in a region including position i (thick red line) near the known target site of VsiRNA1. The scale for the probabilities is indicated on the left side, the scale for the minimal free energy of binding on the right side. At the bottom the protein expression levels in experimental data [SGEK05] are indicated. The isolated 21mer target sequence, displaying the same activity as the wild type mRNA, and 3 mutants are shown. A decreasing optimal free energy of binding is correlated with increasing expression. In the case of the HP5_6 mutant an alternative binding site becomes occupied as the optimal free energy of binding due to this alternative interaction nearly equals ΔG_i at the proposed target site.

posed target site. Since siRNAs can also function as miRNAs [DPS03,ZYC03], the siRNA might act in a miRNA like fashion binding to this alternative target site and contribute to the remaining translational repression of this construct. The incomplete complementarity of the siRNA to the alternative target site should be no obstacle to functionality, since it was shown that miRNAs can be active even if the longest continuous helix with the target site is as short as 4 - 5 basepairs [BSRC05].

Our new accessibility prediction tool can thus be used to identify potential binding sites as well as explain differences in si/miRNA efficiency caused by secondary structure effects.

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. The computation is rigorous, and can thus be used even 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 in-

cluded. The approach presented here therefore overcomes inherent limitations in sampling approaches such as `Sfold` [DL03, DCL04].

Conceptually, it is not hard to extend this approach to other structural features, see also [FHS04]. In practice, however, general purpose implementations are at least tedious. Such practical limitations can be circumvented, however, in the framework of Algebraic Dynamic Programming, as exemplified in `RNAshapes` [GVR04], which allows computations with RNA structures subject to constraints on a coarse grained level.

In our exposition above, all probabilities are conditional probabilities given that the molecules interact at all. Comparison with the partition function of the isolated systems and standard statistical thermodynamics, however, can be used to explicitly compute the concentration dependence of RNA-RNA binding, see e.g. [DZ04]. A more general limitation is our lack of knowledge concerning the energetics of RNA-RNA interactions within 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 [WWF$^+$04] 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, by the Austrian *Gen-AU* bioinformatics integration network and by the German *DFG* Bioinformatics Initiative BIZ-6/1-2.

References

[AZC05] M. Andronescu, Z.C. Zhang, and A. Condon. Secondary structure prediction of interacting RNA molecules. *J. Mol. Biol.*, 345(5):987–1001, 2005.

[BSRC05] J. Brennecke, A. Stark, R.B. Russell, and S.M. Cohen. Principles of microRNA-target recognition. *PLoS Biol.*, 3(3):e85, 2005.

[BSS$^+$03] 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.

[DCL04] 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.

[DL03] Y. Ding and C. E. Lawrence. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res.*, 31:7280–7301, 2003.

[DNS03] D.M. Dykxhoorn, C.D. Novina, and P.A. Sharp. Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.*, 4(6):457–467, 2003.

[DP03] R.M. Dirks and N.A. Pierce. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J. Comput. Chem.*, 24(13):1664–1677, 2003.

[DP04] R.M. Dirks and N.A. Pierce. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *J. Comput. Chem.*, 25(10):1295–1304, 2004.

[DPS03] J.G. Doench, C.P. Petersen, and P.A. Sharp. siRNAs can function as miRNAs. *Genes Dev.*, 17(4):438–442, 2003.

[DZ04] R. A. Dimitrov and M. Zuker. Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys. J.*, 87(1):215–226, 2004.

[EJWT02] S.M. Elbashir, Harborth. J., K. Weber, and T. Tuschl. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, 26(2):199–213, 2002.

[FHS04] Christoph Flamm, Ivo L. Hofacker, and P̃eter F. Stadler. Computational Chemistry with RNA Secondary Structures. *Kemija u industriji*, 53:315–322, 2004. (Proceedings CECM-2 Varaždin 2003).

[GVR04] R. Giegerich, B. Voss, and M. Rehmsmeier. Abstract shapes of RNA. *Nucleic Acids Res.*, 32:4843–4851, 2004.

[HA04] N.O. Hodas and D.P. Aalberts. Efficient computation of optimal oligo-RNA binding. *Nucleic Acids Research*, 32(22):6636–6642, 2004.

[HFS+94] Ivo L. Hofacker, Walter Fontana, Peter F. Stadler, Sebastian Bonhoeffer, Manfred Tacker, and Peter Schuster. Fast Folding and Comparison of RNA Secondary Structures. *Monatsh. Chem.*, 125:167–188, 1994.

[HMA+05] Jörg Hackermüller, Nicole-Claudia Meisner, Manfred Auer, Markus Jaritz, and Peter F. Stadler. The Effect of RNA Secondary Structures on RNA-Ligand Binding and the Modifier RNA Mechanism: A Quantitative Model. *Gene*, 345:3–12, 2005.

[Hof03] I.L. Hofacker. Vienna RNA secondary structure server. *Nucleic Acids Res.*, 31(13):3429–3431, 2003.

[KKFS03] 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.

[KRJ03] A. Khvorova, A. Reynolds, and S. D. Jayasena. Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2):209–16, 2003.

[Mat04] 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.

[McC90] J. S. McCaskill. The equilibrium partition function and base pair binding probabilities for RNA secondary structures. *Biopolymers*, 29:1105–1119, 1990.

[MHU+04] Nicole-Claudia Meisner, Jörg Hackermüller, Volker Uhl, Andras Aszódi, Markus Jaritz, and Manfred 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.

[Mit04] V. Mittal. Improving the efficiency of RNA interference in mammals. *Nat. Rev. Genet.*, 5(5):355–365, 2004.

[MSZT99] 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.

[MT04] G. Meister and T. Tuschl. Mechanisms of gene silencing by double-stranded RNA. *Nature*, 431(7006):343–349, 2004.

[NPGK78] Ruth Nussinov, George Piecznik, Jerrold R. Griggs, and Daniel J. Kleitman. Algorithms for Loop Matching. *SIAM J. Appl. Math.*, 35(1):68–82, 1978.

[OAF+05] 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.

[PRB05] 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.

[RALB+04] . Reynolds A, D. Leake, Q. Boese, Scaringe S., W.S. Marshall, and A. Khvorova. Rational siRNA design for RNA interference. *Nat. Biotechnol.*, 22(3):326–30, 2004.

[RSHG04] M. Rehmsmeier, P. Steffen, M. Hochsmann, and R. Giegerich. Fast and effective prediction of microRNA/target duplexes. *RNA.*, 10(10):1507–17, 2004.

[SGEK05] S. Schubert, A. Grunweller, V.A. 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.

[SHD+03] D.S. Schwarz, G. Hutvagner, T. Du, Z. Xu, N. Aronin, and P.D. Zamore. Asymmetry in the assembly of the RNAi enzyme complex. *Cell.*, 115(2):99–208, 2003.

[TSBB+96] Manfred Tacker, Peter F. Stadler, Erich G. Bornberg-Bauer, Ivo L. Hofacker, and Peter Schuster. Algorithm Independent Properties of RNA Structure Prediction. *Eur. Biophy. J.*, 25:115–130, 1996.

[TSF88] D.H. Turner, N. Sugimoto, and S.M. Freier. RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.*, 17:167–92, 1988.

[WFHS99] 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.

[WWF+04] 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.

[YMT04] K. Yoshinari, M. Miyagishi, and K. Taira. Effects on RNAi of the tight structure, sequence and position of the targeted region. *Nucleic Acids Res.*, 32(2):691–9, 2004.

[ZS81] M. Zuker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.*, 9(1):133–148, 1981.

[Zuk89] M. Zuker. On finding all suboptimal foldings of an RNA molecule. *Science*, 7(244):48–52, 1989.

[Zuk00] M. Zuker. Calculating nucleic acid secondary structure. *Curr. Opin. Struct. Biol.*, 10(3):303–10, 2000.

[Zuk03] M. Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, 31(13):3406–15, 2003.

[ZYC03] Y. Zeng, R. Yi, and B.R. Cullen. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci. USA.*, 100(17):9779–9784, 2003.