

Sfold Frequently Asked Questions (and Answers)

1. How is the algorithm for Sfold different from other established algorithms?

■ The algorithm generates a *statistical* sample of RNA secondary structures from the Boltzmann ensemble of RNA secondary structures. From a statistical mechanics perspective, an RNA molecule may have a population of structures distributed according to a Boltzmann distribution, which gives the probability of a secondary structure I at equilibrium as $(1/U)\exp[-E(I)/RT]$, where $E(I)$ is the free energy of the structure, R is the gas constant, T is the absolute temperature, and U is the partition function for all admissible secondary structures of the RNA sequence. The algorithm samples secondary structures *exactly* and *rigorously* according to the Boltzmann distribution. The sampling method offers an appealing solution to the problem of uncertainties in the folding model and in the free energy parameters. For the details of the algorithm and its unique capabilities listed below, please see our cover article (Ding & Lawrence 2003, *Nucleic Acids Res.* **31**, 7280-7301).

2. What are the unique capabilities of the algorithm?

- The Boltzmann ensemble of RNA secondary structures can be efficiently characterized through significant structural classes by classifying sampled structures.
- The algorithm offers novel probabilistic tools for prediction of RNA target accessibility.
- The algorithm offers novel probabilistic tools for evaluation of RNA/RNA interaction.
- Probabilities of structural motifs such as loops can be easily calculated.
- Free energy distributions are readily available from sampled structures.

3. What is the rationale of Sfold target-accessibility prediction for the design of RNA-targeting nucleic acids?

■ Our prediction of accessibility is based on a statistical sample of the Boltzmann ensemble of secondary structures. This novel approach is appealing for evaluation of target accessibility, because, as noted by researchers from Sirna Therapeutics (formerly Ribozyme Pharmaceuticals, Inc.) “*In the prediction of accessible sites, the identification of a single folded structure for a given target mRNA is not of particular interest. Instead, the objective of this exercise is to assess the **likelihood** of unpaired (or substantially unpaired) sites that could be a ribozyme target ... The ambiguities in thermodynamic parameters – and the possibility that each mRNA exists as a population of different structures – suggest that a **stochastic** approach to the evaluation of accessible sites may be appropriate*” (Christoffersen, McSwiggen & Konings 1994, *J. Mol. Structure (Theochem)* **311**, p. 208). The probability profiling approach in Ding and Lawrence (2001, *Nucleic Acids Res.* **29**, 1034-1046) reveals target sites that are *commonly accessible* for a large number of statistically representative structures for the target RNA. Through assignment of statistical confidence in predictions, this novel approach bypasses the long-standing difficulty in accessibility evaluation due to limited representation of probable structures.

4. Is there experimental evidence that the potency of siRNAs is influenced by target accessibility and secondary structure?

- Target accessibility has long been established as an important factor for the potency of

antisense oligonucleotides and *trans*-cleaving ribozymes. Recently, the importance of target structure and accessibility in determining the potency of siRNAs has been demonstrated, using a number of experimental approaches that include oligo library (Lee *et al.* 2002, *Nat. Biotech.*, **20**, 500-505; pp. 502, 503), oligo array (Bohula *et al.* 2003, *J.Biol. Chem.* **278**, 15991-15997), antisense evaluation of accessibility (Far and Sczakiel (2003, *Nucleic Acids Research*, **31**, 4417-4424)), and by targeting the same sequence in both structured and unstructured sites (Vickers *et al.* 2003, *J. Biol. Chem.* **278**, 7108-7188; p.7114, left column, top paragraph)

5. How about published empirical rules for siRNA duplex thermodynamics and features?

- siRNA duplex unwinding is a unique feature in RNAi pathway, and the published findings from Zamore lab, Amgen group, and recently Dharmacon group have important implication for this step. However, these siRNA duplex rules do not guarantee siRNA function. In the publications cited above on target structure and accessibility, a number of potent siRNAs do not meet the key empirical rules, but their function is explained by target accessibility.

It appears to be a consensus view that after the duplex unwinding the antisense strand in the activated RISC needs to bind to the target sequence through complementary base-pairing for target recognition. This would explain the exquisite specificity by siRNAs. When the target sequence is in a heavily structured (i.e., helical) region, the large energy barrier will likely prevent the formation of the hybrid between the antisense siRNA strand and the target sequences. Zamore lab has shown that a single hydrogen bond can decide which of the duplex strands will be incorporated into RISC. Likewise, we believe that secondary structure, accessibility and thermodynamics at the target site are also important.

6. What is Sfold methodology for siRNA design? How is it different from other on-line design tools?

- The design method is based on reported scientific evidence on factors that appear to be important for siRNA function. More specifically, for siRNA screening, Sfold combines target accessibility prediction, siRNA duplex thermodynamic rules as described by Zamore lab and the Amgen group, and the empirical rules reported by the Dharmacon group. Target accessibility evaluation is a unique feature of Sfold and is expected to improve the chance of success. For detailed description of the design tools and design steps using Sfold, please consult the on-line manual.

7. Have the nucleic acid design methods been experimentally validated?

- We have on-going collaborations with “wet” labs at the Wadsworth Center to test our prediction of target accessibility and our design methods for siRNAs, antisense oligos, and hammerhead ribozymes. Preliminary results are highly encouraging. Summary of preliminary testing data can be found in the PDF file from “Validation” page.

8. What else is planned for Sfold?

- More user-friendly tools are under development for each application module.
- By collaborating with experimental labs, we hope to improve the nucleic acid design methods through experimental feedback. We are currently seeking additional collaborative opportunities for testing and improving siRNA design.
- Incorporation of DNA parameters for other applications such as design of molecular beacons

and PCR primers and amplicons.

- Development of a new module *Stools* for folding large number of short RNAs, for designing control oligos (random, scrambled, mismatched, inverted, etc), for illustrating RNA/RNA interactions, and for providing other auxiliary tools.

9. How can users contribute to the improvement of the software, for the benefit of the scientific community?

- By providing experimental feedback.

- By sharing of newly discovered empirical rules, e.g., for the design of siRNAs. Empirical rules can be easily incorporated into the software for outputting relevant information.

♪♪♪ Updated by Ye Ding, April 15, 2004.