Design of RNA sequences with predefined conformational and kinetic properties

Stefan Hammer, Michael T. Wolfinger, Ivo L. Hofacker, Christoph Flamm

Institute for Theoretical Chemistry, University of Vienna, Austria
Department of Biochemistry and Molecular Cell Biology, Max F. Perutz Laboratories, Vienna, Austria
Center for Integrative Bioinformatics, Max F. Perutz Laboratories, Vienna, Austria

Abstract

The ability to accurately design molecules with predefined properties and functions is one of the key technologies to achieve progress in fields such as Syntetic Biology and Nanotechnology. RNA sequences that form bistable switches can be computed quite easily by using simple heuristics for a combinatorial optimization problem (Flamm et al. 2001). This is possible because theoretical results show that for any two secondary structures many sequences exist that are compartible with both. However, for more than more than two structures this is not the case. The design problem becomes even harder if particular kinetic folding proerties are desired.

We want to design RNA sequences that fold into more than two given secondary structures strictly following predefined conformational and kinetic properties. This will be achieved through an iterative process of sequence design, calculation and evaluation of the kinetic properties, and sequence optimization. The folding landscape is characterised using *RNAsubopt* (Hofacker et al. 1994), *Barriers* (Flamm et al. 2002) and *Treekin* (Wolfinger et al. 2004). The evaluation is done via curve sketching and numerical integration. To garantee a fair and efficient sampling in the space of compatible sequences, the design problem has to be rephrased as a graph coloring problem (Abfalter 2005).

Design Pipeline

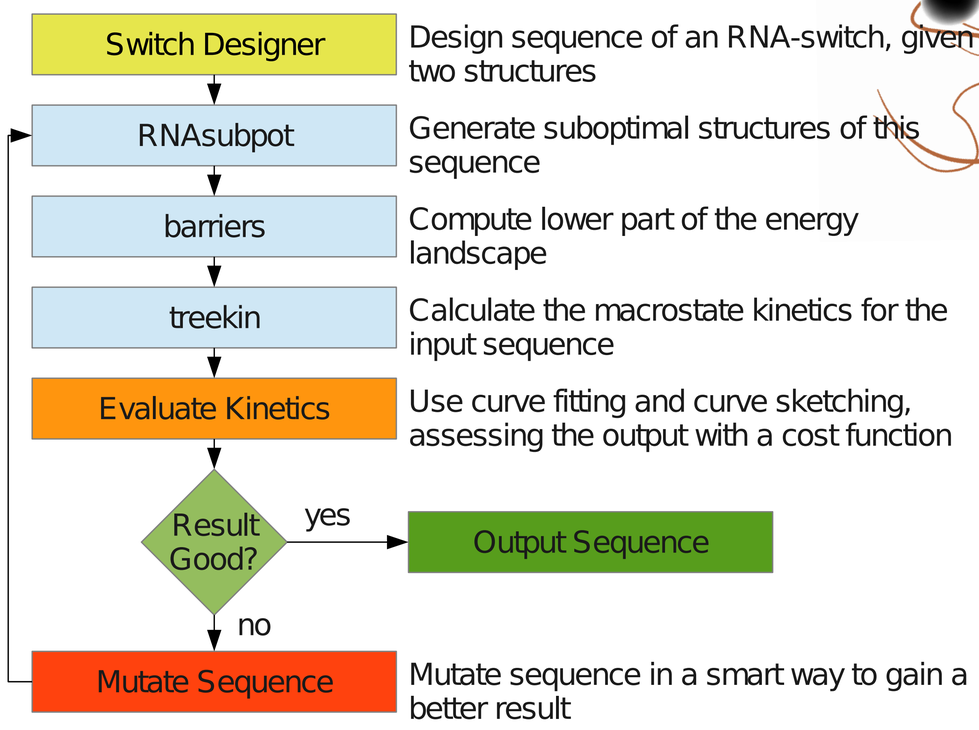

Fig. 3: The design pipline consists of three parts. As a start-sequence we use a bistable RNA molecule and optimise it to become tri- or multistable. The obtained sequence is then evaluated by a cost function after calculating the Energy Landscape and the Kinetics of the system. If the result is good enough it will be exposed to the user, otherwise it will be mutated according to the graph coloring method and resubmitted to the process.

Therefore we use an iterative process consisting of three parts. We use a bistable RNA as a start-sequence and mutate it gradually to get closer and closer to our design goals. To be able to distinguish between better and worse sequences we need to calculate a cost function after evaluating the Energy Landscape and the Kinetics of the system. Accoring to this cost function we decide to output the sequence or to do another round of mutations. The mutations are generated according to the graph coloring method (Abfalter 2005).

INPUT: The input of our newly designed pipeline should be at least three secondary structures in dot-bracket representation and their population probability at a certain relative timepoint.
OUTPUT: The output will be one or more RNA sequences compartible with all given structures which fulfill the design goals as close as possible.

Design of multistable switches

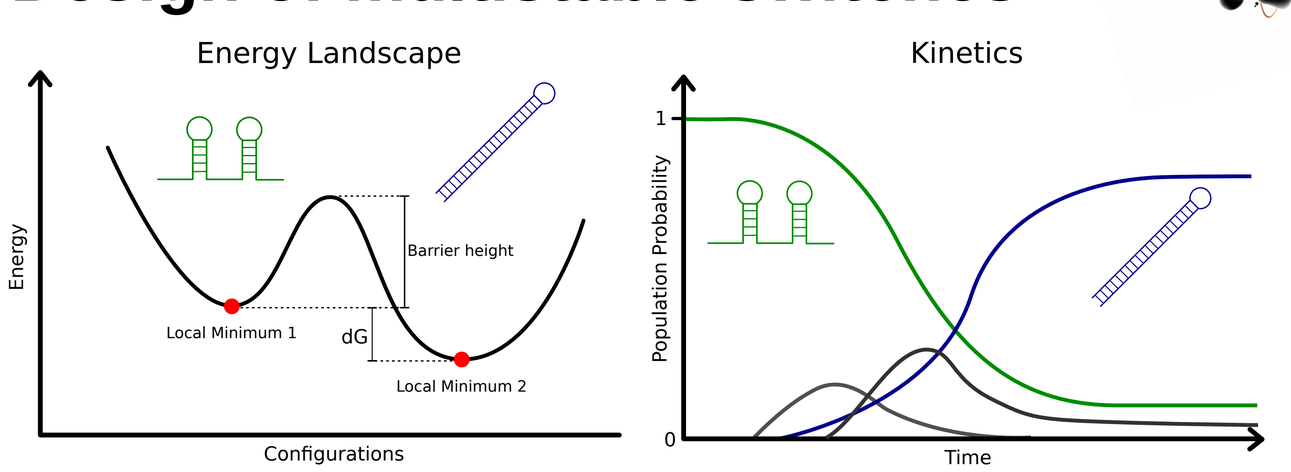

Fig. 1: Schematic representation of the Energy Landscape and the transition between the two states of a bistable RNA switch over time. Notice that other metastable conformations are appearing during the transition.

Flamm et al. (2001) already showed that bistable RNA switches can be easily designed by using simple heuristics (Switch Designer). This is possible because a sequence is always compartible with two secondary structures. Our goal is to design an RNA switch that folds into more than two conformations and is optimised on the level of the Energy Landscape.

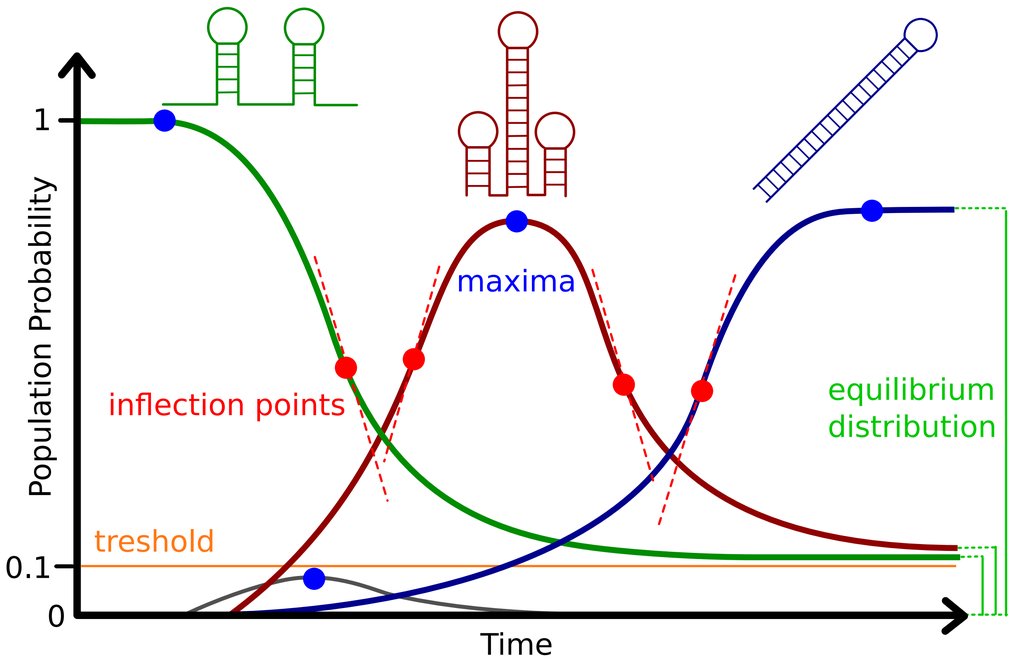

Fig. 2: The kinetics of the system need to be calculated during the iterative process of optimising the sequence. The obtained curves are then evaluated using curve sketing and numerical integration.

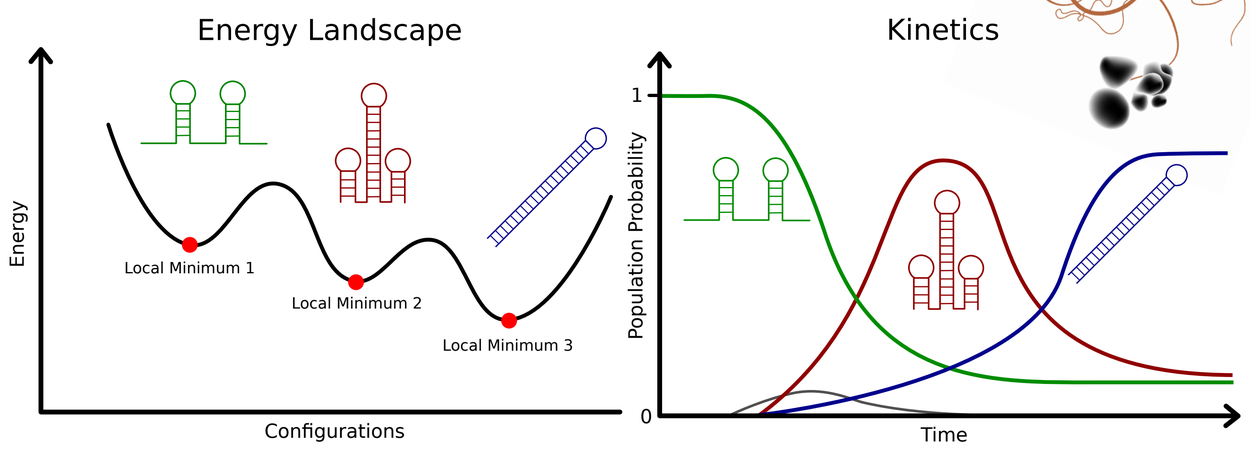

Fig. 4: The resulting sequence folds into more than two given secondary structures and strictly follows predefined conformational and kinetic properties. Upcoming metastable structures should be avoided while the RNA molecule folds into the given conformatins during time.

References

1. Flamm, C., Hofacker, I. L., Maurer-Stroh, S., Stadler, P. F. & Zehl, M. Design of multistable RNA molecules. RNA 7, 254–265 (2001).
2. Wolfinger, M. T., Svrcek-Seiler, W. A., Flamm, C., Hofacker, I. L. & Stadler, P. F. Efficient computation of RNA folding dynamics. J. Phys. Math. Gen. 37, 4731–4741 (2004).
3. Hofacker, I. L. et al. Fast folding and comparison of RNA secondary structures. Monatshefte Für Chem. Chem. Mon. 125, 167–188 (1994).
4. Flamm, C., Hofacker, I. L., Stadler, P. F. & Wolfinger, M. T. Barrier Trees of Degenerate Landscapes. Z. Für Phys. Chem. 216, 155 (2002).
5. Abfalter, I. G. Nucleic Acid Sequence Design As A Graph Colouring Problem. (2005).

FWF Der Wissenschaftsfonds.