

pourRNA - a time- and memory-efficient approach for the guided exploration of RNA energy landscapes - Supplementary Material -

Gregor Entzian¹ and Martin Raden²

¹ University of Vienna, Faculty of Chemistry, Department of Theoretical Chemistry, Währingerstraße 17,
1090 Vienna, Austria

² Bioinformatics Group, University of Freiburg, Georges-Köhler-Allee 106, D-79110 Freiburg, Germany

Contents

A Formal preliminaries with mathematical details	2
B Why is $\max E = 5 \frac{\text{kcal}}{\text{mol}}$ sufficient?	3
C RNA sequences and meta information	4
D Time and memory consumption with multiprocessing	5
E Details to ΔE effects	6
F Exact and approximate kinetics of SL	8
G Results for sv11	10
G.1 Initial structure via co-transcriptional folding	10
G.2 Energy barriers and respective paths (turner-99 parameter)	11
G.3 Energy barriers and respective paths (turner-04 parameter)	13
H Folding Funnels	15

A Formal preliminaries with mathematical details

Within the following, we are investigating the RNA energy landscapes representing non-crossing secondary structures P and single base pair distance. Therein, the structure space $\mathcal{P}$ of an RNA sequence $S \in \{\mathbf{A}, \mathbf{C}, \mathbf{G}, \mathbf{U}\}^n$ of length n comprises all nested structures $P \in [1, n] \times [1, n]^*$ (encoding a set of base pairs (i, j)) for that holds that any position is paired at most once (i.e. $\forall_{(k,l) \neq (r,s) \in P : |\{k, l, r, s\}| = 4}$), base pair indices are ordered (i.e. $\forall_{(i,j) \in P : i < j}$), and no two base pairs are crossing (i.e. $\forall_{(k,l) \in P} \nexists_{(r,s) \in P : k < r < l < s \text{ or } r < k < s < l}$).

If not stated differently, the neighborhood $N(P) \subseteq \mathcal{P}$ of a structure P within the energy landscape is the set of all structures P' that differ in exactly one base pair (i.e. $P' \in N(P) \leftrightarrow |(P \cup P') \setminus (P \cap P')| = 1$). For comparison with other studies, we partially consider 'shift move' neighborhoods, which in addition consider structures as neighbored that can be transformed from one into the other by altering one end of a single base pair. See Fig. A.1 for an example.

The (free) energy $E(P)$ of a structure is determined by the Nearest Neighbor Model (Tinoco *et al.*, 1973) using the parameters from Mathews *et al.* (2004). All energies are given in $\frac{\text{kcal}}{\text{mol}}$ units. A structure P' is considered 'energetically smaller' than P (denoted by $P' \prec P$) if its energy is either lower ($E(P') < E(P)$) or equal but its structure dot-bracket string encoding is lexicographically smaller, which is needed due to the degeneracy of the energy model (see Fig. A.1 and Flamm *et al.* (2002) for further details). A structure $\check{P}$ that is energetically smaller than all its neighbors (i.e. $\forall_{P' \in N(\check{P}) : \check{P} \prec P')$ is called a local minimum.

The triple $(\mathcal{P}, N, E)$ defines the energy landscape studied.

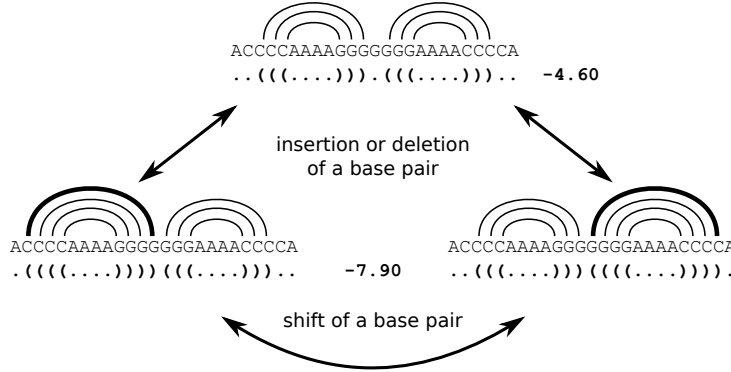


Figure A.1: Visualization of different neighborhood definitions via insertion or deletion of a base pair (aka indel-moves, top path) or the 'shifting' of one base pair end (aka shift-moves, bottom path) for three structures. Each structure is depicted as linear Feynman diagram above the sequence and a respective dot-bracket string below along with the respective energy in $\frac{\text{kcal}}{\text{mol}}$. Furthermore, the example exemplifies the degeneracy of the energy model, since both structures on the bottom have equal energy ($-7.9 \frac{\text{kcal}}{\text{mol}}$) when using the **turner-04** without dangling end contributions. When using lexicographical order of the dot-bracket string representations, the left structure would define a local minimum for both neighborhood definitions while both bottom structures are minima within the insertion-deletion-only neighborhood model for the depicted structure space.

Each gradient basin is associated with a single local minimum structure $\check{P}$ of the energy landscape. The gradient basin $B(\check{P}) \subseteq \mathcal{P}$ is defined recursively and contains all structures whose gradient neighbor is within B . The gradient neighbor $g(P) \in N(P)$ of some structure P (if existent) is the smallest among all energetically smaller neighbors (i.e. it holds $g(P) \prec P$ and $\forall_{P' \in N(P) : g(P) \prec P'}$). Therefore, any local minimum $\check{P}$ does not have a gradient neighbor and is thus the minimal energy structure of a basin. The set of basins $\mathcal{B}$ comprises then a partitioning of $\mathcal{P}$ and we denote with $b \in \mathcal{B}$ a basin for some local minimum. Two basins $b \neq b'$ are neighbored if two of their respective structures are neighbored (i.e. $b \in N(b') \leftrightarrow \exists_{P \in b} :$

$\exists_{P' \in b'} : P \in N(P)$). The energy of a basin is given by its ensemble energy $E(b) = -RT \log(Z_b)$, where Z_b denotes its partition function given by $\sum_{P \in b} w(P)$, i.e. the sum of Boltzmann weights $w(P) = \exp(-E(P)/RT)$ for temperature T and gas constant R .

B Why is $\max E = 5 \frac{\text{kcal}}{\text{mol}}$ sufficient?

For short sequences still accessible for global flooding schemes ($< 100nt$), a bound of $5 \frac{\text{kcal}}{\text{mol}}$ is expected to be sufficient to ensure that all important low energy macro-states are connected. This simply results from the fact, that any two structures P, P' can be connected by first removing step-by-step all exclusive base pairs $P \setminus (P \cap P')$ followed by successively adding of all missing base pairs $P' \setminus (P \cap P')$. In the worst case, all base pairs have to be removed (resulting in the unstructured state with energy $= 0$), while single non-stacked base pairs (along the path) will result in slightly positive values. Since $5 \frac{\text{kcal}}{\text{mol}}$ are above the penalty to introduce a single or a canonical stacking of two base pairs, this ensures connectivity of macro-states with local minimum energy below 0.

C RNA sequences and meta information

Table 1: List of RNA sequences

Name	Sequence	Length	Source
xbix	CUGCGGCUUUGGCUUAGCC	20	(Wolfinger <i>et al.</i> , 2004)
boris1	GACCGGAAGUCCGCCUUC	20	(Fürtig <i>et al.</i> , 2007)
boris2	GAAGGCAACCUUCGGGUUG	20	(Fürtig <i>et al.</i> , 2007)
d25	UCCACGGCUGUAGUGGAUACGGC	25	(Flamm and Hofacker, 2008)
d33	GGGAAUUAUUGUUCUCCUGAGCGGUAGUUCUC	33	(Mann and Klemm, 2011)
ire	CUGUCUCUUGCUUCAACAGUGUUGGACGGAACAG	35	(Mann <i>et al.</i> , 2014)
bhg33	GUGUCGUUUGGAUUAAGGACCUACAACAGGCU	33	(Kucharik <i>et al.</i> , 2014)
d45	GGGCGGUUGGCCUCCGCUAAAUGCGGAAGAUAAUUGUGUCU	45	(Lorenz <i>et al.</i> , 2009)
SL (stem 1)	AACUAAACAUUUUUGAAGACAGUUUCUGUACUUCAUUGGUAUGUAGAGACUUC	56	(Harris <i>et al.</i> , 1995)
sv11	GGGCACCCCUUCGGGGGUCACCUCCGGUAGCUACGCGAGGGUUAAGCGCCUUUCUCCUCCGUGAGCUAACACGCG- GAGUGACCCCGAAAGGGGGUUUCCCA (((((((((((((.....))-))- meta-stable conformation	115	(Biebricher and Luce, 1992)

D Time and memory consumption with multiprocessing

Thread-number-dependent time and memory consumption of *pourRNA*. Figure D.2 show results for sequences **d33** and **ire**, resp., analogously to Fig. 2a from the main manuscript. However, the speed-up here is some times not as good with more threads due to extremely short run-times. Thus, the multi-threading overhead etc. dominates time consumption and provides a less significant signal of the multi-threading potential.

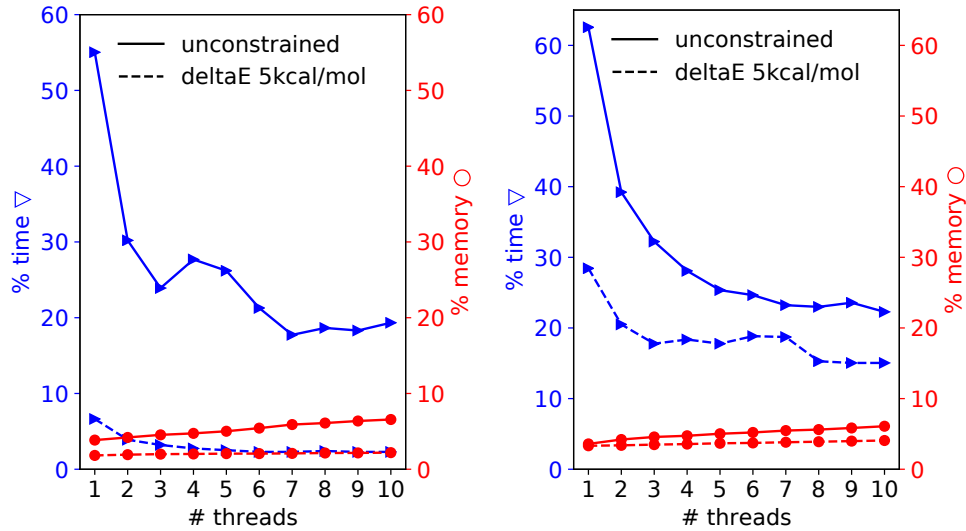


Figure D.2: Time and memory consumption of *pourRNA* in percentage of the *barriers* tool for sequences (left) **d33** and (right) **ire**.

E Details to deltaE effects

The following figures summarize the effects of different deltaE thresholds on the energy landscapes explored by *pourRNA* (1 core, $\text{maxE}=5 \frac{\text{kcal}}{\text{mol}}$). Relative values refer to the unconstrained *pourRNA* results from the main manuscript. The ensemble energy is given by the overall partition function Z via $-RT \log(Z)$. Runtime and memory plots exclude very short sequences due to since respective values are not governed by the sequence features.

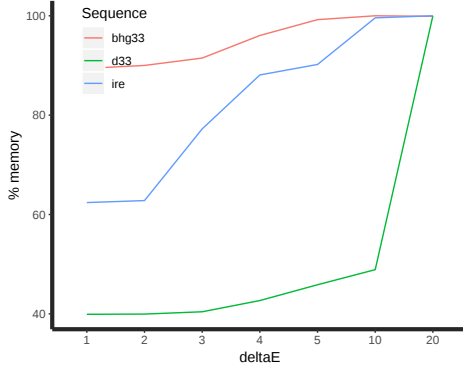


Figure E.3: Memory consumption

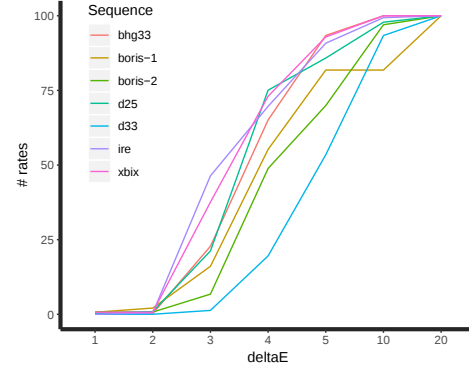


Figure E.6: Rate number reduction

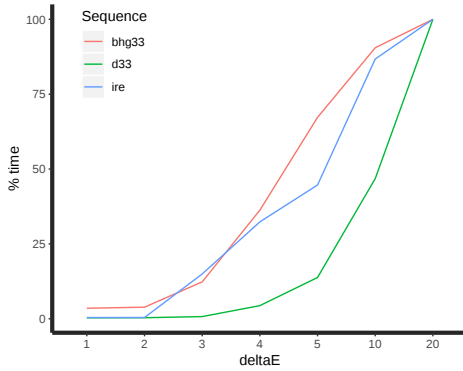


Figure E.4: Time consumption

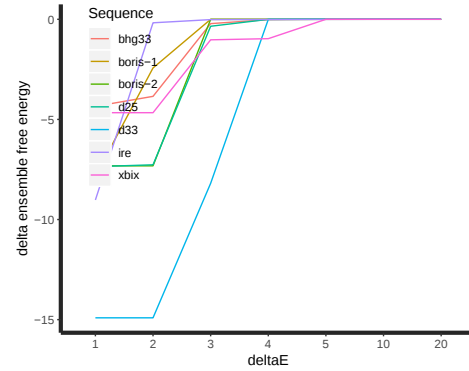


Figure E.7: Ensemble energy reduction

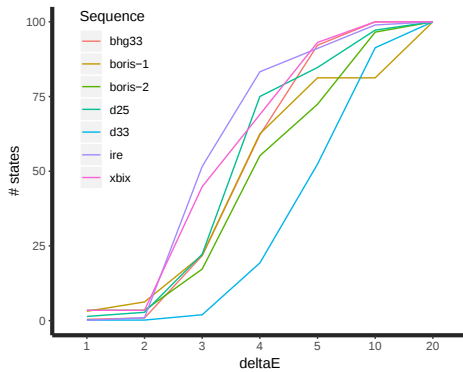


Figure E.5: Macro-state reduction

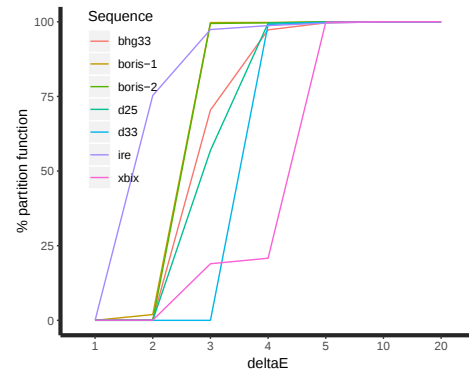


Figure E.8: Partition function reduction

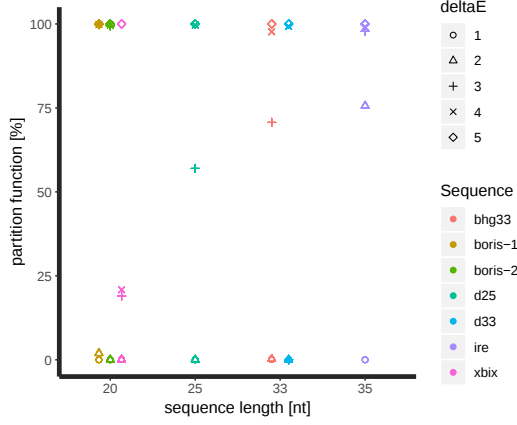


Figure E.9: Partition function reduction and deltaE thresholds for different sequence lengths. There are 3 sequences with 20 nt and 2 sequences with 33 nt. The symbols for these sequences are shifted on the x-axis.

In figure E.9 the effect of the deltaE filter is shown for the sequence lengths 20 nt to 35 nt for different sequences and thresholds. The $\text{deltaE}=3 \frac{\text{kcal}}{\text{mol}}$ threshold results in 100 % partition function for two 20 nt long sequences but for one 20 nt long sequence it is only 24 percent. 100 percent of the partition function is only reached with $\text{deltaE}=5 \frac{\text{kcal}}{\text{mol}}$ for the third 20 nt long sequence. For the longer sequences the coverage of the partition function increases for $\text{deltaE}=3 \frac{\text{kcal}}{\text{mol}}$. However, even for the 35 nt long sequence the threshold has to be adjusted to $\text{deltaE}=4 \frac{\text{kcal}}{\text{mol}}$ or $\text{deltaE}=5 \frac{\text{kcal}}{\text{mol}}$ in order to cover ~ 100 % of the partition function.

Minimal macroscopic rate

The deltaE threshold indirectly defines a lower bound on macroscopic rates considered within the energy landscape. Given that the rate $k_{b \rightarrow b'}$ from basin b to basin b' is defined by $Z_{b,b'}/Z(b)$, we can estimate the minimal rate via the minimal transition state partition function $Z_{b,b'}$. The latter is minimal, if it only covers a single transition state within the deltaE range, i.e. $\bar{Z}_{b,b'} = \exp(-(E_{\text{min}} + \text{deltaE})/RT)$, where E_{min} denotes the energy of the local minimum of b . A lower bound on the basin's partition function $Z(b)$ covers only the respective local minimum, i.e. $\bar{Z}(b) = \exp(-E_{\text{min}}/RT)$. Thus, we get the minimal macroscopic rate $\bar{k}_{b \rightarrow b'} = \bar{Z}_{b,b'}/\bar{Z}(b) = \exp(-(E_{\text{min}} + \text{deltaE})/RT)/\exp(-E_{\text{min}}/RT) = \exp(-\text{deltaE}/RT)$.

Thus, e.g. a deltaE threshold of $5 \frac{\text{kcal}}{\text{mol}}$ corresponds to a minimal macroscopic rate of about 0.0003.

F Exact and approximate kinetics of SL

In the following, we compare the kinetics in terms of population density trajectories computed by *treekin* (Wolfinger *et al.*, 2004) when using exact macroscopic transition rates (Fig. 10(a)) versus approximate rates using a $\text{deltaE}=5 \frac{\text{kcal}}{\text{mol}}$ heuristic (Fig. 10(b)). The plots represent only the ten most prominent basins with their respective local minima (i.e. the ten trajectories with the highest maximum population density). Both computations are using **turner-04** energy parameters from the Vienna RNA package and are based on single indel moves only. The flooding was globally restricted by $\text{maxE}=5 \frac{\text{kcal}}{\text{mol}}$ and the kinetics are reported for the connected component containing the minimum free energy structure.

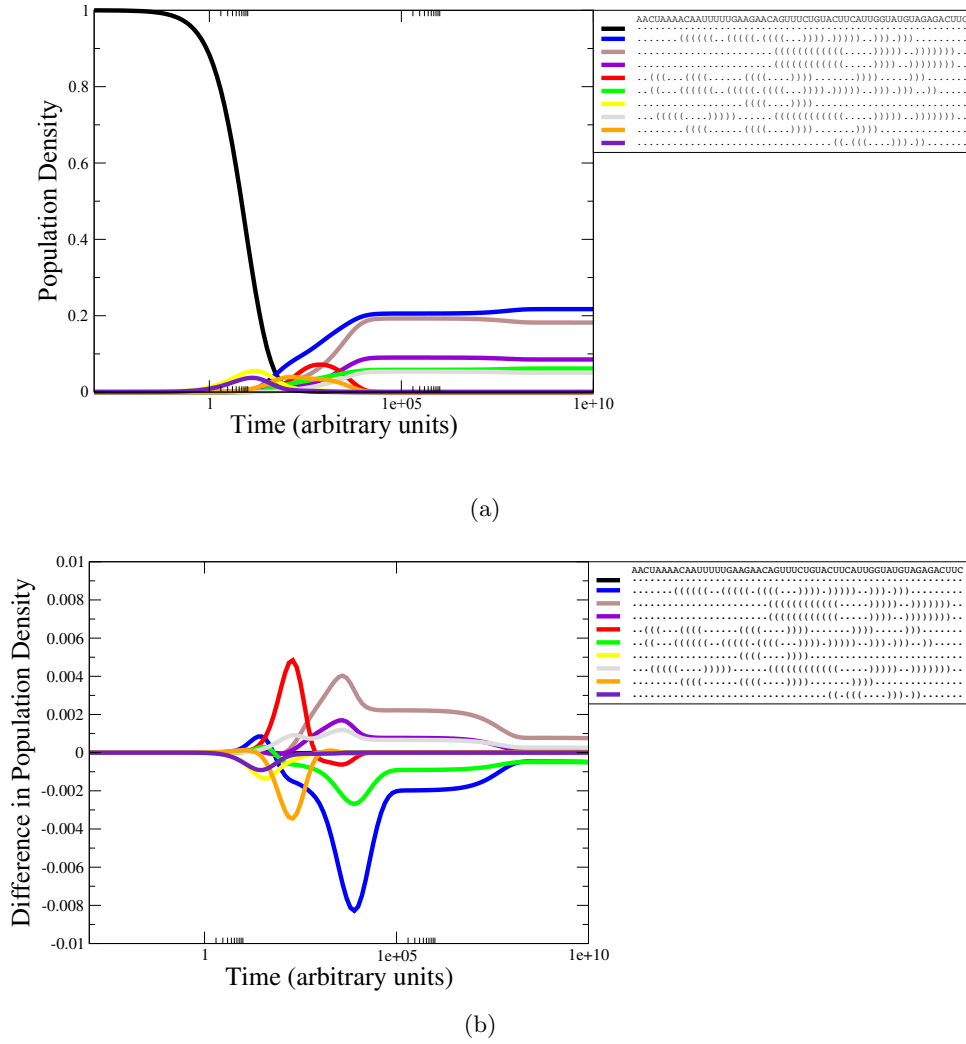


Figure F.10: (a) Folding kinetics of SL when using exact macroscopic transition rates. (b) Deviation of the kinetics with approximate rates resulting from restricted local flooding with $\text{deltaE}=5 \frac{\text{kcal}}{\text{mol}}$ from the exact population densities (PD) in (a), i.e. difference = $\text{PD}(\text{exact}) - \text{PD}(\text{approx})$, which shows underestimation with values above 0.

For comparison, we also computed macroscopic rates via the *Basin Hopping Graph (BHG)* framework. For comparability, we used a maximum energy threshold of $5 \frac{\text{kcal}}{\text{mol}}$ and **turner-04** energy parameters. First, we sampled 1000 structures using the tool *asearch_2.py* (part of the *BHG* framework), which produces Boltzmann samples at different temperatures and maps the structures to local minima. Subsequently, these structures were input to BHG, which connects them with a direct path heuristic. The structures of the direct paths are also mapped to local minima and the procedure is iterated. This took 56 minutes with a peak memory of 43 MB.

Figure F.11 shows that the resulting kinetics deviate much more from the exact kinetics (compare y axis and deviations with Fig. F.10).

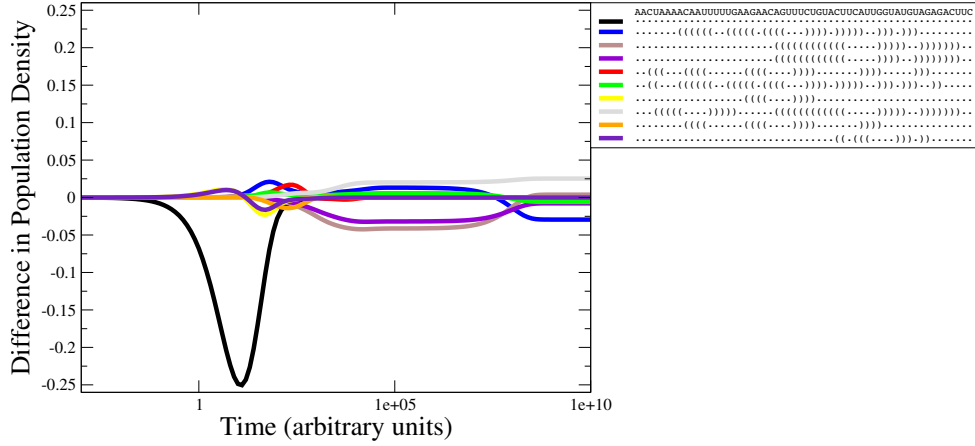


Figure F.11: Difference of the macroscopic folding kinetics of SL when using the tool *Basin Hopping Graph* (with 1000 samples of the tool *asearch*, $5 \frac{\text{kcal}}{\text{mol}}$ global threshold) from the exact population densities.

G Results for sv11

G.1 Initial structure via co-transcriptional folding

In order to show a valid experimental pipeline, we "rediscovered" the meta-stable conformation of **sv11** (see Sec. C). To this end, we ran the co-transcriptional folding simulation webserver KineFold (Xayaphoummine *et al.*, 2005) with the following parameterization:

- co-transcriptional fold: A new base is added every 3 milliseconds (default)
- Simulated molecular time: Suggested (default)
- Pseudoknots: Not allowed
- Entanglements: Non crossing
- Random seed: 58635

The resulting minimum free energy structure is visualized in Fig. G.12 and resembles the structure known from literature (see (Biebricher and Luce, 1992)), visualized with FORNA (Kerpedjiev *et al.*, 2015).

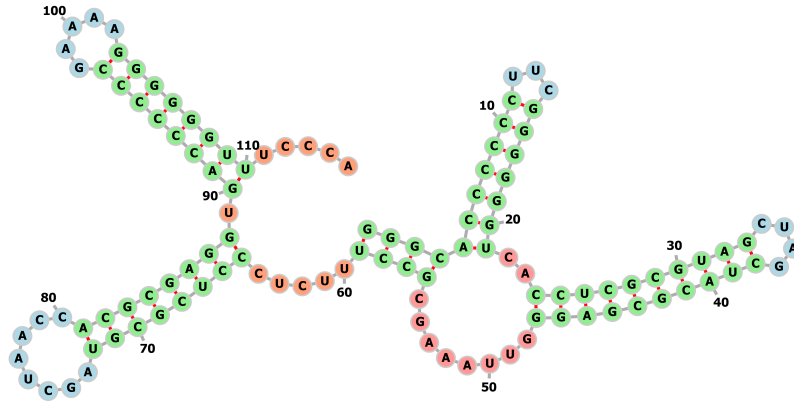


Figure G.12: Structure of **sv11** predicted using co-transcriptional folding by KineFold (rank 1). This is exactly the meta-stable structure of **sv11** taken from (Biebricher and Luce, 1992).

G.2 Energy barriers and respective paths (turner-99 parameter)

Investigation of the refolding pathways

Figure G.13 depicts the energy profile of the refolding pathways identified by *pourRNA* and *BHG*. For *pourRNA*, only the energies of the basins' local minima as well as the traversed local barrier structures are plotted (see Listing 1). Distance on the X-axis corresponds to respective base pair distance and thus microscopic transitions allowing single base pair insertions and deletions.

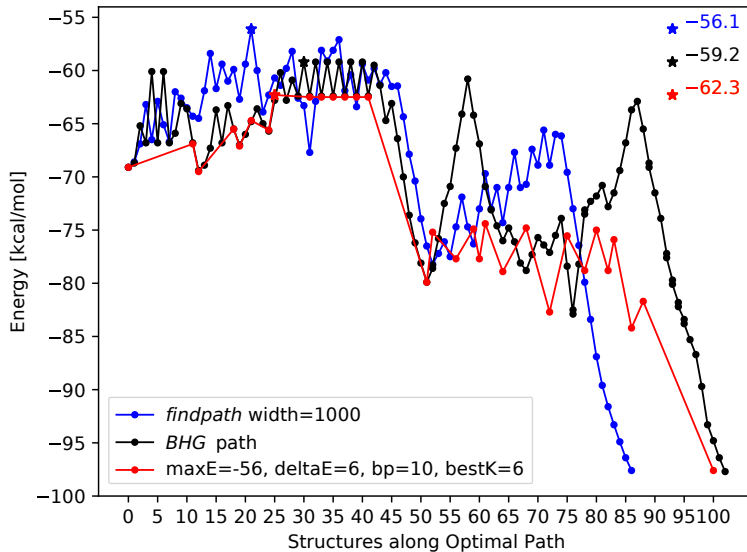


Figure G.13: Energies of minima representatives and saddles along the optimal refolding path of **sv11**'s meta-stable structure to the global optimum (**turner-99** energy parameters). The optimal path is a path with the minimal maximal saddle from all explored paths between the metastable state and the ground state. The red profile with energy barrier $-62.3 \frac{\text{kcal}}{\text{mol}}$ was found via *pourRNA* (using shift moves) while the black profile with barrier energy $-59.2 \frac{\text{kcal}}{\text{mol}}$ was extracted from (Kucharik *et al.*, 2014). Blue provides the path reported by the *findpath* heuristic with a barrier estimate of $-56.1 \frac{\text{kcal}}{\text{mol}}$.

We can see that *pourRNA* rediscovered parts of the path identified by *BHG* since we find the same sequence of local minima forming a plateau for step 25 to 40. This sequence corresponds to a 'melting' of a temporary multi-loop into a single helix stem by 'shifting' base pairs from side loops into the main stem. Due to the degeneracy of the energy model, these structures show the same energy and thus array in a sequence of local minima due to the lexicographical order used to break the degeneracy problem. Notably, *BHG* shows interspacing peaks that are not present in the *pourRNA* profile. These result from the mixed microscopic transition model used in *BHG*: while its definition of local minima (and thus gradient walks and basins) is considering shift moves, *BHG* uses *findpath* to identify local energy barriers between minima, which does *not* consider shift moves. In contrast, here *pourRNA* generally considers shift moves, which renders the local minima correctly as direct neighbors (without a local energy barrier) within the landscape. See Fig. A.1 for a depiction of the missing energy barrier if shift moves are applied.

G.3 Energy barriers and respective paths (turner-04 parameter)

When comparing the *findpath* results using **turner-04** parameters with the results when using **turner-99** parameters, we observe an increased energy barrier estimate for microscopic direct paths to $-55.5 \frac{\text{kcal}}{\text{mol}}$ (compare Tab. 2 and Fig. G.13). This estimate is improved by *pourRNA* to $-58.5 \frac{\text{kcal}}{\text{mol}}$ within 15 minutes using the macroscopic path heuristics.

method (parameters)	$E(\hat{P})$	time	$ \mathcal{B}_p $
<i>findpath</i> (width=1000)	-55.5	<1 m	
<i>pourRNA</i> (deltaE=5, Δ_{bp} =6)	-56.6	3.5 m	7,135
<i>pourRNA</i> (deltaE=5, Δ_{bp} =10)	-57.5	14 m	22,496
<i>pourRNA</i> (deltaE=5)	-57.5	1.9 h	77,480
<i>pourRNA</i> (deltaE=6, Δ_{bp} =5)	-58.5	15 m	9,165
<i>pourRNA</i> (deltaE=8, Δ_{bp} =5)	-58.5	1.74 h	16,226
<i>pourRNA</i> (deltaE=10, Δ_{bp} =0)	-58.5	1.5 h	3,316

Table 2: Barrier estimations (energy of lowest transition states $\hat{P}$ in $\frac{\text{kcal}}{\text{mol}}$ using the **turner-04** energy parameters from Vienna RNA package 2.4.10) and runtime for **sv11** refolding using microscopic direct path search (*findpath*) and *pourRNA*. For *pourRNA* also the number of processed macro-states $b \in \mathcal{B}_p$ is given along with the used threshold for local flooding (deltaE) and macroscopic path exploration (Δ_{bp}).

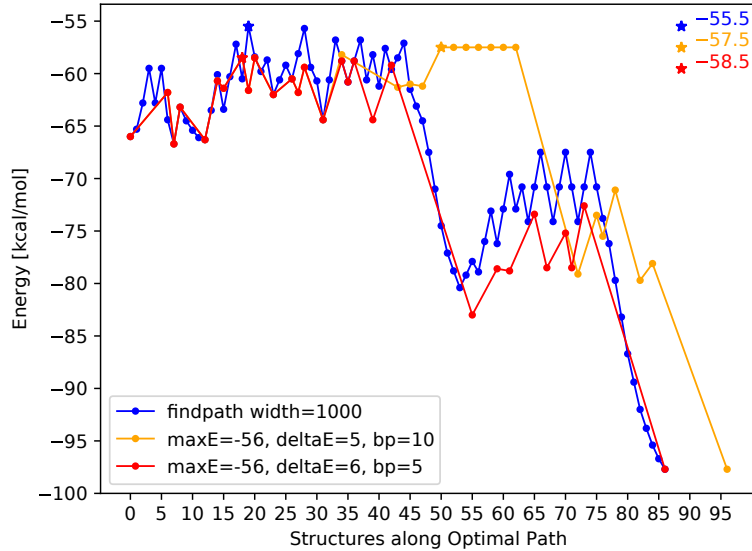


Figure G.14: Energies of minima representatives and saddles along the optimal path (**turner-04** energy parameters). The optimal path is the path with the minimal maximal saddle from all explored paths between the metastable state and the ground state. The RNA **sv11** was used as input of *pourRNA* (with $\text{maxE}=-56 \frac{\text{kcal}}{\text{mol}}$, $\text{deltaE}=6 \frac{\text{kcal}}{\text{mol}}$, $\Delta_{bp}=5$ bp deviation along the direct path between metastable and ground state while using shift moves). This leads to 9,165 local minima, needs 15 m to compute and the path has an energy barrier of $-58.5 \frac{\text{kcal}}{\text{mol}}$ (red curve). This is a better approximation than the *findpath* heuristic with $-55.5 \frac{\text{kcal}}{\text{mol}}$.

H Folding Funnels

Klemm *et al.* (2008) introduced a notion of folding funnels that is based on an iterative recursive aggregation of gradient basins rephrased in the following:

1. The ground state is contained in the funnel F .
2. The local minimum x belongs to the funnel F if a minimum saddle s_x connects x directly to a local minimum in the funnel F .
3. A state z belongs to the funnel if it is connected by a gradient descent path to a local minimum in F .

Given this funnel definition, all states from the funnel are extracted from the landscape and the procedure is iterated until all states are assigned to respective funnels.

Since the definition by Klemm *et al.* is based on the minimum saddle, i.e. the energetically lowest transition state that has lower neighbors in both gradient basins, we provide an adapted definition for basins based on maximal rates in the following:

1. The ground state is contained in the funnel F and defines its center.
2. The gradient basin b belongs to the funnel F if its maximal exit rate $k_{b \rightarrow b'}$ connects b directly to a basin b' with energetically smaller local minimum in the funnel F .
3. A state x belongs to the funnel if it is part of a gradient basin in F .

Note, we also introduced the cascading idea of funnels by Frauenfelder and Leeson (1998) or Kühnl *et al.* (2017), i.e. the assignment considers only basins with lower local minimum, which provides a unique partitioning of the gradient basins into a set of funnels centered at their respective lowest local minima.

This funnel decomposition is depicted in Fig. H.15 for the small artificial xbix RNA also used by Klemm *et al.* (2008). The example considers (for comparability with the previous work) the **turner-99** energy parameters and a micro-state neighborhood based on insertion/deletion of base pairs. Since the previous results were produced using Vienna RNA package v1.8.* (implementing a slightly different energy model compared to recent versions) and Klemm *et al.* (2008) provide no implementation, results are not exactly reproducible. We observe three large funnels centered at the global minimum (1), the meta-stable state (2) and the 4th local minimum (4), which is in accordance with the findings of Klemm *et al.* and agrees with the respective kinetics reported by Wolfinger *et al.* (2004).

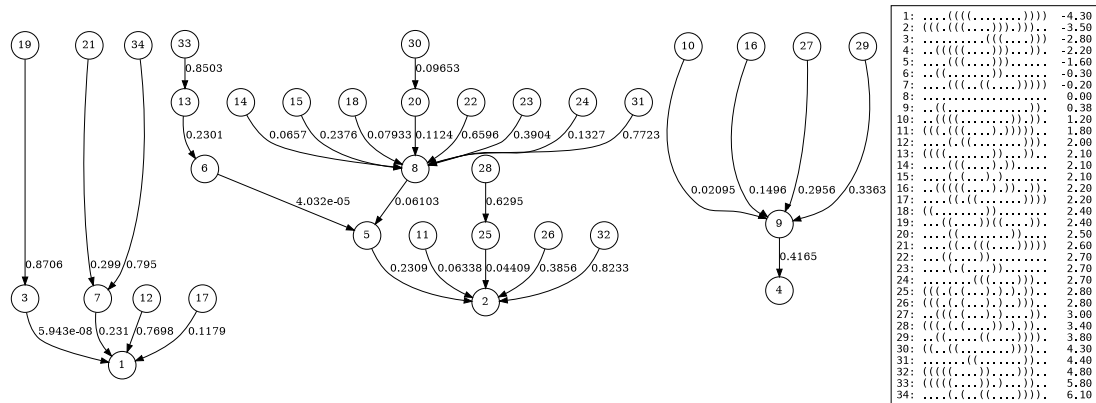


Figure H.15: Folding Funnels in the energy landscape of the xbix RNA with sequence CUGCGGCUUUGGCUCUAGCC (turner-99 energy parameters, indel moves).

pourRNA can be used to directly compute the respective funnel center for any state of the landscape when using appropriate filters. To this end, the state is first mapped to the local minimum of the respective gradient basin (step 3 of the funnel definition). Subsequently, the basin is flooded to identify the macroscopic transition rates to leave the basin. These are reduced to neighbored basins with lower energy (neigh-max-E filter = 0) and the highest among these is picked (step 2 of the funnel definition). This produces the 'macroscopic gradient step' and provides a neighbored basin for which the procedure is iterated until no neighbor with lower energy is found. The local minimum of this basin, with lowest energy among all processed states so far, marks the funnel center (step 1 of the funnel definition) and thus the end of the macroscopic gradient walk. The Listing 3 exemplifies this procedure and provides the respective funnel centers (from the *pourRNA* output) identified for each local minimum of the landscape, which resembles the assignment from Fig. H.15.

Listing 3: Funnel center mapping by *pourRNA* (turner-99 parameters, move-set=0 (indel moves), max-neigh-e=0, filter-best-k=1).

```

1 ....((((.....)))) --> 1 ....((((.....))))
2 (((((((.....)))))).. --> 2 (((((((.....))))))..
3 .....((((.....)))) --> 1 .....((((.....))))
4 ..((((.....)))).... --> 4 ..((((.....))))....
5 ....((((.....)))).... --> 2 ....((((.....))))....
6 ..((((.....)))).... --> 2 ..((((.....))))....
7 ....((((.....)))) --> 1 ....((((.....))))
8 .....((((.....)))) --> 2 .....((((.....))))..
9 ..((((.....)))).... --> 4 ..((((.....))))....
10 ..((((.....)))).... --> 4 ..((((.....))))....
11 (((((((.....)))))).. --> 2 (((((((.....))))))..
12 ....((((.....)))) --> 1 ....((((.....))))
13 (((((((.....)))))).. --> 2 (((((((.....))))))..
14 ....((((.....)))) --> 2 ....((((.....))))..
15 .....((((.....)))) --> 2 .....((((.....))))..
16 ..((((.....)))).... --> 4 ..((((.....))))....
17 ....((((.....)))) --> 1 ....((((.....))))
18 (((((((.....)))))).. --> 2 (((((((.....))))))..
19 ....((((.....)))) --> 1 ....((((.....))))
20 .....((((.....)))) --> 2 .....((((.....))))..
21 ....((((.....)))) --> 1 ....((((.....))))
22 ...((((.....)))).... --> 2 ...((((.....))))....
23 .....((((.....)))) --> 2 .....((((.....))))..
24 ....((((.....)))) --> 2 ....((((.....))))..
25 (((((((.....)))))).. --> 2 (((((((.....))))))..
26 ....((((.....)))) --> 2 ....((((.....))))..
27 ..((((.....)))).... --> 4 ..((((.....))))....
28 (((((((.....)))))).. --> 2 (((((((.....))))))..
29 ..((((.....)))).... --> 4 ..((((.....))))....
30 ....((((.....)))) --> 2 ....((((.....))))..
31 .....((((.....)))) --> 2 .....((((.....))))..
32 (((((((.....)))))).. --> 2 (((((((.....))))))..
33 ....((((.....)))) --> 2 ....((((.....))))..
34 .....((((.....)))) --> 1 .....((((.....))))

```

References

- Biebricher, C. and Luce, R. (1992). In vitro recombination and terminal elongation of RNA by Q beta replicase. *The EMBO Journal*, **11**(13), 5129–5135.
- Flamm, C. and Hofacker, I. L. (2008). Beyond energy minimization: approaches to the kinetic folding of RNA. *Chemical Monthly*, **139**(4), 447–457.
- Flamm, C., Hofacker, I. L., Stadler, P. F., and Wolfinger, M. T. (2002). Barrier trees of degenerate landscapes. *Z.Phys.Chem*, **216**, 155–173.
- Frauenfelder, H. and Leeson, D. T. (1998). The energy landscape in non-biological and biological molecules. *Nature*, **5**, 757–759.
- Fürtig, B., Buck, J., Manoharan, V., Bermel, W., Jäschke, A., Wenter, P., Pitsch, S., and Schwalbe, H. (2007). Time-resolved NMR studies of RNA folding. *Biopolymers*, **86**(56), 360–383.
- Harris, K. A., Crothers, D. M., and Ullu, E. (1995). In vivo structural analysis of spliced leader RNAs in *Trypanosoma brucei* and *Leptomonas collosoma*: a flexible structure that is independent of cap4 methylations. *RNA*, **1**(4), 351–62.
- Kerpedjiev, P., Hammer, S., and Hofacker, I. L. (2015). Forna (force-directed rna): Simple and effective online RNA secondary structure diagrams. *Bioinformatics*, **31**(20), 3377–3379.
- Klemm, K., Flamm, C., and Stadler, P. F. (2008). Funnels in energy landscapes. *The European Physical Journal B*, **63**(3), 387–391.
- Kucharík, M., Hofacker, I. L., Stadler, P. F., and Qin, J. (2014). Basin Hopping Graph: a computational framework to characterize RNA folding landscapes. *Bioinformatics*, **30**(14), 2009–2017.
- Kühnl, F., Stadler, P. F., and Will, S. (2017). Tractable RNA–ligand interaction kinetics. *BMC Bioinformatics*, **18**(12), 424.
- Lorenz, R., Flamm, C., and Hofacker, I. L. (2009). 2D projections of RNA folding landscapes. In *In: German Conference on Bioinformatics 2009*, volume 157 of *Lecture Notes in Informatics*, pages 11–20.
- Mann, M. and Klemm, K. (2011). Efficient exploration of discrete energy landscapes. *Phys. Rev. E*, **83**(1), 011113.
- Mann, M., Kucharik, M., Flamm, C., and Wolfinger, M. T. (2014). Memory efficient RNA energy landscape exploration. *Bioinformatics*, **30**(18), 2584–2591.
- Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004). Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proceedings of the National Academy of Sciences*, **101**(19), 7287–7292.
- Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology*, **246**(150), 40–41.
- Wolfinger, M. T., Svrcek-Seiler, W. A., Flamm, C., Hofacker, I. L., and Stadler, P. F. (2004). Efficient computation of RNA folding dynamics. *Journal of Physics A: Mathematical and General*, **37**(17), 4731–4741.
- Xayaphoummine, A., Bucher, T., and Isambert, H. (2005). Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Research*, **33**(suppl.2), W605–W610.