Shape-based barrier estimation for RNAs

Sergiy Bogomolov[1,◇], Martin Mann[2,◇], Björn Voß[3], Andreas Podelski[1] and Rolf Backofen[2]
[1] *Software Engineering,* [2] *Bioinformatics,* [3] *Genetics and Experimental Bioinformatics at Albert-Ludwigs-Universität Freiburg, Germany*
◇ These authors contributed equally to this work
{bogom,mmann}@informatik.uni-freiburg.de

Abstract:
The ability of some RNA molecules to switch between different metastable conformations plays an important role in cellular processes. In order to identify such molecules and to predict their conformational changes one has to investigate the refolding pathways. As a qualitative measure of these transitions, the barrier height marks the energy peak along such refolding paths. We introduce a meta-heuristic to estimate such barriers, which is an NP-complete problem. To guide an arbitrary path heuristic, the method uses RNA shape representative structures as intermediate checkpoints for detours. This enables a broad but efficient search for refolding pathways. The resulting Shape Triples meta-heuristic enables a close to optimal estimation of the barrier height that outperforms the precision of the employed path heuristic.

1 Introduction

RNA plays a central role in living cells. Numerous RNAs are able to switch between different structures within their life time due to thermodynamics, temperature changes (thermometers), ligand binding (riboswitches) or other signals [FHMS+01]. Such multistable RNAs regulate gene expression directly or are connected to regulatory mechanisms, e.g. splicing [LC93]. For the correct prediction and study of such structural changes it is necessary to identify the lowest energy refolding pathway in the underlying RNA energy landscape. The energy barrier height surmounted along such paths can be used to estimate refolding probabilities [GFW+08] or to study the kinetics of the folding process [WSSF+04].

Maňuch *et al.* have shown that the calculation of the exact barrier height is a hard, NP-complete problem for RNA secondary structure landscapes [MTSC09]. Therefore, exact approaches rely on the full enumeration of the low energy parts of the landscape [SvdPS99, FHSW02, KH05], resulting in exponential runtimes. Heuristics have been introduced to avoid the

exponential behaviour while still providing a reasonable estimate of the barrier height. The first greedy approach by Morgan and Higgs considers direct paths only [MH98] which are of minimal length. Subsequently, the barrier estimation was improved via more advanced direct path heuristics [FHMS$^+$01, TOSY06, GFW$^+$08]. In order to avoid the restriction of direct pathways, heuristics were introduced that allow for minor detours in the landscape [LFH09, DLVHC10]. Such methods revealed the high potential of non-direct pathways.

Our *Shape Triples* approach aims to improve the barrier height approximation of arbitrary path heuristics by splitting the pathway prediction $P_s \rightsquigarrow P_t$ into $P_s \rightsquigarrow R \rightsquigarrow P_t$, where R is a defined checkpoint for a detour. We utilize RNA shapes and their representative structures, the so called *shreps* [GVR04], to define the detour checkpoints R. RNA shapes group structures based on their branching pattern such that the number of shape classes is very small compared to the number of RNA structures. The resulting Shape Triples meta-heuristic, i.e. a high-level strategy that guides other path heuristics [Bla09], enables an efficient and precise estimation of barrier heights within RNA energy landscapes.

To evaluate our method, we show for two bistable RNA molecules the increased precision of the meta-heuristic compared to the employed path heuristic for a large number of refolding paths. We further show that in most cases the exact barrier height can be determined using our Shape Triples approach.

2 Preliminaries

In order to formulate our algorithms and results, we introduce the concept of energy landscapes, the barrier height problem, and their application to RNA. This is followed by an overview of RNA shape abstractions.

Energy Landscapes and Barrier Heights

In order to describe and investigate folding processes, the concept of discrete energy landscapes is applied frequently [Wri32, Sta02, FHSW02]. It is defined by a triple $\langle X, E, N \rangle$, i.e. a finite set of *states* X, an associated *energy function* $E : X \to \mathbb{R}$, and a *neighborhood relationship* $N : X \to \mathcal{P}(X)$, where $\mathcal{P}$ denotes the powerset. The folding process is mainly influenced by the *local minima* $M \subseteq X$ of the landscape defined by $\forall_{m \in M} \forall_{x \in N(m)} : E(m) \leq E(x)$.

A folding trajectory corresponds to a *walk (or path)* $w = (x_1, \ldots, x_l) \in X^l$ of length l within the energy landscape that respects the neighborhood relation $(\forall_i : x_i \in N(x_{i-1}))$. With $W(x_s, x_t)$ we denote the infinite set of *all possible* walks starting in x_s and ending in x_t.

The *barrier height* B denotes the lowest energy peak to make two structures x_s, x_t accessible to each other, i.e.

$$B(x_s, x_t) = \min\{ \max\{ E(x \in w) \mid w \in W(x_s, x_t) \}\} . \tag{1}$$

The barrier height heavily influences the folding probabilities within a certain energy landscape [FFHS00]. It can be used to derive energy landscape abstractions like barrier trees [HS88, FHSW02] and enables studies of folding kinetics [WSSF$^+$04, GFW$^+$08].

The *energy barrier problem* is to determine the exact barrier height B of two given states of an energy landscape.

RNA Secondary Structure Landscapes

In order to investigate the folding behavior of an RNA molecule the energy landscape of its secondary structures can be used [FFHS00, LFH09]. Given the nucleotide *sequence* $S \in \{\texttt{A}, \texttt{U}, \texttt{G}, \texttt{C}\}^n$ of an RNA of length n, a *secondary structure* P is a set of base pairs $\{(i, j) \mid 1 \leq i < j \leq n\}$ such that (a) S_i, S_j form a Watson-Crick (`A-U`, `G-C`) or a `G-U` base pair, with (b) at most one base pair per position, i.e. $\forall_{(i,j),(k,l)} : j \neq k \wedge (i = k \Leftrightarrow j = l)$, such that (c) all pairs are non-crossing, i.e. $\forall_{(i,j),(k,l)} : i < k < j \Leftrightarrow i < l < j$. The free energy of a given structure P can be calculated by a base pair based decomposition into structural elements [ZS81]. We utilize the implementation from the Vienna RNA Package[1] v1.7.2 within the Energy Landscape Library[2] v3.2.0 [MWB07]. All energies are given in $\frac{\text{kcal}}{\text{mol}}$ where calculations use parameters "`-d2 -T 37`". For details of the method applied and the energy parameters we refer to literature [ZS81, Hof03].

The neighborhood within an energy landscape reflects small structural changes along the folding process. To this end we utilize so called *single moves* [FFHS00], i.e. the insertion or deletion of a single base pair. Thus, the neighborhood of a given structure P is defined by $N(P) = \{P' \mid |\texttt{bp}(P) - \texttt{bp}(P')| = 1\}$, using its number of base pairs $\texttt{bp}(P) = |P|$.

The discrete *energy landscape of an RNA* S is thus defined by X as all secondary structures P of S, E as the free energy function defined by Zuker and Stiegler [ZS81], and the single move neighborhood N.

[1] Vienna RNA Package available at `http://www.tbi.univie.ac.at/~ivo/RNA/`
[2] ELL available at `http://www.bioinf.uni-freiburg.de/Software/`

Maňuch *et al.* have shown the NP-completeness of the energy barrier problem in such RNA energy landscapes [MTSC09].

RNA Shape Abstractions

RNA shapes, introduced by Giegerich *et al.* [GVR04], are a coarse grained model of RNA secondary structures. The shape abstraction is a homomorphic mapping of the secondary structure set X of an RNA into a set of compact representations of the different branching pattern covered by X. Five levels of abstraction are introduced and we denote these $\pi_i(P)$, *the shape abstraction of the i-th level* of a given RNA structure P. For details on the method we refer to literature [GVR04, SVR$^+$06]. Throughout this manuscript we use the `RNAshapes`[3] implementation v2.1.5.

Given an RNA energy landscape $\langle X, E, N\rangle$, we denote with Π_i *the set of all shape abstractions of level i of* X, i.e. $\Pi_i = \{\ \pi_i(P) \mid P \in X\}$. Thus each shape $\pi_i \in \Pi_i$ describes a class of structures of X. The structure with minimal energy within the class is called the *shape representative structure* or *shrep* $R(\pi_i)$, i.e. $\forall_{P\in X} : (\pi_i(P) = \pi_i) \rightarrow E(P) \geq E(R(\pi_i))$.

In the following we will utilize the RNA shape abstraction concept to generate a new and efficient meta-heuristic to estimate the barrier height between two RNA structures.

3 Methods

Since we want to present a meta-heuristic that employs an arbitrary path heuristic, we briefly review two existing direct path methods for the energy barrier problem, namely the `MH` heuristic by Morgan and Higgs [MH98] as well as a breadth-first-search (`BFS`) approach [FHMS$^+$01]. Both, the `MH` and `BFS` heuristic, can be implemented in our new RNA Shape Approaches presented afterwards. The exhaustive Shape Network approach exploits the potential of the RNA shape abstraction for the energy barrier problem. This is followed by our efficient *Shape Triples* meta-heuristic that enables a fast and precise barrier approximation.

RNA Direct Path Heuristics

Direct path heuristics find an approximate solution to the energy barrier problem for two RNA structures P_s, P_t. Considering only single moves (base pair insertion/deletion), a *direct path* $\hat{w}$ is a walk $w(P_s, P_t)$ of minimal length, i.e. of *base pair distance* $d(P_s, P_t) = |(P_s \cup P_t) \setminus (P_s \cap P_t)|$

[3] `RNAshapes` available at `http://bibiserv.techfak.uni-bielefeld.de/download/`

[MH98]. In the following the *abbreviation* $B_{DP}(P_s, P_t)$ will be used to denote the barrier height between P_s and P_t estimated by a direct path heuristic.

The MH heuristic: Morgan and Higgs introduced a simple greedy heuristic to explore direct paths [MH98]. It utilizes an iterative conflict-driven scheme of base pair insertions and deletions and evaluates the maximal energy reached within the resulting walk. Applied in several iterations, while storing the path with lowest barrier found, it returns an upper bound on the barrier height. For details on the method refer to the literature [MH98, FHMS$^+$01, GFW$^+$08].

The BFS heuristic: Flamm *et al.* improved the greedy `MH` approach using a limited breadth-first-search (`BFS`) [FHMS$^+$01]. Starting from the initial structure P_s, it enumerates all single moves possible in direct walks towards the target structure P_t. From these walks only the best m candidates are considered for extension in the next iteration. This continues until the full walk length of $d(P_s, P_t)$, and thus the target structure P_t, is reached. `BFS` enables better barrier height approximations compared to `MH` to the cost of increasing runtime correlated with m [GFW$^+$08]. In the following, we denote a `BFS` search with cut-off m with $\texttt{BFS}_m$.

Drawbacks of Direct Paths: Direct path heuristics are fast, but at the cost of precision, since only a small "corridor" of the energy landscape is investigated. Thus, the barrier height estimated via direct paths is usually higher than the exact one, i.e. $B_{\mathrm{DP}}(P_s, P_t) \geq B(P_s, P_t)$ [MH98]. Lorenz *et al.* have shown that lowest barrier pathways often contain detours and that rerouting via non-direct structures can significantly improve barrier height approximations [LFH09].

Shape Approaches

The central idea of our Shape Approaches is to use energy minimal shrep structures as intermediate checkpoints to reroute the path calculation of a given path heuristic, i.e. to go from the start structure P_s via shreps to target P_t. The resulting non-direct detour paths are more likely to enable a precise barrier estimate than the employed path heuristic alone. For simplicity, we exemplify the Shape Approaches employing a direct path heuristic as `MH` or `BFS`.

The Shape Network approach: In order to evaluate the potential of any Shape Approach we utilize the *Shape Network (SN)*, which uses the notion of shapes to create an abstraction of the energy landscape. The Shape Network is a fully connected, labeled graph where each node represents the shrep $R(\pi_i)$ of a shape $\pi_i \in \Pi_i$ of a given fixed shape

abstraction level i. In the following, we ignore the level identifier i and abbreviate $R(\pi_i) = R_\pi$ to ease the presentation. Each edge between two nodes $R_\pi, R_{\pi'}$ is labeled with a barrier height approximation via direct paths $B_{\mathrm{DP}}(R_\pi, R_{\pi'})$ (e.g. using `MH` or `BFS`).

Utilizing a simple variation of the dynamic programming algorithm by Floyd for the shortest path problem [Flo62], we get the *barrier height approximation* $B_F(R_\pi, R_{\pi'})$ *for any two shreps* $R_\pi, R_{\pi'}$ via any path within the Shape Network. Thus, using this estimate we can get an upper bound $B_{\mathrm{SN}}(P_s, P_t)$ of the barrier height between two RNA structures P_s, P_t including detours by

$$B_{\mathrm{SN}}(P_s, P_t) = \min_{\pi,\pi' \in \Pi} \{ \max \left\{ \begin{matrix} B_{\mathrm{DP}}(P_s, R_\pi), \\ B_{\mathrm{F}}(R_\pi, R_{\pi'}), \\ B_{\mathrm{DP}}(R_{\pi'}, P_t) \end{matrix} \right\}, \; B_{\mathrm{DP}}(P_s, P_t) \; \} \qquad (2)$$

The major drawback of the Shape Network approach is the high computational cost to calculate the Shape Network via $|\Pi|^2$ direct path calculations where computation time depends on the heuristic (see direct path section). Afterwards the Floyd algorithm runs efficiently in $O(|\Pi|^3)$ and results in the barrier height approximation B_F between all pairs of shreps. Once calculated, these approximations can be used to estimate the barrier height between any two structures using B_{SN} from Eq. 2 with $O(2|\Pi|)$ path calculations each.

Thus, the Shape Network approach is a useful tool when interested in a vast number of barrier heights, e.g. to calculate a barrier tree representation of the energy landscape's minima [FHSW02]. Beyond that, we can convert the Shape Network itself into an even coarser barrier tree abstraction covering the shrep structures that might reflect general properties of the energy landscape. Finally, the Shape Network approach gives a lower bound for meta-heuristics based on the Shape Approach idea.

The Shape Triples approach: In the following, we will introduce our *Shape Triples (ST)* meta-heuristic which enables a fast and efficient barrier height approximation. It is based on the observation that the majority of the barrier paths within the Shape Network are very short. We get already good upper bounds $B_{\mathrm{ST}}(P_s, P_t)$ on the barrier height when only investigating detours with one intermediate shape representative R_π, i.e.

$$B_{\mathrm{ST}}(P_s, P_t) = \min_{\pi \in \Pi} \{ \max \left\{ \begin{matrix} B_{\mathrm{DP}}(P_s, R_\pi), \\ B_{\mathrm{DP}}(R_\pi, P_t) \end{matrix} \right\}, \; B_{\mathrm{DP}}(P_s, P_t) \; \}. \qquad (3)$$

Thus, our two Shape Approaches yield new barrier height approximations B_{SN} and B_{ST} between the two structures P_s, P_t. These estimates are

```
function GetBST(P_s, P_t, Π)
    B ← B_DP(P_s, P_t)                          ▷ initialization of barrier estimate
    for all (π ∈ Π) do
        if (E(R_π) < B) then                    ▷ low energy shreps only
            B ← min{ B, max{B_DP(P_s, R_π), B_DP(R_π, P_t)}}   ▷ update B if needed
        end if
    end for
    return B                                    ▷ final B_ST(P_s, P_t) estimate
end function
```

Figure 1: Scheme for an efficient calculation of $B_{\mathrm{ST}}(P_s, P_t)$.

related via:

$$B(P_s, P_t) \leq B_{\mathrm{SN}}(P_s, P_t) \leq B_{\mathrm{ST}}(P_s, P_t) \leq B_{\mathrm{DP}}(P_s, P_t). \tag{4}$$

In order to calculate $B_{\mathrm{ST}}(P_s, P_t)$ from Eq. 3 we do not have to consider all shrep structures as possible intermediate checkpoints for detours. Every indirect path using a shrep R with $E(R) > B_{\mathrm{DP}}(P_s, P_t)$ will result in a worse barrier height estimation than already given by B_{DP} (see Eq. 3). Thus we can use an adaptive scheme to reduce the computational cost for calculating B_{ST} that considers only shreps with energy below the best barrier height estimation found so far as given in Fig. 1. The scheme can be further improved when using an energy sorted shape/shrep enumeration: as soon as a shrep exceeds the current barrier estimate the iteration can be terminated. Note, the same applies to the Shape Network approach.

4 Results and Discussion

We investigate the Shape Approaches using the RNA molecules `L45` and `SL` from Tab. 1. `SL` is the spliced leader RNA from *Leptomonas collosoma* taken from [LC93]. It was shown that the ability of this molecule to switch between two metastable structures heavily influences its splicing behavior. `L45` is a bistable artificial RNA taken from [LFH09].

In order to evaluate the methods, we study the barrier height error, i.e. the approximated (Eq. 2/3) minus the exact barrier height (Eq. 1). To this end we pick 5000 random pairs of local minima for `SL` with structural distance ≥ 7 and energy ≤ 0. The exact barrier height is calculated using the exhaustive `barriers` approach [FHSW02].

Figure 2 (left side) evaluates the Shape Approaches compared to the `BFS` direct path heuristics for `SL`. The Shape Network approach performs best

ID	shape i	2	3	4	5	structures
L45	$\lvert\Pi_i\rvert =$	528	68	57	13	$\lvert X\rvert =$ 5,999,391,327
SL	$\lvert\Pi_i\rvert =$	6305	594	336	49	$\lvert X\rvert <$ 1.725×10^{18}
L45	S	GGGCGCGGUUCGCCCUCCGCUAAAUGCGGAAGAUAAAUUGUGUCU				
	P_s	(((((.....)))))(((((.....)))))(((((.....)))))				
	P_t	((((((((((.....(((((.....)))))).....))))))))))				
SL	S	AACUAAAACAAUUUUUGAAGAACAGUUUCUGUACUUCAUUGGUAUGUAGAGACUUC				
	P_s	..((...((((((..(((((.((((...)))).)))))..))).))).))..).....				
	P_t	.....((((((((((((.....))))).)))))))..				

Table 1: RNA shape/structure numbers and sequences S for the energy landscapes investigated. For SL we estimated $|X|$ via sequence length n using the upper bound of $1.07427 \cdot n^{-3/2} \cdot 2.35467^n$ from [CKKS09]. The structures P_s/P_t correspond to the switch structures of the bistable molecules.

among all methods and finds the exact barrier for $\geq$75% of the pairs (SN+$\texttt{BFS}_5$). This shows the potential of detour pathways utilizing RNA shapes. Furthermore, the much simpler Shape Triples heuristic shows only a slightly higher error on average and still outperforms the direct path heuristic.

The figure also compares (on the right) the performance of the Shape Triples approach for different shape levels and direct path heuristics. Here, BFS clearly beats the MH-heuristic and increasing BFS cut-offs lower the error (as in [GFW$^+$08]). More importantly, the Shape Triples approach always yields better results, depicting the robustness of the method and its independence of the direct path method applied. Finally, increased abstraction (shape level) reduces the precision of the method. This is expected since less detours in the landscape are considered (see Tab. 1 for shape numbers). Nevertheless, the differences get less significant when employing a more precise path heuristic like $\texttt{BFS}_5$ (in green).

Table 2 evaluates the Shape Triples approach for the structure pairs from Tab. 1. In most cases B_{ST} matches or is close to the exact barrier height B and improves the upper bound from direct path results (B_{DP}). Note, even for high shape abstraction levels we gain a significant improvement. First experiments reveal that an increase of the BFS cut-off can further improve our B_{ST} results (data not shown).

The number of shapes $|\Pi_i|$ grows slowly exponential with increasing sequence length (see Tab. 1) [LPC08]. Nevertheless, the percentage of shapes considered to calculate B_{ST} drops drastically as shown by $|\Pi_i|$% in Tab. 2. Therefore, even for increasing sequence length, the computation

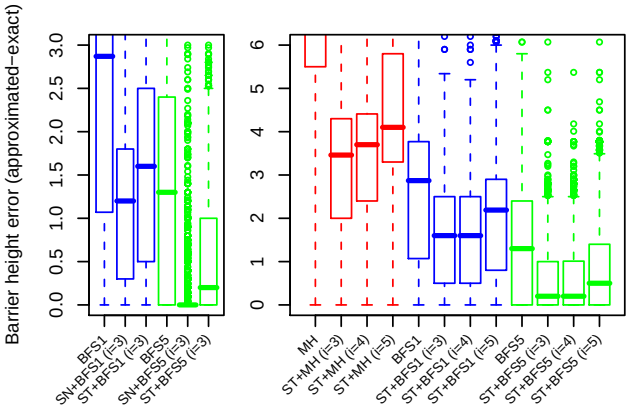

Figure 2: Evaluation of the Shape Approaches for RNA `SL`. (left) Direct path `BFS`-heuristic for cut-off 1 and 5 in comparison to Shape Network and Shape Triples approach at shape level i=3. (right) Performance of the Shape Triples approach when applying different direct path heuristics and shape levels i. Boxes cover 50% of the distribution while solid lines mark the median.

effort of the Shape Triples approach remains low.

We compare our results to the κ,λ-neighborhood approach presented in [LFH09]. There, detours are rerouted through energy minimal structures within the κ,λ-neighborhood, i.e. via energy minimal structures within the structural distances κ and λ to the start and target structures, respectively. Using a $\texttt{BFS}_{100}$ heuristic (R. Lorenz, pers. commun.), Lorenz *et al.* are able to estimate the exact barrier height of -7.5 for `L45` [LFH09][4]. The Shape Triples approach reproduces the same exact barrier height for different shape levels (see Tab. 2) while using a much faster $\texttt{BFS}_5$ with cut-off 5 instead of 100 (see Methods).

5 Conclusion

We have introduced RNA shape based meta-heuristics to estimate the barrier height between RNA structures, an important problem to study multistable RNA molecules. The methods utilize shape representative structures (shreps) as intermediate checkpoints to reroute a given path

[4]Note, in [LFH09] the energy difference $\Delta E = (B(P_s, P_t) - E(P_s))$ is given. Thus, the barrier height was recalculated by $(E(P_s) + \Delta E)$.

ID	$B(P_s, P_t)$			Shape Triples $B_{\mathrm{ST}}(P_s, P_t)$				
				shape i	2	3	4	5
L45	**-7.5**	only $\mathtt{BFS}_5$	-4.87	with $\mathtt{BFS}_5$	**-7.5**	**-7.5**	-6.4	-6.2
		$B_{\mathrm{DP}}(P_s, P_t) \geq$	-4.87	$\lvert\Pi_i\rvert$%	16.7	33.8	31.6	38.5
SL	**0.5**	only $\mathtt{BFS}_5$	2.6	with $\mathtt{BFS}_5$	**0.5**	**0.51**	**0.51**	2.6
		$B_{\mathrm{DP}}(P_s, P_t) \geq$	1.9	$\lvert\Pi_i\rvert$%	5.9	9.4	14.8	18.4

Table 2: Barrier height evaluation for the P_s/P_t structure pairs from Tab. 1. Given is the exact barrier $B(P_s, P_t)$, the estimate via only direct path $\mathtt{BFS}_5$, the lowest barrier for such direct paths $B_{\mathrm{DP}}(P_s, P_t)$, and the Shape Triples approximations $B_{\mathrm{ST}}(P_s, P_t)$ for different shape level using $\mathtt{BFS}_5$. $|\Pi_i|$% denotes the percentage of $|\Pi_i|$ from Tab. 1 used to calculate B_{ST} (see Methods).

heuristic. This enables a broader search in the energy landscape as done by the employed heuristic alone. We have shown that our Shape Triples approach is able to estimate barrier heights close to the optimum using a $\mathtt{BFS}_5$ heuristic. The approach scales with the number of investigated shreps as shown in Fig. 1. Thus, the use of different shape levels enables a trade-off between barrier precision and computational performance (see Tab. 2) where the latter depends on the performance of the individual path heuristic applied.

While being introduced for direct path heuristics only, the method is applicable to any other path heuristic. Thus, we plan to investigate the use of the RNAtabupath [DLVHC10], currently using a different RNA energy scheme, that was shown to yield slightly better results than BFS by allowing for minor detours. When employing RNAtabupath within the Shape Triples approach it may be possible to improve the results even further (see Eq. 4).

We plan to investigate different shrep selection strategies to further speed-up the method. Possible directions are the structural distance to start and target structure or a shape distance based evaluation.

Furthermore, the method is basically not restricted to RNA shapes but open to any sampling of low energy structures of the underlying RNA energy landscape. Thus, any scheme for an efficient calculation of such a set of structures can be used to replace the set of shape representatives in the Shape Triples approach (Fig. 1) and might even improve the results.

Therefore, we consider the Shape Triples meta-heuristic to be a very useful tool to combine results from different algorithmic fields to gain very precise barrier height estimates for arbitrary RNA structures.

References

[Bla09] P.E. Black. metaheuristic. In Dictionary of Algorithms and Data Structures [online], Paul E. Black, ed., U.S. National Institute of Standards and Technology, 30 March 2009. (accessed 15.06.2010) Available from: http://www.itl.nist.gov/div897/sqg/dads/HTML/metaheuristic.html.

[CKKS09] P. Clote, E. Kranakis, D. Krizanc, and B. Salvy. Asymptotics of canonical and saturated RNA secondary structures. *JBCB*, 7(5):869–893, 2009.

[DLVHC10] I. Dotu, W.A. Lorenz, P. Van Hentenryck, and P. Clote. Computing folding pathways between RNA secondary structures. *NAR*, 38(5):1711–1722, 2010.

[FFHS00] C. Flamm, W. Fontana, I.L. Hofacker, and P. Schuster. RNA folding at elementary step resolution. *RNA*, 6(03):325–338, 2000.

[FHMS+01] C. Flamm, I.L. Hofacker, S. Maurer-Stroh, P.F. Stadler, and M. Zehl. Design of multistable RNA molecules. *RNA*, 7(02):254–265, 2001.

[FHSW02] C. Flamm, I.L. Hofacker, P.F. Stadler, and M.T. Wolfinger. Barrier Trees of Degenerate Landscapes. *Z. Phys. Chem.*, 216(2/2002):155–173, 2002.

[Flo62] R.W. Floyd. Algorithm 97: Shortest path. *Commun. ACM*, 5(6):345, 1962.

[GFW+08] M. Geis, C. Flamm, M.T. Wolfinger, A. Tanzer, I.L. Hofacker, M. Middendorf, C. Mandl, P.F. Stadler, and C. Thurner. Folding kinetics of large RNAs. *J of Mol Biol.*, 379(1):160–173, 2008.

[GVR04] R. Giegerich, B. Voss, and M. Rehmsmeier. Abstract shapes of RNA. *NAR*, 32(16):4843–4851, 2004.

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

[HS88] K.H. Hoffmann and P. Sibani. Diffusion in hierarchies. *Phys. Rev. A*, 38(8):4261–4270, 1988.

[KH05] M. Kubota and M. Hagiya. Minimum basin algorithm: An effective analysis technique for DNA energy landscapes. *LNCS*, 3384:202–214, 2005.

[LC93] K.A. LeCuyer and D.M. Crothers. The Leptomonas collosoma spliced leader RNA can switch between two alternate structural forms. *Biochemistry*, 32(20):53015311, 1993.

[LFH09] R. Lorenz, C. Flamm, and I.L. Hofacker. 2D projections of RNA folding landscapes. In *Proc. of GCB'09*, volume 157 of *LNCS*, pages 11–20, 2009.

[LPC08] W.A. Lorenz, Y. Ponty, and P. Clote. Asymptotics of RNA shapes. *J of Comp. Biol.*, 15(1):31–63, 2008.

[MH98] S.R. Morgan and P.G. Higgs. Barrier heights between ground states in a model of RNA secondary structure. *J of Physics A*, 31:3153–3170, 1998.

[MTSC09] J. Maňuch, C. Thachuk, L. Stacho, and A. Condon. NP-Completeness of the Direct Energy Barrier Problem without Pseudoknots. In *DNA Computing*, volume 5877 of *LNCS*, pages 106–115, 2009.

[MWB07] M. Mann, S. Will, and R. Backofen. The Energy Landscape Library–a platform for generic algorithms. *Proc. of BIRD*, 7:83–86, 2007.

[Sta02] P.F. Stadler. Fitness landscapes. In *LNP*, pages 183–204. Springer, 2002.

[SvdPS99] P. Sibani, R. van der Pas, and J.C. Schön. The lid method for exhaustive exploration of metastable states of complex systems. *Comp. Phys. Comm.*, 116(1):17–27, 1999.

[SVR+06] P. Steffen, B. Voss, M. Rehmsmeier, J. Reeder, and R. Giegerich. RNAshapes: an integrated RNA analysis package based on abstract shapes. *Bioinformatics*, 22(4):500–503, 2006.

[TOSY06] T. Takeda, H. Ono, K. Sadakane, and M. Yamashita. A Local Search Based Barrier Height Estimation Algorithm for DNA Molecular Transitions. In *DNA Computing*, volume 3892 of *LNCS*, pages 359–370, 2006.

[Wri32] S. Wright. The Roles of Mutation. In *Congress on Genetics*, page 365, 1932.

[WSSF+04] M.T. Wolfinger, W.A. Svrcek-Seiler, C. Flamm, I.L. Hofacker, and P.F. Stadler. Efficient computation of RNA folding dynamics. *J of Physics A*, 37(17):4731–4741, 2004.

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