

Determinants of Simulated RNA Evolution

Anne Kupczok and Peter Dittrich¹

*Bio Systems Analysis Group, Jena Centre for Bioinformatics and Department of
Mathematics and Computer Science, Friedrich Schiller University Jena, D-07743
Jena, Germany*

Abstract

Models of RNA secondary structure folding are widely used to study evolution in theory and simulation. However, systematic studies of the parameters involved are rare. In this paper we study by simulation how RNA evolution is influenced by three different factors, namely the mutation rate, scaling of the fitness function, and distance measure. We found that for low mutation rates the qualitative evolutionary behavior is robust with respect to the scaling of the fitness function. For efficient mutation rates, which are close to the error threshold, scaling and distance measure have a strong influence on the evolutionary behavior. A global distance measure that takes sequence information additively into account lowers the error threshold. When using a local sequence-structure alignment for the distance, we observed a smoother evolution of the fitness over time. Finally, in addition to the well known error threshold, we identify another threshold of the mutation rate, called divergence threshold, where the qualitative transient behavior changes from a localized to an exploratory search.

Key words: RNA evolution, error threshold, divergence threshold, local sequence-structure alignment

Contents

1	Introduction	2
2	Material and Methods	3
2.1	Basic Model	3
2.2	Studied Factors	5
3	Results	8
3.1	Mutation Rate	8
3.2	Scaling of the Fitness Function	11
3.3	Distance Measure	13
4	Conclusion and Outlook	16

1 Introduction

Models of RNA secondary structure folding are widely used to study evolution in theory and simulation (see Schuster (2001) for an overview), since the secondary structure is recognized as computationally tractable, while being at the same time a realistic phenotype (Fontana and Schuster, 1998; Mathews et al., 1999). In many simulation studies, however, only single experiments as typical examples are presented, e.g., Fontana and Schuster (1998) and Huynen et al. (1996).

In order to get a deeper understanding how simulated RNA evolution depends

¹ Corresponding author. dittrich@minet.uni-jena.de

on chosen parameters, we study by simulation the influence of three factors, namely mutation rate, scaling of the fitness function, and distance measure. Evolutionary behavior is measured by the the change of average and best population fitness over time. Evolutionary success is measured as (average) fitness obtained after a specific number of generations. Furthermore, we investigated the probability that a mutation in a random individual (RNA strand) leads to a quality gain, neutral mutation, or quality decay.

Our results suggest that the qualitative behavior of RNA evolution is robust against various parameter changes, as long as the parameter setting is not close to one of two boundaries, namely the well known error threshold (Eigen and Schuster, 1977) and a threshold where the population switches from a localized to an exploratory behavior. The actual mutation rate at which these transitions occur depends on the chosen scaling and distance measure.

2 Material and Methods

2.1 Basic Model

For the simulation of evolutionary processes we make the following common assumptions: we assume a well-stirred reaction vessel (sympatric population) containing a finite number of RNA sequences that replicate with error. Replication error is modeled by point mutations. Our investigations are restricted to RNA sequences of fixed length l , prohibiting insertion and deletion. The replication rate of an RNA sequence is determined by a fitness function. The fitness function $f : Seq \rightarrow \mathbb{R}_{\geq 0}$ is a mapping that maps an RNA sequence $s \in Seq = \{A, C, G, U\}^l$ to a fitness value $f(s)$. The fitness determines the

rate at which a sequence produces offspring. In our stochastic simulation, the fitness is proportional to the probability that a copy (potentially with error) is created from s during a specific time interval. Thus, the fitness determines the number of expected offspring of s .

As usual in *in silico* RNA evolution, the fitness $f(s)$ is defined as the distance between the sequence s and a chosen fixed target sequence s_{target} , scaled by a scaling function $f_{\text{scale}} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$:

$$f(s) \equiv f_{\text{scale}}(d(s, s_{\text{target}})). \quad (1)$$

The distance d is usually (e.g., Fontana and Schuster, 1998; Fontana et al., 1993b) defined as the structural distance d_{sec} of the secondary structures, e.g., $d(s, s_{\text{target}}) \equiv d_{\text{sec}}(f_{\text{fold}}(s), f_{\text{fold}}(s_{\text{target}}))$, where $f_{\text{fold}} : \text{Seq} \rightarrow \text{Sec}$ is a function that maps a sequence to its secondary structure. In addition to a pure structural distance, we consider also sequence information and a local sequence-structure alignment by Backofen and Will (2004).

The evolutionary dynamics of an initial population $S^0 = \{S_1^0, \dots, S_\lambda^0\}$ of λ randomly generated RNA sequences is simply computed by $S_i^{g+1} \leftarrow \text{mutate}(\text{select}(S^g), p_{\text{mut}})$, where $S_i^g \in \text{Seq}$ denotes the i -th individual at generation g , $i \in \{1, 2, \dots, \lambda\}$, $S_i^g \in \{A, C, G, U\}^l$. Sequences are selected with a probability proportional to their fitness: $\text{Prob}(\text{select}(S^g) = S_i^g) \sim f(S_i^g)$. Each site of an offspring can mutate with probability p_{mut} : $\text{mutate}((s_1, \dots, s_l), p_{\text{mut}}) = (s'_1, \dots, s'_l)$ where $\text{Prob}(s_j = s'_j) = 1 - p_{\text{mut}}$, $\text{Prob}(s_j \neq s'_j) = p_{\text{mut}}$, $j \in \{1, \dots, l\}$, $s_j \in \{A, C, G, U\}$, and l the length of the sequence to be mutated.

2.2 Studied Factors

We studied how the evolutionary behavior depends on the following three factors: mutation rate p_{mut} , scaling function f_{scale} and distance measure d . The latter two are now explained in more detail.

Distance measure d : The distance measure d computes the distance between two sequences. Here, we investigate the following four distance measures:

(1) *Global structure alignment (Glob)*: The conventional *global structure alignment* is a structural distance measure, which calculates the distance solely based on the secondary structures of the two sequences. The functions for computing the secondary structures and the distance between secondary structures are made available by the Vienna RNA package² (Hofacker et al., 1994). The measure is commonly used in many simulation studies of RNA evolution (Fontana and Schuster, 1998):

$$d_1(s, s_{target}) = d_{sec}(f_{fold}(s), f_{fold}(s_{target})) \quad (2)$$

where d_{sec} represents the function `tree_edit_distance` (Shapiro, 1988; Shapiro and Zhang, 1990; Fontana et al., 1993a; Hofacker et al., 1994) and f_{fold} the function `RNAfold` with default parameters (Hofacker et al., 1994; Zuker and Stiegler, 1981; McCaskill, 1990) of the Vienna RNA package (Version 1.4).

(2) *Additive combined sequence and structure alignment (Add)*: Since the function of an RNA does not solely depend on its secondary structure but also on its sequence, we investigate three distance measures that take sequence information into account. The first one, d_2 , considers sequence information simply by adding the (weighted) hamming distance between the sequence s and the

² Available from: www.tbi.univie.ac.at/~ivo/RNA/

target sequence s_{target} to the structural distance:

$$d_2(s, s_{\text{target}}) = [\alpha d_{\text{sec}}(f_{\text{fold}}(s), f_{\text{fold}}(s_{\text{target}})) + (1 - \alpha)d_{\text{hamming}}(s, s_{\text{target}})] \quad (3)$$

where the weight $0 \leq \alpha \leq 1$ is usually chosen as 0.5.

(3) *Multiplicative combined sequence and structure alignment (Mult)*: In addition to the simple additive combined sequence and structure alignment we also defined a distance measure where the hamming distance is multiplied with the structural distance:

$$d_3(s, s_{\text{target}}) = \frac{[d_{\text{sec}}(f_{\text{fold}}(s), f_{\text{fold}}(s_{\text{target}})) + 1][d_{\text{hamming}}(s, s_{\text{target}}) + 1] - 1}{l}. \quad (4)$$

Both measures are fairly unrealistic, but represent the most simple approach to include sequence information. Furthermore they are easy to compute. An additive combination might be justified, since different traits can also respond to selection in an additive way, as shown by Beldade et al. (2002) in selection experiments with butterflies. A multiplicative combination make sense, when both traits are *required* for survival.

(4) *Local sequence-structure alignment (LSSA)*: The local sequence-structure alignment by Backofen and Will (2004) is more elaborated and realistic than the previous measures. The alignment considers sequence and structure information not independently like d_2 and d_3 , but it aligns two sequences simultaneously considering structure and function. Locality in a sequence structure alignment is defined by connectivity via the non-atomic bonds between complementary bases and via the atomic backbone bonds. So, the aligned subsequences need not to be connected in a sequence as in the local sequence alignment. The optimal LSSA for two nested structures can be found by a

dynamic programming algorithm, which takes the similarity of bases and of bonds into account. Because the alignment returns a similarity measure, we have to invert it in order to obtain a distance:

$$d_4(s, s_{\text{target}}) = l \left[1 - \frac{1}{\text{max}} \text{LSSA}(s, f_{\text{fold}}(s), s_{\text{target}}, f_{\text{fold}}(s_{\text{target}})) \right]. \quad (5)$$

The similarity measure *LSSA* represents the function `lssa` provided by the software package *Local Sequence Structure Alignment*³ (Backofen and Will, 2004). The transformation ensures that we get a distance measure in the same range as the other three measures. The maximal similarity is defined as the similarity between two identical copies of the target sequence $\text{max} = \text{LSSA}(s_{\text{target}}, f_{\text{fold}}(s_{\text{target}}), s_{\text{target}}, f_{\text{fold}}(s_{\text{target}}))$. Note that *LSSA* is a local alignment that returns continuous values unlike the function `tree_edit_distance` of the Vienna package returning integers.

Scaling function f_{scale} : The scaling function maps a distance to a fitness value. We considered five different scaling functions, which are illustrated in Fig. 1. All scaling functions decrease monotonously. They differ only in the shape of the slope:

$$(1) \text{ linear} \quad f_{\text{scale}}^1(d) = 100(1 - d/l), \quad (6)$$

$$(2) \text{ exponential} \quad f_{\text{scale}}^2(d) = 100^{1-d/l}, \quad (7)$$

$$(3) \text{ rational} \quad f_{\text{scale}}^3(d) = \frac{1}{0.01 + d/l}, \quad (8)$$

$$(4) \text{ sigmoid} \quad f_{\text{scale}}^4(d) = 100^{1-(d/l)^\sigma}, \quad (9)$$

$$(5) \text{ inverse} \quad f_{\text{scale}}^5(d) = 100 - 100^{d/l} + 1. \quad (10)$$

(FIGURE 1 APPROXIMATELY HERE)

The rational scaling, f_{scale}^3 , is most commonly used in various studies. For the

³ Available from www.bio.inf.uni-jena.de.

Table 1

Mutation rates that resulted in highest fitness in the experiments shown in Fig. 3. Furthermore, best and average fitness obtained after 10000 generations are shown. However, note that fitness values are not comparable between experiments using different distance measures.

Distance measure	Best mutation rate	Best fitness	Average fitness
$d_1(s, s_{\text{target}})$	0.022	71 (± 11.83)	26 (± 4.14)
$d_2(s, s_{\text{target}})$	0.005	16 (± 9.54)	11 (± 6.68)
$d_3(s, s_{\text{target}})$	0.02	52 (± 5.57)	26 (± 4.03)
$d_4(s, s_{\text{target}})$	0.001	3.2 (± 0.35)	3.0 (± 0.33)

$l = 76$, population size $\lambda = 100$) the error threshold is about 0.02 (Fig. 3(a)). When we add sequence information by using the additive sequence structure alignment d_2 instead of d_1 , the error threshold decreases significantly and becomes about 0.005 (Fig. 3(b)). This observation can be explained by the fact that structure *and* sequence information has to be transmitted to the next generation. In other words, information that determines fitness can be destroyed more easily by mutation, because a mutation may destroy the structure, the sequence, or both.

(FIGURE 3 APPROXIMATELY HERE)

Interestingly, when sequence information is considered multiplicatively by using the multiplicative combined sequence and structure alignment d_3 (Fig. 3(c)), the error threshold is the same as in the conventional scenario, where sequence information is ignored (Fig. 3(a)). This phenomenon will be explained in Sec. 3.3.2, where the additive and multiplicative combined sequence and

structure alignment are compared with respect to their independent improvement of sequence and structure information.

When sequence information is considered using the local sequence alignment (LSSA) developed by Backofen and Will, we obtain the smallest error threshold 0.001 (Fig. 3(d)). One reason for this effect is that in the LSSA small deviations from the target structure lead to large decrease of similarity, which is in agreement with our observation that applying LSSA results in the most difficult evolutionary scenario.

At that point it is important to note that our error threshold denotes a value of the mutation rate above which evolution does not progress. Usually, in studies like those by Huynen et al. (1996) and Takeuchi et al. (2005), the error threshold has been measured by assigning a high fitness to the master sequence (or master structure), whereas assigning a much lower fitness to all other sequences. In our study, however, we obtain the error threshold by simulations starting with a randomly initialized population and measuring the fitness gained after a finite time interval, because we are interested in the influence of distance and scaling functions on the evolutionary behavior.

The error threshold found for d_1 (global structure alignment) is in accordance with the recent study by Takeuchi et al. (2005), who found an error threshold around 0.045. Takeuchi et al. (2005) used a larger population size (10000 individuals), which leads to a slightly larger error threshold, because a larger population can carry more unfit mutants (Nowak and Schuster, 1989). According to Nowak and Schuster (1989), the error threshold scales with $1/\sqrt{\lambda}$ with the population size.

Divergence threshold: In addition to the error threshold, we observed a

smaller “divergence threshold” of the mutation rate above which the *average* structural distance to the target structure increases (with increasing mutation rate), whereas the *best* structural distance to the structure decreases (improves). Below this threshold, both, best and average distance to the target structure decrease with increasing mutation rate. Note that in contrast to the conventional error threshold, the divergence threshold is a phenomenon that appears only at finite time, where the population has not yet converged. In Fig. 4 the divergence threshold is illustrated for time $g = 10000$, where the population is still in a transient state. We can see that the divergence threshold (at about $0.001 - 0.003$) and the error threshold (at about 0.02) mark off three different regimes of evolutionary behavior. Note that the characteristics of the exploratory regime between the two thresholds is studied, e.g., by Huynen et al. (1996).

(FIGURE 4 APPROXIMATELY HERE)

3.2 *Scaling of the Fitness Function*

In order to investigate the effect of the scaling of the fitness function we performed for each scaling function simulations for two different mutation rates, $p_{\text{mut}} = 0.001$ and $p_{\text{mut}} = 0.01$, and measured the evolutionary success as the best and average distance obtained after 100,000 generations (conventional global structure alignment d_1 , population size $\lambda = 100$). Table 2 summarizes the results, each obtained as an average over 10 independent runs.

For a low mutation rate $p_{\text{mut}} = 0.001$, far below the optimal mutation rate 0.022 , the behavior is fairly independent from the scaling function. The linear, exponential, and rational scaling functions lead to the same best and average

Table 2

Best and average structural distance obtained after $g = 100000$ generations for five different scaling functions and two mutation rates; averaged over $n = 10$ independent experiments for each parameter setting (standard error in brackets). Note that fitness values cannot be compared between different scalings.

$p_{\text{mut}} =$		0.001		
	Distance d_1		Fitness f	
Scaling	best	avg.	best	avg.
lin $f_{\text{scale}}^1(d_1)$	4.8 (± 0.53)	8.3 (± 0.72)	93.7 (± 0.72)	89.1 (± 0.95)
exp $f_{\text{scale}}^2(d_1)$	6 (± 1.0)	7.7 (± 0.95)	70 (± 4.10)	66 (± 3.84)
rat $f_{\text{scale}}^3(d_1)$	5.6 (± 0.83)	7.3 (± 0.84)	15 (± 2.36)	13 (± 2.21)
sig $f_{\text{scale}}^4(d_1)$	10.8 (± 0.85)	14.6 (± 0.70)	97.5 (± 0.55)	93.2 (± 0.57)
inv $f_{\text{scale}}^5(d_1)$	23 (± 0.86)	31.9 (± 0.62)	96.9 (± 0.22)	92.8 (± 0.33)
$p_{\text{mut}} =$		0.01		
	Distance d_1		Fitness f	
Scaling	best	avg.	best	avg.
lin $f_{\text{scale}}^1(d_1)$	15 (± 1.20)	34.8 (± 0.52)	80 (± 1.58)	54.1 (± 0.69)
exp $f_{\text{scale}}^2(d_1)$	1.4 (± 0.43)	13.9 (± 0.71)	92 (± 2.35)	57 (± 1.85)
rat $f_{\text{scale}}^3(d_1)$	0.6 (± 0.43)	9 (± 1.3)	74 (± 13.22)	44 (± 7.98)
sig $f_{\text{scale}}^4(d_1)$	11 (± 1.4)	27 (± 3.11)	87 (± 9.69)	59 (± 6.68)
inv $f_{\text{scale}}^5(d_1)$	28.6 (± 0.85)	46.9 (± 0.47)	95.3 (± 0.29)	81.0 (± 0.48)

distance to the target structure. The sigmoid scaling function leads also to a relatively good evolutionary performance. Only the “pathological” inverse scaling results in a drastically lower evolutionary success, which is reasonable, since an improvement in distance of a good individual leads only to relatively small improvement in fitness.

For a high and efficient mutation rate $p_{\text{mut}} = 0.01$, which is close to the optimum and error threshold, the scaling function crucially influences the evolutionary behavior. Only the rational and exponential scaling functions perform well, whereas the linear, sigmoid, and inverse scaling function lead to best distance values 10 times worse than with the rational scaling function. In summary, the shape of the scaling function matters when the mutation rate is close to the error threshold. Note that all scaling functions investigated here are strictly monotonous.

3.3 Distance Measure

3.3.1 Fitness Frequency Distribution

For n randomly created sequences we counted the frequency of a positive, negative, and neutral fitness change after every possible one-point-mutation⁴. Figure 5 shows that the different distance measures behave qualitatively in a similar way: the probability to obtain a better individual decreases with the fitness of the parent. On the contrary, the probability to create a less fit offspring increases with the fitness of the parent. Adding sequence information

⁴ The effect of large mutational distances has been studied by Wilke et al. (2003). See Reidys et al. (2001) for a theoretical analysis of neutral networks.

reduces the probability to obtain a neutral mutation (Figs. 5(b)-(d)), as could be expected. The effect is smaller when sequence information is multiplicatively considered (compare Fig. 5(c) and Fig. 5(b)), which may correspond to the lower error threshold of the additive structure and sequence distance. It should be noted that the curve showing the fraction of neutral offspring for LSSA (Fig. 5(d)) poses an unusual peak at fitness 2.4, which does not appear when using the global structure alignment (Figs. 5(a)-(c)).

(FIGURE 5 APPROXIMATELY HERE)

3.3.2 Qualitative Comparison of Additive and Multiplicative Sequence and Structure Alignment (d_2, d_3)

In this section, we compare the distance measures d_2 and d_3 , where sequence information is combined with the conventional global structure alignment by adding and multiplying the hamming distance to the target sequence, respectively. Previously we have shown (Fig. 3) that the error thresholds of additive combination d_2 is lower than the threshold of the global structure alignment d_1 , while the error threshold of the multiplicative combination d_3 is equivalent to the global structure alignment d_1 . In order to investigate this effect in more detail, Fig 6 shows, how structure and sequence contribute to the combined fitness.

(FIGURE 6 APPROXIMATELY HERE)

In Figs. 6(b) and 6(d) we can see that the multiplicative combination quickly optimizes the structural distance, while the hamming distance adapts only very slowly (note the logarithmically scaled time axis). In contrast, when using an additive combination, sequence and structural distance are quickly

optimized simultaneously (Fig. 6(a)). At mutation rate $p_{\text{mut}} = 0.01$, which is just below the error threshold of the multiplicative combination and above the threshold of the additive combination, sequence and structure cannot be optimized using the additive sequence and structure alignment (Fig. 6(c)), as expected. In summary, only the additive combination really improves both, sequence and structure, in parallel; while the multiplicative combination ignores the hamming distance, at first, and thus behaves like d_1 .

3.3.3 Qualitative Behavior of LSSA

When employing the local structure and sequence alignment (LSSA) to measure the distance of an individual to the target sequence, we obtained a quite different quantitative and qualitative evolutionary behavior. First, evolution is more difficult, which is reflected by a low error threshold (0.001 instead of 0.02 with d_1 and d_3 , and 0.005 with d_2). Second, the dynamics of the population's best and average fitness appears smoother compared to the other three distance measures, which are based on the conventional global structure alignment (compare Fig. 7 with Figs. 4(b) and 4(c)). Punctuated equilibria did not become visible: In the dynamics of the best or average fitness (Fig. 7), the typical patterns of punctuated equilibria – namely plateaus of stasis, which are interrupted by short periods of change – did not appear in our experiments with one target sequence and a population size of $\lambda = 100$.

Because LSSA returns continuous values and not integers like the `tree_edit_distance` from the Vienna package, it may be assumed that the smoother behavior is a result of continuous values. We compared evolution on continuous characters (leaving the result of LSSA unchanged) and discrete characters (rounding the

result of LSSA). Figure 7 shows that the smoother behavior retains, even when we round the outcome of LSSA to integers. In contrast, evolution on discrete values of LSSA appears to be more difficult than on continuous.

(FIGURE 7 APPROXIMATELY HERE)

Although the fitness difference between the two experiments shown in Fig. 7 is not large, the evolution has not found a cloverleaf structure after 25000 generations. Among the five experiments we made with the parameters of Fig. 7, a cloverleaf structure was found with continuous values four times and with rounded values only once.

Preliminary studies with sequences of varying length where we also allowed insertion and deletion of nucleotides resulted in higher evolutionary rates, i.e. a faster increasing fitness, because insertions increase the probability for a better local alignment (data not shown). Hence, sequence length increased, when there is no parsimonious pressure that favors smaller sequences.

4 Conclusion and Outlook

We investigated the effect of mutation rate, scaling function, and distance measure on the behavior of simulated RNA evolution. We demonstrated that for low mutation rates the qualitative evolutionary behavior is fairly robust with respect to the scaling of the fitness function, as long as we do not take scaling functions like the inverse function f_{scale}^5 into account, which is still monotonous but where an improvement of a fit genotype results only in a tiny improvement in fitness, such that a new fitter offspring genotype has only a tiny selective advantage and can easily die out.

Changing the distance measure by including sequence information, can drastically increase the difficulty of evolution, which has caused a more than ten-fold decrease in the error threshold in our study. Hence, when taking the model of RNA evolution for quantifying Eigen's Paradox and studying the corresponding problem of the origin of life, we have to be careful concerning the chosen parameters, and be aware that they may have strong quantitative effects. Using the local sequence and structure alignment (LSSA) we even observed a qualitative change, which was characterized by a smoother dynamics of the best and average population fitness. For efficient mutation rates, which are close to the error threshold, scaling and distance measure have a strong influence on the evolutionary behavior.

Finally, in addition to the well known error threshold, we identified another threshold of the mutation rate, called *divergence threshold*, where the qualitative behavior changed from a localized to an exploratory search. As opposed to the error threshold, the divergence threshold becomes visible only during the transient phase of evolution.

In summary, our results support studies like those by Fontana and Schuster (1998) or Huynen et al. (1996), which showed only single experiments and argue that these single experiments are typical and that the displayed behavior is robust against parameter changes. However, the behavior is not arbitrarily robust, which has been quantified by this study with respect to mutation rate, scaling, and distance measure; e.g., it is not enough to demand that the scaling function is monotonous, its shape is important, too.

In the future, the evolution of structure and function of variable length sequences requires further investigations. Sequences of arbitrary length will in-

roduce a new level of complexity, where more aspects have to be considered, such as, how to define deletion and insertion operators for mutation. From our preliminary studies we expect a rich behavioral spectrum, in particular when using a local sequence-structure alignment like LSSA together with variable length sequences. Moving this way towards more realistic evolutionary models will deepen our insight into the complexity of evolutionary dynamics.

Acknowledgments

We are grateful to R. Backofen, F. Centler, I. Hofacker, M. A. Huynen, N. Matsumaru, P. Stadler, N. Takeuchi, S. Will, and the anonymous reviewers. This work was supported by Federal Ministry of Education and Research (BMBF) Grant 0312704A to Friedrich Schiller University Jena.

References

- Backofen, R., Will, S., 2004. Local sequence-structure motifs in RNA. *J. Bioinformatics Comput. Biol.* 2 (4), 681–698.
- Beldade, P., Koops, K., Brakefield, P. M., 2002. Modularity, individuality, and evo-devo in butterfly wings. *Proc. Natl. Acad. Sci. USA* 99 (22), 14262–14267.
- Eigen, M., Schuster, P., 1977. The hypercycle: a principle of natural self-organisation, part A. *Naturwissenschaften* 64 (11), 541–565.
- Fontana, W., Konings, D. A. M., Stadler, P. F., Schuster, P., 1993a. Statistics of RNA secondary structures. *Biopolymers* 33 (9), 1389–1404.
- Fontana, W., Schuster, P., 1998. Continuity in evolution: On the nature of transitions. *Science* 280, 1451–1455.

- Fontana, W., Stadler, P. F., Bornberg-Bauer, E. G., Griesmacher, T., Hofacker, I. L., Tacker, M., Tarazona, P., Weinberger, E., Schuster, P., 1993b. Rna folding and combinatorial landscapes. *Phys. Rev. E* 47 (3), 2083–2099.
- Hofacker, I., Fontana, W., Stadler, P., Bonhoeffer, S., Tacker, M., Schuster, P., 1994. Fast folding and comparison of RNA secondary structures. *Monatshefte Chemie* 125, 167–188.
- Huynen, W. A., Stadler, P. F., Fontana, W., 1996. Smoothness within ruggedness: The role of neutrality in adaptation. *Proc. Natl. Acad. Sci. USA* 93 (1), 397–401.
- Mathews, D. H., Sabina, J., Zuker, M., Turner, D. H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 288 (5), 911–940.
- McCaskill, J. S., 1990. The equilibrium partition function and base pair binding probabilities for RNA secondary structures. *Biopolymers* 29 (6-7), 1105–1119.
- Nowak, M., Schuster, P., 1989. Error thresholds of replication in finite populations mutation frequencies and the onset of muller’s ratchett. *J. Theor. Biol.* 137, 375–395.
- Reidys, C., Forst, C., Schuster, P., 2001. Replication and mutation on neutral networks. *Bull. Math. Biol.* 63 (1), 57–94.
- Schuster, P., 2001. Evolution in silico and in vitro: The RNA model. *Biol. Chem.* 382 (9), 1301–1314.
- Shapiro, B. A., 1988. An algorithm for comparing multiple RNA secondary structures. *Comput. Appl. Biosci.* 4 (3), 387–393.
- Shapiro, B. A., Zhang, K., 1990. Comparing multiple RNA secondary structures using tree comparisons. *Comput. Appl. Biosci.* 6 (4), 309–318.
- Takeuchi, N., Poorthuis, P. H., Hogeweg, P., 2005. Phenotypic error threshold;

- additivity and epistasis in RNA evolution. *BMC Evol. Biol.* 5 (9).
- Wilke, C. O., Lenski, R. E., Adami, C., 2003. Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding. *BMC Evol. Biol.* 3 (3).
- Zuker, M., Stiegler, P., 1981. Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information. *Nucl. Acid. Res.* 9 (1), 133–148.

FIGURES

Figures are scaled as they may appear in the journal.

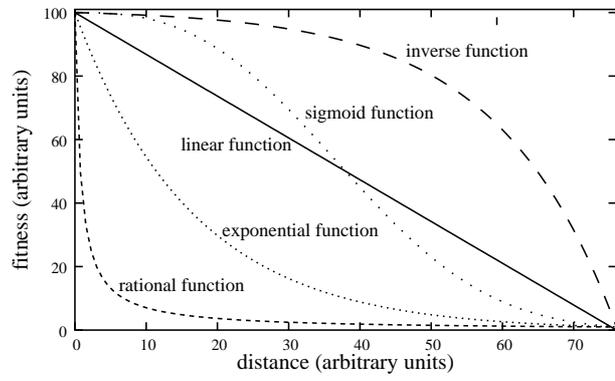
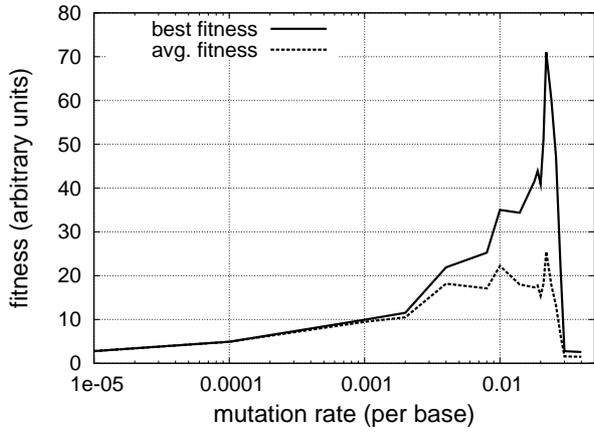
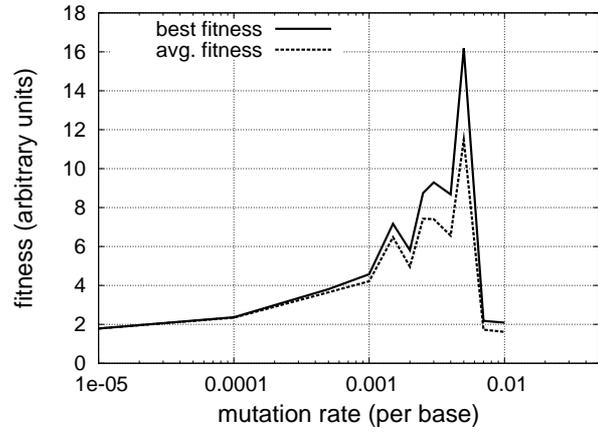


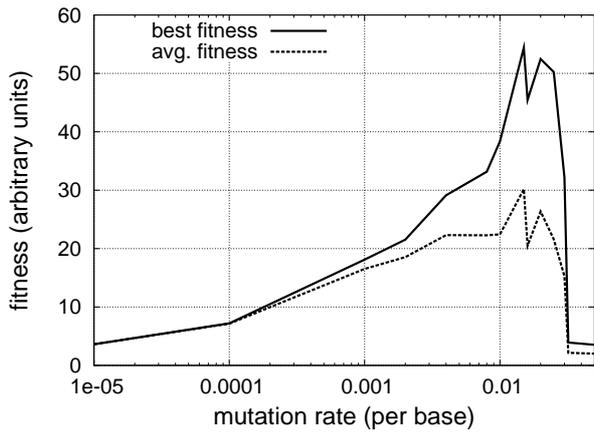
Fig. 1. Illustration of the scaling functions investigated.



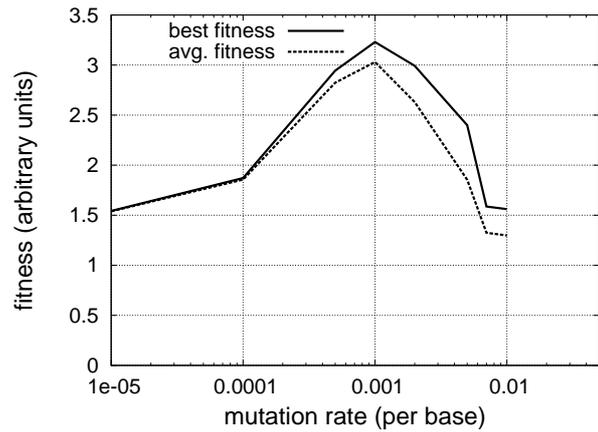
(a) Glob: $d_1(s, s_{\text{target}})$



(b) Add: $d_2(s, s_{\text{target}})$

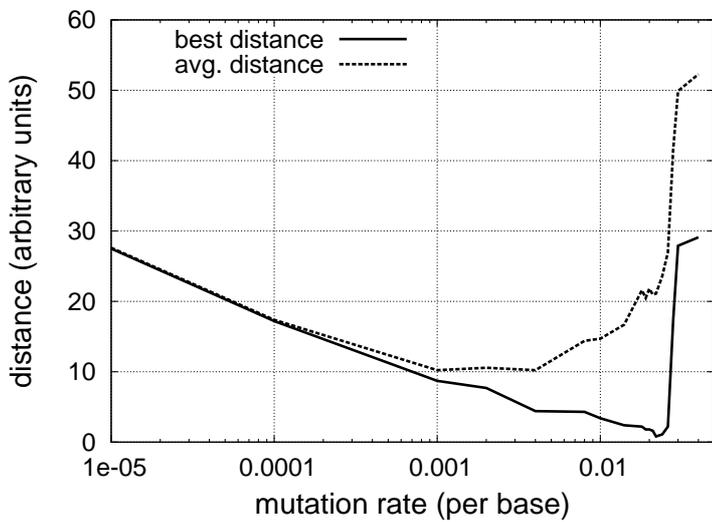


(c) Mult: $d_3(s, s_{\text{target}})$

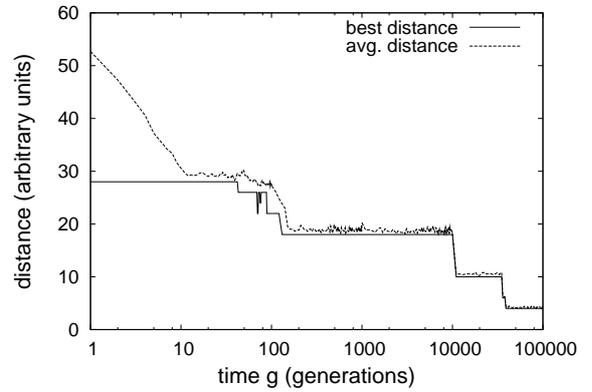


(d) LSSA: $d_4(s, s_{\text{target}})$

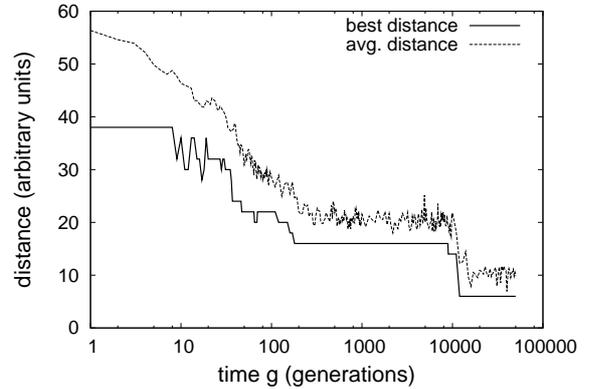
Fig. 3. Best and average fitness after $g = 10000$ generations vs. mutation rate p_{mut} . Population size $\lambda = 100$, fitness function $f(s) = f_{\text{scale}}^3(d_i(s, s_{\text{target}}))$; mean of 20 (a) and 10 (b)-(d) independent experiments, respectively. The fluctuations in a graph are due to the stochastic nature of the simulation model. They indicate the stochastic measuring error. Error bars omitted for clarity.



(a) Best and average distance after $g = 10000$, mean of 20 experiments

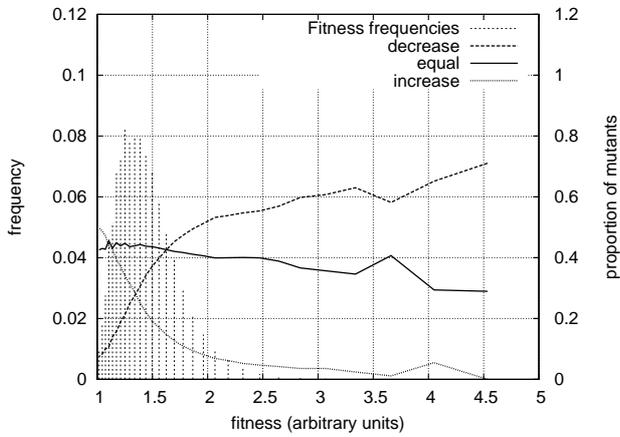


(b) $p_{\text{mut}} = 0.001$

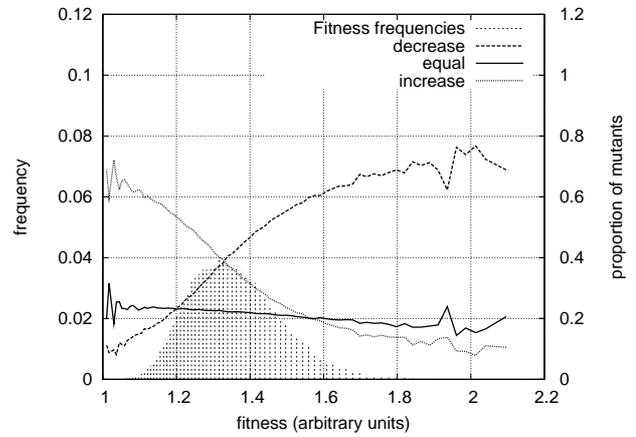


(c) $p_{\text{mut}} = 0.003$

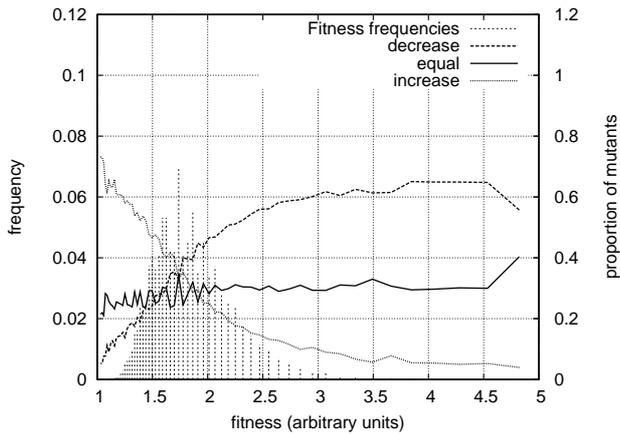
Fig. 4. Illustration of the divergence threshold $p_{\text{mut}} \approx 0.001$ for $g = 10000$: At mutation rate $p_{\text{mut}} \approx 0.001$, the average distance to the target structure increases while the best distance continue to decrease. The qualitative difference can be observed by comparing two typical single experiments with (b) $p_{\text{mut}} = 0.001$ and (c) $p_{\text{mut}} = 0.003$, respectively. Other parameters: “conventional” fitness function $f(s) = f_{\text{scale}}^3(d_1(s, s_{\text{target}}))$, population size $\lambda = 100$.



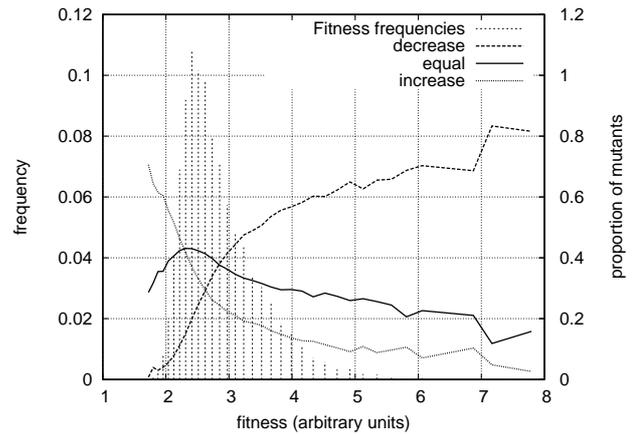
(a) Glob: $d_1(s, s_{\text{target}})$, $n = 100000$



(b) Add: $d_2(s, s_{\text{target}})$, $n = 103031$

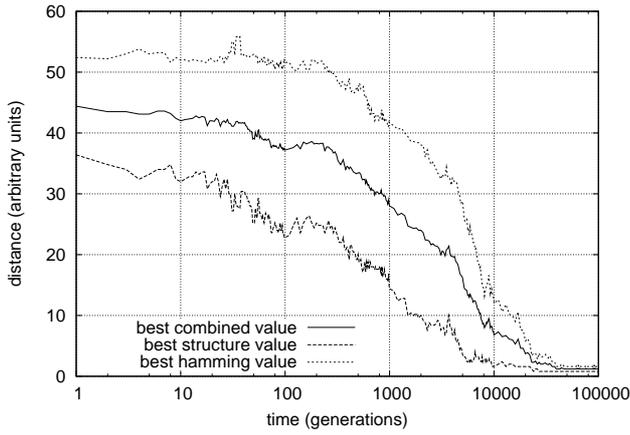


(c) Mult: $d_3(s, s_{\text{target}})$, $n = 87704$

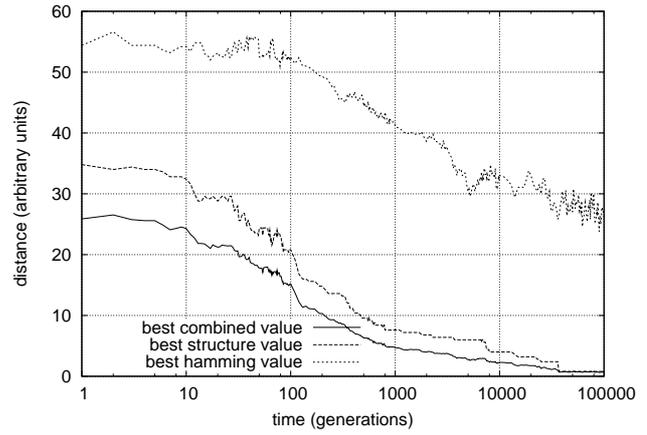


(d) LSSA: $d_4(s, s_{\text{target}})$, $n = 16450$

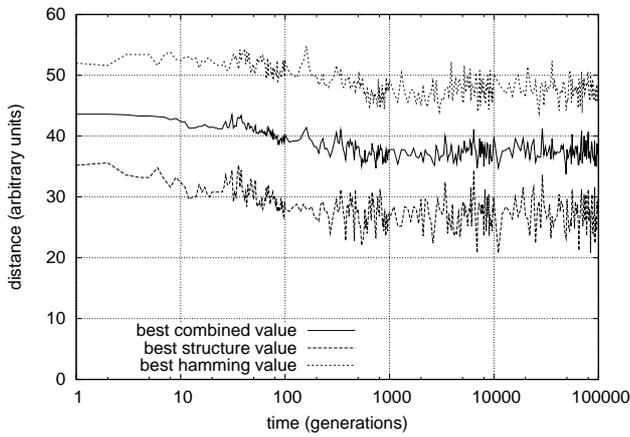
Fig. 5. Estimated probability to obtain an offspring with improved, neutral, or decreased fitness after a one-point-mutation of a randomly created parent vs. the fitness of this parent (horizontal axis). Furthermore a fitness histogram of the randomly generated sequences is plotted. Each figure is based on n samples.



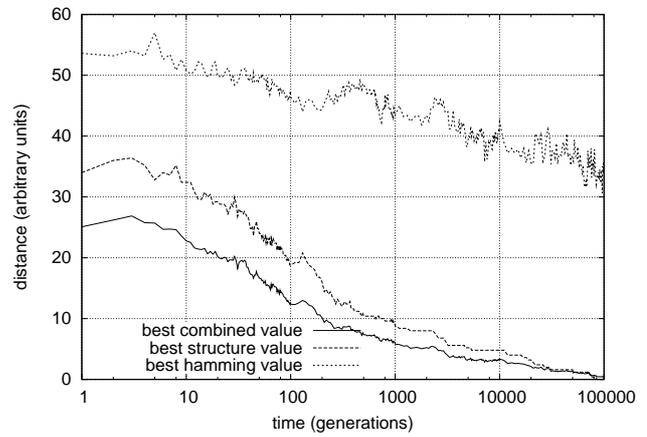
(a) Add: $d_2(s, s_{\text{target}})$, $p_{\text{mut}} = 0.003$



(b) Mult: $d_3(s, s_{\text{target}})$, $p_{\text{mut}} = 0.003$

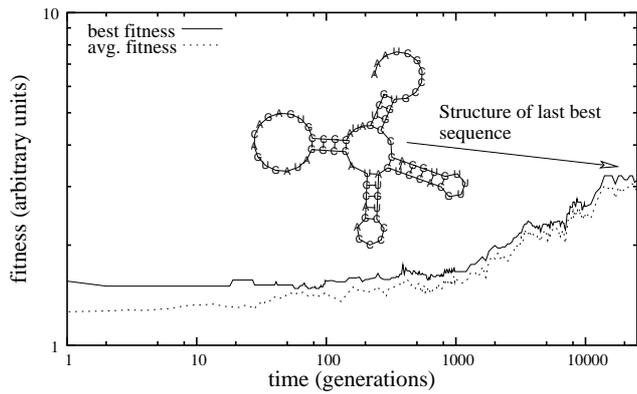


(c) Add: $d_2(s, s_{\text{target}})$, $p_{\text{mut}} = 0.01$

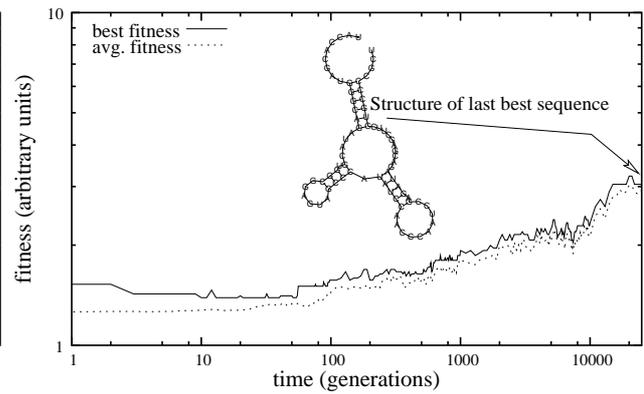


(d) Mult: $d_3(s, s_{\text{target}})$, $p_{\text{mut}} = 0.01$

Fig. 6. Contribution of structure and sequence to fitness. **Left:** additive combination d_2 . **Right:** multiplicative combination d_3 . **Upper:** low mutation rate $p_{\text{mut}} = 0.003$, which was chosen just below the error threshold of the additive combination d_2 . **Lower:** high mutation rate $p_{\text{mut}} = 0.01$, which was chosen just below the error threshold of the multiplicative combination d_3 . The figures show the time evolution of the best population fitness together with the best value contributed by the structural distance and the hamming distance, respectively. The curves are averaged over 5 independent experiments for each parameter setting. Since $p_{\text{mut}} = 0.01$ is above the error threshold of the additive combined sequence structure alignment, there is essentially no adaption in Fig. (c). Other parameters: rational scaling $f_{\text{scale}}^3(d)$, population size $\lambda = 100$.



(a) continuous similarity values



(b) discrete similarity values

Fig. 7. Typical single experiments with local sequence structure alignment (LSSA) and conventional rational scaling. The best structure after $g = 25000$ is shown. (a) Fitness function $f_{\text{scale}}^3(d_4(s, s_{\text{target}}))$. (b) Discretized fitness function $f_{\text{scale}}^3(\text{round}(d_4(s, s_{\text{target}})))$. Other parameters: mutation rate $p_{\text{mut}} = 0.001$, population size $\lambda = 100$, insertion and deletion not allowed.

Captions

Caption 1

Illustration of the scaling functions investigated.

Caption 2

Target structure and sequence used in all experiments (taken from Fontana and Schuster (1998)).

Caption 3

Best and average fitness after $g = 10000$ generations vs. mutation rate p_{mut} . Population size $\lambda = 100$, fitness function $f(s) = f_{\text{scale}}^3(d_i(s, s_{\text{target}}))$; mean of 20 (a) and 10 (b)-(d) independent experiments, respectively. The fluctuations in a graph are due to the stochastic nature of the simulation model. They indicate the stochastic measuring error. Error bars omitted for clarity.

Caption 4

Illustration of the divergence threshold $p_{\text{mut}} \approx 0.001$ for $g = 10000$: At mutation rate $p_{\text{mut}} \approx 0.001$, the average distance to the target structure increases while the best distance continue to decrease. The qualitative difference can be observed by comparing two typical single experiments with (b) $p_{\text{mut}} = 0.001$ and (c) $p_{\text{mut}} = 0.003$, respectively. Other parameters: “conventional” fitness

function $f(s) = f_{\text{scale}}^3(d_1(s, s_{\text{target}}))$, population size $\lambda = 100$.

Caption 5

Estimated probability to obtain an offspring with improved, neutral, or decreased fitness after a one-point-mutation of a randomly created parent vs. the fitness of this parent (horizontal axis). Furthermore a fitness histogram of the randomly generated sequences is plotted. Each figure is based on n samples.

Caption 6

Contribution of structure and sequence to fitness. **Left:** additive combination d_2 . **Right:** multiplicative combination d_3 . **Upper:** low mutation rate $p_{\text{mut}} = 0.003$, which was chosen just below the error threshold of the additive combination d_2 . **Lower:** high mutation rate $p_{\text{mut}} = 0.01$, which was chosen just below the error threshold of the multiplicative combination d_3 . The figures show the time evolution of the best population fitness together with the best value contributed by the structural distance and the hamming distance, respectively. The curves are averaged over 5 independent experiments for each parameter setting. Since $p_{\text{mut}} = 0.01$ is above the error threshold of the additive combined sequence structure alignment, there is essentially no adaption in Fig. (c). Other parameters: rational scaling $f_{\text{scale}}^3(d)$, population size $\lambda = 100$.

Caption 7

Typical single experiments with local sequence structure alignment (LSSA) and conventional rational scaling. The best structure after $g = 25000$ is shown. (a) Fitness function $f_{\text{scale}}^3(d_4(s, s_{\text{target}}))$. (b) Discretized fitness function $f_{\text{scale}}^3(\text{round}(d_4(s, s_{\text{target}})))$. Other parameters: mutation rate $p_{\text{mut}} = 0.001$, population size $\lambda = 100$, insertion and deletion not allowed.