Computational Design of a Circular RNA with Prionlike Behavior

Stefan Badelt**
Christoph Flamm**,†
Ivo L. Hofacker*,**,†,‡
University of Vienna

Keywords
RNA structure, sequence design, self-replication, folding kinetics

Abstract RNA molecules engineered to fold into predefined conformations have enabled the design of a multitude of functional RNA devices in the field of synthetic biology and nanotechnology. More complex designs require efficient computational methods, which need to consider not only equilibrium thermodynamics but also the kinetics of structure formation. Here we present a novel type of RNA design that mimics the behavior of prions, that is, sequences capable of interaction-triggered autocatalytic replication of conformations. Our design was computed with the ViennaRNA package and is based on circular RNA that embeds domains amenable to intermolecular kissing interactions.

1 Introduction

During the last decade, the field of synthetic biology has impressively illustrated that nucleic acids and in particular RNA molecules are reliable materials for the design and implementation of functional circuits as well as nano-scale devices and objects [18, 24, 1, 23]. The reasons for this success are grounded in the facts that for RNA (i) an experimentally measured energy model exists, (ii) regulation at the level of RNA molecules is faster than via the production of proteins, and (iii) design questions are more readily expressed in the discrete framework of binary base pairing than in continuous interactions between, say, the amino acids in proteins.

1.1 RNA Design

RNA molecules have been extensively engineered in the classical context of gene regulation. Successful designs include Boolean networks with miRNAs [27, 34, 47], synthetic RNA switches [22, 40, 17], and artificial ribozymes [43, 13]. In general, the RNA is optimized to switch conformations upon a hybridization interaction, in order to toggle between an ON and an OFF state. Turberfield et al. [38] showed for DNA that such hybridization reactions are reversible by a mechanism called *toehold exchange*, which was later used to design reversible logic circuits [14] as well as transcriptional oscillators [25] and transcription regulators [37]. Recently, multiple toehold switches have been integrated

* Contact author.
** Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17/3, 1090 Vienna, Austria. E-mail: stef@tbi.univie.ac.at (S.B.); xtof@tbi.univie.ac.at (C.F.); ivo@tbi.univie.ac.at (I.L.H.)
† Forschungsverbund Chemistry meets Microbiology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria.
‡ Research Group Bioinformatics and Computational Biology, University of Vienna, Währingerstraße 29, 1090 Vienna, Austria.

Artificial Life 22: 172–184 (2016) doi:10.1162/ARTL_a_00197

in vivo to activate gene expression in response to endogenous RNAs [17]. Cascading of strand-displacement reactions allows for the construction of multi-component chemical reaction networks, exhibiting complex computational and information-processing abilities [50, 41].

DNA and RNA hybridization effects are essential features for algorithmic self-assembly [36, 48] in nanotechnology. RNA nanoparticles have been constructed from smaller self-assembling units [6, 5, 16]. Small RNA motifs, such as complementary kissing hairpins, can be used to assemble complex 1D, 2D, and 3D shapes (for a recent review see [15]). Cayrol et al. [5] have found natural self-assemblies of the small RNA DsrA in *E. coli*, suggesting concentration-dependent RNA regulatory mechanisms via self-assembly.

The inherent complexity of nucleic-acid-based self-assembling systems makes it necessary to optimize hand-crafted designs computationally. Several energy-directed computational methods have been devised for the rational design of nucleic acid molecules that fold into single [20] or multiple [11, 21] predefined conformations or that form ensembles of interacting nucleic acid strands [49, 44]. While these methods carefully model equilibrium properties of the designed RNAs, they provide rudimentary or no support for designing kinetic properties, such as refolding times, which are much more expensive to compute and thus remain a challenge for computational design methods.

1.2 Prions

The *protein-only hypothesis* for the scrapie agent (for a review see [2]) proposes that a prion protein, with an altered (infectious) β-sheet-rich conformation, starts an autocatalytic cascade that uses the normally folded prion proteins as a substrate, converting them to the infectious form. This altered conformation then either self-assembles into fibers, which is the usual phenotype upon prion infection, or catalyzes the refolding of the remaining normally folded prions. A high activation energy between the normal and the infectious conformation prevents spontaneous conversion at detectable rates. The formation of a normal–infectious heteromeric complex lowers the activation energy barrier to convert the normally folded protein into an infectious species. This conversion leads to further recruitment of normally folded proteins in an autocatalytic process. In essence, a single infectious prion protein in a population of normally folded ones is enough to convert the whole population via autocatalytic structure replication into an all-infectious protein population, which self-assembles into long fibers.

1.3 Artificial Life

Prions represent a form of *conformational self-replication* that is so far not observed in the context of RNA biology. Minimal self-switching RNA prions can serve as a model for (i) a new class of riboswitches that provide exponential feedback, or (ii) self-induced nano-units that assemble or disassemble upon stimulation. Importantly, the computational design aspect allows for context-sensitive optimization, which is necessary for applications in synthetic biology and nanotechnology. Previous research in the field of bottom-up synthetic biology has already introduced designs of small self-assembling (for a recent review see [1]), self-replicating [35], and self-polymerizing [33] systems. While all of these results are valuable in showing that such designs are indeed possible, many designs are still done by intuition and hardly adjustable to a context-sensitive implementation such as prospective artificial life forms. Along those lines we recently submitted an experimentally verified computational design of multiple self-polymerizing ribozymes in order to study the dynamics of self-interactive systems [32].

This contribution was motivated by the question of whether RNA molecules can be designed in silico to exhibit the aforementioned prionlike behavior. We show that it is indeed possible to design such an *RNA prion*; whether the suggested sequence really shows the exponential refolding characteristics awaits experimental verification. The RNA prion presented here is a 49-nt-long, circular

RNA, designed as a bistable molecule. It thermodynamically favors one structure (S_1) if present as a monomer and the other structure (S_2) if present as a dimer.

2 Thermodynamics and Kinetics of RNA Prions

RNA secondary structure is a good approximation for the real RNA structure because, in contrast to proteins, RNA secondary structure captures the majority of the folding free energy. Furthermore, a well-established energy model for RNA secondary structures exists that is extensively parameterized via melting experiments [39]. On the computational side, algorithms have been developed for the thermodynamic and kinetic characterization of RNA secondary structures [28]. In particular, the energetically optimal structure as well as properties of the structural ensemble at thermodynamic equilibrium can be computed efficiently. The topology of the discrete folding landscape [12], which has a strong influence on the folding kinetics, can be analyzed in detail. Several approaches model the folding kinetics of RNA secondary structure as a stochastic process with different resolution of the folding landscape [9, 45, 26] (for a review see [10]). Recently these approaches have been extended to operate on folding landscapes that change with time, as in the case of folding during transcription [19].

2.1 Thermodynamics

Let the set $\mathbf{\Omega}$ of RNA secondary structures be restricted to those that are formed from nested isosteric base pairs (GC, AU, GU), which have a minimal hairpin loop size of 3 nt and a maximum interior loop size of 30 nt. These restrictions are broad enough to include the vast majority of known pseudoknot-free RNA conformations, and they define a set of structures for which an experimentally determined energy model E exists [29]. Most importantly from the computational perspective, these definitions allow us to compute the structure of minimum free energy (MFE) and the equilibrium partition function (Z) in $O(n^3)$ time, where n is the sequence length.

For RNA design, a fast computation of the equilibrium partition function Z is of particular interest, since it allows for computing the probability $P(S)$ of forming a secondary structure S and the ensemble free energy G. Let Z be the sum of all Boltzmann-weighted energy contributions in the ensemble of RNA secondary structures $\mathbf{\Omega}$,

$$Z = \sum_{S \in \mathbf{\Omega}} e^{-\frac{E(S)}{RT}}. \tag{1}$$

Then we can compute the probability of any secondary structure as

$$P(S) = \frac{e^{-\frac{E(S)}{RT}}}{Z} \tag{2}$$

and the free energy of the ensemble as

$$G = -kT \cdot \ln(Z). \tag{3}$$

2.2 Kinetic Folding

Let us denote an RNA energy landscape as $\mathscr{L} = (\mathbf{\Omega}, \mathscr{M}, E)$, where $\mathbf{\Omega}$ is the previously introduced set of secondary structures, $\mathscr{M}$ is a move set comprising all possible transitions between structures, and E is an energy function assigning a fitness value for each secondary structure. For an ergodic move set $\mathscr{M}$, we choose the simplest reversible modification of an RNA structure, the opening and closing of a single base pair. If we have the full landscape $\mathscr{L}$, we can calculate folding kinetics as a

continuous-time Markov process where the influx and outflux rates, k_{ij} and k_{ji} respectively, between every two neighboring structures S_i and S_j determine the population $\pi_i(t)$ of a structure S_i at time point t. This transition process can be formulated with the master equation

$$\frac{d\pi_i(t)}{dt} = \sum_{j \neq i} \left(\pi_j(t) k_{ji} - \pi_i(t) k_{ij} \right). \tag{4}$$

An RNA energy landscape as defined above grows exponentially with sequence length, making exact folding simulations infeasible for molecules longer than 30 to 40 nucleotides. However, refolding times between two conformations correlate with the smallest energy barrier β separating them. Formally, any folding path $\mathcal{P}$ has a saddle point with energy $E_{\mathcal{P}} = \max_{s \in \mathcal{P}} E(s)$, and the energy barrier separating two minima S_1 from S_2 is determined by the lowest among all possible paths $\mathcal{P}_{S_1 \to S_2}$:

$$\beta_{S_1 \to S_2} = \min_{\mathcal{P}_{S_1 \to S_2}} E_{\mathcal{P}} - E(S_1) = \min_{\mathcal{P}} \max_{s \in \mathcal{P}} E(s) - E(S_1). \tag{5}$$

Finding the best folding path has been shown to be a NP-hard problem [30]. However, there exist fast heuristics to compute direct (shortest) paths between two structures, such as the `findpath` [11] method available in the ViennaRNA package, as well as indirect paths using, for example, `RNAtabupath` [7]. For sequences shorter than about 100 nt, paths with a minimal energy barrier can be computed exactly with `RNAsubopt` [46] and `barriers` [12]. The first program computes a list of all RNA secondary structures within a certain energy range; `barriers` processes a sorted list of these conformations to compute all local minima and the saddle points connecting them by a flooding algorithm.

2.3 Landscapes of RNA Prions

RNA prions are bistable molecules that favor one structure (S_1) if present as a monomer and the other structure (S_2) upon dimerization. To avoid spontaneous refolding between S_2 and S_1, the energy barrier $\beta_{S_2 \to S_1}$ has to be very high (see Figure 1).

In our design, S_2 forms two 10-nt hairpins that are prone to hybridize via a kissing interaction as reported for the HIV-DIS loop [8], a highly conserved stem–loop sequence found in many retroviruses. Importantly, S_2 should not only stabilize other molecules in S_2 conformation at high concentrations, but actively lower the energy barrier to refold S_1 into S_2 (see Figure 1).

If these landscape properties are fulfilled, it ensures that an initial population of *only* S_1 molecules will not refold into S_2 unless the refolding is triggered by an external mechanism. However, as soon as a small population of molecules in S_2 conformation is present, they can catalyze the refolding of S_1 molecules into S_2.

3 Design of an RNA Prion

RNA molecules with complex energy landscapes can be designed in a two-step process. First, a fast heuristic is used to generate and select promising candidates according to a cost function $\mathcal{C}$ that specifies thermodynamic aspects of the design objective. Second, minimal refolding barriers are computed for the set of candidate molecules and contribute to the final ranking of molecules (see Figure 2 for the prion design pipeline).

3.1 Cost Function

The cost function used for RNA prion design consists of two parts: $\mathcal{C} = \mathcal{C}_{\mathcal{M}} + \mathcal{C}_{\mathcal{D}}$. Here $\mathcal{C}_{\mathcal{M}}$ is the cost function to optimize bistable molecules (see Figure 1) without considering the refolding barrier:

$$\mathcal{C}_{\mathcal{M}} = E(S_1) - G + \alpha((E(S_1) - G_1) + (E(S_2) - G_2)) + \alpha(E(S_1) - E(S_2) + \epsilon)^2 \tag{6}$$

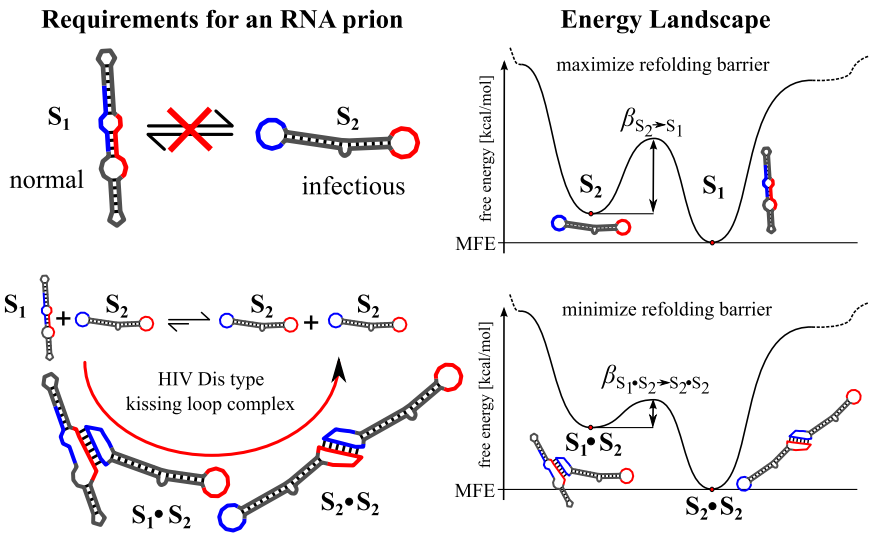

Figure 1. Schematic requirements for an RNA prion. Red and blue parts of the molecule are complementary and can form an intermolecular hybridization reaction. *Upper panel*: S_1 and S_2 are stable conformations that do not refold into each other, since the conformations are separated by a high energy barrier in the energy landscape. *Lower panel*: S_2 destabilizes S_1 with a HIV-Dis type kissing loop interaction. The energy barrier for the $S_1 \bullet S_2$ complex to refold into the $S_2 \bullet S_2$ complex is low and therefore allows spontaneous refolding.

where G is the ensemble free energy (see Equation 3), G_1 and G_2 are constrained ensemble free energies over all structures that form base pairs exclusively possible for S_1 and S_2 respectively (see Figure 3), ϵ specifies the desired energy difference between S_1 and S_2, and α is a constant to weight the terms. Thus, the first term ensures that S_1 is the ground state, while the next two terms optimize for sequences that have few alternative structures in the vicinity of S_1 and S_2.

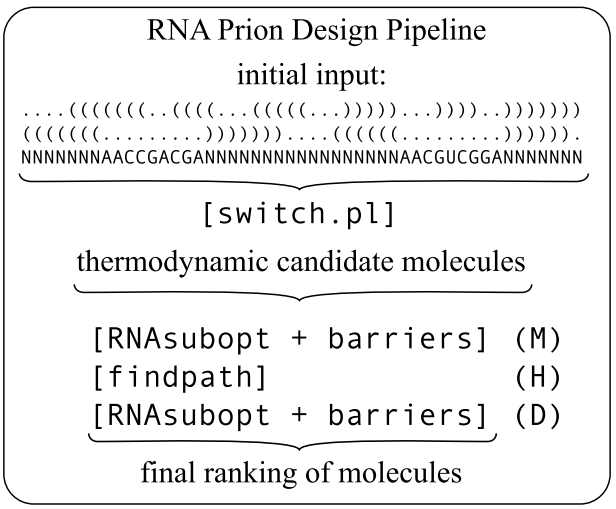

Figure 2. Programs used to design RNA prions. The initial input consists of two secondary structures in dot–bracket notation (i.e., every unpaired nucleotide is represented by a dot, and base pairs are shown as matching parentheses; see Figure 6 for the same structures shown in a graphical representation) and a sequence constraint for the kissing interaction. `switch.pl` returns a set of candidate molecules, which are analyzed with `RNAsubopt` and `barriers` to compute the minimum barrier height for monomer refolding (M) and dimer refolding (D). `findpath` is used to approximate the initial hybridization interaction of two monomers in structures S_1 and S_2 (H).

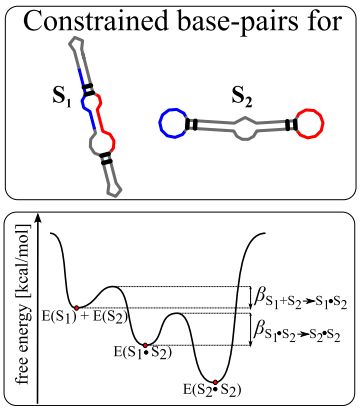

Figure 3. *Upper box*: Constraints for the monomer landscape. Among all structures that can form, the base pairs shown (S_1 or S_2) should be the best. The base pairs are chosen such that no structures can fulfill both constraints. *Lower box*: Optimization parameters for the dimer landscape. The energy of the hybridization complex $S_1 \bullet S_2$ should be half way between the sum of the monomer energies and the kissing complex $S_2 \bullet S_2$. The barrier heights β determine the ranking of final candidates.

$\mathcal{C}_\mathcal{D}$ optimizes the dimer landscape, so that the energy of the transition state formed by the initial intermolecular hybridization interaction ($S_1 \bullet S_2$) lies approximately half way between the energy of single molecules $E(S_1) + E(S_2)$ and the ground state of the kissing dimer interaction ($S_2 \bullet S_2$):

$$\mathcal{C}_\mathcal{D} = \left(\frac{E(S_2 \cdot S_2) + (E(S_1) + E(S_2))}{2} - E(S_1 \cdot S_2) \right)^2. \tag{7}$$

Finally, we rank the molecules by the difference of refolding barriers Δβ for monomers and dimers. Since refolding of dimers is a two-step process composed of (i) the initiation of the kissing action and (ii) the subsequent intramolecular refolding (see Figure 3), we only consider the rate-limiting step, that is, the one with the larger of the two barriers:

$$\Delta\beta = \beta_{S_2 \to S_1} - \max\left(\beta_{S_1 + S_2 \to S_1 \cdot S_2}, \beta_{S_1 \cdot S_2 \to S_2 \cdot S_2}\right). \tag{8}$$

3.2 Energy Evaluation

RNA kissing interactions go beyond normal pseudoknot-free secondary structures as defined above, since they comprise non-nested base pairs, and their energies can therefore not be computed from the standard energy model. While we expect the energies of the intermolecular helix to be well described by standard energy parameters, it is less predictable how the (mostly entropic) contribution of the hairpin loops involved in the kiss will change. Thermodynamic stabilities of kissing interactions similar to the HIV–DIS loop, which features a 6-nt intermolecular interaction with a 9-nt hairpin loop, were studied in detail by Weixlbaumer et al. [42]. By varying the loop sequence they demonstrated that normal stacking energies can indeed be used for the intermolecular helix. More importantly, they found kissing interactions to be surprisingly more stable than other hybridization structures. Based on the measurements in [42], we compute the energy of kissing hairpins as the energy of the intermolecular helix without energy bonuses for single-base stacking of adjacent unpaired bases (*dangling ends*) and add a loop energy of −4.2 kcal/mol. In contrast, other intermolecular hybridization structures are penalized by the usual intermolecular initiation energy of +4.1 kcal/mol, also confirmed by Weixlbaumer et al. [42].

These energy parameters enable us to accurately evaluate the energy of $S_2 \bullet S_2$, but, unfortunately, they are not sufficient for creating a consistent energy model for all intermediate conformations

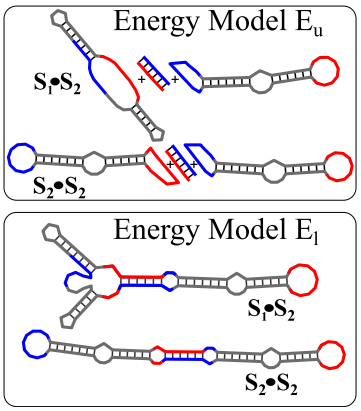

Figure 4. Two energy models to score the hybridization reactions also depicted in Figure 1. E_u is the sum of energies for S_1, S_2, and the hybridization reaction; E_l transforms the molecule into a hybrid that can be scored with standard energy parameters.

along the refolding paths $P_{S_1+S_2 \to S_1 \cdot S_2}$ and $P_{S_1 \cdot S_2 \to S_2 \cdot S_2}$. We therefore approximate the energy barrier $\beta_{S_1+S_2 \to S_1 \cdot S_2}$ with the `findpath` heuristic and compute the best energy barrier $\beta_{S_1 \cdot S_2 \to S_2 \cdot S_2}$, using the `RNAsubopt–barriers` approach with two different energy models that serve as an upper and a lower bound (see Figure 4).

The upper bound dimer energy model E_u to find the lowest energy barrier $\beta_{S_1 \cdot S_2 \to S_2 \cdot S_2}$ is computed from the energy of the structure formed from monomers 1 and 2—$E(M_1)$ and $E(M_2)$, respectively—and the energy of the duplex interaction is stabilized by a bonus of −4.2 kcal/mol:

$$E_u(D) = E(M_1) + E(M_2) + E\left(\text{dup}_{12}\right) - 4.2. \tag{9}$$

This approach implicitly adds a penalty for *unpaired* loop regions, which are actually involved in the hybridization interaction, and therefore underestimates the actual energy of $S_2 \bullet S_2$.

Alternatively, the kissing interaction can be modeled as a regular intramolecular helix. If both molecules are cut after the 5′-AA of the interacting loops and connected to the beginning of the respective other strand, this results in a circular hybrid monomer that can be evaluated by standard energy parameters. Since the free-energy bonus for kissing interactions (−4.2 kcal/mol) approximately compensates for the entropic penalty of intermolecular interactions (+4.1 kcal/mol), $S_2 \bullet S_2$ is closer to the energy according to Weixlbaumer et al. [42]. Using this energy model, we increment the degree of loops involved in the kissing interaction as if a regular helix were formed, and thereby increase the entropic loop penalty. The interior loop of S_1 becomes a multi-loop, and the hairpin loop of S_2 becomes an interior loop. We have

$$E_l(D) = E\left(\text{Hybrid}_{M1 \cdot M2}\right). \tag{10}$$

All optimization steps are based on the energy model E_u, for analysis of the best sequences both energy models were applied, and we await experimental feedback to decide which model is better.

The barrier for the initiation of the intermolecular hybridization $\beta_{S_1+S_2 \to S_1 \cdot S_2}$ is computed by first finding the best path for opening the competing helix in S_1 and second computing the energy barrier for the intermolecular hybridization reaction using the standard penalty of +4.1 kcal/mol. The bonus energy of −4.2 kcal/mol is added once the full kissing region is formed, to be consistent with the energy model E_u. The energy barrier for the initiation of the duplex formation is therefore usually either the last base pair in the process of unfolding the competing helix, or the first inserted base pair towards the kissing interaction.

3.3 switch.pl

switch.pl [11] of the ViennaRNA package [28] was used to design bistable molecules and was modified to support the novel cost function composed of Equations 6 and 7 as well as the folding of circular RNAs. The algorithm first builds a dependence graph in order to efficiently and fairly sample RNA sequences that are compatible with both structural constraints. Since switch.pl can only design bistable and not tristable sequences, the structural constraint for the kissing interaction was specified indirectly as a sequence constraint. This reduces the number of candidate molecules, but ensures that they always have an experimentally validated, stable kissing interaction.

The chosen sequences for the kissing interaction have a similar free hybridization energy to that of the best kissing interaction examples shown in Weixlbaumer et al. [42], but differ by point mutations to be compatible with structural constraints for S_1 and S_2. Importantly, (i) the kissing hairpins cannot form intramolecular base pairs that would compete with the formation of an intermolecular kissing interaction, and (ii) the two complementary regions forming the kissing interaction should not form an intramolecular helix in S_1, which would make them inaccessible for intramolecular interactions. The asymmetric design shown in Figure 1 allows S_2 to open a shorter helical region that has a less stable free energy than the subsequent formation of the kissing interaction.

Alternatively, one could use RNAdesign [21], which can build dependence graphs for multiple structural constraints and therefore increases the search space, allowing for novel kissing interactions. This more general design attempt will be implemented upon experimental feedback for the design presented here.

3.4 RNAsubopt, barriers, findpath

The candidate molecules computed by switch.pl are subsequently ranked by the difference of barrier heights for single-molecule refolding and kissing-dimer refolding. Computation for monomer refolding is straightforward, by computing the suboptimal structures for the monomer using RNAsubopt followed by evaluation of the minimal barrier height $\beta_{S_2 \to S_1}$ with barriers. For the kissing-dimer interaction, suboptimal structures were computed for both energy models described above. When using the energy model E_u (see Equation 9), suboptimals were computed with the constraint that the region involved in the kissing interaction of S_1 (red in Figure 1) is unpaired (i.e., involved in the kissing interaction); according to the energy model E_l (see Equation 10), suboptimal structures were computed for the hybrid, with the constraint that the kissing region is paired, while the part corresponding to molecule 2 was kept constant in conformation S_2. The barrier for the initiation of the intramolecular hybridization is computed using findpath for opening the competing helix in S_1 and subsequently adding the duplex energies for formation of the kissing interaction.

4 Results

Out of all possible sequences that are compatible with the sequence and structure constraints shown in Figure 2, 158 different sequences were returned from switch.pl using 1000 individual runs with each 2×10^6 optimization steps. After postprocessing with RNAsubopt, barriers, and findpath to compute the final ranking of the generated designs, 69 sequences either turned out not to fold exactly into the ground state structures specified as input for switch.pl, or had a higher barrier for dimer refolding than for monomer refolding and were therefore excluded from further analysis. In the remaining pool of 89 sequences, 23 showed differences between monomer and dimer refolding barriers higher than 4 kcal/mol according to E_u (see Equation 9) for dimers. These sequences were visually inspected, and we selected a candidate (ACCUGGGAACCGGCGACCCAGGUUUUCGGAACCAACGUCGGAGGUUCCU) for demonstration of prion behavior, that has (i) a very high refolding barrier for monomer refolding with +16.70 kcal/mol, and (ii) an energy landscape with as few as possible competing local minima to S_1 and S_2. The dimer refolding barriers are +11.60 and +8.60 kcal/mol for E_u and E_l, respectively. In comparison, the molecule

presented in Badelt et al. [3] has the same barrier for monomer refolding (+16.70 kcal/mol), but +13.60 and +9.80 kcal/mol for prion-induced refolding.

Figure 5 shows the equilibrium between the two stable conformations S_1 and S_2 as a function of RNA concentration. The concentrations of monomers $[M]$ and dimers $[D]$ in equilibrium can be computed by the equilibrium partition function Z (see Equation 1):

$$K = \frac{[D]}{[M]^2} = \frac{Z_D}{Z_M^2}. \tag{11}$$

The equilibrium partition function of the monomer (Z_M) can be computed directly using the McCaskill algorithm [31] implemented in `RNAfold`; Z_D can be approximated as

$$Z_D = Z_{c1} \cdot Z_{c2} \cdot Z_{\text{dup}} \tag{12}$$

with Z_{c1} and Z_{c2} denoting the partition functions of two monomers under the constraint that the blue (c1) or red (c2) interaction region (see Figure 3) is unpaired and thus available for forming an intermolecular (kissing) interaction. Z_{dup} is the partition function of the intermolecular duplex formed between the two molecules. This model follows the assumption that dimerization can only involve an interaction between the strands of the kissing interaction.

Since we are interested in the conformations formed upon monomerization and dimerization, we divided the total partition function Z_M into three parts: Z_{S_1}, Z_{S_2}, and Z_o. Here Z_{S_1} and Z_{S_2} contain all conformations constrained to form base pairs that can only be formed in structure S_1 or S_2, respectively, whereas Z_o contains all other conformations, that is, conformations that are not compatible with both constraints. Constraints are chosen such that (i) the helices formed by S_1 and S_2 are preserved and (ii) there are no structures fulfilling both constraints (see Figure 3). We computed the relative concentration of S_1 in monomers and dimers as

$$[S_1] = \frac{Z_{S_1}}{Z_M} \cdot [M] + \left(\frac{Z_{S_1 + c1}}{Z_{c1}} + \frac{Z_{S_1 + c2}}{Z_{c2}} \right) \cdot [D] \tag{13}$$

where Z_{S_1+c1} stands for a partition function that has both the constraint to fold into structure 1 (S_1) and the constraint to be unpaired in interaction region 1 ($c1$). Relative concentrations of S_2 were computed accordingly and can be seen in Figure 5.

Figure 6 shows the energy profiles of the best refolding paths between S_1 and S_2, either for a single RNA monomer or for an RNA engaged in kissing interaction with another molecule. Since S_1 is the thermodynamically favored state in monomers, we show the refolding path from S_2 (−13.70 kcal/mol) to S_1 (−16.00 kcal/mol) in the top panel. The barrier of this refolding path is 16.70 kcal/mol, making a non-induced switching of conformations unlikely.

The bottom panel of Figure 6 shows the energy profile for a scenario where an intermolecular interaction is first formed between one molecule in conformation S_1 and a second in S_2, followed by intramolecular refolding of the first molecule from S_1 into S_2. S_2 is now the favored conformation, since it is stabilized by the kissing interaction. In contrast, S_1 is destabilized, since one helix cannot be formed together with the intermolecular duplex. Theoretically, there would be a second possible duplex interaction that required S_1 to open eight base pairs in two helices, but since this interaction is not thermodynamically favored, it is not depicted in Figure 6.

For the initiation of the kissing interaction, all competing intramolecular base pairs of S_1 have to open first, and then the intermolecular base pairs can form. The energy barrier for this interaction is 8.90 kcal/mol and leads to a new local minimum conformation at −34.50 kcal/mol.

For the intramolecular refolding from S_1 to S_2, we compare the two energy models discussed above. The energy model E_u returns a barrier of 11.60 kcal/mol, while the energy model E_l results in

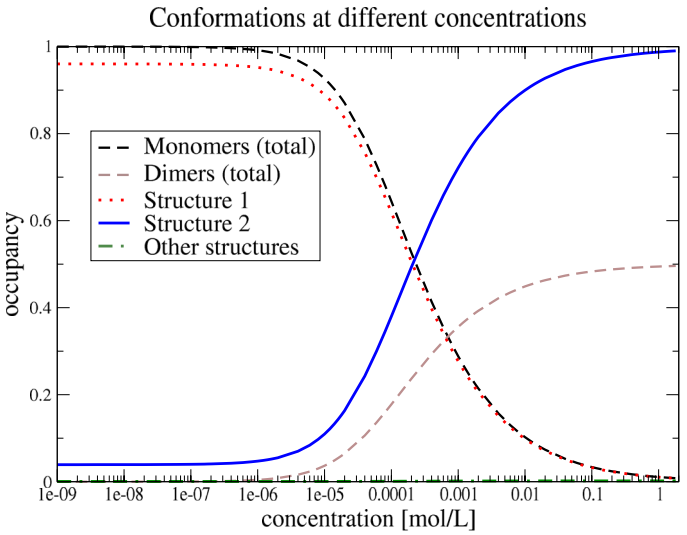

Figure 5. Conformational switching upon change of concentration. The transition from monomer to dimer conformations between 1 μM and 10 mM goes together with a switch from structure S_1 to S_2.

a barrier of 8.60 kcal/mol. Note also that the folding path itself is different, due to the different modeling of involved loop regions.

5 Conclusion

In this contribution we have shown that the computational design of RNA molecules that exhibit prion-like behavior is feasible, and that the computational machinery is developed enough for a rigorous

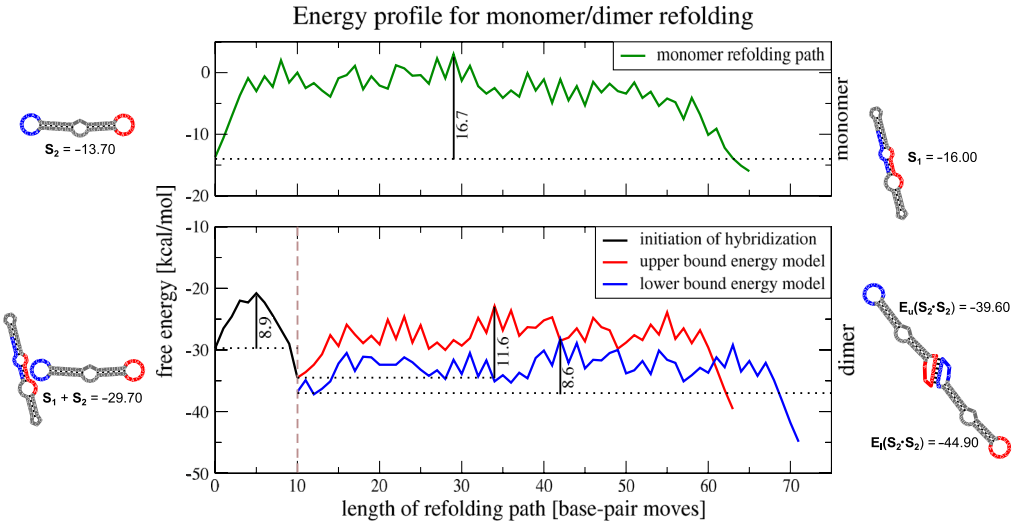

Figure 6. Energy profiles along the refolding path between structures S_1 and S_2. *Top panel*: refolding of a monomer; *bottom panel*: refolding while interacting with a second molecule. Blue- and red-colored regions are designed to form intermolecular base pairing. The lower panel shows a comparison between two energy models that differ in the energy contribution of the loop regions involved in the intermolecular pairing. In either case, the relative energy of refolding is lower than for the monomer.

analysis of the behavior of the resulting sequences. As in the original prion system, the misfolded conformation forms, via a kissing interaction, a heterodimeric complex with the native conformation. This interaction destabilizes the native conformation and triggers refolding into the misfolded conformation. Hence, we demonstrated, at least in silico, that RNA molecules possess the necessary structural capabilities for conformational replication. The calculations show that the kissing interaction drastically lowers the activation energy for refolding. Furthermore, the misfolded conformation can oligomerize. In principle, the oligomerization could inhibit the exponential growth characteristics of the misfolded conformation. Experimental results from fiber-forming dynamic combinatorial libraries [4] show that mechanical forces lead to fiber breaking, restoring the exponential growth characteristics. One difficulty in our computational design is the lack of energy parameters for complex interaction structures that occur as folding intermediates. The design process is, however, flexible and can incorporate feedback from wet lab results. We therefore envision that practical RNA designs should be refined in a few rounds of wet lab testing and adaptation of the computational models.

Conformational replication constitutes a novel regulatory mechanism possessing highly nonlinear dynamic characteristics. This type of behavior is necessary for the construction of signal-enhancing molecular circuits. Such amplifying devices are usually hard to construct with RNA. Our design could easily be coupled with interaction sites for an external signal molecule that triggers the initial refolding event. Such a device could detect a single molecular event and translate it into a large, easily detected signal. In a prospective artificial metabolism, the infectious conformation may be used to trigger transcriptional or translational riboswitches, and, if it were regulating its own transcription, the self-switching mechanism could be used for time-delayed feedback loops. The aspect of self-assembly to long fibers—which initially does not seem of particular interest, since it limits exponential growth—may provide a basic model towards the design of an RNA-based cytoskeleton. However, the current challenge is the synthesis of the RNA molecule in one particular conformation in vivo, which, for example, may be approached using techniques involving the tRNA ligase.

Acknowledgments

This work was supported in part by the FWF International Programme I670, the DK RNA program FG748004, the EU-FET grant RiboNets 323987, and the COST Action CM1304 "Emergence and Evolution of Complex Chemical Systems."

References

1. Afonin, K. A., Lindsay, B., & Shapiro, B. A. (2013). Engineered RNA nanodesigns for applications in RNA nanotechnology. *RNA Nanotechnology*, *1*, 1–15.
2. Aguzzi, A., Sigurdson, C., & Heikenwaelder, M. (2008). Molecular mechanisms of prion pathogenesis. *Annual Review of Pathology: Mechanisms of Disease*, *3*, 11–40.
3. Badelt, S., Flamm, C., & Hofacker, I. L. (2014). Computational design of a circular RNA with prion-like behaviour. In H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), *Artificial life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems* (pp. 565–568). Cambridge, MA: MIT Press.
4. Carnall, J. M. A., Waudby, C. A., Belenguer, A. M., Stuart, M. C. A., Peyralans, J. J.-P., & Otto, S. (2010). Mechanosensitive self-replication driven by self-organization. *Science*, *327*(5972), 1502–1506.
5. Cayrol, B., Nogues, C., Dawid, A., Sagi, I., Silberzan, P., & Isambert, H. (2009). A nanostructure made of a bacterial noncoding RNA. *Journal of the American Chemical Society*, *131*(47), 17270–17276.
6. Chen, C., Sheng, S., Shao, Z., & Guo, P. (2000). A dimer as a building block in assembling RNA. *Journal of Biological Chemistry*, *275*(23), 17510–17516.
7. Dotu, I., Lorenz, W. A., Van Hentenryck, P., & Clote, P. (2010). Computing folding pathways between RNA secondary structures. *Nucleic Acids Research*, *38*(5), 1711–1722.
8. Ennifar, E., Paillart, J.-C., Marquet, R., Ehresmann, B., Ehresmann, C., Dumas, P., & Walter, P. (2003). HIV-1 RNA dimerization initiation site is structurally similar to the ribosomal A site and binds aminoglycoside antibiotics. *Journal of Biological Chemistry*, *278*(4), 2723–2730.

9. Flamm, C., Fontana, W., Hofacker, I. L., & Schuster, P. (2000). RNA folding at elementary step resolution. *RNA*, *6*, 325–338.
10. Flamm, C., & Hofacker, I. L. (2008). Beyond energy minimization: Approaches to the kinetic folding of RNA. *Monatshefte für Chemie*, *139*(4), 447–457.
11. Flamm, C., Hofacker, I. L., Maurer-Stroh, S., Stadler, P. F., & Zehl, M. (2001). Design of multi-stable RNA molecules. *RNA*, *7*, 254–265.
12. Flamm, C., Hofacker, I. L., Stadler, P. F., & Wolfinger, M. T. (2002). Barrier trees of degenerate landscapes. *Zeitschrift für Physikalische Chemie*, *216*, 155–173.
13. Frommer, J., Appel, B., & Müller, S. (2015). Ribozymes that can be regulated by external stimuli. *Current Opinion in Biotechnology*, *31*, 35–41.
14. Genot, A. J., Bath, J., & Turberfield, A. J. (2011). Reversible logic circuits made of DNA. *Journal of the American Chemical Society*, *133*(50), 20080–20083.
15. Grabow, W. W., & Jaeger, L. (2014). RNA self-assembly and RNA nanotechnology. *Accounts of Chemical Research*, *47*(6), 1871–1880.
16. Grabow, W. W., Zakrevsky, P., Afonin, K. A., Chworos, A., Shapiro, B. A., & Jaeger, L. (2011). Self-assembling RNA nanorings based on RNAI/II inverse kissing complexes. *Nano Letters*, *11*(2), 878–887.
17. Green, A. A., Silver, P. A., Collins, J. J., & Yin, P. (2014). Toehold switches: De-novo-designed regulators of gene expression. *Cell*, *159*(4), 925–939.
18. Guo, P. (2010). The emerging field of RNA nanotechnology. *Nature Nanotechnology*, *5*, 833–842.
19. Hofacker, I. L., Flamm, C., Heine, C., Wolfinger, M. T., Scheuermann, G., & Stadler, P. F. (2010). BarMap: RNA folding on dynamic energy landscapes. *RNA*, *16*, 1308–1316.
20. Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, S., Tacker, M., & Schuster, P. (1994). Fast folding and comparison of RNA secondary structures (the Vienna RNA Package). *Monatshefte für Chemie*, *125*(2), 167–188.
21. Höner zu Siederdissen, C., Hammer, S., Abfalter, I., Hofacker, I. L., Flamm, C., & Stadler, P. F. (2013). Computational design of RNAs with complex energy landscapes. *Biopolymers*, *99*(12), 1124–1136.
22. Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R., & Collins, J. J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnology*, *22*(7), 841–847.
23. Ishikawa, J., Furuta, H., & Ikawa, Y. (2013). RNA tectonics (tectoRNA) for RNA nanostructure design and its application in synthetic biology. *WIREs RNA*, *4*, 651–664.
24. Khalil, A. S., & Collins, J. J. (2010). Synthetic biology: Applications come of age. *Nature Reviews Genetics*, *11*, 367379.
25. Kim, J., & Winfree, E. (2011). Synthetic in vitro transcriptional oscillators. *Molecular Systems Biology*, *7*(1), 465.
26. Kucharík, M., Hofacker, I. L., Stadler, P. F., & Qin, J. (2014). Basin hopping graph: A computational framework to characterize RNA folding landscapes. *Bioinformatics*, *30*(14), 2009–2017.
27. Leisner, M., Bleris, L., Lohmueller, J., Xie, Z., & Benenson, Y. (2010). Rationally designed logic integration of regulatory signals in mammalian cells. *Nature Nanotechnology*, *5*(9), 666–670.
28. Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA package 2.0. *Algorithms for Molecular Biology*, *6*(1), 26.
29. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., & 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.
30. Maňuch, J., Thachuk, C., Stacho, L., & Condon, A. (2011). NP-completeness of the energy barrier problem without pseudoknots and temporary arcs. *Natural Computing*, *10*(1), 391–405.
31. McCaskill, J. S. (1990). The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*, *29*, 1105–1119.
32. Petkovic, S., Badelt, S., Block, S., Flamm, C., Delcea, M., Hofacker, I. L., & Müller, S. (2015). Sequence-controlled RNA self-processing: Computational design, biochemical analysis and visualization by AFM. *RNA Biology*, *21*, 1249–1260.

33. Pieper, S., Vauleon, S., & Müller, S. (2007). RNA self-processing towards changed topology and sequence oligomerization. *Biological Chemistry*, *388*(7), 743–746.

34. Rinaudo, K., Bleris, L., Maddamsetti, R., Subramanian, S., Weiss, R., & Benenson, Y. (2006). A universal RNAi-based logic evaluator that operates in mammalian cells. *Nature Biotechnology*, *25*(7), 795–801.

35. Robertson, M. P., & Joyce, G. F. (2014). Highly efficient self-replicating RNA enzymes. *Chemistry & Biology*, *21*(2), 238–245.

36. Rothemund, P. W. K., Papadakis, N., & Winfree, E. (2004). Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biology*, *2*(12), e424.

37. Subsoontorn, P., Kim, J., & Winfree, E. (2012). Ensemble Bayesian analysis of bistability in a synthetic transcriptional switch. *ACS Synthetic Biology*, *1*(8), 299–316.

38. Turberfield, A. J., Mitchell, J., Yurke, B., Mills Jr., A. P., Blakey, M., & Simmel, F. C. (2003). DNA fuel for free-running nanomachines. *Physical Review Letters*, *90*(11), 118102.

39. Turner, D. H., & Mathews, D. H. (2010). NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Research*, *38*, D280–D282.

40. Wachsmuth, M., Sven Findeiß, S., Weissheimer, N., Stadler, P. F., & Mörl, M. (2013). *De novo* design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Research*, *41*(4), 2541–2551.

41. Wang, F., Lu, C.-H., & Willner, I. (2014). From cascaded catalytic nucleic acids to enzyme–DNA nanostructures: Controlling reactivity, sensing, logic operations, and assembly of complex structures. *Chemical Reviews*, *114*, 2881–2941.

42. Weixlbaumer, A., Werner, A., Flamm, C., Westhof, E., & Schröder, R. (2004). Determination of thermodynamic parameters for HIV-1 DIS type loop–loop kissing complexes. *Nucleic Acids Research*, *32*, 5126–5133.

43. Wieland, M., Benz, A., Klauser, B., & Hartig, J. S. (2009). Artificial ribozyme switches containing natural riboswitch aptamer domains. *Angewandte Chemie*, *121*(15), 2753–2756.

44. Wolfe, B. R., & Pierce, N. A. (2014). Sequence design for a test tube of interacting nucleic acid strands. *ACS Synthetic Biology*, *4*(10), 1086–1100.

45. Wolfinger, M. T., Svrcek-Seiler, A. W., Flamm, C., Hofacker, I. L., & Stadler, P. F. (2004). Efficient computation of RNA folding dynamics. *Journal of Physics A: Mathematical and General*, *37*, 4731–4741.

46. Wuchty, S., Fontana, W., Hofacker, I. L., & Schuster, P. (1999). Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers*, *49*(2), 145–165.

47. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. (2011). Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science*, *333*(6047), 1307–1311.

48. Yin, P., Choi, H. M., Calvert, C. R., & Pierce, N. A. (2008). Programming biomolecular self-assembly pathways. *Nature*, *451*(7176), 318–322.

49. Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. (2011). NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, *32*(1), 170–173.

50. Zhang, D. Y., & Seelig, G. (2011). Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry*, *3*, 103–113.