Predicting Secondary Structural Folding Kinetics for Nucleic Acids

Peinan Zhao,[†] Wen-Bing Zhang,[†*] and Shi-Jie Chen[‡*]
[†]Department of Physics, Wuhan University, Wuhan, China; and [‡]Department of Physics and Astronomy and Department of Biochemistry, University of Missouri, Columbia, Missouri

ABSTRACT We report a new computational approach to the prediction of RNA secondary structure folding kinetics. In this approach, each elementary kinetic step is represented as the transformation between two secondary structures that differ by a helix. Based on the free energy landscape analysis, we identify three types of dominant pathways and the rate constants for the kinetic steps: 1), formation; 2), disruption of a helix stem; and 3), helix formation with concomitant partial melting of a competing (incompatible) helix. The third pathway, termed the tunneling pathway, is the low-barrier dominant pathway for the conversion between two incompatible helices. Comparisons with experimental data indicate that this new method is quite reliable in predicting the kinetics for RNA secondary structural folding and structural rearrangements. The approach presented here may provide a robust first step for further systematic development of a predictive theory for the folding kinetics for large RNAs.

INTRODUCTION

RNAs are quite dynamic and are prone to the formation of multiple metastable structures due to the formation of various stable basepairs and base stacks. Extensive kinetic experiments, such as temperature-jump, single molecule, and time-resolved NMR spectroscopy experiments, have shown that RNA (1–6) and DNA (7–9) often involve multiple intermediates and pathways. Functional studies suggest that the functional structures of RNA can be different from the minimum free energy structure. For instance, the active state of SV11, replicated by Qβ replicase (10,11), is a metastable conformation that acts as a template for Qβ replicase. In contrast, the lowest energy structure is functionally inactive. By melting and rapid quenching, the molecule can be reconverted from the inactive stable state to the active metastable state (12). In addition, experiments suggest that alternative conformations of the same RNA sequence perform different functions (13–15). The capability of RNA molecules to form multiple (metastable) conformations for different functions is probably used by nature to regulate versatile functions of RNA. Therefore, prediction of the native state cannot provide all the information about function. We also need to understand how the metastable states are formed and how the transitions between the different metastable states occur, as well as their pathways, rate-limiting steps, and timescales.

RNA folding kinetics is directly tied to RNA biological functions. The functions of ribozymes (16,17), anti-HIV RNA aptamers (18–20), gene expression regulators such as miRNA, siRNA, and riboswitches (21–27) and other RNAs are often kinetically controlled. For instance, self-induced riboswitches can regulate gene expression by limiting the folding of functional structures to certain time windows (28,29). The hok/sok system of plasmid R1 (28) regulates the plasmid maintenance through mRNA conformation rearrangements into different functional forms during different biochemical reactions. In the realistic timescales for biochemical reactions, folding of the RNA functional structures are often determined by the kinetics instead of by equilibrium thermodynamics (29,30). More recently, experimental studies suggest that kinetic control plays a critical role in a nucleic-acid-based nanomechanical switch/sensor (31). Therefore, the ability to accurately predict the kinetic rates, transition states, and pathways, especially for RNAs of realistic chain length, is vital for a quantitative understanding of the biological functions of many RNA molecules and is essential for rational design of new nanomechanical devices.

Experimental findings (32,33) suggest that most RNAs fold through a hierarchical pathway: Secondary structure are formed rapidly, resulting in a state in which much of the helices and loops are formed but these secondary structural elements lack stable tertiary contacts, and the subsequent slow-folding of the three-dimensional tertiary structure would consolidate the secondary structures through the formation of tertiary interactions. In the hierarchical folding scenario, the structure serves as a scaffold for the tertiary structural folding. Therefore, secondary structural folding kinetics would be important to determine the overall tertiary structural folding kinetic pathways of the molecule (34–41). Predicting how RNAs fold at the secondary structural level is a prerequisite for the prediction of the folding kinetics at the tertiary structural level. This article addresses the development of a physics-based predictive model for RNA secondary structural folding kinetics.

Theoretical studies based on large-scale atomistic molecular dynamics simulations (42,43) and other simulational methods (44–50) have provided detailed microscopic trajectories for the formation of the intermediates on the folding

Submitted October 21, 2009, and accepted for publication December 29, 2009.

*Correspondence: wbzhang@whu.edu.cn or chenshi@missouri.edu

Editor: Kathleen B. Hall.

0006-3495/10/04/1617/9 $2.00

doi: 10.1016/j.bpj.2009.12.4319

pathways. The simulational approach has the advantage of being able to provide atomistic details for the transition states, the kinetic intermediates, and the folding trajectories for the specific sequences studied. However, owing to the restrictions of the computational efficiency, the method is limited to short sequences and short timescales and is often complicated by the issue of incomplete conformational sampling. To reduce the effective number of the conformations, several computational methods (51–55) have been developed based on coarse-grained kinetic moves (basic kinetic steps), such as the cooperative formation and disruption of an entire helix stem. However, these simplified kinetic moves could miss the important folding pathways for some sequences/structures, as explained below in this article.

Recently, based on the preequilibrated conformational macrostates (56) and the kinetic master equation, a kinetic cluster method (57,58) was used to predict RNA folding kinetics, which includes the folding rates, pathways, and kinetic intermediates. The results are quite reliable, as validated by direct tests against experimental measurements and all-atom computer simulations for model systems (43). However, the method is also limited to short chains due to the rapid increase of the number of conformational preequilibrated macrostates for longer sequences.

In this study, we develop a (to our knowledge) new computational method, which combines the master equation and the free energy landscape, to study the RNA secondary structural folding kinetics. The motivation to develop such a new theory is to predict the folding pathways and rates based on a reduced effective ensemble of conformations, so the kinetic predictions for long sequences with large conformational ensembles become computationally viable. The kinetic moves in the model are based on the addition and deletion of helix stems. In contrast to the previous similar models (50–55), the current new method allows partial addition and deletion of helices according to the free energy profiles. This would cause new types of kinetic pathways for conformational switches, including low-barrier pathways that cannot be predicted by the previous similar models. Based on the free energy landscape, we estimate the rate constant for each pathway and find the dominant pathways from the lowest-barrier (fastest) routes.

THEORY AND METHODS

Master equation

Consider an ensemble of conformational states. The populational kinetics $p_i(t)$ for each state i at time t can be described by the equation (i.e., the master equation)

$$dp_i(t)/dt = \sum_j \left[k_{j\to i}\, p_j(t) - k_{i\to j}\, p_i(t)\right],$$

where $\sum_j$ denotes the sum over all the conformations and $k_{j\to i}$ and $k_{i\to j}$ are the rate constants for the respective transitions. The above master equation has an equivalent matrix form,

$$d\mathbf{p}(t)/dt = \mathbf{M}\cdot\mathbf{p}(t),$$

where $\mathbf{p}(t)$ is the vector form of the populations of the states and $\mathbf{M}$ is the rate matrix defined as $M_{ij} = k_{i\to j}$ for $i \neq j$ and $M_{ij} = -\sum_{l\neq i} k_{il}$ for $i = j$.

By diagonalizing the rate matrix, we can predict the populational kinetics for a given initial condition as

$$\mathbf{p}(t) = \sum_{m=1} C_m \mathbf{n}_m e^{-\lambda_m t}, \quad (1)$$

where $-\lambda_m$ and $\mathbf{n}_m$ are the m^{th} eigenvalue and eigenvector of the rate matrix $\mathbf{M}$, and C_m is the coefficient as determined from the initial condition $\mathbf{p}(0)$.

Helices as kinetic building blocks

RNA secondary structure is stabilized mainly by the base-stacking interactions. Because a single (unstacked) basepair is not stable and can quickly unfold, we define an elementary kinetic step for RNA secondary structural change to be the formation/disruption of a stack or a stacked basepair (57). The transition-state theory gives the kinetic rate constants for the formation (k_+) and the disruption (k_-) of a base stack as

$$k_+ = k_0 e^{-\frac{\Delta G_+}{k_B T}},\ k_- = k_0 e^{-\frac{\Delta G_-}{k_B T}},$$

where the prefactor k_0 is fitted from the experimental data and is equal to $6.6\times10^{12}\ s^{-1}$ for the formation/disruption of an AU basepair and $6.6\times10^{13}\ s^{-1}$ for a GC basepair (58), k_B is the Boltzmann constant, T is the temperature, and $\Delta G_\pm$ is the free energy barrier for the respective transition. We assume that the barrier for the formation of a stack is caused by the reduction in entropy ΔS, i.e., $G_+ = T\Delta S$. If the stack closes a loop, the formation of the stack is accompanied by concomitant entropic decrease for loop closure; therefore, the kinetic barrier for loop closure is $G_+ = T\Delta S = T(\Delta S_{loop} + \Delta S_{stack})$, where ΔS_{loop} is the entropy of the loop and ΔS_{stack} is the entropy of the stack. We assume that the barrier for the disruption of a basepair is caused by the energetic (enthalpic) cost ΔH to break the hydrogen bonding and the base-stacking interactions, i.e., $G_- = \Delta H_{stack}$, where ΔH_{stack} is the enthalpy of the stack. In summary, the rates for the formation and disruption of a stack are

$$k_+ = k_0 e^{-\frac{\Delta S_{stack}}{k_B}},\ k_- = k_0 e^{-\frac{\Delta H}{k_B T}}, \quad (2)$$

respectively, and the rates for formation and disruption of a loop-closing stack (and the loop) are

$$k_+^{loop} = k_0 e^{-\frac{\Delta S_{loop}+\Delta S_{stack}}{k_B}},\ k_-^{loop} = k_0 e^{-\frac{\Delta H}{k_B T}}. \quad (3)$$

In general, the formation of the (first) loop-closing stack with rate constant k_+^{loop} is much slower than the subsequent addition of base stacks with rate constant k_+ (e.g., for a 4-nt loop closed by an AUUA base stack, k_+^{loop} and k_+ are equal to $8.6\times10^4\ s^{-1}$ and $2.3\times10^8\ s^{-1}$, respectively). Therefore, we regard the formation of the loop-closing stack of a helix as the nucleation step for the formation of the stem.

The folding condition (temperature) is determined by the sequence of the whole chain. For example, a GC-rich hairpin-forming sequence has a higher folding temperature than an AU-rich sequence. Therefore, the formation of a single loop-closing stack (such as AU-stacks) may not be sufficient to stabilize the loop. Instead, it may require several (e.g., three) consecutive basepairs to stabilize a loop. Therefore, in our conformational enumeration for the (stable) secondary structures, we neglect helices that are too short (with $\leq$2 basepairs) and keep only stable helices (with $\geq$3 basepairs). We develop a kinetic theory for secondary structural folding by treating helices as elementary kinetic building blocks. The transformation of two structures involves the deletion and formation of different helices, as described below. In the following, we will first classify the relationships between two helices (as the initial and final helices for a kinetic move).

Relationships between helices

A helix can be defined by its four terminal nucleotides $[x_1, x_2, x_3, x_4]$ (Fig. S1 in the Supporting Material), where(x_2, x_3) and (x_1, x_4) are the basepairs at the terminus of the helix stem. We consider the formation of helix B $[b_1, b_2, b_3, b_4]$ in a structure that contains helix A $[a_1, a_2, a_3, a_4]$ and the formation of helix A in a structure containing helix B. Based on the compatibility between the two helices, we classify the following three types of relationships among helices (see Fig. 1):

1. *Compatible*: If helix A and helix B contain no overlapping nucleotides (see Fig. 1 *a*), then helix B (respectively, A) can be formed without the disruption of helix A (respectively, B). Because the primary concern of the present form of the theory is with the (pseudo-knot-free) secondary structures, pseudoknotted stems (which can be compatible), are not considered here.
2. *Partially compatible*: If helix A and helix B have partial overlap in the nucleotides (see Fig. 1 *b*), then the formation of helix B (respectively, A) would involve partial disruption of helix A (respectively, B).
3. *Incompatible*: Helix A and helix B completely overlap with each other so that the formation of a helix requires the complete disruption of the other helix (see Fig. 1 *c*).

For a given structure containing n helices $\{h_1, h_2, \cdots, h_i, \cdots, h_n\}$, if helix h_{n+1} is compatible with all the helices h_i $(1 \le i \le n)$, then a new structure $\{h_1, h_2, \cdots, h_i, \cdots, h_n, h_{n+1}\}$ with $n + 1$ helices can be generated by adding the new helix h_{n+1} to the n-helix structure. However, if helix h_{n+1} is compatible with all the helices and partially compatible with helix h_i, the process of adding the helix h_{n+1} would involve an ensemble of $(n + 1)$-helix conformations. Helices h_{n+1} and h_i are (partially) compatible (incompatible) because some basepairs in helix h_i forbid the formation of certain basepairs in helix h_{n+1}. Therefore, the disruption of such incompatible basepairs in h_i would allow the formation of basepairs in h_{n+1}. As a result, the disruption of the different numbers of the incompatible basepairs in h_i would lead to an ensemble of conformations that contain partially melted helix h_i and partially formed helix h_{n+1}.

Move set and the transition rate for a kinetic move

Adding/deleting a helix

In this section, we develop a theory to compute the rate constant for the addition and deletion of a helix. The added helix can be: 1), compatible; 2), partially compatible; or 3), incompatible with the original structure. We first consider a single pathway for the formation of helix through the nucleation/zipping process. We will then extend the theory to treat multiple pathways.

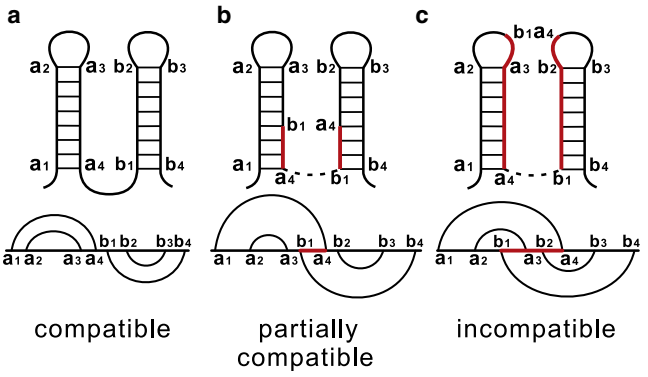

FIGURE 1 The relationship between two helices A $[a_1, a_2, a_3, a_4]$ and B $[b_1, b_2, b_3, b_4]$ can be classified into three types: (*a*) compatible, (*b*) partially compatible, and (*c*) incompatible. For the compatible helices, the formation of helix B does not require unfolding of helix A. For incompatible and partially compatible (incompatible) helices, the formation of helix B requires complete and partial unfolding of helix A, respectively.

The 0^{th} order approximation. For a typical free energy landscape for the nucleation/zipping process (see Fig. 2 *a*), because the rate k_+^{loop} for the formation of the first loop-closing stack is much smaller than the rate k_+ for adding a new base stack to an existing stack, once the first stack is formed, zipping of the helix would be fast. Therefore, as the lowest order approximation, the folding rate can be estimated as the rate of loop-closing k_+^{loop}.

The first order approximation. The above 0^{th} order approximation overestimates the folding rate. As the population flows into state 2 (with only one stack) from state 1 (open state) (see Fig. 2 *a*), a fraction K'_1 of the population will go backward to state 1 and the rest fraction of $K_1 = 1 - K'_1$ of the population would proceed forward to form the next stack in the helix,

$$K_1 = \frac{k_{f2}}{k_{f2} + k_{b1}}; \qquad K_1' = \frac{k_{b1}}{k_{f2} + k_{b1}} = 1 - K_1, \tag{4}$$

where k_{f2} is the rate of adding the second stack to the first stack (state 2 → state 3) and k_{b1} is the rate of breaking the loop-closing stack in state 2 (state 2 → state 1). Therefore, the overall rate can be estimated as

$$k_F^{(1)} = k_+^{\text{loop}} K_1.$$

We call $k_F^{(1)}$ the first-order approximation to the rate because the rate is estimated based on the disruption/formation of the first base stack (state 2 → state 1 rebounce; see Fig. 2 *a*).

The second order approximation. Because the rebounce occurs also for other subsequent states after the formation of the first (loop-closing) base stack (state 2), to improve the estimation of the rate, we further consider the breaking of the second base stack (from the loop) in the folding process. This leads to a backward (rebounce) populational fraction of K_2' (state 3 → state 2) and a forward populational fraction of K_2 (state 3 → state 4), respectively (see also Fig. 2 *a*):

$$K_2 = \frac{k_{f3}}{k_{f3} + k_{b2}}; \qquad K_2' = \frac{k_{b2}}{k_{f3} + k_{b2}} = 1 - K_2. \tag{5}$$

A fraction K_2' of the population in state 3 would rebound to state 2. A fraction of K_1' (K_1) of the rebounded population would flow to state 1 (state 3). Subsequently, $K_2'K_1'$ $(K_2'K_1)$ of the fractional population would flow back to state 1 (reenter state 3). Such iterative processes continue, resulting in a net

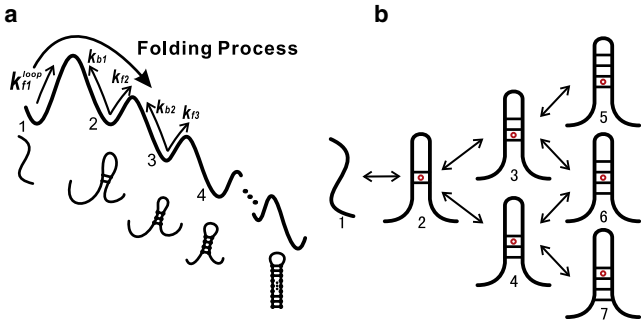

FIGURE 2 (*a*) A schematic free energy landscape for a pathway from the open chain to a helix. Transition from state 1 to state 2 is the formation of the loop-closing base stack. The values k_f and k_b are calculated from Eqs. 2 and 3. (*b*) Multiple pathways for the formation of a helix after the first (nucleation) stack formed.

fractional population K that reaches state 4 (containing three stacks) from state 2 (containing the loop-closing stack only):

$$K = K_1\left(1 - K_2' K_1' \sum_{n=0}^{\infty} \left(K_2' K_1\right)^n\right) = K_1\left(1 - K_2' K_1' \frac{1}{1 - K_2' K_1}\right). \tag{6}$$

Assuming the zipping of the full helix from the three-stack state 4 is a fast downhill process, we can estimate the overall folding rate as

$$k_{\mathrm{F}}^{(2)} = k_{+}^{\mathrm{loop}} K.$$

We call $k_{\mathrm{F}}^{(2)}$ the second-order approximation for the folding rate because the disruption of the second base stack (state 3 → state 2 rebounce) is accounted for. The unfolding rate can be estimated from the detailed balance condition,

$$k_{\mathrm{U}}^{(2)} = k_{+}^{\mathrm{loop}} K e^{-\frac{\Delta G}{k_{\mathrm{B}}T}},$$

where ΔG is the free energy difference between the open chain and the helix.

Multiple pathways. In general, after the first stack is formed, there may be multiple ways to add the next base stack. Each pathway may further branch out as folding proceeds. For the growth of a continuous stretch of helix, we consider the addition of a base stack from two ends, each leading to a different branch of the bifurcated pathways (e.g., pathways 1 → 2 → 3, and 1 → 2 → 4 in Fig. 2 *b*). The rate of each pathway can be estimated from the above methods. For example, for pathway 1 → 2 → 3, the corresponding K_1, K'_1, K_2, and K'_2 are

$$K_1 = \frac{k_{23}}{k_{23} + k_{21} + k_{24}},$$
$$K_1' = \frac{k_{21}}{k_{23} + k_{21} + k_{24}},$$
$$K_2 = \frac{k_{35} + k_{36}}{k_{32} + k_{35} + k_{36}},$$
$$K_2' = \frac{k_{32}}{k_{35} + k_{36} + k_{32}},$$

where k_{ij} is the rate from state i to j. So the overall folding rate along the 1 → 2 → 3 pathway is $k_{12}K$, where K is determined by Eq. 6. For the folding scenario shown in Fig. 2 *b*, The folding rate from the given nucleation stack (see the base stack in state 2) is the sum of the rates for the different pathways. The overall folding rate is the sum of the rates over the different nucleation base stacks.

The second-order approximation can give quite accurate estimations for the rates (and the populations kinetics), and the first-order approximation overestimates the folding rate (see Supporting Material).

Arm-by-arm exchange

Tunneling pathway. If two helices are incompatible, they cannot coexist in the same structure. What are the kinetic pathways for the transformation between two incompatible helices? Previous approaches assumed a two-step transition (deletion-addition pathway) by fully unwinding the existing helix followed by subsequent folding of the new helix from the open state. However, in this study, we find that a complete disruption of the helix is not necessary for the fast formation of a new (incompatible) helix. We use Fig. 1 *c* to illustrate the calculation of the rate for the $A \rightarrow B$ transition. The graph shows two incompatible helices A [a_1, a_2, a_3, a_4] and B [b_1, b_2, b_3, b_4]. The final formation of helix B requires the complete melting of helix A. Consider the formation of helix B from a structure that contains helix A. The partial melting of helix A to the partially unfolded state [$a_2 + a_3 \times - b_2 + 1$, a_2, a_3, $b_2 - 1$] would be sufficient to allow the initiation of the folding of helix B, namely, the formation of the loop-closing stack ($b_2 - 1$, b_2, b_3, $b_3 + 1$) for helix B. After the formation of the loop-closing stack in helix B, the elongation of helix B would proceed step by step after further melting of helix A, until helix A is completely melted. The remaining base stacks of helix B are formed in a zipping process after helix A is completely unfolded.

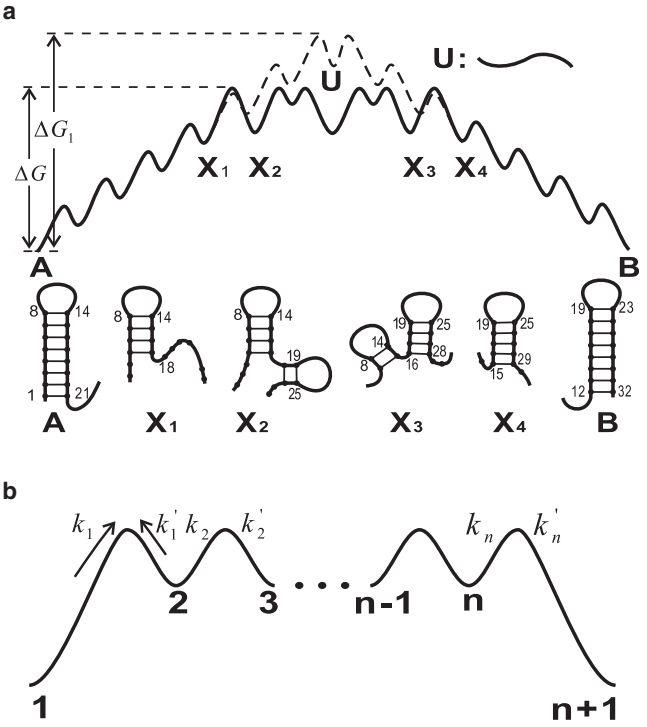

FIGURE 3 (*a*) The free energy landscape for the transition between two incompatible helices. Helices A and B contain nucleotides (basepairs) that are incompatible with each other. X_{i} is the state where helix A is partially melted with the some of the incompatible basepairs disrupted and helix B is partially formed. U is the open state. (*b*) A schematic free energy profile for the arm-by-arm exchange process. The values k_1 and k'_{n} are calculated from Eq. 7; k'_1, k_2, ⋯, k_{n} are calculated from Eq. 2.

The above pathway involves a low kinetic barrier (see Fig. 3 *a*) because

1. The formation of the first base stack (nucleation base stack) of helix B requires only partial instead of full disruption of helix A; and
2. The free energy increase for the further disruption of helix A is compensated by the free energy decrease for the formation of base stacks in helix B.

We call such a pathway the tunneling pathway (a tunnel in the free energy mountain). The tunneling path has a much lower energy barrier than the deletion-addition pathway, which requires the chain to reach the open state (the top of the free energy landscape). The tunneling pathway may be the dominant pathway for the $A \rightarrow B$ transition. This is similar to the Morgan-Higgs saddle-point approach (59), in which the saddle-point height is estimated as the highest point along the path. However, the free energy landscape suggests that multiple high free energy points along the path may exist, as described below. Because the pathway involves the growth of helix B in exchange with the breaking of helix A, we call this move set (i.e., transformation between two incompatible helices) an arm-by-arm exchange process.

According to the free energy landscape (Fig. 3 *a*), we can classify the tunneling pathway for the helix A → helix B transition into three stages:

Process 1: *Partial melting of helix A and nucleation of helix B*. From state A (full helix) to state X_1 (Fig. 3 *a*), each step involves breaking of one stack of helix A. From state X_1 to state X_2, the first (loop-closing) stack of helix B is

formed. This stage ($A \rightarrow X_1 \rightarrow X_2$) is generally an uphill process on the free energy landscape.

Process 2: Exchanging. From state X_2 to state X_3 (Fig. 3 *a*), helix B is folded while helix A is unfolded until the last (loop-closing) base stack in A is disrupted. In this process, the free energy oscillates: the free energy increases when a base stack in A is disrupted and decreases when a base stack in B is closed.

Process 3: Zipping. From state X_4 to state B (Fig. 3 *a*), helix B grows through a zipping process. This is a downhill process on the free energy landscape.

Rate constant for the tunneling pathway

Processes 1 and 3 above can be treated as two-state transitions without significant accumulation of intermediate states (see Fig. 3 *b*). Here we show the calculation for the rate for process 1 ($A \rightarrow X_2$). The rate for process 3 ($X_4 \rightarrow B$) can be computed in the similar way.

The rate constant $k_{A\leftrightarrow X_2}$ for transitions $A \leftrightarrow X_2$ can be calculated as

$$\begin{aligned} k_{A\rightarrow X_2} &= k_0\, e^{-\frac{\Delta G}{k_B T}} = e^{-\frac{\Delta G_{A,X_1}}{k_B T}}\, k_{X_1\rightarrow X_2}, \\ k_{X_2\rightarrow A} &= k_{A\rightarrow X_2}\, e^{-\frac{\Delta G_{A,X_2}}{k_B T}}, \end{aligned} \tag{7}$$

where ΔG is the free energy barrier (see Fig. 3 *a*), $\Delta G_{A,X_1}(\Delta G_{A,X_2})$ is the free energy difference between state A and state X_1 (X_2), and $k_{X_1\rightarrow X_2}$ is the rate for the transition $X_1 \rightarrow X_2$ (formation of the loop-closing stack of helix B),

$$k_{X_1\rightarrow X_2} = k_0\, e^{\frac{\Delta S_{loop}+S_{stack}}{k_B}},$$

where ΔS_{loop} is the loop entropy and ΔS_{stack} is the entropy of the loop-closing stack in helix B.

The population p_i of each state i in process 2 is determined by the following master equation:

$$\begin{cases} \dfrac{dp_1}{dt} = -k_1 p_1 + k_1' p_2 \\ \dfrac{dp_2}{dt} = k_1 p_1 + k_1' p_2 - k_2 p_2 + k_2' p_3 = 0 \\ \dfrac{dp_3}{dt} = k_2 p_2 - k_2' p_3 - k_3 p_3 + k_3' p_4 = 0 \\ \vdots \\ \dfrac{dp_n}{dt} = k_{n-1} p_{n-1} - k_{n-1}' p_n - k_n p_n + k_{n+1}' p_{n+1} = 0 \\ \dfrac{dp_{n+1}}{dt} = k_n p_n - k_n' p_{n+1}. \end{cases} \tag{8}$$

Here, k_1 and k'_1 are the rate constants for the transitions $A \rightarrow X_2$ and $X_2 \rightarrow A$, respectively, k_j and k'_j ($2 \le j \le n-1$) are the rate constants for the intermediate transitions in process 2, and k_n and k'_n are the rate constants for the transitions $X_3 \rightarrow B$ and $B \rightarrow X_3$, respectively.

All the intermediate states involved in process 2 have high free energies. Therefore, we assume no significant populational aggregation for these states during the kinetic process. We treat these states with steady-state approximation ($dp_i/dt \sim 0$). The steady-state solution for Eq. 8 gives the transition rates between state 1 (= state A in Fig. 3 *a*) and state (n+1) (= state B in Fig. 3 *a*):

$$\begin{aligned} k_{1\rightarrow n+1} &= \frac{\prod_i^n k_i}{\sum_{j=0}^{n-1}\left(\prod_{i=1}^{j} k_i' \prod_{m=j+2}^{n} k_m\right)},\ k_{n+1\rightarrow 1} \\ &= k_{1\rightarrow n+1} e^{-\frac{\Delta G_{1,(n+1)}}{k_B T}}, \end{aligned} \tag{9}$$

where $\Delta G_{1,\,(n+1)}$ is the free energy difference between state 1 and state (n+1). It gives quite accurate estimations for the rate constant (see Supporting Material).

Two-arm by two-arm exchange

The above tunneling pathway can be generalized to transitions between structures with multiple incompatible helices. As shown in Fig. 4, along the tunneling pathway $A \rightarrow E$, folding of the new helices (h_3 and h_4) is initiated after partial melting of the existing helices (h_1 and h_2). After the nucleation of the two new helices, the process is similar to that of arm-by-arm exchange, where each of the new helices closes a new base stack after an existing incompatible base stack is disrupted. The process continues until the original two (incompatible) helices are completely disrupted. After that, the two new helices elongate, following a downhill zipping pathway. The transition rate for such a two-arm by two-arm exchange process can be calculated in the same way as that for arm-by-arm exchange, except that now each step involves melting/formation of two stacks (one stack for each helix), whereas the arm-by-arm exchange involves one stack in each step. It gives quite accurate estimations for the rate constant (see Supporting Material).

RESULTS AND DISCUSSION

The tests above with the exact solutions of the master equations suggest the validity of our theory for the rate constant for kinetic moves (addition and deletion of a helix). In this section, we extend the test of the theory by predicting the overall folding kinetics from the sequence. We then apply the theory to predict the kinetics for the conformational switch for nucleic acids and compare the theoretical predictions with the experimental data.

Test with the original stack-based kinetic theory

To test the validity of our helix-based kinetic theory, where each kinetic move is the addition or deletion of a helix, we predict the folding kinetics for a 27-nt model system and compare the results with the predictions from the original stack-based kinetic theory, where each kinetic move is the formation or disruption of a base stack. The 27-nt sequence

5′ AUAGGUUAUAUAUCACGUAUAGCCUAU 3′

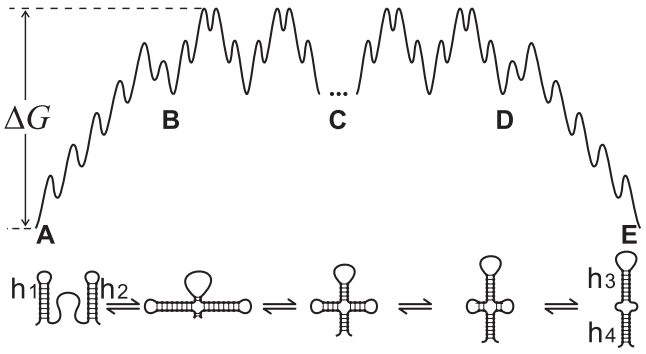

FIGURE 4 The free energy landscape of the two-arm-by-two-arm exchange pathway. State A contains helices h_1 and h_2, state B has a few stacks of helices h_1 and h_2 melted and one stack of helix h_3 and h_4 each formed, state C represents a state where h_1 and h_2 are partially unfolded, state D is the state in which only one stack of helix h_1 and h_2 each remains, and state E is the state with helices h_3 and h_4.

has 1949 stack-based conformations (i.e., conformations defined through base stacks). In the stack-based theory, the kinetic move set is the formation/disruption of a stack with the rate constant given by Eq. 1. For an initially unfolded chain, the populational kinetics solved from the exact master equation based on the 1949 × 1949 rate matrix shows a kinetic intermediate with 40% fractional population (Fig. 5). In contrast, in the helix-based conformational model (i.e., conformations defined by the constituent helix stems), the 27-nt sequence can form 11 structures. Using the rate constant theories described above for the creation/annihilation of a helix, we construct the 11 × 11 rate matrix. We find that the populational kinetics predicted from the 11 × 11 rate matrix is in good agreement with the predictions from the original 1949 × 1949 rate matrix (see Fig. 5). The result suggests that our helix-based kinetic theory may be a reliable approximation to the original theory based on the complete stack-based conformational ensemble.

Applications to conformational switches of nucleic acids and comparisons to experimental results

Recently, Viasnoff et al. (31) measured the kinetics of the conformational switch for a designed 87-nt DNA sequence (see Fig. 6),

$$\mathrm{TPGTTGP'(T)_{20}PCTTCP'T},$$

where

$$\begin{aligned} \mathrm{P} &= 5'\,\mathrm{CGACCTCAGCATCG\ and\ P'} \\ &= 5\,\mathrm{CGATGCTGAGGTCG}. \end{aligned}$$

The sequence allows the formation of two sets of helix stems: (h_1, h_2) and (h_3, h_4), where

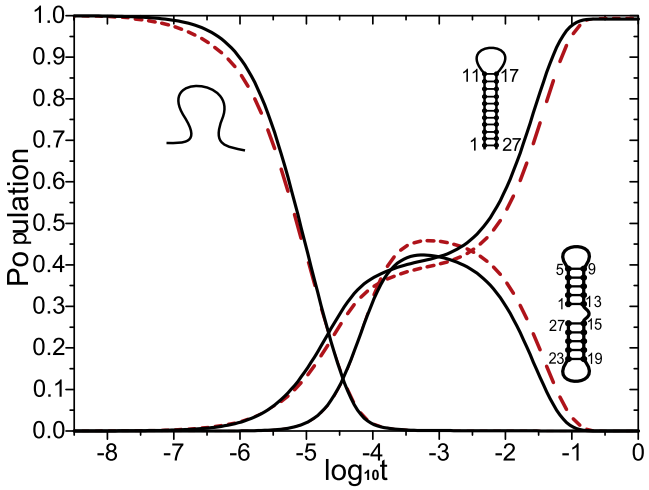

FIGURE 5 The populational kinetics for the folding of the 27-nt sequence: 5′AUAGGUUAUAUAUCACGUAUAGCCUAU 3′. The dashed lines are from the exact master equation with the complete conformational ensemble; the solid lines are from our helix-based kinetic model.

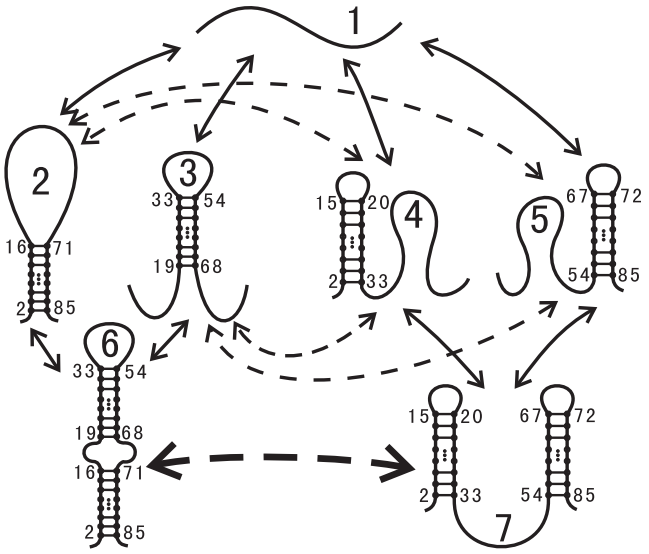

FIGURE 6 The reduced states and the transition network of the 82-nt DNA sequence. State 6 is the native structure. The solid line denotes the deletion-addition pathway, the dashed line denotes the arm-by-arm exchange pathway, and the thick dashed line denotes the two-arm-by- two-arm exchange pathway.

$$h_1 = [2,\ 16,\ 71,\ 85],\ h_2 = [19,\ 33,\ 54,\ 68],$$

$$h_3 = [2,\ 15,\ 20,\ 33]\ \text{and}\ h_4 = [54,\ 67,\ 72,\ 85].$$

The conformational switch is between the elongated structure (h_1, h_2) (state 6 in Fig. 6) and the contracted structure (h_3, h_4) (state 7 in Fig. 6). Under 1 M NaCl condition, the elongated structure is the thermodynamically stable structure with a free energy of $\Delta G = -39.6$ kcal/mol as compared to $\Delta G = -36.3$ kcal/mol of the contracted structure. Under fast renaturation, it was found that ~83% of the population is folded into the contracted conformation and the contracted structure lasted several weeks before folding to the final thermal equilibrium state (the elongated structure).

The 87-nt chain has 91,429,581 stack-based conformations. In contrast, in the helix-based conformational model, only seven structures exist (Fig. 6). The transition rate between any two (of the seven) helix-based conformational states can be calculated from the theory developed above (such as the tunneling pathway model for the arm-by-arm exchange). Based on the rate 7 × 7 rate matrix, we compute the populational kinetics, the pathways, and the folding rate (see Fig. 7 *a*).The theoretical predictions indicate that the population of the contracted structure emerges as a kinetic intermediate with a significant fractional population of 87%. The result is in good agreement with the experiment. In addition, the predicted lifetime of the intermediate states is ~10^6 s (about several weeks), which is also consistent with the experiment.

To extract the pathways from the rate matrix, we compute the net populational flux, which is the probability $P_{i\to j}$ for

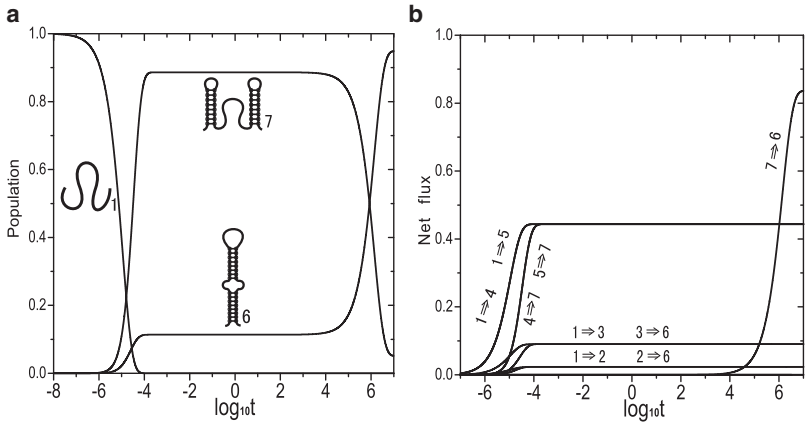

FIGURE 7 (*a*) Populational kinetics of the open state, intermediate state, and the native state of the 87-nt sequence. (*b*) The net populational fluxes among the seven states in Fig. 6. 1 → 3 and 1 → 5, 3 → 7, and 5 → 7, coincide with each other. The net flux between other states are very small and thus not shown in the figure.

the molecule to switch from state i to state j during the time period $0 \rightarrow t$,

$$P_{i\rightarrow j}(t) = \int_0^t \left(P_i(t')k_{i\rightarrow j} - P_j(t')k_{j\rightarrow i}\right)dt,$$

where $P_i(t)$ and $P_j(t)$ are the populations of states i and j at time t (Fig. 7 *b*), respectively. From the unfolded state, there are four pathways ($1 \rightarrow 2$, $1 \rightarrow 3$, $1 \rightarrow 4$, and $1 \rightarrow 5$) to initiate the folding process. Our result for the net flux shows that the majority of the populational flux flows into states 4 and 5 (each occupies ~44%). States 4 and 5 contain one of the two helices of the intermediate state (the contracted structure). After a short time delay, the fluxes of $4 \rightarrow 7$ and $5 \rightarrow 7$ dominate, corresponding the formation of the intermediate state.

Why does the chain fold to the misfolded intermediate ($1 \rightarrow 4$ and $1 \rightarrow 5$) instead of to the native state ($1 \rightarrow 2$ and $1 \rightarrow 3$) in the initial stage? The low initial fluxes for pathways $1 \rightarrow 2$ and $1 \rightarrow 3$ can be explained as the following. The hairpin loops of the two (native) helices in states 2 and 3 (54 nts and 20 nts, respectively) are much larger than the 4-nt loops for the two (nonnative) helices in states 4 and 5. From the experimental measurements (60) and the theoretical modeling (61), the loop entropy depends on the loop size n as $\Delta S_{loop} \propto k_B \ln n^{-1.8}$. The corresponding loop-closing rate is $k_f^{loop} \propto n^{-1.8}$ (Eq. 3). The rate for the formation of a helix with m stacks and a loop of size l is

$$\sim k_f \propto \sum_{i=0}^{m-1} (l + 2i)^{-1.8}.$$

Therefore, we have

$$k_{1\rightarrow 4,5} \propto \sum_{i=0}^{12} (4 + 2i)^{-1.8},$$

$$k_{1\rightarrow 2} \propto \sum_{i=0}^{13} (54 + 2i)^{-1.8},$$

and

$$k_{1\rightarrow 3} \propto \sum_{i=0}^{13} (20 + 2i)^{-1.8}.$$

The fraction of the population to the intermediate nonnative helix along the pathway $1 \rightarrow 4$ or $1 \rightarrow 5$ is

$$\sim \frac{k_{1\rightarrow 4}}{k_{1\rightarrow 4} + k_{1\rightarrow 5} + k_{1\rightarrow 2} + k_{1\rightarrow 3}} \simeq 45\%,$$

and the fraction to the native helices is

$$\frac{k_{1\rightarrow 3}}{k_{1\rightarrow 4} + k_{1\rightarrow 5} + k_{1\rightarrow 2} + k_{1\rightarrow 3}} \simeq 7\%$$

and

$$\frac{k_{1\rightarrow 2}}{k_{1\rightarrow 4} + k_{1\rightarrow 5} + k_{1\rightarrow 2} + k_{1\rightarrow 3}} \simeq 2\%,$$

respectively. The results from the simple analysis agree with the data from the complete computations with the master equation.

The conversion from the misfolded intermediate to the final native state proceeds along three pathways: the direct transition through the two-arm-by-two-arm exchange (path $7 \rightarrow 6$) and two other nondirect transitions through melting of one arm first (path $7 \rightarrow 5$ and $7 \rightarrow 4$) followed by an arm-by-arm exchange ($4 \rightarrow 2$, $4 \rightarrow 3$, $5 \rightarrow 2$, $5 \rightarrow 3$). The calculated flux curves indicate that the direct pathway is the dominant pathway. This is because the rate for the direct pathway ($k_{7\rightarrow 6} = 1.5 \times 10^{-7}\ s^{-1}$) is larger than that of the nondirect pathways ($k_{7\rightarrow 4} = k_{7\rightarrow 5} = 1.8 \times 10^{-8}\ s^{-1}$), which involve a high-barrier process to unfold 12 stacks.

SUMMARY

To treat the large conformational ensemble for the RNA secondary structural folding kinetics, we have developed a new helix-based conformational model for folding kinetics. Each kinetic move in the new model is the annihilation or creation of a helix stem. For a given secondary structure, we find a low-barrier pathway for the formation of new helix incompatible with the existing structure. The pathway involves unfolding-induced-folding, namely, each step of the partial unfolding (= the disruption of a basepair a) in

the existing structure is followed by the formation of a basepair b of the new helix if basepairs a and b are incompatible. We call this the tunneling pathway because it represents a low-barrier tunneling process in the free energy profile. The tunneling pathway has a much faster rate than the conventional folding pathway that requires a high-barrier complete unfolding of the chain. Tests against experimental data and exact master equation solutions with complete conformational ensemble suggest that the new RNA folding kinetic theory is quite reliable in the predictions of folding kinetics.

Because the new theory is based on helix-based conformations, the number of conformational states is significantly reduced as compared with the original basestack-based conformational states. Therefore, the theory have the potential to treat large RNA molecules. The application of the theory to the kinetics of the conformational switch for a 87-nt DNA chain leads to the prediction of a long-lived kinetic intermediate (as observed in the experimental measurement). Furthermore, the theory shows that

1. The formation of the misfolded intermediate is due to the slow rate for the closure of the large hairpin loops of the native helices;
2. The long lifetime for the intermediate is due to the slow unfolding rate for the disruption of the long helix stems in the intermediate state; and
3. The pathway for the conformational switch from the misfolded kinetic intermediate to the final native state is a two-arm-by-two-arm exchange process.

The application and experimental test suggest that the new theory developed here may be useful for the analysis and rational design of the kinetic properties for large RNAs.

This theory may also provide insights about folding with chain elongation during transcription. Because the rate for a basepair formation (k_+^{loop} for a loop-closing basepair and k_+ otherwise, which, for a 4-nt loop, are equal to 8.6×10^4 s^{-1} and 2.3×10^8 s^{-1}, respectively) is larger than the transcription speed (~10–200 nt/s) (62), during the time interval (0.005–0.1 s) before the next base transcribed, the newly transcribed base may form basepairs with other bases that are already transcribed. The newly formed basepair can either be added to an already formed helix or branch out as a nucleus of further folding. The populational kinetics of the folded structures depends on the elongation speed and the transition rate. If the intermediate structure formed during the transcription process is not native, whether the final native structure is formed would be determined by the competition between the transcription rate and the rate for the transition between the intermediate state and the native state, which, equivalently, is the competition between the transcription time and the lifetime of the intermediate state.

Further development of the theory requires consideration of more complex pathways. For instance, in the two-arm-by-two-arm exchange pathway (in Fig. 4), depending on sequence context, the newly folded helices h_3 and h_4 might not grow in a symmetric manner. Many other low-barrier branched-out nonsymmetric pathways for the two-arm-by-two-arm exchange process might exist. Moreover, pseudoknotted motifs are frequently occurring in large RNA folds, so it is therefore useful to include the pseudoknots in the further development of the theory.

SUPPORTING MATERIAL

Two figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00087-1.

This work was partly supported by the Program for New Century Excellent Talents at Wuhan University under grant No. NCET-06-0623, the National Natural Science Foundation of China under grants No. 10774115 and No. 30670487 (to W.-B. Z.), and the National Science Foundation grant No. MCB-0920411 (to S.-J. C.).

REFERENCES

1. Nagel, J. H., C. Flamm, …, C. W. Pleij. 2006. Structural parameters affecting the kinetics of RNA hairpin formation. *Nucleic Acids Res.* 34:3568–3576.
2. Micura, R., and C. Höbartner. 2003. On secondary structure rearrangements and equilibria of small RNAs. *ChemBioChem.* 4:984–990.
3. Fürtig, B., J. Buck, …, H. Schwalbe. 2007. Time-resolved NMR studies of RNA folding. *Biopolymers.* 86:360–383. doi:10.1002/bip.20761.
4. Harlepp, S., T. Marchal, …, D. Chatenay. 2003. Probing complex RNA structures by mechanical force. *Eur Phys. J. E. Soft Matter.* 12: 605–615.
5. Jean, J. M., and K. B. Hall. 2001. 2-Aminopurine fluorescence quenching and lifetimes: role of base stacking. *Proc. Natl. Acad. Sci. USA.* 98:37–41.
6. Liphardt, J., B. Onoa, …, C. Bustamante. 2001. Reversible unfolding of single RNA molecules by mechanical force. *Science.* 292:733–737.
7. Bonnet, G., O. Krichevsky, and A. Libchaber. 1998. Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. USA.* 95:8602–8606.
8. Ansari, A., S. V. Kuznetsov, and Y. Shen. 2001. Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. USA.* 98:7771–7776.
9. Wallace, M. I., L. Ying, …, D. Klenerman. 2001. Non-Arrhenius kinetics for the loop closure of a DNA hairpin. *Proc. Natl. Acad. Sci. USA.* 98:5584–5589.
10. Biebricher, C. K., S. Diekmann, and R. Luce. 1982. Structural analysis of self-replicating RNA synthesized by Qβ replicase. *J. Mol. Biol.* 154:629–648.
11. Biebricher, C. K., and R. Luce. 1992. In vitro recombination and terminal elongation of RNA by Qβ replicase. *EMBO J.* 11:5129–5135.
12. Zamora, H., R. Luce, and C. K. Biebricher. 1995. Design of artificial short-chained RNA species that are replicated by Qβ replicase. *Biochemistry.* 34:1261–1266.
13. Baumstark, T., A. R. Schröder, and D. Riesner. 1997. Viroid processing: switch from cleavage to ligation is driven by a change from a tetraloop to a loop E conformation. *EMBO J.* 16:599–610.
14. Perrotta, A. T., and M. D. Been. 1998. A toggle duplex in hepatitis δ virus self-cleaving RNA that stabilizes an inactive and a salt-dependent pro-active ribozyme conformation. *J. Mol. Biol.* 279:361–373.
15. Schultes, E. A., and D. P. Bartel. 2000. One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science.* 289:448–452.

16. Kruger, K., P. J. Grabowski, …, T. R. Cech. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell*. 31:147–157.

17. Bartel, D. P., and J. W. Szostak. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science*. 261:1411–1418.

18. Joyce, G. F. 1989. Amplification, mutation and selection of catalytic RNA. *Gene*. 82:83–87.

19. Ellington, A. D., and J. W. Szostak. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature*. 346:818–822.

20. Tuerk, C., and L. Gold. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. 249:505–510.

21. Mironov, A. S., I. Gusarov, …, E. Nudler. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell*. 111:747–756.

22. Nahvi, A., N. Sudarsan, …, R. R. Breaker. 2002. Genetic control by a metabolite binding mRNA. *Chem. Biol*. 9:1043–1049.

23. Winkler, W. C., and R. R. Breaker. 2003. Genetic control by metabolite-binding riboswitches. *ChemBioChem*. 4:1024–1032.

24. Bartel, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 116:281–297.

25. Nudler, E., and A. S. Mironov. 2004. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci*. 29:11–17.

26. He, L., and G. J. Hannon. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet*. 5:522–531.

27. Winkler, W. C., A. Nahvi, …, R. R. Breaker. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*. 428:281–286.

28. Gerdes, K., and E. G. H. Wagner. 2007. RNA antitoxins. *Curr. Opin. Microbiol*. 10:117–124.

29. Nagel, J. H. A., A. P. Gultyaev, …, C. W. Pleij. 1999. Metastable structures and refolding kinetics in hok mRNA of plasmid R1. *RNA*. 5: 1408–1418.

30. Morgan, S. R., and P. G. Higgs. 1996. Evidence for kinetic effects in the folding of large RNA molecules. *J. Chem. Phys*. 105:7152–7157.

31. Viasnoff, V., A. Meller, and H. Isambert. 2006. DNA nanomechanical switches under folding kinetics control. *Nano Lett*. 6:101–104.

32. Brion, P., and E. Westhof. 1997. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct*. 26:113–137.

33. Tinoco, Jr., I., and C. Bustamante. 1999. How RNA folds. *J. Mol. Biol*. 293:271–281.

34. Bartley, L. E., X. Zhuang, …, D. Herschlag. 2003. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J. Mol. Biol*. 328:1011–1026.

35. Treiber, D. K., and J. R. Williamson. 2001. Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol*. 11:309–314.

36. Woodson, S. A. 2000. Recent insights on RNA folding mechanisms from catalytic RNA. *Cell. Mol. Life Sci*. 57:796–808.

37. Brenowitz, M., M. R. Chance, …, K. Takamoto. 2002. Probing the structural dynamics of nucleic acids by quantitative time-resolved and equilibrium hydroxyl radical "footprinting". *Curr. Opin. Struct. Biol*. 12:648–653.

38. Thirumalai, D., N. Lee, …, D. Klimov. 2001. Early events in RNA folding. *Annu. Rev. Phys. Chem*. 52:751–762.

39. Chen, S. J. 2008. RNA folding: conformational statistics, folding kinetics, and ion electrostatics. *Annu. Rev. Biophys*. 37:197–214.

40. Sosnick, T. R., and T. Pan. 2003. RNA folding: models and perspectives. *Curr. Opin. Struct. Biol*. 13:309–316.

41. Uhlenbeck, O. C. 1995. Keeping RNA happy. *RNA*. 1:4–6.

42. Sorin, E. J., Y. M. Rhee, …, V. S. Pande. 2003. Insights into nucleic acid conformational dynamics from massively parallel stochastic simulations. *Biophys. J*. 85:790–803.

43. Sorin, E. J., M. A. Engelhardt, …, V. S. Pande. 2002. RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop. *J. Mol. Biol*. 317:493–506.

44. Flamm, C., W. Fontana, …, P. Schuster. 2000. RNA folding at elementary step resolution. *RNA*. 6:325–338.

45. Isambert, H., and E. D. Siggia. 2000. Modeling RNA folding paths with pseudoknots: application to hepatitis δ virus ribozyme. *Proc. Natl. Acad. Sci. USA*. 97:6515–6520.

46. Danilova, L. V., D. D. Pervouchine, …, A. A. Mironov. 2006. RNAKinetics: a web server that models secondary structure kinetics of an elongating RNA. *J. Bioinform. Comput. Biol*. 4:589–596.

47. Lin, J. C., and D. Thirumalai. 2008. Relative stability of helices determines the folding landscape of adenine riboswitch aptamers. *J. Am. Chem. Soc*. 130:14080–14081.

48. Pincus, D. L., C. Hyeon, and D. Thirumalai. 2008. Effects of trimethylamine *n*-oxide (TMAO) and crowding agents on the stability of RNA hairpins. *J. Am. Chem. Soc*. 130:7364–7372.

49. Schmitz, M., and G. Steger. 1996. Description of RNA folding by "simulated annealing". *J. Mol. Biol*. 255:254–266.

50. Gultyaev, A. P., F. H. van Batenburg, and C. W. Pleij. 1995. The computer simulation of RNA folding pathways using a genetic algorithm. *J. Mol. Biol*. 250:37–51.

51. Martinez, H. M. 1984. An RNA folding rule. *Nucleic Acids Res*. 12:323–334.

52. Mironov, A. A., L. P. Dyakonova, and A. E. Kister. 1985. A kinetic approach to the prediction of RNA secondary structures. *J. Biomol. Struct. Dyn*. 2:953–962.

53. Tang, X., B. Kirkpatrick, …, N. Amato. 2004. Using motion planning to study RNA folding kinetics. *In* Proceedings of the International Conference on Computational Molecular Biology (RECOMB). ACM Press, San Diego, CA.

54. Xayaphoummine, A., T. Bucher, …, H. Isambert. 2003. Prediction and statistics of pseudoknots in RNA structures using exactly clustered stochastic simulations. *Proc. Natl. Acad. Sci. USA*. 100:15310–15315.

55. Geis, M., C. Flamm, …, C. Thurner. 2008. Folding kinetics of large RNAs. *J. Mol. Biol*. 379:242–261.

56. Konishi, Y., T. Ooi, and H. A. Scheraga. 1982. Regeneration of ribonuclease A from the reduced protein. Rate-limiting steps. *Biochemistry*. 21:4734–4740.

57. Zhang, W. B., and S. J. Chen. 2002. RNA hairpin-folding kinetics. *Proc. Natl. Acad. Sci. USA*. 99:1931–1936.

58. Zhang, W. B., and S. J. Chen. 2006. Exploring the complex folding kinetics of RNA hairpins: I. General folding kinetics analysis. *Biophys. J*. 90:765–777.

59. Morgan, S. R., and P. G. Higgs. 1998. Barrier heights between ground states in a model of RNA secondary structure. *J. Phys. A. Math. Gen*. 31:3153–3170.

60. Serra, M. J., and D. H. Turner. 1995. Predicting thermodynamic properties of RNA. *Methods Enzymol*. 259:242–261.

61. Chen, S. J., and K. A. Dill. 2000. RNA folding energy landscapes. *Proc. Natl. Acad. Sci. USA*. 97:646–651.

62. Pan, T., and T. Sosnick. 2006. RNA folding during transcription. *Annu. Rev. Biophys. Biomol. Struct*. 35:161–175.