Effects of flanking regions on HDV cotranscriptional folding kinetics

YANLI WANG, ZHEN WANG, TAIGANG LIU, SHA GONG, and WENBING ZHANG
Department of Physics, Wuhan University, Wuhan, Hubei 430072, P.R. China

ABSTRACT
Hepatitis delta virus (HDV) ribozyme performs the self-cleavage activity through folding to a double pseudoknot structure. The folding of functional RNA structures is often coupled with the transcription process. In this work, we developed a new approach for predicting the cotranscriptional folding kinetics of RNA secondary structures with pseudoknots. We theoretically studied the cotranscriptional folding behavior of the 99-nucleotide (nt) HDV sequence, two upstream flanking sequences, and one downstream flanking sequence. During transcription, the 99-nt HDV can effectively avoid the trap intermediates and quickly fold to the cleavage-active state. It is different from its refolding kinetics, which folds into an intermediate trap state. For all the sequences, the ribozyme regions (from 1 to 73) all fold to the same structure during transcription. However, the existence of the 30-nt upstream flanking sequence can inhibit the ribozyme region folding into the active native state through forming an alternative helix Alt1 with the segments 70–90. The longer upstream flanking sequence of 54 nt itself forms a stable hairpin structure, which sequesters the formation of the Alt1 helix and leads to rapid formation of the cleavage-active structure. Although the 55-nt downstream flanking sequence could invade the already folded active structure during transcription by forming a more stable helix with the ribozyme region, the slow transition rate could keep the structure in the cleavage-active structure to perform the activity.

Keywords: HDV ribozyme; cotranscriptional; folding kinetics; pathway; pseudoknot

INTRODUCTION
RNA can carry out numerous biological functions, such as translating genetic information into proteins (Skog et al. 2008; Mercer et al. 2009), regulating gene expression (Grundy and Henkin 1998; Batye et al. 2004), and catalyzing biochemical processes (Cheah et al. 2007; Neupane et al. 2011; Li and Breaker 2013; Lin and Thirumalai 2013; Reining et al. 2013; Hoffmann et al. 2014) by forming specific secondary and tertiary structures. RNA pseudoknots are examples of minimal structural motifs in structured RNAs with tertiary interactions. They have been found to play important roles in a wide range of biological functions, from ribosomal frameshifting (Gesteland and Atkins 1996; Kim et al. 1999; Giedroc et al. 2000; Plant et al. 2003; Cornish et al. 2005; Plant and Dinman 2005) to human telomerase RNA (hTR) activity (Comolli et al. 2002; Theimer et al. 2003; Chen and Greider 2005; Marrone et al. 2005). Many ribozymes (Lehnert et al. 1996; Zarrinkar and Williamson 1996; Ferré-D’Amare et al. 1998; Treiber et al. 1998; Pan and Woodson 1999; Russell et al. 2000) form a well-defined 3D enzymatic shape with pseudoknots. The HDV ribozyme is a small, single-stranded RNA satellite of hepatitis B virus (HBV), which can enhance the virulence of HBV infections. HDV replicates by a double rolling-circle model and the nascent RNA is processed into monomers by self-cleavage of the genomic or antigenomic ribozyme (Lai 1995; Taylor et al. 1996; Taylor 2006; Kapral et al. 2014). Pseudoknot structure is also important for HDV ribozyme function (Perrotta and Been 1990, 1991; Wadkins et al. 1999). A detailed understanding of the folding pathway of this ribozyme may provide insight into replication of HDV and help identify targets for therapeutics.

The folding of functional RNA structures is often coupled with the transcription process (Sharma et al. 2010; Ameur et al. 2011; Tilgner et al. 2012; Brugiolo et al. 2013; Hamperl and Cimprich 2014). For instance, the functional native structure of the tetrahymena group I intron may form within the timescale of transcription, which is much faster than the
refolding of the complete chain in vitro (Brehm and Cech 1983; Wu and Tinoco 1998; Treiber and Williamson 2001; Heilman-Miller and Woodson 2003). It has been proposed that natural RNAs can effectively avoid the formation of misfolded structures during the cotranscriptional folding process (Kramer and Mills 1981; Treiber and Williamson 1999; Wong et al. 2007), but the mechanism is still not completely clear. In vivo, during the transcription elongation, because the upstream has already been folded, this will further influence the folding pathways of the downstream section. Recently, a few experiments have studied the self-cleavage activity of the HDV ribozyme during transcription (Chadalavada et al. 2000, 2002, 2007; Diegelman-Parente and Bevilacqua 2002). A clear and detailed understanding of the kinetic process of RNA folding including the folding pathways during transcription is crucial for uncovering the mechanism of RNA functions.

Many theoretical methods have been used to study the RNA folding kinetics. A few works have attempted to improve the energy parameters of pseudoknots (Cao and Chen 2005, Zhang et al. 2008). Molecular dynamic and Monte Carlo (MC) simulation approaches have been used to study the transition states, kinetic intermediates, and folding trajectories for a few specific sequences (Gultyaev 1991; Flamm et al. 2000; Isambert and Siggi 2000; Sorin et al. 2002; Krasovska et al. 2005; Yingling and Shapiro 2005; Danilova et al. 2006; Lin and Thirumalai 2008; Veeraraghavan et al. 2010), but the methods are limited to short timescales as a result of the restrictions of the computational efficiency and incomplete conformational sampling. Though coarse-grained models have already been developed to study RNA folding (Cho et al. 2009; Denesyuk and Thirumalai 2011; Shi et al. 2011), they are still limited to the simple H-pseudoknot folding without long life intermediate states. Based on coarse-grained kinetic moves, which can effectively reduce the number of conformations, several computational methods have been developed (Mironov et al. 1985; Geis et al. 2008; Tang et al. 2008; Hofacker et al. 2010); however, these simplified kinetic moves would miss the important folding pathways for some sequences/structures. Cao and Chen (Cao and Chen 2006) investigated RNA pseudoknot folding and unfolding kinetics with a combined master equation and kinetic cluster approach, but the method is also limited to short chains with small conformational equilibrated macro-states.

In this study, we combine the methods of predicting the folding kinetics of the pseudoknots (Chen et al. 2014), the cotranscription folding kinetics theory (Zhao et al. 2011; Gong et al. 2015a) of the secondary structures, and the transition node approximation method (Gong et al. 2015b) for long RNAs to establish a systematic method to predict the cotranscription folding kinetics of a long RNA chain with pseudoknots. Then we further studied how the flanking regions affect the HDV self-cleavage function during transcription.

RESULTS

Cotranscriptional folding for 99-nt HDV ribozyme can avoid forming trap intermediates

To test the validity of the transition node approximation, we first predict the cotranscriptional folding kinetics of the 99-nt sequence (which contains 84-nt ribozyme sequence and 15-nt downstream flanking sequence) for the HDV ribozyme at a transcription rate of 15 nt/sec using the original theory (Zhao et al. 2011) and the approximation method. Applying the transition node approximation reduces the number of conformations from 1884 to 142 at the end of transcription. As shown in Figure 1, the results show that from step 1 to step 99, the population kinetics of the main states are almost identical for the two methods. This indicates that the approximation effectively reduces the number of conformations and can reliably predict the cotranscriptional folding kinetics of longer nascent RNA chains. Hence, we would use the new method to study the cotranscriptional folding kinetics of HDV with different lengths.

For the 99-nt sequence, as the chain grows, the nascent RNA chain folds to the native state C8 through a series of discrete intermediate states (C0-C8) (Fig. 1): (i) When the first 13 nucleotides (nt) are released by the RNAP, the hairpin structure C1 is formed. (ii) At step 29, structure C1 quickly converts to structure C2 by adding a new hairpin. (iii) From step 37, most of the structure C2 quickly transits to structure C3, and a small fraction of structure C2 converts to structure C4. Besides, structure C3 can slowly transit to structure C4. (iv) At step 63, structure C3 almost all converts to structure C5, which only stays several steps. (v) From step 64, a part of helix P4 begins to form, so the structures C5 and C4 quickly convert to structures C6 and C7 by adding helix P4, respectively. At the same time, structure C6 can slowly transit to structure C7. (vi) From the 81st step, structure C6 quickly converts to native structure C8, and a small fraction transits to C7. At the end of the transcription, the inactive structure C7 only occupies ~10.18% and the active structure C8 occupies ~78.37%.

The cotranscriptional folding of the wild 99-nt HDV ribozyme can effectively avoid the meta-stable intermediate C7, which occupies ~50% of the population and lasts ~30 min and then transits to the native state in the refolding process (Chadalavada et al. 2000, 2002). The cotranscriptional folding also shows bifurcation folding behavior from step 34, one pathway as C2-C3-C5-C6-C8 would directly fold to the native state, another pathway folds to the intermediate state along C2-C4-C7. Although structure C2 begins to directly convert to structures C3 and C4 at the same time from step 34, the transition rate from C2 to C3 k_{C2-C3} \approx 1.48 \times 10^2 \text{ sec}^{-1} is larger than that from C2 to C4, k_{C2-C4} \approx 3.09 \times 10^0 \text{ sec}^{-1}, which is not a zipping process due to a bulge loop. Even though the structure C4 (ΔG = −21.19 kcal/mol) is more stable than structure C3 (ΔG = −19.31 kcal/mol), when the sequence grows to
step 39, most of the population structure C2 fluxes to structure C3 instead of C4, which indicates that the cotranscriptional folding behavior is a nonequilibrium folding kinetics. However, although the structure C3 could transit to C4 through a helix exchanging pathway (helix P1 exchange to helix AltP1), according to the steady-state approximation, the rates between the two structures could be obtained through Equation 8 as follows: \( k_{C3 \rightarrow C4} \approx 0.04 \text{ sec}^{-1} \), \( k_{C4 \rightarrow C3} \approx 1.65 \times 10^{-1} \text{ sec}^{-1} \). Along the two pathways C3–C5–C6–C8 and C4–C7, the transitions are very fast through adding a new helix at each transition step, so most of the population would transit to the native state C8 from C3, and only a small fraction of the population would transit to the intermediate state C7 from C4. Structure C6 could also transit to structure C7 along the pathway as the transition from C3 to C4 through helix AltP1 exchanging with helix P1 and has the same transition rates. So from step 39 to 81, at which structure C8 is more stable than C7, the population flow from the upper pathway to the lower pathway at the time window \((81 - 39)/\upsilon\) sec could be approximated as \( P_{\text{lower}} = p_3(1 - e^{-k_{3 \rightarrow 4}t}) \approx 10.06\% \). To further explore the effects of the transcription rate on the cotranscriptional folding kinetics, the cotranscriptional folding behavior with transcription rate 40 nt/sec has been studied. As shown in Figure 1C, the only difference is that the intermediate state C7 has an even lower population. This is because the time window \(42/\upsilon\) sec from the upper pathway to the lower pathway is shorter at the transcription rate 40 nt/sec than that at 15 nt/sec, the population for structure C7 decreases from 10.18% to 3.93%. Although at step 99 the population of structure C8 (~78.38%) decreases to 53.83%, the other population is in structure C6, and would
transit to C8, \( \sim 0.18 \) sec. The results indicate that accelerating the transcriptional rate can lead to more conformation flux to native structure. As the transcription rate further increases to 200 nt/sec, the lower pathway almost vanishes and at the end of transcription, the intermediate structure only occupies \( \sim 1.87\% \) (see Fig. 2A). Although the structures (C4 and C7) on the slow pathway have the lowest free energy during the transcription until step 81, as the cotranscription folding is a nonequilibrium kinetic process, even at the slowest transcription rate, they could only occupy a small portion of the population, which is far less than its equilibrium population (\( \sim 90\% \)). At a higher transcription rate, there would be less time to get equilibration, so the intermediate state would get less population.

Transcription often pauses at U-rich sites (Gusarov and Nudler 1999; Artsimovich and Landick 2000). Recently, it was found that pausing at a special site could lead to different folding structures (Helmling et al. 2018). Here, we use the 99-nt HDV ribozyme sequence as a model to test the effects of transcription pausing on the folding. Suppose the pausing time is 45 sec at step 39; the folding results are different from that without pausing (Fig. 2B). At the end of transcription, the inactive intermediate structure C7 occupies \( \sim 80\% \) of the population. Transcription pausing would provide extra time for the system to get equilibration, as the intermediate structures (C4 and C7) have the lowest free energy, so more population would go to the intermediate structure.

The 30-nt upstream fragment can lead to HDV ribozyme inactivity by forming an Alt1 helix

To explore how the flanking sequence affects the genomic HDV self-cleavage during transcription, the cotranscription folding kinetics of the 129-nt HDV ribozyme (from –30 to 99 fragment, with a 30-nt upstream) were also studied. As the chain increases, the nascent chain folds into a series of discrete intermediate states and at the end of transcription, the structure S12 occupies \( \sim 87.1\% \) and the structure C13 is \( \sim 10.2\% \). Both of them do not contain the pseudoknot structure (Fig. 3). So the 30-nt upstream sequence leads to the HDV in cleavage-inactive states. The experimental results showed that for this sequence, the reaction was not complete after 24 h and helix Alt1 was formed (Chadalavada et al. 2000; Diegelman-Parente and Bevilacqua 2002). Our results are consistent with the experiments. Although there are little differences among the folded structures before nucleotide 63 is transcribed for the 129- and 99-nt sequences, they all fold into the same structures, denoted by S10 and C6, respectively, when nucleotides 64–81 are transcribed. For the 99-nt sequence, structure C6 would transit to the cleavage active state C8. For the 129-nt sequence, the structure S10 could also transit to the cleavage active state S14 with the same rate as that from C6 to C8. But from the nucleotide 82 transcribed, due to the ability of the upstream segment to form an Alt1 helix with nucleotides 79–86, structure S10 could also transit to structure S12 by adding the Alt1 helix. As structure S10 transits to S12 through adding the Alt1 helix, the transition rate from S10 to S12 \( k_{S10\rightarrow S12} = 4.13 \times 10^3 \) sec\(^{-1}\) is much larger than that from S10 to S14. Also, the free energy of S12 \( \Delta G = -52.70 \) kcal/mol) is much lower than that of S14 \( \Delta G = -43.45 \) kcal/mol). So, most of the population would transit to the cleavage inactive state S12.

The 54-nt upstream region can restore HDV ribozyme activity by forming a self-structure P(–1) helix

For the 153-nt HDV ribozyme (from –54 to 99, with a 54-nt upstream segment), the nascent chain folds into a series of discrete intermediate states, and at the end of transcription, the active structure X12 occupies \( \sim 83.1\% \), while the inactive structure X13 occupies only 5.63% (Fig. 4), which is different from that of the 30-nt upstream flanking sequence but is similar to that of the 99-nt sequence. For the 153-, 129-, and 99-nt sequences, when the nucleotide 73 is released, the ribozyme regions (from nucleotide 1 to nucleotide 73) of the three sequences fold to the same structures denoted as C6, S10, and X10, respectively. The main difference is that for the 153-nt sequence, structure X10 includes a stable helix P(–1), which consists of the upstream sequence (from –54 to –17). Because the early formed P(–1) helix is stable, exchanging this helix with Alt1 is not only thermodynamically unfavorable but also kinetically unfavorable. So the P(–1) helix can prevent the 30-nt upstream segment from pairing with nucleotides 79–86 to form an Alt1 helix. Structure
X10 would transit to the cleavage active state X12 as the similar ways of the structure C6 transits to structure C8 for the 99-nt sequence. For this sequence, the experiments (Chadalavada et al. 2000; Diegelman-Parente and Bevilacqua 2002) detected that the RNA was nearly completely cleaved during transcription and detected the hairpin P(−1). Our prediction is in agreement with the experimental results.

**The cotranscriptional folding is important for the HDV ribozyme with a 55-nt downstream region to perform cleavage activity**

To further study the effects of the downstream flanking sequence on the cotranscriptional folding kinetics, we studied the cotranscriptional folding kinetics of the 140-nt (85-nt ribozyme region and 55-nt downstream fragment) sequence. When nucleotide 110 is released, it folds as the same structures as that of the 99 nt, the active structure C8 occupies ~85% of the population, and the inactive structure C7 occupies ~12% of the population (Fig. 5). When nucleotide 113 is transcribed, the downstream segment would fold to a hairpin AltP5, and then C8 would transit to structure C10 by adding the helix AltP5. If there is further elongation to nucleotide 128, the downstream segment would fold to a more stable helix P5, and the structure would transit from C10 to C14 through helix exchanging. Meanwhile, a small portion of the population would transit to C12 from C7 as that from C8 to C14. The self-cleavage structure C14 occupies ~81.5%, and the inactive structure C13 occupies ~14.7%. But when the downstream segments 129 to 132 are transcribed, this...
segment could form base pairs with the segments 80 to 83 of the ribozyme region, structure C12 would transit to C13. But this helix could not coexist with the P2 helix of active structure C14. Although the free energy of structure C13 ($\Delta G = -56.78 \text{ kcal/mol}$) is lower than that of C14 ($\Delta G = -53.80 \text{ kcal/mol}$), which occupies most of the population before C13 could be formed due to the cotranscription folding, as the transition rate from C14 to C13 $k_{C14 \rightarrow C13} \approx 0.04 \text{ sec}^{-1}$ is much slower than the self-cleavage rate $k_c = 40 \text{ min}^{-1}$ (Diegelman-Parente and Bevilacqua 2002), most of the native structure is self-cleaved before the native structure C14 transits to the alternative structure C13. For this sequence the experimental results showed that the RNA was almost completely cleaved during transcription and inferred the existence of helix P5 (Diegelman-Parente and Bevilacqua 2002; Chadalavada et al. 2007). Our prediction is consistent with the experimental results.

**DISCUSSION**

The HDV ribozyme, a human pathogen, has genomic and antigenomic versions of a small ribozyme embedded in its 1.7-kb RNA genome (Lai 1995; Lazinski and Taylor 1995; Karayiannis 1998). By virtue of their roles in rolling-circle viral replication, these small ribozymes have evolved to function cotranscriptionally in the presence of flanking sequence. To study the cotranscriptional folding kinetics, we have extended the newly developed RNA cotranscription folding
kinetics theory (Zhao et al. 2010, 2011; Gong et al. 2015a) and the transition node approximation method (Gong et al. 2015b) with pseudoknots (Chen et al. 2014). In this approach, the transition rates for an elementary step are as follows: (i) formation, (ii) disruption of a helix stem, and (iii) helix formation with concomitant partial melting of an incompatible helix, which are calculated with the free energy landscape. Based on the creation/disruption and exchange of helices, we investigated the cotranscription folding behavior of the HDV ribozyme at discrete steps.

The cotranscriptional folding kinetics for the 99-nt HDV ribozyme sequence, two upstream flanking sequences, and one downstream flanking sequence are studied with this new model. Our results indicate that this method is reliable for calculating the cotranscriptional folding kinetics of long mRNA chains. According to our calculations, we found that the cotranscriptional folding kinetics for the 99-nt HDV ribozyme is different from its refolding, which folds into an intermediate trap state, while during transcription, it can effectively avoid the trap intermediates and quickly fold to the cleavage-active state. It is dramatically different from its refolding kinetics. However, the existence of the 30-nt upstream flanking sequence can inhibit the ribozyme region folding into the active native state through forming an alternative helix Alt1 with the segments 70–90. The longer upstream flanking sequence 54 nt itself forms a stable hairpin structure that sequesters the formation of Alt1 helix and leads to rapid formation of cleavage-active structure. Although the 55-nt downstream flanking sequence could invade the already formed active structure during transcription by forming a more stable helix with the ribozyme region, the slow transition rate could keep the structure in the cleavage-active structure to perform the activity.

Our method, based on the transition node approximation, can efficiently reduce the conformation ensemble to study the cotranscriptional folding kinetics of long RNA chains. The transcription rate and transcription pausing, which can be regulated by RNAP, concentration of ribonucleoside triphosphate (rNTP), proteins and other cellular environments, could also be incorporated into our methods. However, the current theory has limitations: The free energy parameters of this model are of RNA at 1 M NaCl solution condition, and the effect of the Mg$^{2+}$ ions on these parameters is neglected. But Mg$^{2+}$ can significantly stabilize the tertiary interactions (Draper 2004; Draper et al. 2005; Grilley et al. 2006; Tan and Chen 2010, 2011) and thus it may possibly alter the folding pathways. Furthermore, we could incorporate the tertiary interaction into this method. First, we obtain all the tertiary structures formed during transcription from the predicted secondary structures; a few methods have been developed to predict the tertiary structures from the secondary structures (Das and Baker 2007; Parisien and Major 2008; Popenda et al. 2012; Zhao et al. 2012; Xu et al. 2014). Then we could get the tertiary interaction through a molecular dynamical simulation (Wang et al. 2016). Hence, this method should be improved to enhance its efficiency as well as incorporate the Mg$^{2+}$ effects and the effects of other cofactors, such as divalent metal ions.

**FIGURE 5.** (A) The population kinetics of main states (C0–C14) formed during the transcription at a transcription rate of 15 nt/sec for HDV ribozyme with 55-nt downstream segment. (B) The folded structures and the transition pathways during the transcription. C0 to C7 are the same as in Figure 1.
counterion-specific binding and tertiary interaction, into the model.

MATERIALS AND METHODS

Master equation

Assuming that there are Ω states in the RNA conformation space, the population kinetics \( P_i(t) \) for each state \( i \) in the conformation space at time \( t \) can be described by the master equation:

\[
\frac{dP_i(t)}{dt} = \sum_j \left[ k_{j \rightarrow i} P_j(t) - k_{i \rightarrow j} P_i(t) \right],
\]

where \( \Sigma \) denotes the sum over all the conformations, and \( k_{j \rightarrow i} (k_{i \rightarrow j}) \) is the rate constant for the transition from state \( i \) to state \( j \) (state \( j \) to state \( i \)). The master equation can be written as the following matrix form:

\[
\frac{dP(t)}{dt} = M \cdot P(t),
\]

where \( P(t) = (P_1(t), P_2(t), \ldots, P_\Omega(t)) \) is the fractional population vector, and \( M \) is the \( \Omega \times \Omega \) rate matrix with elements \( M_{ij} = k_{i \rightarrow j} \) for \( i = j \) and \( M_{ij} = -\sum_{i \neq j} k_{j \rightarrow i} \) for \( i \neq j \).

Solving the master equation can yield the probability kinetics given the initial folding condition \( t > 0 \) as the following equation:

\[
P(t) = \sum_{m=1}^{\infty} C_m \eta_m e^{-\lambda_m t},
\]

where \( \eta_m \) and \( -\lambda_m \) are the \( m \)-th eigenvalue and eigenvector of rate matrix \( M \), respectively, and \( C_m \) is the coefficient that is dependent on the initial condition.

Generating conformation space

In our model, the rates for formation \( (k_+ \) and disruption \( (k_- \) of a base stack can be calculated as (Zhao et al. 2010): \( k_+ = k_0 e^{-\Delta G_ \text{form}/k_B T} \) and \( k_- = k_0 e^{-\Delta G_ \text{disrupt}/k_B T} \), respectively, where \( k_0 \) is a prefactor, \( k_B \) is the Boltzmann constant, \( T \) is the temperature, \( \Delta G_ \text{form} \) and \( \Delta G_ \text{disrupt} \) are the energy barriers for the formation and disruption of a stack. Recently, it has been validated that the kinetic barrier for the formation of a base stack is concomitant entropic decrease: \( \Delta G_\text{form} = T \Delta S_\text{form} \), and that for the disruption of a base stack is the enthalpic cost: \( \Delta G_\text{disrupt} = \Delta H_\text{disrupt} \), from molecular dynamic simulation (Wang et al. 2016). Thus, the rates for formation and disruption of a base stack (not closing the loop) can be written as follows:

\[
k_+ = k_0 e^{-\Delta S_\text{form}/k_B T}, \quad k_- = k_0 e^{-\Delta H_\text{disrupt}/k_B T}.
\]

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

\[
k_+ = k_0 e^{-(\Delta S_\text{loop} + \Delta S_\text{stack})/k_B T},
\]

\[
k_- = k_0 e^{-\Delta H_\text{loop}/k_B T},
\]

where \( -\Delta S_\text{loop} \) is the entropy change of the loop, \( -\Delta S_\text{stack} \) and \( -\Delta H_\text{loop} \) are the entropy and enthalpy changes upon formation or disruption of the stack, respectively. The prefactor \( k_0 \) is equal to 6.6 \times 10^{2} \text{sec}^{-1} \) for the formation/disruption of a GC base pair and 6.6 \times 10^{3} \text{sec}^{-1} \) for an AU base pair (Zhang and Chen 2009; Zhao et al. 2010, 2011, 2014; Chen and Zhang 2012; Chen et al. 2014).

The rate for the formation of a base stack is usually larger than that of disrupting the stack, except the loop closing stack under the folding condition; hence, once the first few stacks in a helix are closed and stabilized, zipping of the subsequent stacks in the helix would be fast and it can quickly slip into the fully folded helix (Zhao et al. 2010). This suggests that it is proper to use the helices as building blocks for the study of the overall (i.e., slower) folding kinetics. In our model, the conformation space is constructed upon the helix-based building blocks and the kinetic move is the addition or deletion of a helix or an exchange between two helices (Chen et al. 2014).

In our model (Zhao et al. 2010), RNA structures are constructed by helices, which consist of consecutive base stacks. There are three types of relationships between two helices (Fig. 6): 1. Compatible: The two helices have no overlapping nucleotides with each other (Fig. 6A and B for the secondary structure and the pseudoknot structure, respectively). 2. Partially compatible: The two helices have partial overlapping nucleotides with each other (Fig. 6C and D for the secondary structure and the pseudoknot structure, respectively). 3. Incompatible: The two helices overlapping with each other (Fig. 6E).

Each structure must consist of compatible or partially compatible helices. If helix \( H_{m+1} \) is compatible with all the helices \( H_1 (1 \leq i \leq m) \) of the structure containing \( m \) helices \( \{H_1, H_2, \ldots, H_i, \ldots, H_m\} \), then a new structure with \( m+1 \) helices \( \{H_1, H_2, \ldots, H_i, \ldots, H_m, H_{m+1}\} \) can be formed by adding the new helix \( H_{m+1} \) to the \( m \)-helix structure. However, if helix \( H_{m+1} \) is partially compatible with helix \( H_m \) and compatible with all other helices, adding the helix \( H_{m+1} \) would involve an ensemble of \( (m+1) \)-helix conformations, which contain partially melted helix \( H_m \) and partially melted helix \( H_{m+1} \). Because some base pairs in helix \( H_m \) would prohibit the formation of certain base pairs in helix \( H_{m+1} \), the disruption of such incompatible base pairs in \( H_m \) would allow the formation of base pairs in \( H_{m+1} \).

The free energy of conformations without pseudoknots is calculated by the nearest-neighbor model (Xia et al. 1998; Mathews et al. 1999). The free energy of pseudoknots related conformations is calculated with the model proposed by Chen et al. (2014) based on the model of Eddy (Rivas and Eddy 1999). In the model, the stability of the stacks does not change whether they are involved in a pseudoknot or not, while the free energy of the loop in a pseudoknot is calculated as follows:

\[
G_{ps} = 0.83G_{\text{in}} + 0.2n_f + 0.1n_p, \quad \text{for } n_f \leq 9,
\]

\[
G_{ps} = 0.83G_{\text{in}} + 0.2(9 + \log(n_f/9)) + 0.1n_p, \quad \text{for } n_f > 9,
\]

FIGURE 6. Relationship of two helices, compatible: (A) and (B), partially compatible: (C) and (D), incompatible: (E).
where $G_p$ is the free energy of a pseudoknot loop, $G_s$ is the free energy of loops before the pseudoknot is formed, $n_f$ is the number of free bases in the pseudoknot, and $n_p$ is the number of paired bases in the pseudoknot.

**Kinetic move set and rate constant calculation**

**Rate of adding or deleting a helix**

As under the folding condition, the zipping pathway is the most probable pathway of forming a helix (Zhao et al. 2010). For the zipping pathway (Fig. 7A), as the rate for forming the first loop-closing stack is much smaller than that for adding a new stack to the existing stack, so the formation of the first stack is the rate-limited for formation of the helix. Hence, the free energy landscape shows a downhill profile after the formation of the third base stack. The rate $k_f$ of the helix formation (along one specific pathway) could be approximated as the rate for the formation of the three-stack state (Zhao et al. 2010):

$$k_f = k_{3→2}K_1 \left(1 - K_2K_4 \sum_{n=0}^{∞} (K_2K_4)^n\right)$$

$$= k_{3→2}K_1 \left(1 - K_2K_4 \frac{1}{1 - K_2K_4}\right), \quad (5)$$

where $k_{3→2}$ denotes the transition rate from state $i$ to $j$, $K_i$ and $K_j$ are the forward and reverse probability of state $i$:

$$K_1 = \frac{k_{2→3}}{k_{2→3} + k_{2→1}}, \quad K_2 = \frac{k_{3→4}}{k_{3→4} + k_{3→2}}, \quad K_4 = \frac{k_{3→2}}{k_{3→4} + k_{3→2}}. \quad (6)$$

After the first base stack could be formed anywhere inside the helix for a given RNA molecular, the rate $k_f$ for formation of a helix is the sum of the rates along all the pathways (Fig. 7B) with the different first (nucleation) base stacks. The rate for deleting the helix can be estimated from the detailed balance condition:

$$k_d = k_f e^{-\Delta G/kT}, \quad (7)$$

where $\Delta G$ is the free energy difference between the two structures.

**Rate of exchanging between two helices**

If two helices are incompatible, they cannot coexist in the same structure. The most probable pathway of the conversion from helix $A$ to helix $B$ is the tunneling pathway (Fig. 8), where after the first two base stacks of helix $A$ are disrupted, in each subsequent step, disruption of a stack in helix $A$ is accompanied by formation of a stack in helix $B$. Based on the tunneling pathway, the rate for helix exchange can be calculated as (Zhao et al. 2010):

$$k_{A→B} = \frac{\prod_{j=0}^{n} k_{i→i+1}}{\prod_{i=0}^{∞} k_{i→i+1}} \cdot \frac{k_{B→A}}{k_{B→A}}$$

$$= k_{A→B} e^{-\Delta G_A/kT}, \quad (8)$$

where $k_{i→i+1}$ is the rate constant for the formation (disruption) of a base stack in $A$ ($B$), and $k_{B→A}$ is the rate constant for the disruption (formation) of a base stack in $A$ ($B$). $\Delta G_{A,B}$ is the free energy difference between structure $A$ and structure $B$.

**Cotranscriptional folding kinetics**

In our model (Zhao et al. 2011), releasing one nucleotide by RNA polymerase (RNAP) to freely form possible structures can be regarded as a transcriptional step. If the transcription speed of a RNA sequence is $v$ nucleotides per seconds, the (real) time window for $M$-th nucleotide (nt) to be (newly) transcribed to the $(M+1)$-th nucleotide will be $1/v$ sec. Then the $M$-nt chain samples, the conformation space $\Omega$ in the $1/v$ sec and its population distribution is relaxed from $[p_1(M)_{\text{beg}}, p_2(M)_{\text{beg}}, \ldots, p_M(M)_{\text{beg}}]$ to $[p_1(M)_{\text{end}}, p_2(M)_{\text{end}}, \ldots, p_M(M)_{\text{end}}]$, where the $p_i(M)_{\text{beg}}$, $p_i(M)_{\text{end}}$ are the populations of state $i$ at the beginning and the end of the step $M$, respectively. This is defined as the $M$-nt step. For each step, the population kinetics is calculated in the same manner.

The beginning population of the $M$-th step can be inherited from the ending population of the $(M-1)$-th step. According to the possible changes of the structures upon the extension of the chain by one nucleotide, the structures in the current step can be classified as four types (Fig. 9). Types a and b: The newly transcribed $M$-th nucleotide does not pair with any nucleotides. The new $M$-nt chain can retain the same structure as the $(M-1)$-nt chain. Type c: The newly transcribed nucleotide can pair with an upstream nucleotide to elongate a helix by one base pair. Because the zipping of a new stack (base pair) is much faster than transcribing a nucleotide (Zhao et al. 2011), the two structures can be recognized as "directly inherited" and thus have the same population. Type d: the newly
A transcribed nucleotide can pair with another nucleotide to form a new helix, which cannot be formed in the previous step. In this category, all structures in the current step containing the new helix have a population of zero at the beginning of the $M$-th step. Then the population distribution at the beginning of step $M$ can be summarized by

$$p(M)_{\text{begin}} = p(M-1)_{\text{end}} \quad \text{for } a, b, \text{ and } c;$$

$$p(M)_{\text{begin}} = 0 \quad \text{for } d.$$  

The folding kinetics for the RNA chain during the transcription can be computed by applying this method from the first step to the end of transcription.

**Transition node approximation**

All the possible structures of the $M$-th step can be divided into two ensembles (Fig. 9): $\Omega_{\text{old}}$, in which the states are of the $a$, $b$, and $c$ types and with nonzero initial population; and $\Omega_{\text{new}}$, in which the states (type $d$) contain the newly formed helix at this step and with zero initial population. Most of the initial population at each transcriptional step is mainly concentrated in a few stable and metastable states (Zhao et al. 2010), which comprise the subensemble $\Omega_{\text{old}}^h$. The threshold for the initial population of states in $\Omega_{\text{old}}^h$ is set to be 0.03, which can ensure that above 90% of the population is occupied by states in $\Omega_{\text{old}}^h$ at each transcription step for HDV. If the new states are much more unstable than these old states in $\Omega_{\text{old}}^h$, it is impossible for them to obtain a population. Assuming that the free energy of the most unstable states in $\Omega_{\text{old}}^h$ is $\max E$, then the new conformation ensemble $\Omega_{\text{new}}$ can be divided into two sub-ensembles: $\Omega_{\text{new}}^h$, in which all the free energy of each conformation is lower than $(\max E + 2)$ kcal/mol; and $\Omega_{\text{new}}^r$, which contains the rest conformations in $\Omega_{\text{new}}$. As the free energy of these structures in $\Omega_{\text{new}}^h$ is at least +2 kcal/mol higher than those in $\Omega_{\text{old}}^h$, so they can occupy far less than 1% of the population even at equilibrium. Thus, for the newly formed conformations, only those in $\Omega_{\text{new}}^h$ are possible to have the population accumulation. Although states in $\Omega_{\text{new}}^r$ are unlikely to contribute to the population aggregation, some conformations may be on the main pathways for population flow. So the reduced assembly conformation space consists of structures in $\Omega_{\text{old}}^h \cup \Omega_{\text{new}}^h$, and parts of $\Omega_{\text{new}}^r$.

Because the transcriptional time scale of HDV is smaller than 15 sec, if the rate of one transition pathway is much less than (0.0044) sec$^{-1}$, there almost will be no RNAs folded via this pathway. Hence, if the transition rates along the direction of the population flow slower than (0.0044) sec$^{-1}$, they are unlikely to contribute to the folding. Thus, we only consider the pathways through states in $\Omega_{\text{old}}^h$ that can transit to states in $\Omega_{\text{new}}^h$, with corresponding transition rates greater than (0.0044) sec$^{-1}$ and if there is no pathway between the two states, then we will relax the criteria to nine nodes or more to avoid isolate states in the transition network. After all the possible transitions are searched, the structures in $\Omega_{\text{old}}^h, \Omega_{\text{new}}^h$, together with the structures in $\Omega_{\text{new}}^r$ located in the saved transition network. Assuming that the free energy of the most unstable states in $\Omega_{\text{old}}^h$ is $\max E$, then the new conformation ensemble $\Omega_{\text{new}}$ can be divided into two sub-ensembles: $\Omega_{\text{new}}^h$, in which all the free energy of each conformation is lower than $(\max E + 2)$ kcal/mol; and $\Omega_{\text{new}}^r$, which contains the rest conformations in $\Omega_{\text{new}}$. As the free energy of these structures in $\Omega_{\text{new}}^h$ is at least +2 kcal/mol higher than those in $\Omega_{\text{old}}^h$, so they can occupy far less than 1% of the population even at equilibrium. Thus, for the newly formed conformations, only those in $\Omega_{\text{new}}^h$ are possible to have the population accumulation. Although states in $\Omega_{\text{new}}^r$ are unlikely to contribute to the population aggregation, some conformations may be on the main pathways for population flow. So the reduced assembly conformation space consists of structures in $\Omega_{\text{old}}^h \cup \Omega_{\text{new}}^h$, and parts of $\Omega_{\text{new}}^r$.

Because the transcriptional time scale of HDV is smaller than 15 sec, if the rate of one transition pathway is much less than (0.0044) sec$^{-1}$, there almost will be no RNAs folded via this pathway. Hence, if the transition rates along the direction of the population flow slower than (0.0044) sec$^{-1}$, they are unlikely to contribute to the folding. Thus, we only consider the pathways through states in $\Omega_{\text{old}}^h$ that can transit to states in $\Omega_{\text{new}}^h$, with corresponding transition rates greater than (0.0044) sec$^{-1}$ and if there is no pathway between the two states, then we will relax the criteria to nine nodes or more to avoid isolate states in the transition network. After all the possible transitions are searched, the structures in $\Omega_{\text{old}}^h, \Omega_{\text{new}}^h$, together with the structures in $\Omega_{\text{new}}^r$ located in the saved transition network. Assuming that the free energy of the most unstable states in $\Omega_{\text{old}}^h$ is $\max E$, then the new conformation ensemble $\Omega_{\text{new}}$ can be divided into two sub-ensembles: $\Omega_{\text{new}}^h$, in which all the free energy of each conformation is lower than $(\max E + 2)$ kcal/mol; and $\Omega_{\text{new}}^r$, which contains the rest conformations in $\Omega_{\text{new}}$. As the free energy of these structures in $\Omega_{\text{new}}^h$ is at least +2 kcal/mol higher than those in $\Omega_{\text{old}}^h$, so they can occupy far less than 1% of the population even at equilibrium. Thus, for the newly formed conformations, only those in $\Omega_{\text{new}}^h$ are possible to have the population accumulation. Although states in $\Omega_{\text{new}}^r$ are unlikely to contribute to the population aggregation, some conformations may be on the main pathways for population flow. So the reduced assembly conformation space consists of structures in $\Omega_{\text{old}}^h \cup \Omega_{\text{new}}^h$, and parts of $\Omega_{\text{new}}^r$.
ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China under grant numbers 11574234 and 31270761 (to W.Z.) and grant number 31600592 (to S.G.).

REFERENCES

Ameur A, Zaghloul A, Halvardson J, Wetterborn A, Gylensten U, Cavelier L, Feuk L. 2011. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol* 18: 1435–1440.

Artisimovich I, Landick R. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc Natl Acad Sci* 97: 7090–7095.

Batey RT, Gilbert SD, Montange RK. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* 432: 411–415.

Brehm SL, Cech TR. 1983. The fate of an intervening sequence RNA: excision and cyclization of the *Tetrahymena* ribosomal RNA intervening sequence in vivo. *Biochemistry* 22: 2390–2397.

Brugliolo M, Herzel L, Neugebauer KM. 2013. Counting on co-transcriptional splicing. *F1000prime Rep* 9: 9.

Cao S, Chen S. 2005. Predicting RNA folding thermodynamics with a reduced chain representation model. *RNA* 11: 1884–1897.

Cao S, Chen ST. 2006. Predicting RNA pseudoknot folding thermodynamics. *Nucleic Acids Res* 34: 2634–2652.

Chadalavada DM, Knudsen SM, Nakano S, Bevilacqua PC. 2000. A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme. *J Mol Biol* 301: 349–367.

Chadalavada DM, Senchak SE, Bevilacqua PC. 2002. The folding pathway of the genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots. *J Mol Biol* 317: 559–575.

Chadalavada DM, Cerrone-Szakal AL, Bevilacqua PC. 2007. Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. *RNA* 13: 2189–2201.

Cheah MT, Wachtler A, Sudarsan N, Breakefield RR. 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* 447: 497–500.

Chen J, Greider CW. 2005. Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci* 102: 8080–8085.

Chen J, Zhang W. 2012. Kinetic analysis of the effects of target structure on siRNA efficiency. *J Chem Phys* 137: 225102.

Chen J, Gong S, Wang Y, Zhang W. 2014. Kinetic partitioning mechanism of HDV ribozyme folding. *J Chem Phys* 140: 25102.

Cho SS, Pincus DL, Thirumalai D. 2009. Assembly mechanisms of RNA pseudoknots are determined by the stabilities of constituent secondary structures. *Proc Natl Acad Sci* 106: 17349–17354.

Comolli LR, Smirnov I, Xu L, Blackburn EH, James TL. 2002. A molecular switch underlies a human telomerase disease. *Proc Natl Acad Sci* 99: 16998–17003.

Cornish PV, Hennig M, Giedroc DP. 2005. A loop 2 cytidine-stem 1 minor groove interaction as a positive determinant for pseudoknot-stimulated 1 ribosomal framshefing. *Proc Natl Acad Sci* 102: 12694–12699.

Danilova LV, Pervouchine DD, Favorov AV, Mironov AA. 2006. RNAKinetics: a web server that models secondary structure kinetics of an elongating RNA. *J Bioinform Comput Biol* 4: 589–596.

Das R, Baker D. 2007. Automated de novo prediction of native-like RNA tertiary structures. *Proc Natl Acad Sci* 104: 14664–14669.

Denesyuk NA, Thirumalai D. 2011. Crowding promotes the switch from hairpin to pseudoknot conformation in human telomerase RNA. *J Am Chem Soc* 133: 11858–11861.

Diegelman-Parente A, Bevilacqua PC. 2002. A mechanistic framework for co-transcriptional folding of the HDV genomic ribozyme in the presence of downstream sequence. *J Mol Biol* 324: 1–16.

Draper DE. 2004. A guide to ions and RNA structure. *RNA* 10: 335–343.

Draper DE, Grilley D, Soto AM. 2005. Ions and RNA folding. *Annu Rev Biophys Biomol Struct* 34: 221–243.

Ferré-D’Amare AR, Zhou K, Doudna JA. 1998. Crystal structure of a hepatitis delta virus ribozyme. *Nature* 395: 567–574.

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

Geis M, Flamm C, Wollinger MT, Tanzer A, Hofacker IL, Middendorf M, Mandl C, Stadler PF, Thurner C. 2008. Folding kinetics of large RNAs. *J Mol Biol* 379: 160–173.

Gesteland RF, Atkins JJ. 1996. Recoding: dynamic reprogramming of translation. *Annu Rev Biochem* 65: 741–768.

Giedroc DP, Theimer CA, Nixon PL. 2000. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J Mol Biol* 298: 167–185.

Gong S, Wang Y, Zhang W. 2015a. Kinetic regulation mechanism of pbeU riboswitch. *J Chem Phys* 142: 15103.

Gong S, Wang Y, Zhang W. 2015b. The regulation mechanism of ytl and mef riboswitches. *J Chem Phys* 143: 45103.

Grilley D, Soto AM, Draper DE. 2006. Mg2+–RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc Natl Acad Sci* 103: 14003–14008.

Grundy FJ, Henkin TM. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol Microbiol* 30: 737–749.

Guliyeva AP. 1991. The computer simulation of RNA folding involving pseudoknot formation. *Nucleic Acids Res* 19: 2489–2494.

Gusarov I, Nudler E. 1999. The mechanism of intrinsic transcription termination. *Mol Cell* 3: 495–511.

Hamperl S, Cimprich KA. 2014. The contribution of co-transcriptional RNA: DNA hybrid structures to DNA damage and genome instability. *DNA Repair* 19: 84–94.

Heilman-Miller SL, Woodson SA. 2003. Effect of transcription on folding of the *Tetrahymena* ribozyme. *RNA* 9: 722–733.

Helming C, Klotzner DP, Sochor F, Mooney RA, Wacker A, Landick R, Fürig B, Heckel A, Schwabe H. 2018. Life times of metastable states guide regulatory signaling in transcriptional riboswitches. *Nat Commun* 9: 944.

Hofacker IL, Flamm C, Heine C, Wollinger MT, Scheuermann G, Stadler PF. 2010. BarMap: RNA folding on dynamic energy landscapes. *RNA* 16: 1308–1316.

Hoffmann S, Otto C, Doose G, Tanzer A, Langenberger D, Christ S, Kunz M, Holdt LM, Teupser D, Hackermüller J, et al. 2014. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol* 15: R34.

Isambert H, Siggia ED. 2000. Modeling RNA folding paths with pseudoknots: application to hepatitis delta virus ribozyme. *Proc Natl Acad Sci* 97: 6515–6520.

Kapral GI, Jain S, Noeske J, Doudna JA; Richardson DC, Richardson JS. 2014. New tools provide a second look at HDV ribozyme structure, dynamics and cleavage. *Nucleic Acids Res* 42: 12833–12846.

Karayiannis P. 1998. Hepatitis D virus. *Rev Med Virol* 8: 13–24.

Kim Y, Su L, Maas S, O’Neill A, Rich A. 1999. Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. *Proc Natl Acad Sci* 96: 14234–14239.

Kramer FR, Mills DR. 1981. Secondary structure formation during RNA synthesis. *Nucleic Acids Res* 9: 5109–5124.

Krasovska MV, Sefčíková J, Spáčková N, Šponer J, Walter NG. 2005. Structural dynamics of precursor and product of the RNA enzyme from the hepatitis delta virus as revealed by molecular dynamics simulations. *J Mol Biol* 351: 731–748.

Lai MM. 1995. The molecular biology of hepatitis delta virus. *Annu Rev Biochem* 64: 259–286.

Lazinski DW, Taylor JM. 1995. Regulation of the hepatitis delta virus ribozymes: to cleave or not to cleave? *RNA* 1: 225.

www.rnajournal.org
Lehnert V, Jaeger L, Michele F, Westhof E. 1996. New loop–loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the Tetrahymena thermophila ribozyme. *Chem Biol* 3: 989–1009.

Li S, Breaker RR. 2013. Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. *Nucleic Acids Res* 41: 3022–3031.

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

Lin J, Thirumalai D. 2013. Kinetics of allosteric transitions in S-adenosylmethionine ribozymes are accurately predicted from the folding landscape. *J Am Chem Soc* 135: 16641–16650.

Marrone A, Walne A, Doka I. 2005. Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr Opin Genet Dev* 15: 249–257.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288: 911–940.

Mercer TR, Dinger ME, Mattick JS. 2009. Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10: 155–159.

Mironov AA, Dyakonova LP, Kister AE. 1985. A kinetic approach to the prediction of RNA secondary structures. *J Biomol Struct Dyn* 2: 953–962.

Neupane K, Yu H, Foster DA, Wang F, Woodside MT. 2011. Single-molecule force spectroscopy of the add adenine riboswitch reveals folding to regulatory mechanism. *Nucleic Acids Res* 39: 7677–7687.

Pan J, Woodson SA. 1999. The effect of long-range loop-loop interactions on folding of the *Tetrahymena* self-splicing RNA. *J Mol Biol* 294: 955–965.

Parisien M, Major F. 2008. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452: 51.

Perrotta AT, Been MD. 1990. The self-cleaving domain from the genomic RNA of hepatitis delta virus: sequence requirements and the effects of denaturant. *Nucleic Acids Res* 18: 6821–6827.

Perrotta AT, Been MD. 1991. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* 350: 434–436.

Plant EP, Dinman JD. 2005. Torsional restraint: a new twist on frame-shifting pseudoknots. *Nucleic Acids Res* 33: 1823–1833.

Plant EP, Jacobs KL, Harger JW, Meskauskas A, Jacobs JL, Baxter JL, Petrov AN, Dinman JD. 2003. The 9-A solution: how mRNA pseudoknots promote efficient programmed −1 ribosomal frameshifting, *RNA* 9: 168–174.

Popenda M, Szachniuk M, Antczak M, Purzycka KJ, Lukasiak P, Bartol N, Blazejcz W, Adamiak RW. 2012. Automated 3D structure composition for large RNAs. *Nucleic Acids Res* 40: e112.

Reining A, Noziminov S, Schlepckow K, Buhr F, Furtig B, Schwalbe H. 2013. Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature* 495: 355–359.

Rivas E, Eddy SR. 1999. A dynamic programming algorithm for RNA structure prediction including pseudoknots. *J Mol Biol* 285: 2053–2068.

Russell R, Millett IS, Doniach S, Herschlag D. 2000. Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nat Struct Mol Biol* 7: 367–370.

Schultes EA, Bartel DP. 2000. One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science* 289: 448–452.

Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Russell R, Millett IS, Doniach S, Herschlag D. 2000. Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Biophys J* 101: 176–187.

Taylor JM. 2006. Structure and replication of hepatitis delta virus RNA. *Curr Top Microbiol Immunol* 307: 1–23.

Taylor DR, Lee SB, Romano PR, Marshak DR, Hinnebusch AG, Esteban M, Mathews MB. 1996. Autophosphorylation sites participate in the activation of the double-stranded-RNA-activated protein kinase PKR. *Mol Cell Biol* 16: 6295–6302.

Theimer CA, Finger LD, Trantirek L, Feigon J. 2003. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci* 100: 449–454.

Tilgner H, Knowles DG, Johnson R, Davis CA, Chakrabortty S, Djebali S, Curado J, Snyder M, Gingeras TR, Guigo R. 2012. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for IncRNAs. *Genome Res* 22: 1616–1625.

Treiber DK, Williamson JR. 1999. Exposing the kinetic traps in RNA folding. *Curr Opin Struct Biol* 9: 339–345.

Treiber DK, Williamson JR. 2001. Beyond kinetic traps in RNA folding. *Curr Opin Struct Biol* 11: 309–314.

Treiber DK, Rook MS, Zarrinkar PP, Williamson JR. 1998. Kinetic intermediates trapped by native interactions in RNA folding. *Science* 279: 1943–1946.

Yingling YG, Shapiro BA. 2005. Dynamic behavior of the telomerase ribonucleic acid secondary structures. *Proc Natl Acad Sci* 102: 1470–1475.

Zhao P, Zhang W, Chen S. 2010. Predicting ion binding properties for RNA tertiary structures. *J Am Chem Soc* 132: 125107.

Zhao Y, Huang Y, Zhou G, Wang Y, Man J, Xiao Y. 2012. Automated and fast building of three-dimensional RNA structures. *Sci Rep* 2: 734.