Molecular dynamics study of ways of RNA base-pair formation

Qiangqiang Duan, Peng Tao, Jun Wang, and Yi Xiao*
School of Physics, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China

(Received 11 December 2019; revised 14 July 2020; accepted 21 August 2020; published 8 September 2020)

Base pairing is a fundamental phenomenon in RNA structure and function. However, although there have been considerable recent advances, some important aspects of base-pair formation are still unknown, including the ways of base-pair formation and detailed roles of metal ions. Here we show that base pairs can form through four different ways: stabilizing, bridging, rotating, and shifting. Among them the stabilizing and bridging ways involve direct binding of metal ions while the rotating and shifting ways do not in most cases. Furthermore, we find that the formations of base pairs in different positions of the hairpin stem may adopt different ways.

DOI: 10.1103/PhysRevE.102.032403

I. INTRODUCTION

Noncoding RNAs perform many cellular functions through base pairing with other RNA or DNA and within themselves to form tertiary structures [1,2]. Although such base pairing is a well-known feature of RNA, the pairing mechanism is a long-standing unsolved problem [3–21]. Recently, Xu *et al.* [22] studied the elementary steps of RNA base pairing through a single nucleotide folding by using a hybrid approach (molecular dynamics simulation, kinetic Monte Carlo simulation, and master equation methods) and revealed the pathways and transitions of base pairing. They also showed that metal ion binding around the base region preceded the formation of a stable base pair and ion dissociation from the base region will cause base-pair destabilization. Wang *et al.* [23,24] studied the kinetic mechanism of base-pair opening and closing through simulating a single terminal base pair near its melting temperature. They obtained thermodynamic and kinetic parameters of a single base pair through molecular dynamics simulation and calculated the transition rates of base-pair opening and closing. These studies provided valuable insights into the mechanism of base-pair formation. However, some important aspects of base-pair formation are still unclear, including the ways of base-pair formation and how metal ions are involved in them.

Molecular dynamics (MD) simulation is a useful approach to study the detailed mechanism of RNA folding [8,11,15,25–42]. However, successful simulations of RNA folding were limited to a few small RNA hairpins and were done only by using sampling-enhanced replica exchange molecular dynamics [15,35,40,43,44], but not conventional molecular dynamics. One of the main reasons for this is that the current RNA force field underestimated the interactions between bases [43]. In spite of this difficulty in the simulation of the folding of entire RNA molecules, we found that conventional molecular dynamics simulation gave many events of base-pair formation. This provides an opportunity to use these events to analyze the detailed mechanism of base-pair formation.

Here we report our analysis of base-pair formations from simulated 233-μs folding trajectories of three RNA hairpins by using conventional molecular dynamics. Previous works studied the folding of a single base or a pair of bases in a stem under the condition that other base pairs of the stem were formed [22–24]. In this work, we go a step further and our simulations start from conformations in which no base pairs are formed. Thus, we can investigate possible ways of base-pair formation and whether the pairing way of the firstly formed base pair is different from those of subsequently formed ones. It is expected that metal ions should participate in the formation of base pairs because of the strong electronegativity in the backbone of the RNA molecule. However, we found four ways of base-pair formation in which two of them involve the ions directly, but two do not.

II. MATERIALS AND METHODS

Three RNA hairpins. Three RNAs were used in our study and their native structures are hairpins [Fig. 1(a)]. The first one (PDB ID: 1ZIH) has 12 nucleotides with a sequence GGGCG-CAAGCCU [11,45,46] and its native structure forms three Watson-Crick base pairs (4C-9G, 10C-3G, and 11C-2G) and a wobble base pair (1G-12U). The second one (PDB ID: 2AHT) also has 12 nucleotides (sequence: GGGUGUAAACCU) and its native structure forms three Watson-Crick base pairs (9A-4U, 10C-3G, and 11C-2G) and a wobble base pair (1G-12U). Different from 1ZIH, 2AHT has an A-U base pair. The third hairpin (PDB ID: 2KOC) has 14 nucleotides with a sequence GGCACUUCGGUGCC and its native structure contains five Watson-Crick base pairs (5C-10G, 4A-11U, 3C-12G, 13C-2G, and 14C-1G) and also has an A-U base pair. Our analyses mainly focus on the 1ZIH and the other two, 2AHT and 2KOC, are used to check if the ways of base-pair formation for 1ZIH are the same as for other RNAs.

Simulation details. Two MD simulation sets were generated in this work (major set: set I; minor set: set II,

*yxiao@hust.edu.cn

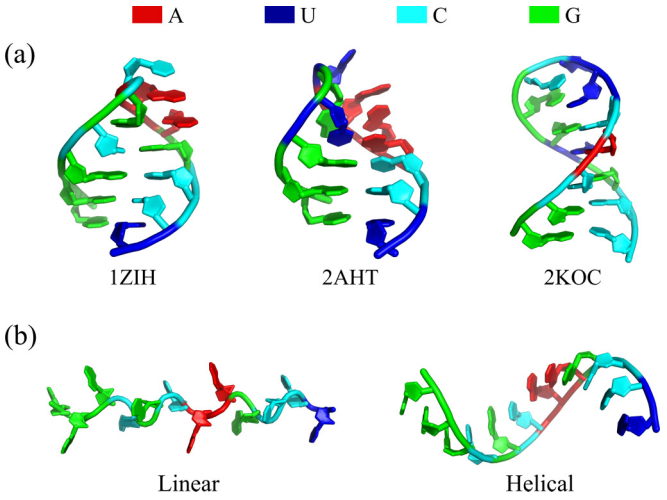

FIG. 1. (a) The cartoon representation of the native structures of three RNA hairpins, and the corresponding PDB IDs are shown at the bottom of the structures. (b) Two starting conformations of 1ZIH: single-strand linear conformation (left) and single-strand helical conformation (right). The nucleotides A, U, C, and G are colored in red, blue, cyan, and green, respectively.

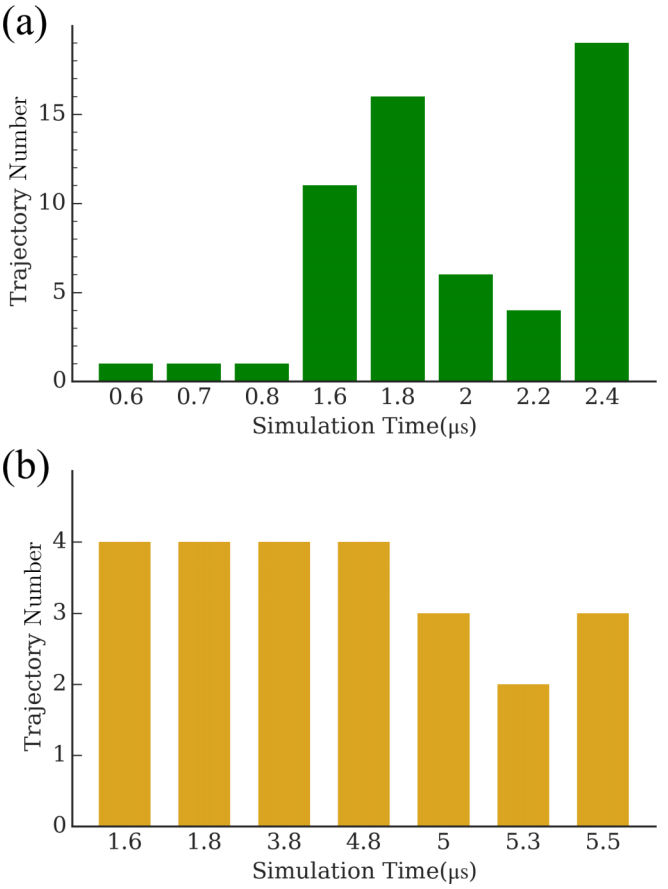

FIG. 2. The distribution of the number of simulations vs their lengths for simulations in (a) set I and (b) set II.

Table I). All MD simulations were performed using the AMBER14 [47] package and the force field used for the RNA molecule was AMBER ff99bsc0+χOL3 [48,49], consisting of a sugar-phosphate parameter refinement and a pucker flipping modification. The system was solvated in a TIP3P water cubic box with a water buffer of 12 Å.

In set I, the starting conformations of the three RNAs are single-strand extended ones generated from *tleap* in the AMBER14 software package [Fig. 1(b)]. For 1ZIH, the distribution of the number of simulations vs their lengths is shown in Fig. 2. For 2AHT and 2KOC, 100 and 40 200-ns-trajectories were simulated, respectively. All simulated systems were first neutralized by K^+ and then 0.3 mol/l extra KCl were added (denoted as K+KCl). For the force field parameters of K^+ and Cl^-, the hydration free energy (HFE) set by Li *et al.* was used, which is specifically parametrized for the TIP3P water to reproduce the hydration free energies using the 12-6 Lennard-Jones (LJ) nonbonded model (denoted as ions1lsm_hfe, Tables I and II) [50]. This parameter set improved the transferability of previous ones to ion-pair solutions.

To see the dependence of the simulation results on initial conformations, ion types, and ion parameter sets, we also performed additional simulations for 1ZIH (set II) with initial single-strand linear or helical conformation, different ion buffers, ion parameter sets, and/or initial ion positioning. The initial single-strand helical conformation was generated from NAB [51] in the AMBER14 package [Fig. 1(b)]. The additional ion buffers include Na^+ neutralized and 0.3 mol/l

TABLE I. Simulation sets with different initial conformation under varying cation buffers and ion parameter sets.

MD set	PDB ID	Initial conformation	Ion buffer	Ion parameter set	Number of trajectories	Timescale (μs)	Total (μs)
Set I	1ZIH	Linear	K+KCl	ions1lsm_hfe	59	114.9	142.9
	2AHT	Linear	K+KCl	ions1lsm_hfe	100	20	
	2KOC	Linear	K+KCl	ions1lsm_hfe	40	8	
Set II	1ZIH	Linear	K+KCl	ionsjc	4	20.5	90.1
	1ZIH	Linear	Na+NaCl	ionsjc	4	15.2	
	1ZIH	Helical	K+KCl	ions1lsm_hfe	4	7.2	
	1ZIH	Helical	K+KCl	ionsjc	4	19.2	
	1ZIH	Helical	Na+NaCl	ions1lsm_hfe	4	6.4	
	1ZIH	Helical	Na+NaCl	ionsjc	4	21.6	
Total					227	233	233

TABLE II. The LJ potential parameters of three kinds of ions in two parameter sets, including the van der Waals radius R_{min} and the depth of the potential well ε.

Ions	Parameter set	R_{min} (Å)	ε (kcal/mol)
Na^+	ionsjc	1.369	0.087
	ions1lsm_hfe	1.475	0.037
K^+	ionsjc	1.705	0.193
	ions1lsm_hfe	1.719	0.151
Cl^-	ionsjc	2.513	0.035
	ions1lsm_hfe	2.252	0.602

extra NaCl salt solvation (Na+NaCl). The additional ion force field parameter set is an older set parametrized for the TIP3P water using the 12-6 LJ-type nonbonded model by Joung and Cheatham [52] (ionsjc).

Before MD production, each trajectory was prepared with two steps of minimization and one heating step: (1) Minimize the buffer water and ions for 4000 steps, with atoms in RNA hairpin fixed; (2) minimize the whole system without any restraints for 10 000 steps; (3) heat the system from 0 to 300 K with an *NVT* ensemble, and the seed for the pseudorandom number generator was based on the simulation running date and time. After that, a 200-ps preproduction stage was

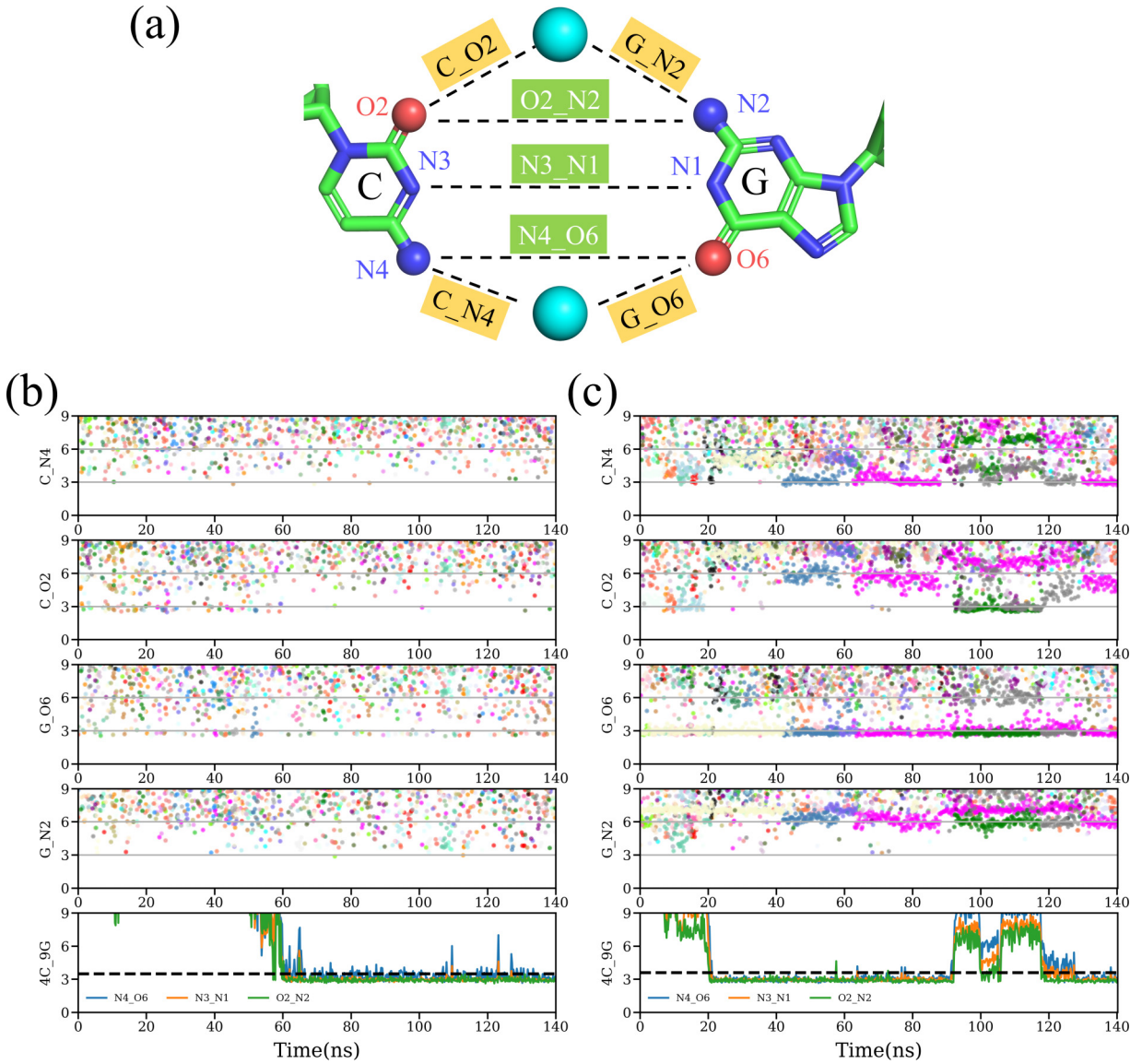

FIG. 3. (a) Seven important distances (dashed line) are analyzed in this work, including four distances (yellow rectangle) of each K^+ (cyan sphere) to the atoms N4 and O2 of cytosine (C_N4 and C_O2) and the atoms N2 and O6 of guanine (G_N2 and G_O6), and three distances of the atoms N4, N3, and O2 of cytosine to the atoms O6, N1, and N2 of guanine (N4_O6, N3_N1, and O2_N2, green rectangle). (b) An example of C_N4, C_O2, G_N2, and G_O6 changing over time (first four panels) in a fragment of a folding trajectory of 1ZIH, which is rescaled from 0 to 140 ns. Different K^+ ions are represented by dots with different colors. The last panel records the time courses of N4_O6, N3_N1, and O2_N2 distances of the C-G base pair. The black dashed line represents the distance cutoff of 3.6 Å. (c) Another example of the time courses of these distances in a fragment of a folding trajectory of 1ZIH shows the different dynamics of K^+ ions compared to the example in (b).

equilibrated using the parallel engine PMEMD.MPI running on a CPU.

In the production stage, an integration time step of 2 fs was employed. The SHAKE algorithm, which binds hydrogen to its neighbor heavy atoms, was used to enable the 2-fs time step. The temperature was kept by coupling to a Langevin heat bath (300 K) of 2 ps^{-1} frequency and the pressure was maintained around 1 bar by a Berendsen barostat with isotropic position scaling. A cutoff of nonbonded interactions was set to 12 Å, and the long-range electrostatic interactions were calculated by using the particle mesh Ewald [53] method with a grid spacing of 1 Å. The trajectory was restored every 10 000 steps, which means 20 ps per frame. The program PMEMD.CUDA [54] was used to generate all the trajectories, resulting in a rate of about 100 ns/day on the Nvidia GTX780 for the 1ZIH system.

Trajectory analysis. Here the standard for the formation of a base pair is that the three hydrogen bonds of a C-G base pair and the two hydrogen bonds of an A-U(G-U) base pair are formed. There are two criteria to define hydrogen bonds. The first criterion for the occurrence of a hydrogen bond is a maximum donor-acceptor distance of 3.5 Å and a minimum donor-proton-acceptor angle of 120°, which was used to extract the fragments of trajectories (usually 100 ns; see Fig. S1 in the Supplemental Material [55] for the entire trajectories of the base pairing) that include the formations of the base pairs. After that, a simpler distance-dependent criterion was applied to locate and visualize the formation of the base pairs; namely, all the donor-acceptor distances of the hydrogen bonds of these base pairs should be smaller than 3.6 Å. This is because this criterion is based on the stricter previous one and can also simplify subsequent analysis. All the hydrogen bonds and the distances were analyzed by CPPTRAJ [56] and all structure snapshots were generated by PYMOL [57].

III. RESULTS

We obtained 85 and 72 events of base-pair formation from the simulations of set I and set II, respectively. We shall analyze the 85 pairing events in set I in detail (see Figs. S2–S4 in the Supplemental Material [55] for each event).

The behavior of ions in base-pair formation. To see if the ions are involved in the base-pair formation, we calculated the distances of each ion to the atoms N4 and O2 of cytosine (denoted as C_N4 and C_O2) and the atoms N2 and O6 of guanine (denoted as G_N2 and G_O6) during the folding [Fig. 3(a)]. We find that the behaviors of ions can be divided into two types. Figure 3(b) shows a typical example for one type in which no ions bind to the bases directly and only some ions approach closely (<3.6 Å) to the atoms O6 of guanine and O2 of cytosine but leave quickly. Figure 3(c) is a typical example of another type in which there is always one, but not necessarily the same ion, that keeps binding (<3.6 Å) to the atom O6 of the guanine and O2 of cytosine. However, these types of ion behaviors happen for the atom O6 before and after the formation of base pairs, which is defined by three distances. That is to say, the distances between the atom O2 of cytosine and N2 of guanine (O2_N2), between the atom N3 of cytosine and N1 of guanine (N3_N1), and between the atom N4 of cytosine and O6 of guanine (N4_O6) are smaller than 3.6 Å. However, for atom O2, the bindings of ions only happen before the formation of base pairs and rarely happen after the formation of the base pairs. These phenomena are further confirmed by statistics of all events of base-pair formation (Fig. 4).

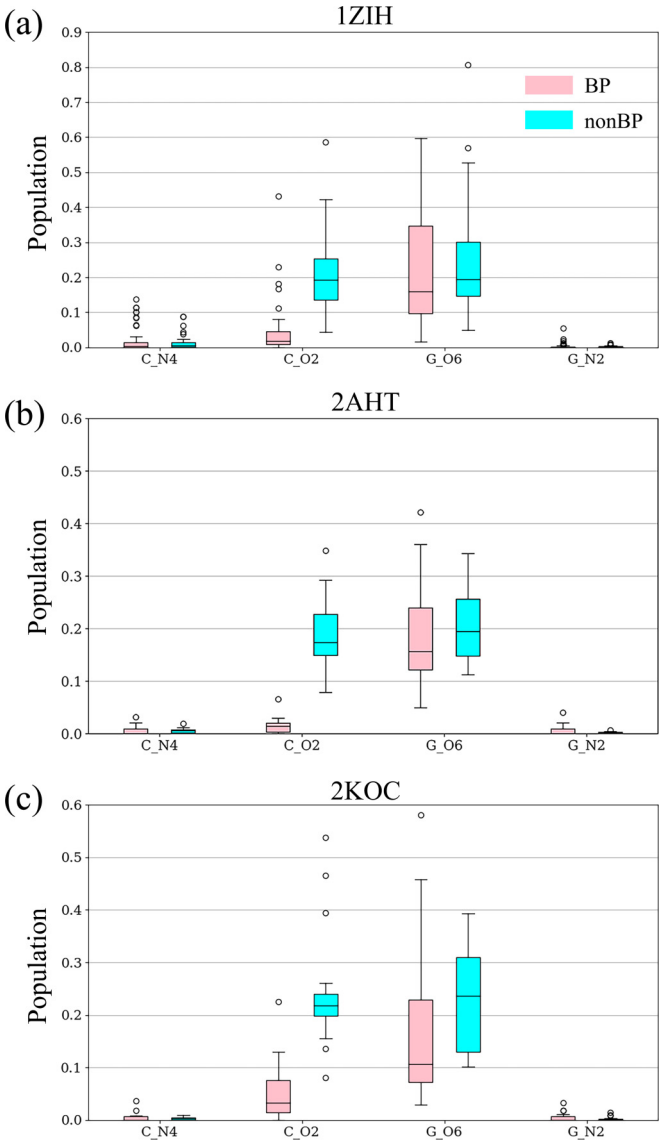

FIG. 4. The standard boxplots of the populations for four distances C_N4, C_O2, G_O6, and G_N2 less than 3.6 Å before (nonBP, pink) and after (BP, cyan) base-pair formation for all events of the base-pair formation of (a) 1ZIH, (b) 2AHT, and (c) 2KOC, respectively. The open circles represent singular values.

Ways of base-pair formation. We observed four ways of base pairing: stabilizing, bridging, rotating, and shifting (Fig. 5). The four pairing ways are defined in the following. First, we calculated the distances between all K^+ (or Na^+) and the specific atoms (N4 and O2 of C, and N2 and O6 of G) of the base pair. If there is not any K^+ binding to these atoms (the distance <3.6 Å) during the formation of the base pair, then the pairing way is assigned to be rotating. If there is K^+ binding to these atoms, then we check whether the K^+ keeps binding to one of the bases to limit its movement before and

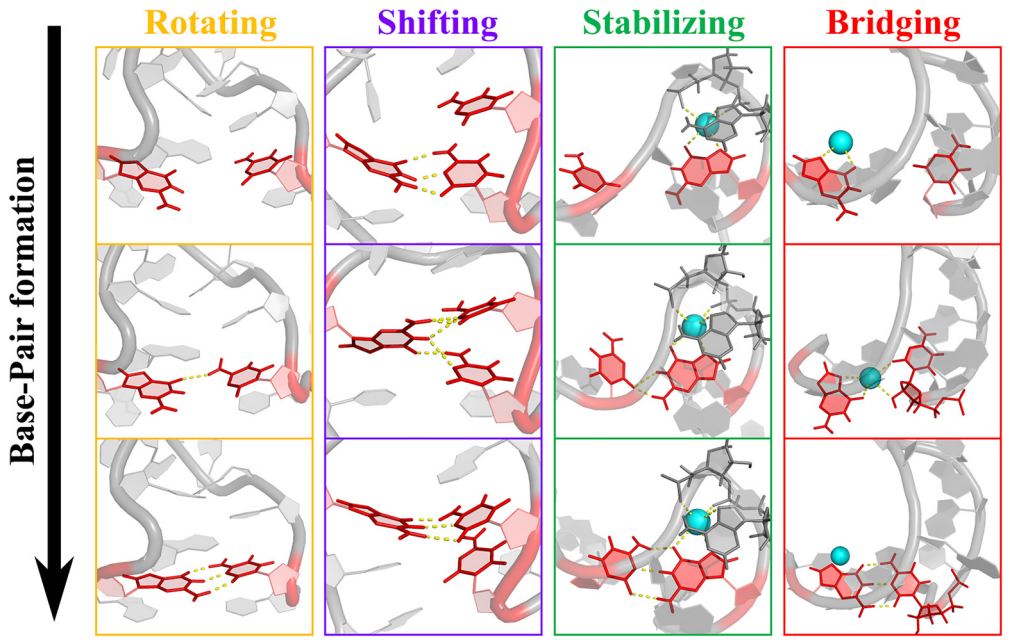

FIG. 5. Four representative ways of base-pair formation observed in this study. The balls in cyan are K^+ ions. The yellow dashed line represents that the distance between two atoms is smaller than 3.6 Å.

after the pairing. If it does, we define the way as stabilizing. If not, then we check whether the K^+ keeps binding to the two bases, or linking the two bases before the pairing, and disassociates after the pairing. If it does, the way is defined as bridging. Finally, if the formation of one base pair comes from a base pair (share one base in former base pair) that already exists (formed by the above three ways), we define the way as shifting.

Among the four ways, two of them (stabilizing and bridging ways) involve direct binding of the ions while the other two (rotating and shifting ways) do not in most cases. More specifically, among 43 rotating and 15 shifting ways, only 11 and 3 of them involve direct binding of the ions. However, in these situations, the movements of these ions are disordered and the bound ions rapidly disassociate after binding to the bases, which are treated as noise in this study since this kind of binding also exists even without pairing.

In the stabilizing way, the ion binds to one of the bases to limit the movement of the base, which makes the formation of the base pair easier. This is very similar to the single nucleotide folding in previous work where only one of the partners in a base pair can fold freely [22].

In the bridging way, the ion first binds to one base before pairing and then links to another base. Finally, the departure of the ion provides the opportunity for the formation of the base pair. The ion plays the role of a bridge that connects the two bases and force them to approach each other. This also makes the pair form easily.

In the shifting way, a base pair is firstly formed and then one of the bases will shift to form a new base pair with the neighbor base of its original partner. This is a possible mechanism of conformational adjustment, by which a first formed non-native base pair can transform into a native one.

In the rotating way, no ions directly bind to the two bases of G and C. The two bases can form a base pair by rotating their relative orientations from various directions. However, in this way, ions may play indirect roles in base pairing, e.g., binding to the loop region to bend the RNA strand, which can bring two bases closer (Fig. 6).

To make these four pairing ways easier to understand, we also provide an example for each way (Fig. 7). See Figs. S2–S4 in the Supplemental Material [55] for the formations of all

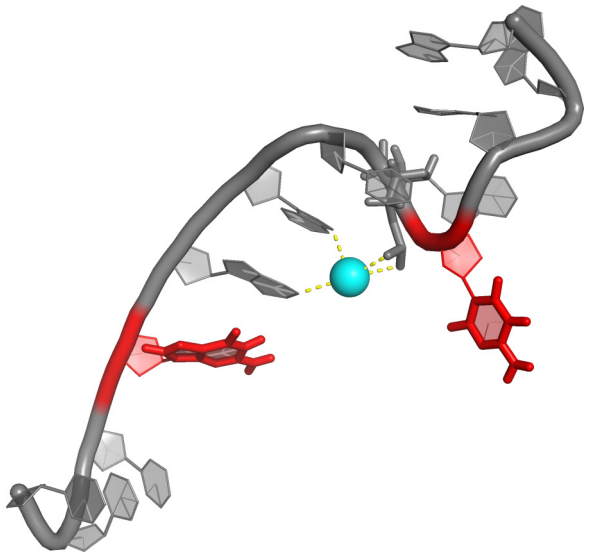

FIG. 6. An example of an indirect effect of ions on base-pair formation. The ball in cyan is a K^+ ion. In this example, the ion binds to the loop region of 1ZIH (not the base pair) to stabilize the local structure; thus it may be helpful to the formation of the base pair.

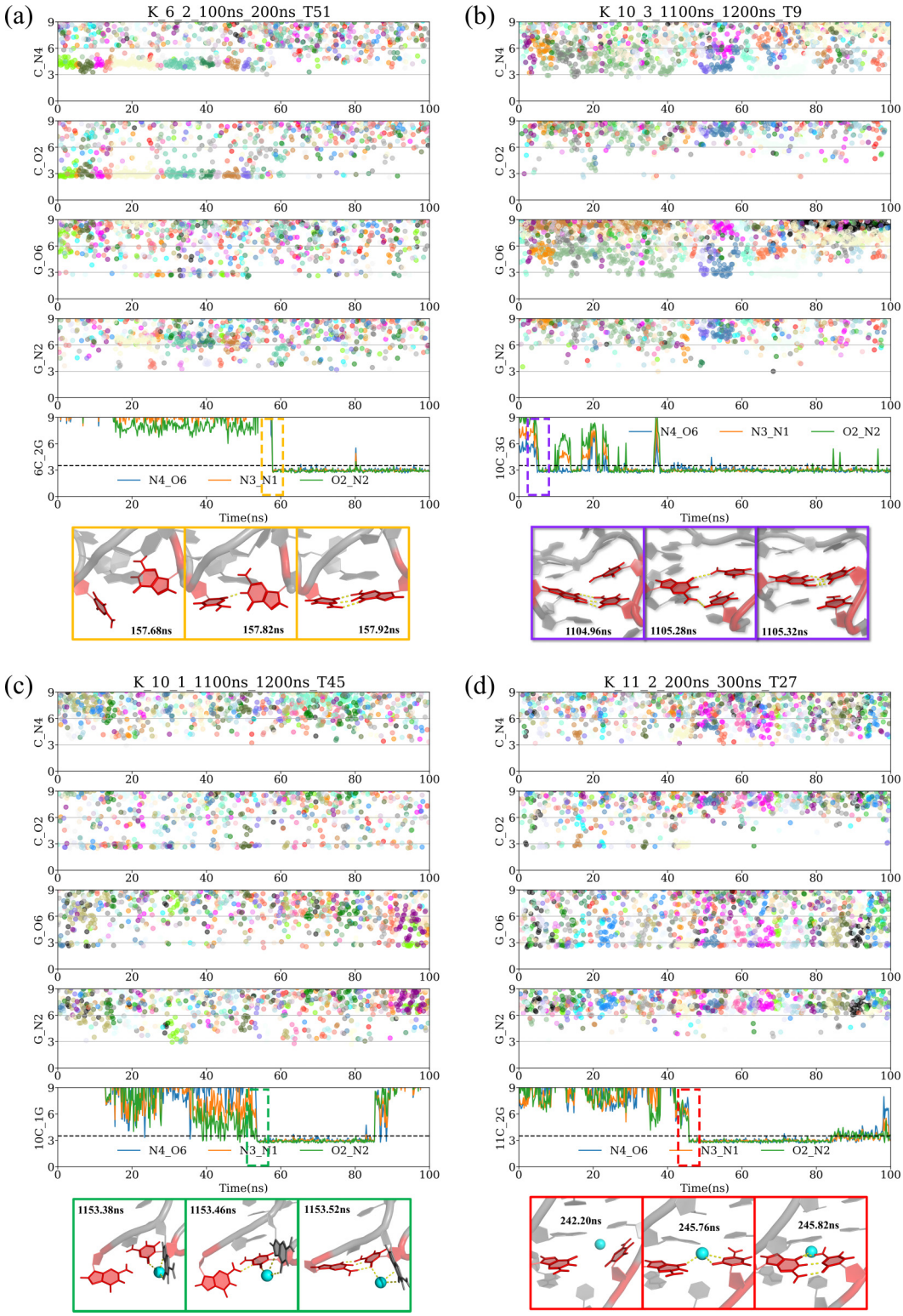

FIG. 7. Four examples of the formations of four base pairs in the (a) rotating way, (b) shifting way, (c) stabilizing way, and (d) bridging way, respectively. In subplot (a), the title "K_6_2_100ns_200ns_T51" means that this subplot records the formation of 6G-2C (K+KCl buffer) during the simulation time range from 100 to 200 ns in the 51st trajectory. The other three subplots (b–d) are the same as (a). In addition, the structural snapshots of the formation of these base pairs are extracted from the trajectories and the simulation times of these snapshots are also labeled.

TABLE III. The number of occurrences of four ways of base pairing in set I.

	Stabilizing	Bridging	Rotating	Shifting	Total
1ZIH	8	11	27	7	53
2AHT	5	0	6	5	16
2KOC	2	1	10	3	16
Total	15	12	43	15	85

the 85 base pairs. Table III shows the number of occurrences of each type of pairing way of 1ZIH, 2AHT, and 2KOC. Notably, the ways of base-pair formations of these three RNAs are the same and the occurrence of the rotating pairing ways is much more than the other three pairing ways. This may be because many non-native base pairs are formed in this way. Furthermore, we found the same ways of base-pair formation for set II (Table IV) and this indicates that our results do not depend on whether the initial conformation is linear or helical, the ion is K^+ or Na^+, and the ion force-field parameter set is new or old.

Folding pathways of the stem. Besides the events of base-pair formation, we also obtained four events of stem formation of the hairpin 1ZIH under different cation conditions and initial conformations (see Fig. S5 in the Supplemental Material [55] for these events). Since the end base pair G-U consists of only two hydrogen bonds, which is easy to be broken down or reformed, we consider the stem to be formed if the three C-G pairs (4C-9G, 10C-3G, and 11C-2G) have formed. These events enable us to investigate whether the base pairs in different positions of the hairpin stem have different pairing ways. For the five events, the pairing ways of the three C-G base pairs are listed in Table V. It is noted that the first one and last three events were simulated starting from linear and helical single-strand conformation, respectively.

It can be seen from Table IV that the order of base pairing in the hairpin is from the inner to the outer in most cases. However, whenever the order of base pairing—in most cases the firstly formed base pairs—adopted the stabilizing pairing way or shifting pairing way if a non-native pair formed firstly, the secondly formed native pairs adopted the rotating pairing way and the thirdly formed native pairs adopted the bridging pairing way. In particular, three 4C-9G adopted stabilizing ways, three 10C-3G adopted rotating ways and three 11C-2G adopted bridging ways. These results indicate that usually the formation of the middle pair does not have the direct participation of ions while the formation of the inner and outer pairs does.

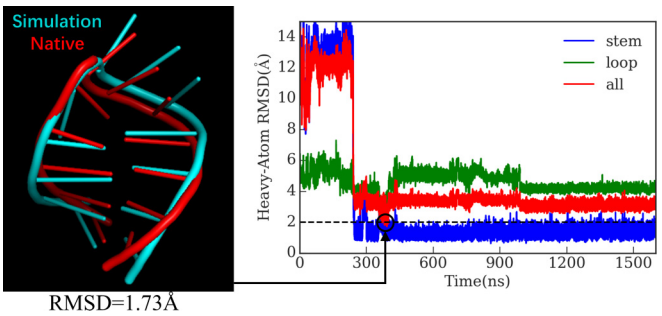

FIG. 8. The heavy-atom RMSDs of the entire hairpin (red), the stem (blue), and the loop (green) change over time in an example trajectory. The left inset shows the structural alignment with the lowest RMSD (1.73 Å) during the simulation.

Although we have observed four folding events of the stem, the folding of the entire hairpin is still difficult. Figure 8 shows an example of the time course of the heavy-atom root-mean-square deviations (RMSDs) of the entire hairpin (residue number: 1–12), the stem (residue number: 1–4 and 9–12), and the loop (residue number: 3–10) to the experimentally determined structure. It can be seen that the stem already forms around 250 ns; however, the loop region is trapped into a non-native conformation except for a transient native folding (RMSD $<$ 2 Å) around 400 ns, resulting in a native configuration with the lowest RMSD of 1.73 Å. This result may reflect a general defect of current RNA force fields in describing the flexibility of the loop regions.

IV. DISCUSSION

The results above show that there are mainly four ways of base pairing: stabilizing, bridging, rotating, and shifting. The results also show that the cations may directly participate in the formation of the base pair or may not. For the folded events of the hairpin, the formations of the inner and outer native C-G base pairs (4C-9G and 11C-2G) usually have the direct participation of cations and adopt stabilizing and bridging ways while the formations of the middle C-G base pair usually do not have direct participation of cations. In most cases, the inner base pair forms firstly. These results hold for both Na^+ and K^+ ions.

TABLE IV. The number of occurrences of four ways of base pairing in set II (1ZIH).

Initial structure	Ions buffer	Force field	Stabilizing	Bridging	Rotating	Shifting
Helical	Na+NaCl	Ionsjc	7	1	8	5
		Ions1lsm_hfe	1	2	3	1
	K+KCl	Ionsjc	7	8	1	–
		Ions1lsm_hfe	6	2	–	–
Linear	Na+NaCl	Ionsjc	1	–	8	–
	K+KCl	Ionsjc	6	2	2	1
Total			28	15	22	7

TABLE V. The pairing pathways of the three base pairs of 1ZIH, namely, the order of the formation of 4C-9G, 10C-3G, and 11C-2G. The corresponding ways of base-pair formation are shown inside the parentheses.

Ion buffer	Firstly formed nonnative pair	Firstly formed native pair	Secondly formed native pair	Thirdly formed native pair
K+KCl		4C-9G (Stabilizing)	10C-3G (Rotating)	11C-2G (Bridging)
K+KCl		4C-9G (Stabilizing)	10C-3G (Rotating)	11C-2G (Bridging)
Na+NaCl		4C-9G (Stabilizing)	10C-3G (Rotating)	11C-2G (Bridging)
Na+NaCl	11C-3G (Rotating)	4C-9G, 10C-3G, and 11C-2G (Shifting, formed simultaneously)		

Our results show that the cations not only neutralize the electronegativity of the backbone but also compensate the decreasing of the conformational entropy by binding to the bases. This binding can also stabilize one of the bases that are in pairing and plays a role that is very similar to previous single base folding with other base pairs and the partner base fixed.

It is noted that the probability of successful simulation of base pairing is very low. One of the main reasons for this may be that current force fields underestimated the hydrogen-bond interactions due to the missing of charge polarization effect [43]. The results above show that the binding of metal ions to the bases is also important to the formations of base pairs, especially because of the strong electronegativity of RNA backbone [58].

V. CONCLUSION

Here we reported our analysis of events of base-pair formation extracted from our simulated folding trajectories of three small RNA hairpins and found four ways of base-pair formation in the free folding situation: stabilizing, bridging, rotating, and shifting. The first two ways involve the ions directly and the last two mainly not. Our results may be helpful to the understanding of the base-pairing mechanism in RNA structure formation and RNA-DNA interactions.

ACKNOWLEDGMENT

This work is supported by the NSFC under Grants No. 11874162 and No. 31570722.

Q.D. and P.T. contributed equally to this work.

[1] G. Pljevaljcic, D. P. Millar, and A. A. Deniz, Biophys. J. **87**, 457 (2004).
[2] J. Wang, K. Mao, Y. Zhao, C. Zeng, J. Xiang, Y. Zhang, and Y. Xiao, Nucleic Acids Res. **45**, 6299 (2017).
[3] M. F. Hagan, A. R. Dinner, D. Chandler, and A. K. Chakraborty, Proc. Natl. Acad. Sci. USA **100**, 13922 (2003).
[4] F. Colizzi and G. Bussi, J. Am. Chem. Soc. **134**, 5173 (2012).
[5] S. Bottaro, F. Di Palma, and G. Bussi, Nucleic Acids Res. **42**, 13306 (2014).
[6] S. J. Chen, Annu. Rev. Biophys. **37**, 197 (2008).
[7] F. Colizzi, C. Perez-Gonzalez, R. Fritzen, Y. Levy, M. F. White, J. C. Penedo, and G. Bussi, Proc. Natl. Acad. Sci. USA **116**, 22471 (2019).
[8] N. A. Denesyuk, N. Hori, and D. Thirumalai, J. Phys. Chem. B **122**, 11860 (2018).
[9] N. A. Denesyuk and D. Thirumalai, J. Phys. Chem. B **117**, 4901 (2013).
[10] K. Sarkar, D. A. Nguyen, and M. Gruebele, RNA **16**, 2427 (2010).
[11] E. J. Sorin, M. A. Engelhardt, D. Herschlag, and V. S. Pande, J. Mol. Biol. **317**, 493 (2002).
[12] F. Wang, L. Z. Sun, P. Cai, S. J. Chen, and X. Xu, Biophys. J. **117**, 1674 (2019).
[13] Y. Wang, T. Liu, T. Yu, Z. J. Tan, and W. Zhang, RNA **26**, 470 (2020).
[14] G. Zhou, J. Loper, and S. Geman, BMC Bioinf. **20**, 666 (2019).
[15] G. Zuo, W. Li, J. Zhang, J. Wang, and W. Wang, J. Phys. Chem. B **114**, 5835 (2010).
[16] S. R. Choi, N. H. Kim, H. S. Jin, Y. J. Seo, J. Lee, and J. H. Lee, Comput. Struct. Biotechnol. J. **17**, 797 (2019).
[17] P. Sharma, M. Chawla, S. Sharma, and A. Mitra, RNA **16**, 942 (2010).
[18] J. Abi-Ghanem, C. Rabin, M. Porrini, E. Dausse, J. J. Toulme, and V. Gabelica, Chem. Phys. Chem. **18**, 2782 (2017).
[19] W. Stephenson, P. N. Asare-Okai, A. A. Chen, S. Keller, R. Santiago, S. A. Tenenbaum, A. E. Garcia, D. Fabris, and P. T. Li, J. Am. Chem. Soc. **135**, 5602 (2013).
[20] D. P. Aalberts and W. K. Jannen, RNA **19**, 475 (2013).
[21] M. Mosayebi, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. Doye, J. Phys. Chem. B **118**, 14326 (2014).
[22] X. Xu, T. Yu, and S.-J. Chen, Proc. Natl. Acad. Sci. USA **113**, 116 (2015).
[23] Y. Wang, S. Gong, Z. Wang, and W. Zhang, J. Chem. Phys. **144**, 115101 (2016).
[24] Y. Wang, Z. Wang, Y. Wang, T. Liu, and W. Zhang, J. Chem. Phys. **148**, 045101 (2018).
[25] L. Bao, J. Wang, and Y. Xiao, Phys. Rev. E **99**, 012420 (2019).
[26] J. Wang and Y. Xiao, Phys. Rev. E **94**, 040401(R) (2016).
[27] D. Thirumalai and C. Hyeon, Biochemistry **44**, 4957 (2005).
[28] H. T. Nguyen, N. Hori, and D. Thirumalai, Proc. Natl. Acad. Sci. USA **116**, 21022 (2019).
[29] A. Mandic, R. L. Hayes, H. Lammert, R. R. Cheng, and J. N. Onuchic, J. Phys. Chem. B **123**, 1505 (2019).
[30] L. Jin, Y. L. Tan, Y. Wu, X. Wang, Y. Z. Shi, and Z. J. Tan, RNA **25**, 1532 (2019).
[31] P. Ebrahimi, S. Kaur, L. Baronti, K. Petzold, and A. A. Chen, Methods **162–163**, 96 (2019).

[32] N. Hori, N. A. Denesyuk, and D. Thirumalai, J. Phys. Chem. B **122**, 11279 (2018).
[33] M. Havrila, P. Stadlbauer, P. Kuhrova, P. Banas, J. L. Mergny, M. Otyepka, and J. Sponer, Nucleic Acids Res. **46**, 8754 (2018).
[34] L. Z. Sun and S. J. Chen, Methods Mol. Biol. **1632**, 1 (2017).
[35] P. Kuhrova, R. B. Best, S. Bottaro, G. Bussi, J. Sponer, M. Otyepka, and P. Banas, J. Chem. Theory Comput. **12**, 4534 (2016).
[36] M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, and J. M. Bujnicki, Nucleic Acids Res. **44**, e63 (2016).
[37] Y. Bian, J. Zhang, J. Wang, J. Wang, and W. Wang, PLoS One **10**, e0129089 (2015).
[38] D. Chakraborty, R. Collepardo-Guevara, and D. J. Wales, J. Am. Chem. Soc. **136**, 18052 (2014).
[39] M. Faber and S. Klumpp, Phys. Rev. E **88**, 052701 (2013).
[40] A. A. Chen and A. E. Garcia, Proc. Natl. Acad. Sci. USA **110**, 16820 (2013).
[41] J. C. Lin, C. Hyeon, and D. Thirumalai, J. Phys. Chem. Lett. **3**, 3616 (2012).
[42] Z. Gong, Y. Zhao, C. Chen, and Y. Xiao, PLoS One **7**, e45239 (2012).
[43] P. Kuhrova, V. Mlynsky, M. Zgarbova, M. Krepl, G. Bussi, R. B. Best, M. Otyepka, J. Sponer, and P. Banas, J. Chem. Theory Comput. **15**, 3288 (2019).
[44] G. R. Bowman, X. Huang, Y. Yao, J. Sun, G. Carlsson, L. J. Guibas, and V. S. Pande, J. Am. Chem. Soc. **130**, 9676 (2008).
[45] E. J. Sorin, Y. M. Rhee, B. J. Nakatani, and V. S. Pande, Biophys. J. **85**, 790 (2003).
[46] E. J. Sorin, Y. M. Rhee, and V. S. Pande, Biophys. J. **88**, 2516 (2005).
[47] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**, 1668 (2005).
[48] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, Biophys. J. **92**, 3817 (2007).
[49] M. Zgarbova, M. Otyepka, J. Sponer, A. Mladek, P. Banas, T. E. Cheatham, and P. Jurecka, J. Chem. Theory Comput. **7**, 2886 (2011).
[50] P. Li, L. F. Song, and K. M. Merz, Jr., J. Chem. Theory Comput. **11**, 1645 (2015).
[51] T. J. Macke and D. A. Case, in *Molecular Modeling of Nucleic Acids* (ACS Publications, Washington, DC, 1998), p. 379.
[52] I. S. Joung and T. E. Cheatham, J. Phys. Chem. B **112**, 9020 (2008).
[53] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[54] R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand, and R. C. Walker, J. Chem. Theory Comput. **9**, 3878 (2013).
[55] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.102.032403 for plots with the entire trajectories of the base pairing, the ionic distribution around each base pair, and the snapshots of the formations of all base pairs and the stem of 1ZIH.
[56] D. R. Roe and T. E. Cheatham III, J. Chem. Theory Comput. **9**, 3084 (2013).
[57] S. Yuan, H. S. Chan, S. Filipek, and H. Vogel, Structure **24**, 2041 (2016).
[58] D. E. Draper, RNA **10**, 335 (2004).