Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

Biophysical *Journal*
Article

Biophysical Society

Simulation Study of the Plasticity of k-Turn Motif in Different Environments

Haomiao Zhang,[1] Haozhe Zhang,[1] and Changjun Chen[1,*]
[1]Biomolecular Physics and Modeling Group, School of Physics, Huazhong University of Science and Technology, Wuhan, Hubei, China

ABSTRACT The k-turn is a widespread and important motif in RNA. According to the internal hydrogen bond network, it has two stable states, called N1 and N3. The relative stability between the states changes with the environment. It is able to accept different conformations in different environments. This is called the "plasticity" of a molecule. In this work, we study the plasticity of k-turn by the mixing REMD method in explicit solvent. The results are concluded as follows. First, N1 and N3 are almost equally stable when k-turn is in the solvent alone. The molecule is quite flexible as a hinge. However, after binding to different proteins, such as the proteins L7Ae and L24e, k-turn falls into one global minimum. The preferred state could be either N1 or N3. On the contrary, the other nonpreferred state becomes unstable with a weaker binding affinity to the protein. It reveals that RNA-binding protein is able to modulate the representative state of k-turn at equilibrium. This is in agreement with the findings in experiments. Moreover, free energy calculations show that the free energy barrier between the N1 and N3 states of k-turn increases in the complexes. The state-to-state transition is greatly impeded. We also give a deep discussion on the mechanism of the high plasticity of k-turn in different environments.

SIGNIFICANCE We study the plasticity of k-turn alone and in the complexes with different proteins. The simulation results confirm that proteins L7Ae and L24e modulate the representative state of k-turn at equilibrium. We find that the free energy surface of k-turn changes when it is bound to different proteins, and it is also different to a free k-turn. Some critical intermediate states between the N1 and N3 states are determined. The simulation is performed by our recently published mixing replica exchange molecular dynamics method. The method combines the advantages of the existing biased and unbiased sampling methods. It can do the free energy calculation on any collective variables, no matter if they are involved in the bias potential or not.

INTRODUCTION

RNA has become an increasingly important chemical species for the understanding of the gene expression in life science. The messenger RNA is responsible for carrying the genetic information (1), the transfer RNA brings amino acids to the growing polypeptide chain (2), and the ribosomal RNA (rRNA) constitutes the ribosome and catalyzes the formation of the peptide bonds (3). In the three-dimensional structures of these molecules, there are some specific building blocks named recurrent RNA motifs (4,5). Among these motifs, kink-turn (k-turn) is the most common element in RNA (6). It was first observed in the crystal structure of 5′-stem loop of human U4 small nuclear RNA (7). Soon after that, Steitz et al. found six k-turns in the 2S rRNA of *Haloarcula marismortui* and two in the 16S rRNA of *Thermus thermophilus* (8).

The k-turn motif is critical to the protein and RNA structure and function. For example, by binding to the L7Ae protein family, it can stabilize the folded structure of the protein (9–11), and it determines the folding process of S-adenosylmethionine riboswitch (12). It is also believed that k-turn is involved in the other RNA functions, such as the translation and modification process, the spliceosome assembly, and the control of gene expression (13).

According to (14), a standard structure of k-turn is illustrated in Fig. 1 *a*. From the 5′ end to the 3′ end, the nucleotides on the loop are named L1, L2, and so on. Nucleotides on the strand that have a bulge take a suffix b and those on the other strand take a suffix n. The position numbers of the nucleotides increase outwards from the loop. They are negative and positive in the 5′ direction and 3′ direction,

Submitted March 23, 2020, and accepted for publication August 12, 2020.
*Correspondence: cjchen@hust.edu.cn
Editor: Wilma Olson.
https://doi.org/10.1016/j.bpj.2020.08.015

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

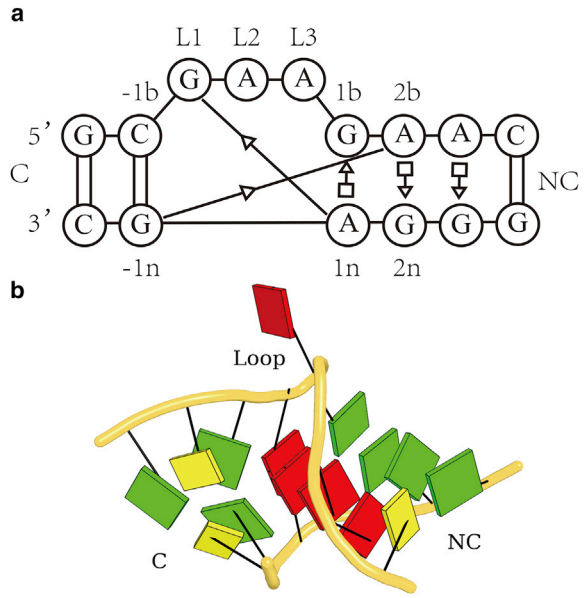

FIGURE 1 (*a*) The secondary structure of a typical k-turn motif in RNA. It contains three fragments: a canonical helix (C), a loop region (L), and a noncanonical helix (NC). Following the standard nomenclature in (14), the nucleotides on the loop has a prefix L. The nonloop nucleotides on the loop strand have a suffix b and those on the other strand take a suffix n. The figure is generated by RNAVIEW (65). (*b*) The tertiary structure of a k-turn drawn by Pymol (38) and DSSR (36). Different nucleotides are shown by different colored squares (A: *red*, C: *yellow*; and G: *green*). To see this figure in color, go online.

respectively. For convenience, we use this nomenclature in our article from now on. Interactions between the nucleotides are described by the Leontis-Westhof terminology (15).

Fig. 1 *b* gives the tertiary structure of a standard k-turn. It is commonly composed of three distinct elements: a noncanonical helix (NC), a canonical helix (C), and an internal loop (L). The two helical axes are not coplanar, and they form a V-shaped bend. Usually, the bending angle is between 45 and 55°. The tandem A:G basepairs with the *trans*-sugar-Hoogsteen interaction (tSH) (A(1n):G(1b), A(2b):G(2n)) are believed to be the core of the k-turn structure (16). Additionally, there are two important hydrogen bonds formed between the two basepairs (G(L1):A(1n), G(−1n):A(2b)) on the interface of the minor grooves of the NC and C helices.

Currently, many properties of k-turn have been reported by the simulations or experiments. The k-turn is a dimorphic motif, which has a less-bent structure in the free solution. But with ions like Mg^{2+}, the structure is tightly kinked (17). In 2007, Lilley et al. reported that the 2′-hydroxyl group of G(L1) forms a hydrogen bond with the N1 atom of A(1n). If the hydroxyl group is removed, the metal ion-induced folding process will be prevented (14). The k-turn motif can also fold when it binds to the L7Ae protein with a high affinity (18). A high temperature unfolding simulation indicates that the complex follows an induced fit mechanism upon the binding process (19). And moreover, k-turn exists within the riboswitch. The RNA tertiary interactions in riboswitch make the k-turn more stable (20). Simulations in the solution present that k-turn is flexible in RNA like a molecular hinge (6,21). By using the umbrella sampling method, it is found that different A-minor interactions are coupled with the k-turn's kinked-to-extended transition (22).

In general, two stable structures of k-turn are classified according to the special hydrogen bond between the nucleotide residue A(2b) and G(−1n). In one structure, N1, the N1 atom of A(2b) forms a hydrogen bond with the O2′ atom of G(−1n) (A-minor 1 motif). And in the other structure, N3, the first N1 atom of the bond is replaced by the N3 atom in A(2b) (23) (A-minor 0 motif). As shown in Fig. 2, the N1 class of k-turn contains three hydrogen bonds (H1, H3, and H4 in Fig. 2 *a*), and N3 contains four (H1, H2, H5, and H6 in Fig. 2 *b*) between G(−1n), A(2b), and G(2n). Although with a different hydrogen bond network, both of the two structures are stable to k-turn. The dominant structure of k-turn depends on its surrounding environment. For example, a typical k-turn named *H. marismortui* Kt-7 (HmKt-7) accepts the N1 structure in the ribosome. But when it is placed into a S-adenosylmethionine riboswitch, HmKt-7 changes to the N3 structure normally. Moreover, if HmKt-7 binds to the L7Ae protein as a free RNA fragment, it also has a N3 structure (23). After a statistical analysis of different biomolecules aside from the ribosome, all the 16 determined HmKt-7 structures are in N3 (16).

As discussed above, k-turn prefers different conformations in different environments. However, the relative stability between the N1 and N3 class has not yet been studied in depth by simulations. The detailed transition process is not very clear either. In this article, we carry out different simulations for k-turn in various environments, including a free k-turn in explicit solvent, a k-turn binding to the L7Ae protein, and a k-turn binding to the L24e protein. The simulation results provide the evidence for the plasticity of the molecule.

MATERIALS AND METHODS

Simulation protocols

There are six different simulations in our work. Their simulation protocols are listed in Table 1. They are divided into two categories: regular molecular dynamics (MD) simulations (simulation (a), (c), and (e)) and mixing replica exchange molecular dynamics (REMD) simulations (simulation (b), (d), and (f)). Regular MD shows the structural stability of the k-turn in a free or protein-bound state. And mixing REMD (24,25) (discussed in the next section) provides the widely distributed sampling data for the study of the state transitions.

The details of the simulations are given as follows. The first simulation (a) is a 200-ns regular MD simulation of a single k-turn from the N1 or N3 structure independently. The N1 structure is extracted from the *H. marismortui* large ribosomal subunit (26) (Protein Data Bank,

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

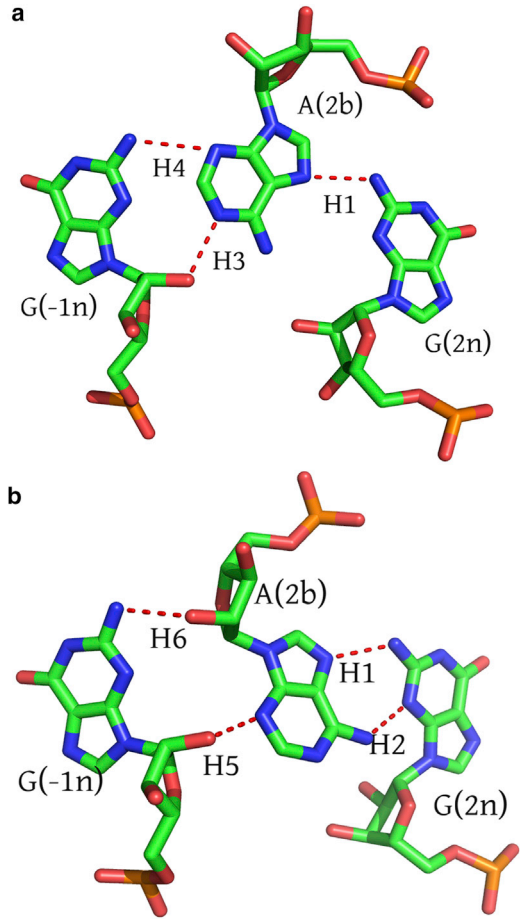

FIGURE 2 Different hydrogen bond networks of k-turn in the N1 and N3 classes. (*a*) N1 contains three hydrogen bonds, which are marked by H1, H3, and H4 in the figure, respectively. H1 is the bond between N7 atom of A(2b) (acceptor) and N2 atom of G(2n) (donor). H3 is the bond between N1 atom of A(2b) (acceptor) and O2′ atom of G(−1n) (donor). H4 is the bond between N3 atom of A(2b) (acceptor) and N2 atom of G(−1n) (donor). (*b*) N3 contains four hydrogen bonds, including H1, H2, H5, and H6. H1 is the same as that in N1. H2: N3 atom of G(2n) (acceptor) and N6 atom of A(2b) (donor). H5: N3 atom of A(2b) (acceptor) and O2′ atom of G(−1n) (donor). H6: O2′ atom of A(2b) (acceptor) and N2 atom of G(−1n) (donor). The N3 to N1 transition can also be called the A-minor 0/1 transition (64). To see this figure in color, go online.

PDB: 3cc2, nucleotide residues 77–82 and 92–100 from chain 0). The N3 structure is extracted from the k-turn/L7Ae complex (9) (PDB: 4bw0, nucleotide residues 3–8 and 15–23). The second 1000-ns mixing REMD simulation (b) is also performed for a single k-turn, which starts from the N3 structure. In the simulation, an adaptive bias potential is applied on the biased replica of the system for the fast sampling purpose (27). When the simulation is over, we perform the free energy calculation and the transition path construction from the sampling data of the unbiased REMD replicas (see below for the details of the Materials and Methods). The third and fourth simulations (c) and (d) are carried out for the k-turn/L7Ae complex (PDB 4bw0) by the regular MD and mixing REMD method, respectively. And similarly, the fifth and sixth simulations (e) and (f) are for the k-turn/L24e complex from the ribosome (26) (PDB: 3cc2). These different kinds of systems provide the proof for the plasticity of the k-turn motif in different environments.

All the systems in Table 1 are solvated in a periodic box of TIP3P waters (28). The free k-turn system has 0.2 M Mg^{2+} ions and neutralized by Cl^{-}. And the k-turn/L7Ae and k-turn/L24e complex are directly neutralized by Na^{+} ions. The simulation software is AMBER 18 (29). The RNA and protein force field are f99bsc0 plus χOL3 (30,31) and ff14SB (32), respectively. Before the formal simulation, all the systems are minimized by two steps. The first is a 10,000 steps minimization for only the solvent molecules, and the second is another 10,000 steps for all the molecules (solute plus solvent). After the minimization, the systems are heating up to 300 K within 160 ps and then equilibrating further for 40 ps at a constant pressure (1 bar). All the production simulations (a)–(f) are performed in the NVT ensemble. The simulation method is Langevin dynamics with a friction constant of 10 ps^{-1} (33). The time step is 2.0 fs. SHAKE algorithm (34) is used to fix the bond length involving the hydrogens. Particle mesh Ewald (35) is employed to treat long-range electrostatic interactions.

In this work, we analyze the nucleotide-nucleotide interactions in RNA by DSSR (36) and the nucleotide-protein interaction by SNAP (37). All the three-dimensional molecular structures in this article are produced by Pymol (38).

Mixing REMD simulation

Mixing REMD is an efficient simulation method for biomolecules (24,25). It combines the advantages of the biased and unbiased sampling strategies. In mixing REMD, three different kinds of replicas are used in the simulation, including the biased replicas, equilibrium replicas, and REMD replicas. The biased replicas do the fast sampling work in the collective variable space by an adaptive bias potential independently (27). At a fixed time step, their conformations are copied to the equilibrium replicas for relaxation. And then the states are transferred further to the REMD replicas. The exchanges of the REMD replicas in the temperature space are the same as that in a standard REMD simulation (39).

Actually, mixing REMD is not an independent sampling method. It is only a mixture of existing methods. But this mixture does make the simulation more convenient. On one hand, it only applies a bias potential or force to the biased replicas, which can be built by various methods like an adaptively biased sampling method (27,40,41), metadynamics (42,43), or an adaptive biasing force (44,45). The adaptive potential is quite useful for the exploration of the conformations along a particular collective variable or reaction coordinate. But unlike previous adaptive sampling methods, in mixing REMD, the bias potential is not involved in the free

TABLE 1 Protocols for Different Simulations in this Work

System	Box Length (Å)	Box Angle (°)	Temperature (K)	Simulation Time (ns)	Sampling Method
k-turn[(a)]	62.43, 62.43, 62.43	109.47, 109.47, 109.47	300	200	Regular MD
k-turn[(b)]			300, 302.5, 305	1000	Mixing REMD
k-turn/L7Ae[(c)]	78.67, 76.85, 68.90	90.0, 90.0, 90.0	300	200	Regular MD
k-turn/L7Ae[(d)]			300, 302.5, 305	1000	Mixing REMD
k-turn/L24e[(e)]	81.54, 81.54, 81.54	109.47, 109.47, 109.47	300	200	Regular MD
k-turn/L24e[(f)]			300, 302.5, 305	1000	Mixing REMD

The superscripts (a)–(f) in the first column are referred to as the shorthand of the corresponding simulations.

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

energy calculation. It does not need to be extremely convergent at the end of the simulation. On the other hand, mixing REMD also has some unbiased replicas (REMD replicas) for the production of the unbiased sampling data in the canonical ensembles. The sampling is independent of the predefined collective variables. So, after a mixing REMD simulation, it is convenient to perform the free energy calculation in any additional collective variable space. For example, in our previous work (25), we study the folding pathway of a tetraloop (UUCG). The simulation is accelerated on two collective variables (root mean-square deviation (RMSD) and the number of the hydrogen bonds in the native structure), whereas the free energy surface is calculated on the other two (eRMSD (46) and C1′-C1′ distance). To do the simulation, we have developed a fast sampling and analysis tool (FSATOOL) (25). Its sampling module works together with AMBER (29). For more details on the implementation and the performance of the mixing REMD method, the readers are referred to our previous articles (24,25).

As to the mixing REMD simulations of k-turn in this work (simulation (b), (d), and (f) in Table 1), there are one biased replica at 305 K, one equilibrium replica at 305 K, and three REMD replicas at 300, 302.5, and 305 K, respectively. For the biased replica, we use an adaptive bias potential with two particular collective variables (27). The first collective variable (CV1) is the sum of the two distances: HO2′ atom of G(−1n) and N1 atom of A(2b) and H22 atom of G(−1n) and N3 atom of A(2b). These two distances correspond to the two hydrogen bonds H3 and H4 in Fig. 2 *a*. So, the minimum of CV1 represents the N1 structure of k-turn (*upper-left blue circle* in Fig. 3). The second collective variable (CV2 for short) is also the sum of two distances: HO2′ atom of G(−1n) and N3 atom of A(2b) and H22 atom of G(−1n) and O2′ atom of A(2b). These two distances correspond to the two hydrogen bonds H5 and H6 in Fig. 2 *a*. The minimum of CV2 represents the N3 structure of k-turn (*lower-right red triangle* in Fig. 3). The adaptive bias potential updates continuously in the simulation. It increases the sampling ability of the biased replica in a two-dimensional collective variable space spanned by CV1 and CV2 and then helps the unbiased REMD replicas produce sufficient sampling data on the two critical N1 and N3 states.

However, there is one more thing that must be considered. The fast-growing bias potential contains multiple atom-to-atom distance functions. These distances may increase largely at the late stage of the simulation and break the two strands of k-turn. To overcome this problem, we add an additional restraint potential to the biased replica at the border of the collective variable space (*colored region* in Fig. 3). It must be noted that the equilibrium and REMD replicas are not affected by the restraint potential as well as the bias potential. For all the three mixing REMD simulations, the average exchange rate of the REMD replicas in the temperature space and the round-trip time (47) are provided in Table S1. The convergence of the sampling data (Kullback-Leibler divergence) (48) is shown in Fig. S4.

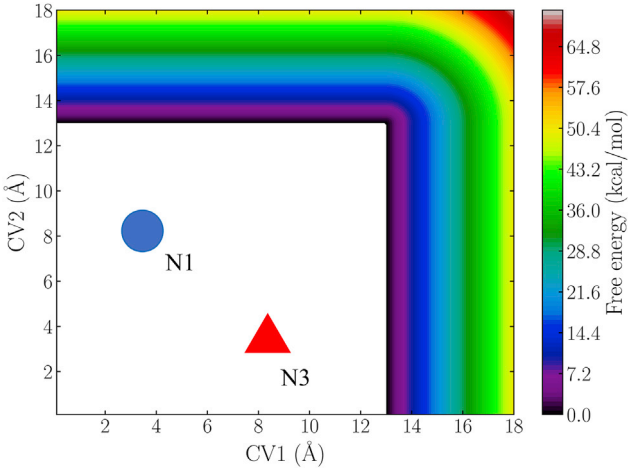

FIGURE 3 The initial distribution of the bias potential in the two-dimensional collective variable space. The two collective variables, CV1 and CV2, are defined in the main text. The upper-left blue circle and the lower-right red triangle mark the positions of the N1 and N3 states of k-turn, respectively. The colored region at the border (CV1 or CV2 larger than 13 Å) indicate the existence of the restraint potential. To see this figure in color, go online.

Markov state model

Markov state model (MSM) (49,50) is a popular trajectory analysis tool for the biomolecules. It is able to reveal the long-term dynamics from lots of short trajectories. By the model, one can easily obtain the dominant transition path and the corresponding transition rate between some important states. Generally, a standard MSM analysis contains the following steps. First, to reduce the degrees of freedom, all the sampling data in the trajectories are projected into a low-dimensional space. This can be done by the principal component analysis (PCA) (51) or time-lagged component analysis methods (52). Second, all the sampling data are clustered to a set of microstates by k-means (53), k-means++ (53,54), or k-medoids methods (55). After this, a transition probability matrix is constructed between these microstates with a suitable lag time. Third, based on the transition probability matrix, the microstates are further coarse grained to a few representative macrostates by the Robust Perron Cluster Cluster Analysis (PCCA+) (56) or Bayesian Agglomerative Clustering Engine (BACE) methods (57). Finally, according to the transition path theory (TPT) (49,58), a flux network is built between the macrostates from the start to the end state. The main transition paths can also be extracted from the network.

In this work, we perform the analysis on the sampling data at 300 K by the MSM module in FSATOOL (25). The data are first clustered to the microstates by k-means++ (53,54) in a two-dimensional space spanned by CV1 and CV2 and then coarse grained to the macrostates by Robust Perron Cluster Analysis (56). The transition network is forced to satisfy the detailed balance condition by the maximal likelihood algorithm (59,60). Based on the MSM analysis results, we confirm the fact again that the environment does affect the relative stability and the transition time between the N1 and N3 state of k-turn.

MM/GBSA calculation

MM/GBSA has been shown to be a reliable method for the calculation of the binding free energy of protein-RNA complexes (61). In this work, we perform the MM/GBSA calculation for the k-turn/L7Ae and the k-turn/L24e complexes by MMPBSA.py (62) in AMBER 18 (29). The GB model used in the calculation is GB neck (igb = 7) (63), and the salt concentration is 0.2 M. The calculation provides the binding affinity of k-turn to different proteins in either the N1 or the N3 state.

RESULTS AND DISCUSSION

In this section, we show the simulation results of k-turn in different environments. The details of the simulation protocols are given in Table 1.

Simulation of a single k-turn in explicit solvent

It has been found that k-turn behaves like a molecular hinge in many systems (6,21). To illustrate this feature, we first perform two 200-ns MD simulations at 300 K for a single k-turn in water. One starts from the N1 conformation, and the other starts from N3 (Table 1, simulation (a)). The molecule has 15 nucleotide residues. As shown in Fig. 1, its

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

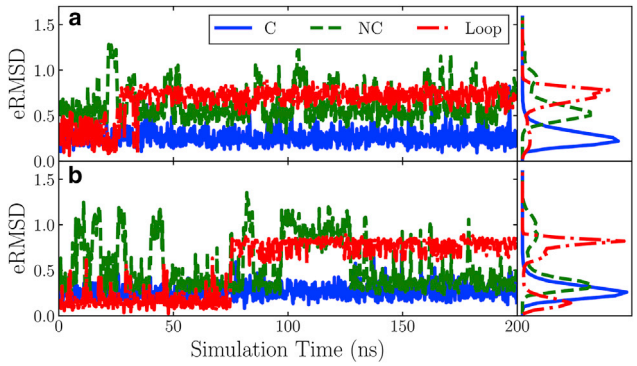

FIGURE 4 The changes of eRMSD of the three fragments of k-turn in the regular MD simulations (simulation (a) in Table 1). (*a*) Simulation starts from the N1 conformation. (*b*) Simulation starts from the N3 conformation. The canonical helix C, non-canonical helix NC, and central loop region L are plotted by blue solid lines, green dashed lines, and red dash-dotted lines, respectively. All the eRMSD calculations in this article are performed by plumed (66). The probability distributions of the eRMSD of the fragments are shown on the right side of the figure. To see this figure in color, go online.

structure can be divided into three fragments: a canonical helix C, a loop region L, and a non-canonical helix NC. The changes of their eRMSD (46) to the N1 and N3 conformation are presented in Fig. 4, *a* and *b*, respectively. As a function to measure distances between different three-dimensional RNA structures, eRMSD includes the positions of three carbon atoms on the 6-membered ring of each nucleobase in a specified order. For pyrimidines, the atoms are C2, C4, and C6, and for purines, the atoms are C2, C6, and C4 (46). The reference N1 and N3 structures for the eRMSD calculation are extracted from the k-turn/L24e and k-turn/L7Ae complexes, respectively (see the details in Materials and Methods). The figure shows that in the most simulation time, the central loop L of k-turn (*red dash-dotted lines*) has a larger structure variation to the initial structure than the other two helices C and NC at the ends (*blue solid lines and green dashed lines*), no matter if the molecule stays in the N1 or N3 state. Such a rigid-flexible-rigid structure allows a k-turn to link different rigid parts of RNA in a soft way and helps the construction of large RNA structures.

Although the N1 and N3 classes of k-turn have the similar hinge-like structures, their hydrogen bond networks are different to each other. As shown in Fig. 2 before, N1 contains three hydrogen bonds (H1, H3, and H4) and N3 has four (H1, H2, H5, and H6). The changes of the six hydrogen bonds (H1–H6) in the two regular simulations are plotted in Fig. 5. When a bond is formed at a particular time step, it adds a short vertical line in the figure. Fig. 5 *a* shows the result of the simulation at N1. It is found that the three hydrogen bonds owned by the N1 class, i.e., H1, H3, and H4, are quite stable in the whole simulation time. As to the other bonds (H2, H5, and H6), H2 has a much smaller occupancy (0.37), and H5 and H6 are not formed in the simulation. On the contrary, in the simulation of k-turn at N3 (Fig. 5 *b*), the related bonds, H1, H2, H5, and H6, are well kept. And the other two bonds, H3 and H4, do not show up from the beginning to the end. The simulation results present that both of N1 and N3 are the stable states for a free k-turn. This is crucial to the maintenance of the RNA structures in vivo. Furthermore, the detailed basepair interactions for k-turn at N1 and N3 are shown in Figs. S1 and S2, respectively.

From Fig. 5 *b*, we find that the strength of the hydrogen bonds H5 and H6 are weaker than H1 and H2 at the N3 state. This can be explained by their positions in the k-turn structure. As shown in Figs. 1 and 2, H5 and H6 connect the two remote helices, C and NC. And differently, H1 and H2 are both in the same NC helix. Moreover, we analyze the simulation trajectory at N3 by PCA (51). Those nucleotides with the largest displacements are A(L3), G(1b), A(3b), C(4b), and G(4n) (marked in Fig. 5 *c* by the *red arrows*). They are all on the L loop and the NC helix. Consequently, their large displacements shall decrease the stabilities of the related hydrogen bonds.

The above simulations show that both of the N1 and N3 states of k-turn are stable. Even if the simulation time is extended to five times longer (1000 ns), a transition between the two states cannot still be observed (Fig. S3). This is different than a previous study (64). In that study, the N3 to N1 transition (A-minor 0/1 transition) is observed in the regular MD simulations. This can be explained by the use of the different force field. In our work, the force field is AMBER ff99bsc0 plus χOL3. And in a previous study (64), it is AMBER ff99. As introduced in the (30), the ff99bsc0 force field is a refinement of the ff99 force field. It provides correct representation of the α/γ concerted rotation in nucleic acids. And the χOL3 force field (31) updates the glycosidic torsion χ based on the quantum chemical calculations.

To improve the sampling efficiency, an additional simulation is performed for the single k-turn molecule by the mixing REMD method (24,25) (simulation (b) in Table 1). The simulation is accelerated by an adaptive bias potential. Because the molecule is free and flexible, such a growing bias potential (force) makes it move even faster. To protect the global structure, we use a restraint potential at the border of the collective variable space (Fig. 3). Besides this, we also add a very weak positional restraint potential to the heavy atoms in the sugar-phosphate backbone of k-turn (force constant 0.01 kcal/mol/Å^2).

With the simulation data, we calculate the free energy surface at 300 K and show it in Fig. 6 *a*. The collective variables are CV1 and CV2. Each of them contains two hydrogen bonds that only exist in the N1 or N3 states, respectively. The free energy surface presents three stable states of the k-turn: N1, N3, and an intermediate state I. The structure of state I is shown in Fig. 6 *b*. After analyzing the data in the two-dimensional collective variable space by

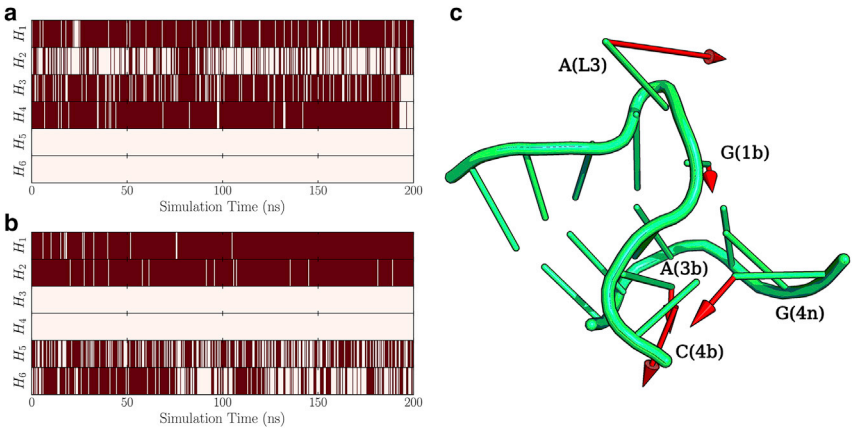

FIGURE 5 The changes of the six hydrogen bonds in the simulations of a single k-turn motif. (*a*) Data of the simulation at N1. (*b*) Data of the simulation at N3. The labels H1 to H6 on the vertical axes represent the marked hydrogen bonds in Fig. 2. (*c*) Illustration of the nucleotides with the largest displacements in the simulation of a k-turn at N3. The calculation is performed by PCA (51). The directions and the magnitudes of the displacements are marked by the red arrows. To see this figure in color, go online.

MSM (49,50) and TPT (49,58), it is confirmed that k-turn has to pass the state I on the dominant transition path connecting the N1 and N3 states. On the transition path from the N3 state, the nucleotide A(2b) rotates anticlockwise slightly from the initial state, which breaks the original hydrogen bond H5 and H6 and forms the H4 bond. The nucleotide G(2n)'s sugar pucker switches from the C2′-endo conformation to the C3′-endo conformation simultaneously. And at the same time, one more nucleotide-nucleotide interaction between G(2n) and G(1b), tSH, appears in the state. Among all the snapshots in state I, the tSH interaction has an occupancy of 0.9. Such a stable interaction makes state I appear to be a free energy minimum. When the k-turn passes the intermediate state, its nucleotide A(2b) continues to rotate anticlockwise. It breaks the hydrogen bond H2 and forms H3 instead. At the end of the transition, the k-turn reaches the final N1 state.

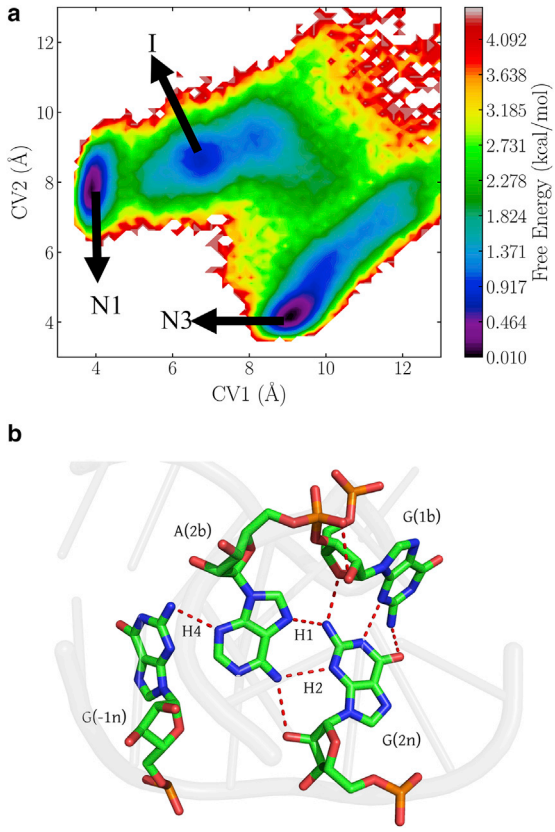

FIGURE 6 (*a*) Free energy surface of a single k-turn in a two-dimensional collective variable space spanned by CV1 and CV2. Each of them includes two particular hydrogen bonds in the N1 or N3 structures of k-turn. The data are from a mixing REMD simulation (simulation (b) in Table 1). (*b*) The hydrogen bonds of the intermediate state I on the free energy surface (marked by the *red dashed lines*). To see this figure in color, go online.

Simulation of a k-turn binding with the L7Ae protein

When binding to the protein L7Ae (9), the k-turn appears to be folded into the N3 state. There are a lot of hydrogen bonds at the binding interface of the k-turn/L7Ae complex. The most popular ones are listed in Table 2 and drawn in Fig. 7 *a*. It shows that the nucleotide residues A(L3), G(2n), and G(1b) on the loop region L and the non-

TABLE 2 Hydrogen Bonds at the Interface of the k-Turn/L7Ae Complex

Nucleotide Residue	Amino Acid Residue[a]	H-Bond Number	Type[b]	Average Distance (Å)
A(L3)	THR-58	2	P/A	2.68
	ALA-119	1	P/A	2.77
G(2n)	LYS-55	2	B/A	3.18
	ASN-59	1	B/A	2.77
	GLU-60	1	S/A	2.58
G(1b)	LYS-56	1	B/A	2.87
	ASN-59	1	B/A	2.91
	GLU-60	3	B/A	2.86

[a]The number after each residue name is the index of the residue in the protein.

[b]The residue-residue interaction types are classified by SNAP (37). B/A, base/amino acid; P/A, phosphate/amino acid; S/A, sugar/amino acid.

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

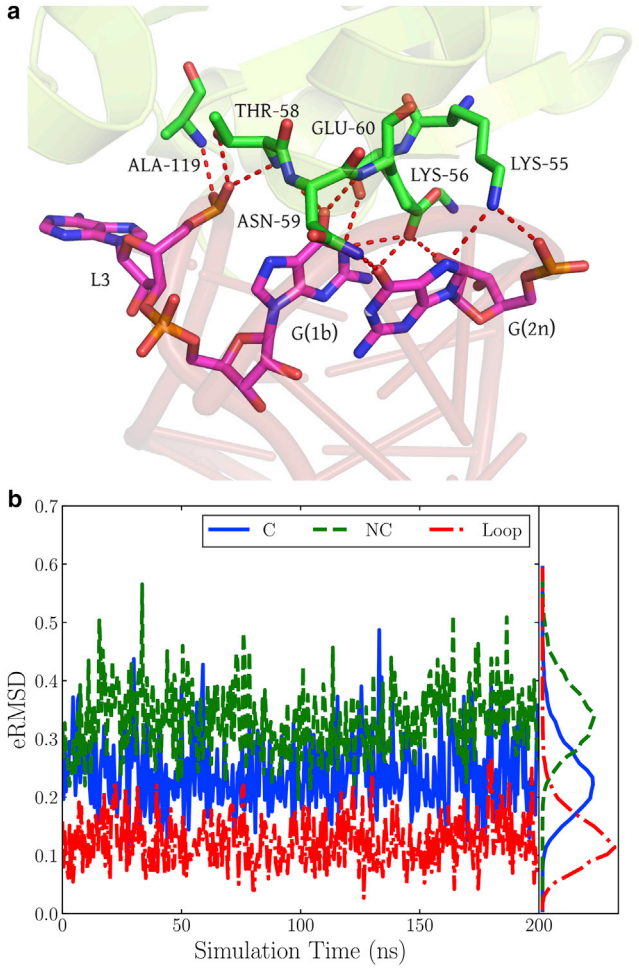

FIGURE 7 (*a*) Presentation of the mostly important hydrogen bonds at the binding interface of the k-turn/L7Ae complex. In the figure, the k-turn and the L7Ae protein are drawn as a light red and green cartoon model, respectively. Those key residues in the formation of the hydrogen bonds are highlighted by the stick models. All the hydrogen bonds are plotted by the red dashed lines. (*b*) The changes of eRMSD of the three fragments of k-turn in a regular MD simulation for the complex (simulation (c) in Table 1). The data of the canonical helix C, non-canonical helix NC, and loop region L are plotted by the blue solid lines, green dashed lines, and red dash-dotted lines, respectively. The probability distributions of the eRMSD of the fragments are shown on the right side of the figure. To see this figure in color, go online.

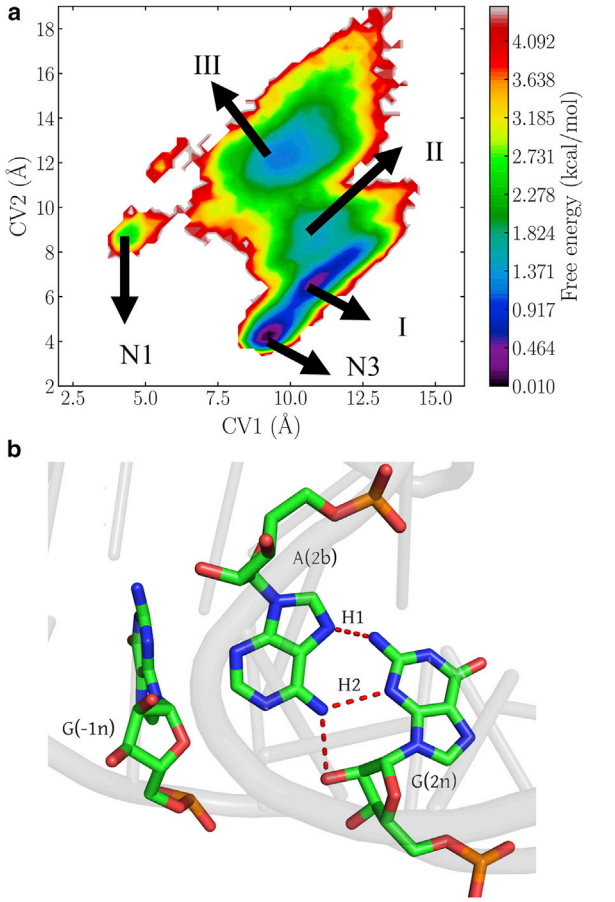

FIGURE 8 (*a*) Free energy surface of k-turn in a two-dimensional collective variable space spanned by CV1 and CV2. The data are from a mixing REMD simulation of the k-turn/L7Ae complex (simulation (d) in Table 1). (*b*) The hydrogen bonds of the intermediate state II on the free energy surface (marked by the *red dashed lines*). To see this figure in color, go online.

canonical helix NC are the key residues to the binding of the complex. And moreover, in a 200-ns regular MD simulation for the k-turn/L7Ae complex (simulate (c) in Table 1), the eRMSD of the three fragments of k-turn (N helix, NC helix, and L loop) fluctuate within a range of 0.5 (Fig. 7 *b*). This is much smaller than a free k-turn (Fig. 4). It reveals the fact that the protein L7Ae helps keep the k-turn's initial structure by forming the hydrogen bonds at the interface.

Because of the stabilization effect of the binding protein, it is difficult for a k-turn to transfer from the N3 state to the N1 state. We have to explore the transitions by the mixing REMD method with an adaptive bias potential (simulate (d) in Table 1). The calculated free energy surface in the collective variable space is given in Fig. 8 *a*. Besides the N1 and N3 state, there are three more free energy minima on the surface, the intermediate states I, II, and III. Clearly, the N3 state in the native structure is the most stable one among all the states. Compared with the free energy surface of a single k-turn (Fig. 6 *a*), N1 is no longer a meaningful free energy minimum for the complex (Fig. 8 *a*). The whole free energy surface presents a funnel-like landscape. Moreover, according to the MSM (49,50) and TPT (49,58) analysis results, the dominant transition path from N3 to N1 is N3-I-II-N1, which accounts for 70% of the total flux. The intermediate state III is an exception. It stays out of the main path because of the absence of the critical hydrogen bonds. As to the intermediate state II on the path, its structure is shown in Fig. 8 *b*. This state keeps the same hydrogen bonds H1 and H2 as in the initial N3 state. But the other two native hydrogen bonds H5 and H6 are broken. This is similar to the intermediate state I of the single k-turn (Fig. 6 *b*), except the tSH between G(2n) and G(1b).

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

TABLE 3 Hydrogen Bonds at the Interface of the k-Turn/L24e Complex

Nucleotide Residue	Amino Acid Residue[a]	H-Bond Number	Type[b]	Average Distance (Å)
A(1n)	ARG-56	1	P/A	3.19
	ASN-58	2	P/A	3.10
G(−1n)	ARG-56	2	P/A	3.30
	ALA-59	1	P/A	2.89
G(1b)	ASP-120	2	B/A	2.84

[a]The number after each residue name is the index of the residue in the protein.
[b]The residue-residue interaction types are classified by SNAP (37). B/A, base/amino acid; P/A, phosphate/amino acid.

Because of the missing interaction, the state II is not as stable as before.

Simulation of a k-turn binding with the L24e protein

In the ribosome, the k-turn that is binding to the protein L24e has a N1 structure. The primary hydrogen bonds at the binding interface are listed in Table 3 and drawn in Fig. 9 *a*. The most involved nucleotide residues are G(−1n), A(1n), and G(1b). Surprisingly, the residues on the loop region do not participate in the formation of the hydrogen bonds deeply. In a regular MD simulation of the k-turn/L24e complex (simulation (e) in Table 1), the eRMSD of the C helix and the loop of the k-turn fluctuate within 0.4 in the most simulation time. This indicates that the initial structure of k-turn can also be kept in the ribosome, even if it stays in a different N1 state.

Then, we perform a mixing REMD simulation for the k-turn/L24e complex (simulation (f) in Table 1). The free energy surface is shown in Fig. 10 *a*, which is also a funnel-like landscape. It consists of four free energy minima, including two intermediate states I and II. Unlike in the previous case, the global minimum of the surface switches to the N1 state for this complex. The MSM (49,50) and TPT (49,58) analysis show that both of the two intermediate states are on the primary transition path from N1 to N3. The structure of the intermediate state II is given in Fig. 10 *b*. Its nucleotide residue A(2b) rotates ∼60° from its initial structure. The native hydrogen bonds H3 and H4 have been broken in this state, and the new bond H2 is formed. So, this intermediate state has the same hydrogen bonds (H1 and H2) as that in the previous case (intermediate II in Fig. 8 *b*). The k-turn follows a similar transition path between N1 and N3 in different complexes.

CONCLUSIONS

The k-turn motif has two available conformations in biomolecules, i.e., the N1 and N3 classes. These two classes are mainly distinguished by different hydrogen bonds between the nucleotide residues G(2b) and A(−1n). In this work, we show that the free energies of the two states change with different environments. When in the solvent as a free molecule, their free energies are comparable to each other. The k-turn acts like a molecular hinge as in the folding of a riboswitch (11). But if binding to the protein L7Ae, the k-turn only has one global free energy minimal N3. And with the protein L24e in ribosome, N1 becomes the most stable one for k-turn. These calculations confirm the experimental structures of the complexes.

This free energy change can be explained simply by the hydrogen bond interactions in the complex. For example, for the first k-turn/L7Ae complex in its native N3 state,

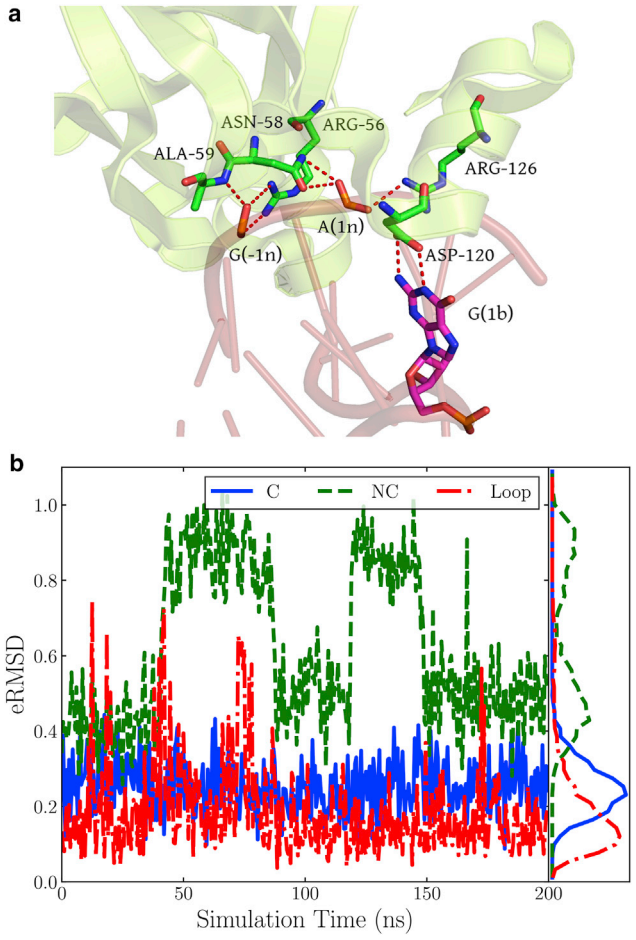

FIGURE 9 (*a*) Presentation of the mostly important hydrogen bonds at the binding interface of the k-turn/L24e complex. In the figure, the k-turn and the L24e protein are drawn as a light red and green cartoon model, respectively. Those key residues in the formation of the hydrogen bonds are highlighted by the stick models. All the hydrogen bonds are plotted by the red dashed lines. (*b*) The changes of eRMSD of the three fragments of k-turn in a regular MD simulation of the complex (simulation (e) in Table 1). The data of the canonical helix C, non-canonical helix NC, and loop region L are plotted by the blue solid lines, green dashed lines, and red dash-dotted lines, respectively. The probability distributions of the eRMSD of the fragments are shown on the right side of the figure. To see this figure in color, go online.

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

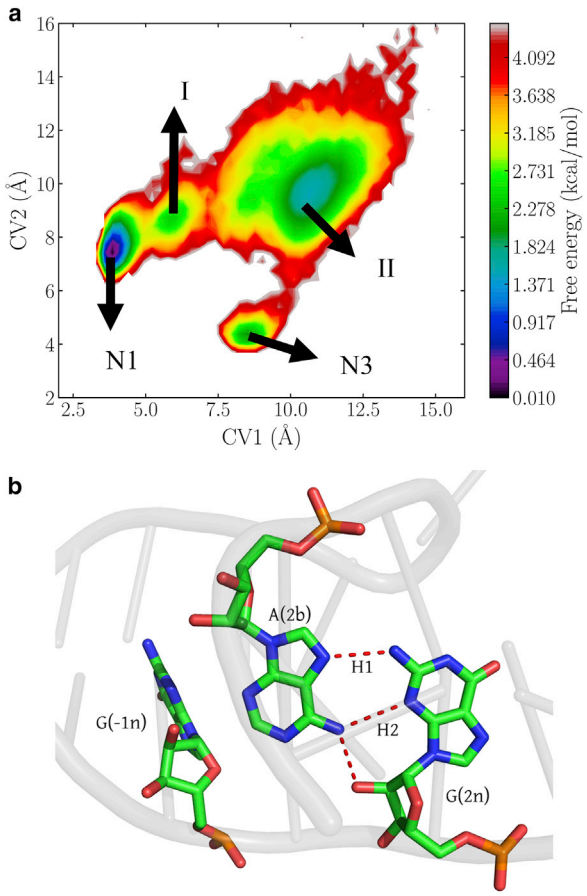

FIGURE 10 (*a*) The free energy surface of k-turn in a two-dimensional collective variable space spanned by CV1 and CV2. The data are from a mixing REMD simulation of the k-turn/L24e complex (simulation (f) in Table 1). (*b*) The hydrogen bonds of the intermediate state II on the free energy surface (marked by the *red dashed lines*). To see this figure in color, go online.

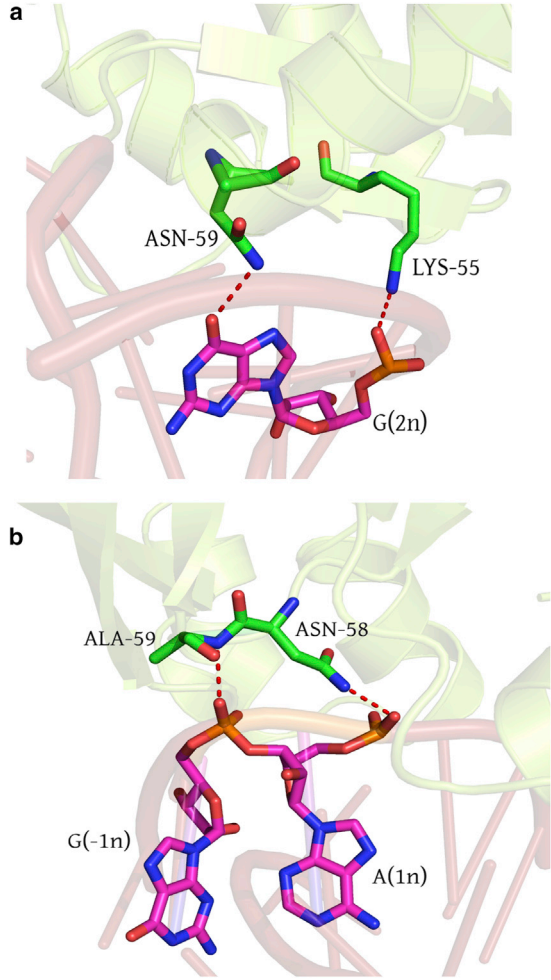

FIGURE 11 (*a*) The hydrogen bonds with the nucleotide residue G(2n) at the binding interface of the k-turn/L7Ae complex. The k-turn is in the non-preferred N1 state (given in Fig. 8 *a*). (*b*) The hydrogen bonds with the nucleotide residues G(−1n) and A(1n) at the binding interface of the k-turn/L24e complex. The k-turn is in the nonpreferred N3 state (given in Fig. 10 *a*). In the figure, the k-turn and the proteins are drawn as the light red and green cartoon models, respectively. All the hydrogen bonds are plotted by the red dashed lines, and the related residues are highlighted by the stick models. To see this figure in color, go online.

the nucleotide residue G(2n) has four hydrogen bonds at the binding interface (Fig. 7 *a*; Table 2), and in its nonpreferred N1 state, it only has two (Fig. 11 *a*). The second k-turn/L24e complex is in a similar situation. The nucleotide residues G(−1n) and A(1n) form six hydrogen bonds in the native N1 state. But in the nonpreferred N3 state, they are only involved in two hydrogen bonds (Fig. 11 *b*). Fewer hydrogen bond interactions shall decrease the binding affinity of the complexes in their nonpreferred states. To confirm this conclusion, we do an additional MM/GBSA calculation (61) for the two complexes in either the N1 or the N3 state (100 ns, 1000 snapshots). The calculation is performed by MMPBSA.py (62) in AMBER 18 (29). The results in Table 4 show that both of the two complexes have a lower binding free energy (higher binding affinity) in their native states.

Furthermore, from the calculations, we find that the binding proteins can not only change the free energy differences between the N1 and N3 state but also the free energy barriers on the transition path. Compared with the barrier on the free energy landscape of a single k-turn (Fig. 6 *a*), both of the barriers of the complexes increase from 2.5 to 4.0 kcal/mol approximately (Figs. 8 *a* and 10 *a*). The barriers greatly slow down the state-to-state transitions. According to the MSM (49,50) and TPT (49,58) anaysis, the N1→N3 transition time increase 37 times for the k-turn/L24e complex, and the N3→N1 transition time increases 57 times for the k-turn/L7Ae complex. Besides the barriers, the free energies of the intermediate states on the transition path increase too. In these states, the k-turn loses the structural flexibility as well as some internal interactions. For example, the rotation of the nucleotide residue A(2b) is limited in the intermediate state, and the internal tSH interaction between G(2n) and G(1b) has been lost.

Please cite this article in press as: Zhang et al., Simulation Study of the Plasticity of k-Turn Motif in Different Environments, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2020.08.015

TABLE 4 The Binding Free Energies of the k-Turn/L7Ae Complex and the k-Turn/L24e Complex

Complex	k-Turn/L7Ae		k-Turn/L24e	
State of k-Turn	N1 (Nonpreferred)	N3 (Native)	N1 (Native)	N3 (Nonpreferred)
ΔG_{gas}	57.66 (138.43)	−310.08 (126.95)	1363.95 (98.27)	1172.90 (74.45)
ΔG_{solv}	−110.53 (134.37)	230.58 (124.22)	−1412.25 (96.88)	−1218.08 (73.68)
ΔG_{total}	−52.86 (8.24)	−79.49 (10.26)	−48.30 (6.78)	−45.18 (5.08)

The k-turn stays in either the N1 or the N3 state. All the calculations are performed by MMPBSA.py (62) in AMBER 18 (29). The units of the data in the table are kcal/mol. The data in the parentheses are the SD.

Our simulations give a reasonable explanation of the plasticity of k-turn in association with different proteins. This may help further understand the role of k-turn in many RNA-related processes (13).

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.08.015.

AUTHOR CONTRIBUTIONS

Haomiao Zhang designed the research, performed the research, analyzed data, and wrote the article. Haozhe Zhang analyzed data. C.C. designed the research and wrote the article.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under grant 31770773.

REFERENCES

1. Crick, F. 1970. Central dogma of molecular biology. *Nature.* 227:561–563.
2. Sharp, S. J., J. Schaack, …, D. Söll. 1985. Structure and transcription of eukaryotic tRNA genes. *CRC Crit. Rev. Biochem.* 19:107–144.
3. 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.
4. Sarver, M., C. L. Zirbel, …, N. B. Leontis. 2008. FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. *J. Math. Biol.* 56:215–252.
5. Petrov, A. I., C. L. Zirbel, and N. B. Leontis. 2013. Automated classification of RNA 3D motifs and the RNA 3D motif atlas. *RNA.* 19:1327–1340.
6. Rázga, F., J. Koca, …, N. B. Leontis. 2005. Hinge-like motions in RNA kink-turns: the role of the second a-minor motif and nominally unpaired bases. *Biophys. J.* 88:3466–3485.
7. Vidovic, I., S. Nottrott, …, R. Ficner. 2000. Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol. Cell.* 6:1331–1342.
8. Klein, D. J., T. M. Schmeing, …, T. A. Steitz. 2001. The kink-turn: a new RNA secondary structure motif. *EMBO J.* 20:4214–4221.
9. Huang, L., and D. M. Lilley. 2013. The molecular recognition of kink-turn structure by the L7Ae class of proteins. *RNA.* 19:1703–1710.
10. Moore, T., Y. Zhang, …, H. Li. 2004. Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure.* 12:807–818.
11. Hamma, T., and A. R. Ferré-D'Amaré. 2004. Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution. *Structure.* 12:893–903.
12. Heppell, B., and D. A. Lafontaine. 2008. Folding of the SAM aptamer is determined by the formation of a K-turn-dependent pseudoknot. *Biochemistry.* 47:1490–1499.
13. Lilley, D. M. J. 2012. The structure and folding of kink turns in RNA. *Wiley Interdiscip. Rev. RNA.* 3:797–805.
14. Liu, J., and D. M. J. Lilley. 2007. The role of specific 2′-hydroxyl groups in the stabilization of the folded conformation of kink-turn RNA. *RNA.* 13:200–210.
15. Leontis, N. B., and E. Westhof. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA.* 7:499–512.
16. Huang, L., and D. M. J. Lilley. 2018. The kink-turn in the structural biology of RNA. *Q. Rev. Biophys.* 51:e5.
17. Goody, T. A., S. E. Melcher, …, D. M. Lilley. 2004. The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA.* 10:254–264.
18. Turner, B., and D. M. Lilley. 2008. The importance of G.A hydrogen bonding in the metal ion- and protein-induced folding of a kink turn RNA. *J. Mol. Biol.* 381:431–442.
19. Ye, W., J. Yang, …, H. F. Chen. 2013. Kink turn sRNA folding upon L7Ae binding using molecular dynamics simulations. *Phys. Chem. Chem. Phys.* 15:18510–18522.
20. Schroeder, K. T., P. Daldrop, and D. M. Lilley. 2011. RNA tertiary interactions in a riboswitch stabilize the structure of a kink turn. *Structure.* 19:1233–1240.
21. Rázga, F., N. Spackova, …, J. Sponer. 2004. Ribosomal RNA kink-turn motif–a flexible molecular hinge. *J. Biomol. Struct. Dyn.* 22:183–194.
22. Curuksu, J., J. Sponer, and M. Zacharias. 2009. Elbow flexibility of the kt38 RNA kink-turn motif investigated by free-energy molecular dynamics simulations. *Biophys. J.* 97:2004–2013.
23. Daldrop, P., and D. M. Lilley. 2013. The plasticity of a structural motif in RNA: structural polymorphism of a kink turn as a function of its environment. *RNA.* 19:357–364.
24. Zhang, H., Q. Gong, …, C. Chen. 2019. Combining the biased and unbiased sampling strategy into one convenient free energy calculation method. *J. Comput. Chem.* 40:1806–1815.
25. Zhang, H., Q. Gong, …, C. Chen. 2020. FSATOOL: a useful tool to do the conformational sampling and trajectory analysis work for biomolecules. *J. Comput. Chem.* 41:156–164.
26. Blaha, G., G. Gürel, …, T. A. Steitz. 2008. Mutations outside the anisomycin-binding site can make ribosomes drug-resistant. *J. Mol. Biol.* 379:505–519.
27. Babin, V., C. Roland, and C. Sagui. 2008. Adaptively biased molecular dynamics for free energy calculations. *J. Chem. Phys.* 128:134101.
28. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
29. Case, D. A., I. Y. Ben-Shalom, …, P. A. Kollman. 2018. AMBER 2018. University of California, San Francisco, CA.

30. Pérez, A., I. Marchán, …, M. Orozco. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.* 92:3817–3829.
31. Zgarbová, M., M. Otyepka, …, P. Jurečka. 2011. Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* 7:2886–2902.
32. Maier, J. A., C. Martinez, …, C. Simmerling. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11:3696–3713.
33. Pastor, R. W., B. R. Brooks, and A. Szabo. 1988. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* 65:1409–1419.
34. Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.
35. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N•log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
36. Lu, X.-J., H. J. Bussemaker, and W. K. Olson. 2015. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.* 43:e142.
37. Lu, X. J., and W. K. Olson. 2003. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31:5108–5121.
38. DeLano, W. L. 2002. Pymol: an open-source molecular graphics tool. *CCP4 Newslett. Protein Crystallogr.* 40:82–92.
39. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.
40. Moradi, M., V. Babin, …, C. Sagui. 2009. Conformations and free energy landscapes of polyproline peptides. *Proc. Natl. Acad. Sci. USA.* 106:20746–20751.
41. Babin, V., and C. Sagui. 2010. Conformational free energies of methyl-alpha-L-iduronic and methyl-beta-D-glucuronic acids in water. *J. Chem. Phys.* 132:104108.
42. Barducci, A., G. Bussi, and M. Parrinello. 2008. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* 100:020603.
43. Barducci, A., M. Bonomi, and M. Parrinello. 2011. Metadynamics. *WIREs Comput. Mol. Sci.* 1:826–843.
44. Comer, J., J. C. Gumbart, …, C. Chipot. 2015. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B.* 119:1129–1151.
45. Darve, E., and A. Pohorille. 2001. Calculating free energies using average force. *J. Chem. Phys.* 115:9169–9183.
46. Bottaro, S., F. Di Palma, and G. Bussi. 2014. The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res.* 42:13306–13314.
47. Katzgraber, H. G., S. Trebst, …, M. Troyer. 2006. Feedback-optimized parallel tempering Monte Carlo. *J. Stat. Mech.* 2006:P03018.
48. Kullback, S., and R. A. Leibler. 1951. On information and sufficiency. *Ann. Math. Stat.* 22:79–86.
49. Noé, F., C. Schütte, …, T. R. Weikl. 2009. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. USA.* 106:19011–19016.
50. Chodera, J. D., and F. Noé. 2014. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* 25:135–144.
51. Wold, S., K. Esbensen, and P. Geladi. 1987. Principal component analysis. *Chemometr. Intell. Lab.* 2:37–52.
52. Pérez-Hernández, G., F. Paul, …, F. Noé. 2013. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* 139:015102.
53. Hartigan, J. A., and M. A. Wong. 1979. Algorithm AS 136: a k-means clustering algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* 28:100–108.
54. Arthur, D., and S. Vassilvitskii. 2007. k-means++: the advantages of careful seeding. *In* Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, pp. 1027–1035.
55. Park, H. S., and C. H. Jun. 2009. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* 36:3336–3341.
56. Röblitz, S., and M. Weber. 2013. Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. *Adv. Data Anal. Classif.* 7:147–179.
57. Bowman, G. R. 2012. Improved coarse-graining of Markov state models via explicit consideration of statistical uncertainty. *J. Chem. Phys.* 137:134111.
58. Weinan, E., and E. Vanden-Eijnden. 2006. Towards a theory of transition paths. *J. Stat. Phys.* 123:503–523.
59. Bowman, G. R., K. A. Beauchamp, …, V. S. Pande. 2009. Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.* 131:124101.
60. Prinz, J.-H., H. Wu, …, F. Noé. 2011. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* 134:174105.
61. Chen, F., H. Sun, …, T. Hou. 2018. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. *RNA.* 24:1183–1194.
62. Miller, B. R., III, T. D. McGee, Jr., …, A. E. Roitberg. 2012. MMPBSA.py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 8:3314–3321.
63. Mongan, J., C. Simmerling, …, A. Onufriev. 2007. Generalized born model with a simple, robust molecular volume correction. *J. Chem. Theory Comput.* 3:156–169.
64. Réblová, K., J. E. Šponer, …, J. Šponer. 2011. A-minor tertiary interactions in RNA kink-turns. Molecular dynamics and quantum chemical analysis. *J. Phys. Chem. B.* 115:13897–13910.
65. Yang, H., F. Jossinet, …, E. Westhof. 2003. Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.* 31:3450–3460.
66. Tribello, G. A., M. Bonomi, …, G. Bussi. 2014. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* 185:604–613.