BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Combined Mechanism of Conformational Selection and Induced Fit in U1A−RNA Molecular Recognition

Ikuo Kurisaki,[†,‡] Masayoshi Takayanagi,[†,‡,§] and Masataka Nagaoka*[,†,‡]

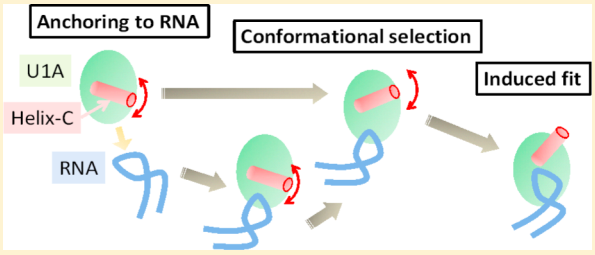
[†]Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

[‡]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Honmachi, Kawaguchi 332-0012, Japan

[§]Venture Business Laboratory, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

Supporting Information

ABSTRACT: In this study, we demonstrate that U1A−RNA molecular recognition is mediated by a combined mechanism of conformational selection and induced fit. The binding of U1A to RNA has been discussed in the context of induced fit that involves the reorientation of the α-helix in the C-terminal region (Helix-C) of U1A to permit RNA access only when U1A correctly recognizes RNA. However, according to our molecular dynamics simulations, even in the absence of RNA, Helix-C spontaneously reoriented to permit RNA access. Nonetheless, such a conformational change was still incomplete. Helix-C was often partially or even fully unfolded and in an infrequent RNA-accessible conformation, which can be detected using state-of-the-art nuclear magnetic resonance methodology. These results suggest that the formation of an energetically stabilized complex is promoted by specific interactions between U1A and RNA. In conclusion, in the recognition of RNA by U1A protein, we propose a combined mechanism that requires the reorientation of Helix-C and the subsequent contact with RNA through conformational selection, although the stabilization of the U1A−RNA complex is caused by induced fit. We further propose a modification to the conventional assumption regarding the mechanism of U1A−RNA molecular recognition.



The formation of protein−RNA complexes is a fundamental process that mediates RNA metabolism.[1] The failure of the formation of these complexes has highly detrimental effects on cells.[2] Therefore, defining the mechanism of the formation of protein−RNA complexes (protein−RNA molecular recognition) is a central issue in basic biological and medical research. Analysis of the binding between U1A and RNA can serve to address this issue. U1A possesses two characteristics typical of an RNA-binding protein: (1) sequence-specific RNA recognition and (2) a large conformational change that occurs upon RNA binding (induced fit). These characteristics are critical for understanding protein−RNA molecular recognition and its functional outcome. Thus, the U1A−RNA system has been extensively studied[3−13] to establish the molecular basis of the sequence-specific recognition of RNA by U1A.[6−8,12]

Although the induced fit model that may explain the binding of U1A to RNA has been discussed for more than 15 years, it remains poorly understood.[5,12−15] The induced fit of U1A was characterized by determining the reorientation of its C-terminal region upon binding to RNA[5,12] [Figure 1 (C-terminal region, red)]. Thus, apo-U1A folds into the closed form in which the α-helix in the C-terminal region (Helix-C, consisting of residues 91−98) covers part of the RNA-binding surface, i.e., Phe56 (Figure 1A). In contrast, when U1A binds to RNA, it folds into the open form, and Helix-C moves away, uncovering the buried Phe56 to permit RNA access [Figure 1B (RNA, transparent blue)].

Tang and Nilsson performed simulations to evaluate the formation of the U1A−RNA complex at the atomic level.[14] According to the results of their molecular dynamics (MD) simulation study and experimental studies, they proposed a three-step reaction leading to the formation of the U1A−RNA complex (i.e., the TN conjecture). First, U1A and RNA are attracted to each other through electrostatic interactions. Second, U1A is anchored to RNA through the interactions of Glu19 and Arg52 with RNA bases. Third, Helix-C reorients toward the open form upon binding to the changed conformation of RNA "in response to the correct recognition of the RNA".[14] In the third step, U1A makes loose contact with RNA to maintain the closed form and gradually interacts specifically with the RNA upon the rearrangement of Helix-C to permit RNA access, leading to the formation of the final U1A−RNA complex. Later, Williamson proposed the U1A−RNA cofolding scenario (mutually induced fit,[15] i.e., the W conjecture) as follows. Helix-C and RNA are initially unfolded partially, and Helix-C refolds into an α-helix that is coupled with the folding of the RNA during binding. Although these

Received: December 23, 2013
Revised: April 20, 2014
Published: May 14, 2014

   dx.doi.org/10.1021/bi401708q | *Biochemistry* 2014, 53, 3646−3657

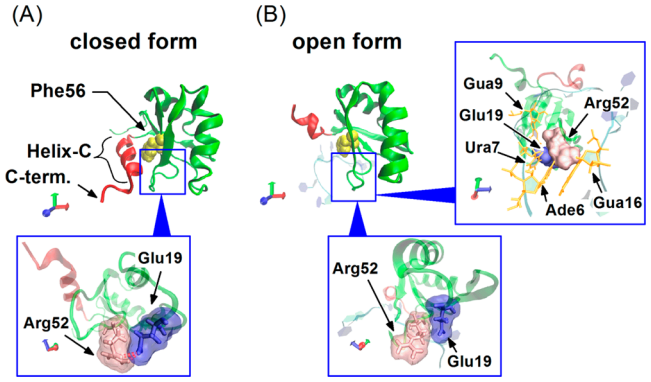


Figure 1. Conformational change of U1A upon binding to RNA. (A) Apo-U1A in the closed form (PDB entry 1FTH). (B) U1A in the open form (PDB entry 1URN) bound to RNA. In panels A and B, Glu19 and Arg52 are shown as blue and pink sticks and/or surfaces, respectively, in each inset. In panel A, hydrogen bonds between Glu19 and Arg52 are represented by red dotted lines in the inset. In panel B, the bound RNA is shown in transparent blue and blue, and RNA residues interacting with Glu19 and Art52 (i.e., Ade6, Ura7, Glu9, and Gua16) are shown as orange sticks in the right inset. The C-terminal region (residues 89−102) and others (residues 1−88) are colored red and green, respectively. Phe56 in the RNA-binding surface is shown as yellow van der Waals spheres.

hypotheses are considered landmarks for subsequent studies of U1A−RNA molecular recognition,[6,8,10,13,16] they remain unaltered. This is because neither experimental nor computer simulation studies have directly shown Helix-C reorientation at the atomic level.

In a subsequent X-ray crystallographic study by Rupert et al., an alternative open form of U1A was resolved in the absence of RNA binding[17] (Figure 2A). The structure was determined

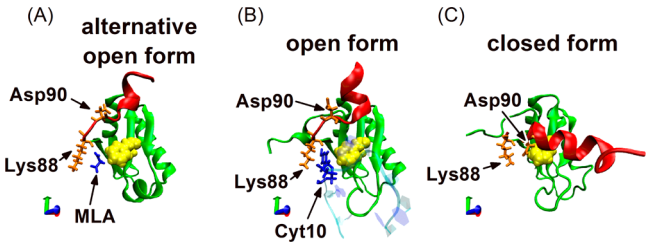


Figure 2. U1A structures determined using X-ray crystallography and NMR. (A) Apo-U1A in the alternative open form (PDB entry 1NU4) in the presence of a high concentration of MAL. (B) U1A in the open form (PDB entry 1URN) bound to RNA. (C) Apo-U1A in the closed form (PDB entry 1FTH). Lys88 and Asp90 are shown as orange sticks. In panel B, Cyt10 is shown as blue sticks. In panel C, MAL stacking on Lys88 is shown as blue sticks. U1A, the C-terminal region, Phe56, and RNA are colored as in Figure 1.

from crystals formed in the presence of a high concentration of malonate (MAL), an organic acid possessing two negative charges contributed by its two carboxyl groups. In the alternative open form, an MAL molecule is present at the position occupied by the RNA residue, the cytosine (Cyt10), in the U1A−RNA complex[12] (Figure 2A,B). This is the only MAL molecule present around the RNA-binding surface of U1A, suggesting the importance of Cyt10 for the stabilization of the open form. According to these observations, Rupert et al. suggested that the MAL mimics RNA to stabilize the alternative open form. Moreover, Cyt10 is stacked on Lys88. Considering that in the closed form, Lys88 is located in the vicinity of Asp90, which connects Helix-C to the remaining part of U1A (Figure 2C), we assume that stacking of Cyt10 with Lys88 stabilizes the open form by preventing Asp90 from interacting with Lys88.

On the other hand, according to the conformational flexibility of the C-terminal region of apo-U1A, Rupert et al. also suggested that Helix-C could reorient spontaneously in the absence of RNA binding and claimed that U1A assumes the open form not because of induced fit but because of conformational selection.[17] Moreover, under thermal equilibrium, a protein in an aqueous solution can assume multiple conformations, ranging from the unfolded to native forms. Furthermore, a protein assumes the active conformation not because of specific protein−protein or protein−ligand interactions but because of the influence of environmental factors such as salt concentration, pressure, and pH.[18−21] Therefore, it is possible that apo-U1A assumes the active conformation, the open form, in aqueous solution independent of specific interactions with RNA, i.e., through conformational selection.

Keeping these conditions in mind, we made such an assumption that the reorientation of the C-terminal region does not necessarily occur "in response to the correct recognition of the RNA"[14] but in such an aqueous solution with MAL, or even in an aqueous solution with or without MAL. To verify this assumption, we performed several MD simulations of the U1A conformation in the presence and absence of MAL. We found that such a reorientation does not necessarily occur in the presence of MAL, suggesting that U1A converts spontaneously from the closed form to the open form, i.e., through conformational selection. However, the observed conformations differed from the open form with respect to conformational stability and/or the conformation of Helix-C. This indicates that neither induced fit nor conformational selection explains the mechanism of U1A's conformational change. Therefore, we here address the modification of the TN conjecture to establish a complete understanding of U1A−RNA molecular recognition.

MATERIALS AND METHODS

Structures and Force Field Parameters. The atomic coordinates of U1A were obtained from the Protein Data Bank.[22] The apo-U1A structure in the closed form was generated from the averaged conformation of 43 NMR structures (PDB entry 1FHT[5]). The missing Met1 was compensated, and the residues from Glu103 to Glu117 were truncated. Molecular modeling was performed using the LEaP module of AmberTools.[23]

To calculate the forces acting among atoms, Amber force field 99SB, the TIP3P water model, and JC ion parameters adjusted for the TIP3P water model were applied for amino acid residues, water molecules, and ions,[24−27] respectively. The force field parameters of MAL were determined as follows. The atomic coordinates of MAL were obtained from the U1A−MAL complex (PDB entry 1NU4[17]). The missing hydrogen atoms were compensated using UCSF Chimera.[28] Both carboxyl groups of MAL were deprotonated under the physiological condition (i.e., pH 7.4). The molecular structure was energetically minimized, and the restrained electrostatic potential charge was calculated at the Hartree−Fock/6-31G* level using Gaussian03.[29] The atomic charges of MAL are listed in Table 1. The other force field parameters (bond, angle, dihedral, and van der Waals parameters) were assigned

dx.doi.org/10.1021/bi401708q | Biochemistry 2014, 53, 3646−3657

according to the general Amber force field (GAFF) 1.4 parameters.[30,31]

Table 1. Atomic Charges of Malonate

atom name[a]	atom type[a]	RESP charge (au)
C1	c2	0.92868
O1A	o	−0.88818
O1B	o	−0.88818
C2	c3	−0.22788
H22	hc	−0.03838
H23	hc	−0.03838
C3	c2	0.92868
O3B	o	−0.88818
O3A	o	−0.88818
total		−2.00000

[a]Atom name and type follow the style of GAFF.[26,27]

For the apo-U1A structure, the compensated residues were energetically minimized using a molecular mechanics simulation, where 500 steps of the steepest descent method were followed by 1500 steps of the conjugate gradient method. Subsequently, hydrogen atoms were also minimized. These calculations were performed in vacuum, where the nonbonded cutoff was set to 99 Å.

MD Simulations. All MD simulations were performed under the periodic boundary condition, where the rectangular box was set so that the minimal distance from the box's face to the molecular surface was 10 Å. The apo-U1A structure in the closed form was solvated by adding TIP3P model water molecules[25] and was electrically neutralized by adding Cl^- as counterions, and we have termed it the $U1A^{closed}$ system. We also prepared the $U1A^{closed}$–MAL system by adding 20 MAL molecules and 40 Na^+ ions to the $U1A^{closed}$ system initially on the surfaces of the periodic boundary box (refer to Table 2 for the molecular components of systems), where the *addions2* command in the LEaP module of AmberTools[23] was employed, and equilibrated them through 50 ns *NPT* MD simulations to be distributed uniformly in the periodic boundary box. This new system uses a sodium malonate concentration of 0.2 M. Electrostatic interaction was treated using the particle mesh Ewald method, where the real space cutoff was set to 9 Å. Vibrational motions associated with hydrogen atoms were frozen using the SHAKE algorithm. The translational center-of-mass motion was removed by every 1000 steps to maintain the whole system around the origin, avoiding the overflow of coordinate information from the MD trajectory format. MD trajectories were recorded at intervals of 1 ps for the subsequent analysis.

For the $U1A^{closed}$ system, we executed 25 sets of MD simulations. Each MD simulation consisted of five sequential simulations as follows: (i) *NVT* (0.1 K, 10 ps) → (ii) *NVT* (1–300 K, 20 ps) → (iii) *NVT* (300 K, 20 ps) → (iv) *NPT* (300 K, 1 bar, 100 ps) → (v) *NVT* (300 K, 20 ns). The initial atomic velocities were randomly assigned from the Maxwellian distribution at 0.1 K. The first four simulations were executed for the temperature and density relaxation using a Langevin thermostat with a 1 ps^{-1} collision coefficient and a Berendsen barostat[32] with a 2 ps coupling constant. In the second simulation, the thermostat temperature was increased linearly from 1 to 300 K. The last 20 ns *NVT* simulation was a production run for conformational sampling and was performed using the Berendsen thermostat[32] with a 5 ps coupling constant. The pressure value averaged over the time period between 12 and 20 ns was 13.0 bar on average over the 25 MD trajectories (95% confidence interval of 0.9–25.1 bar). Then, we assumed that the system was equilibrated under a pressure of 13 bar during this period. The first *NVT* simulation and the four remaining simulations used time steps of integration of 1 and 2 fs, respectively.

For the $U1A^{closed}$–MAL system, we also executed 25 sets of MD simulations. The initial atomic coordinates of each MD simulation were generated from the snapshot structure acquired at 5 ns from each of the 25 MD simulations of the $U1A^{closed}$ system. Na^+ ions and MAL molecules were added to each of the 25 snapshot structures as described above. The initial atomic velocities were randomly assigned from the Maxwellian distribution. Before the five sequential simulations, which were similar to those performed in the $U1A^{closed}$ systems, had been performed, the energies of water molecules, Na^+ ions, and MAL molecules were rapidly quenched, where the system was tightly coupled with the Berendsen thermostat[32] with a 0.001 ps coupling constant. The quenched MD simulation includes the following relaxation states: (i) *NVT* (1 K, 2.5 ps) → (ii) *NVT* (1–300 K, 2.5 ps) → (iii) *NPT* (300 K, 1 bar, 100 ps). The first MD simulation is for quenching the kinetic energy of the system. The second and third MD simulations are for the relaxation of water molecules, Na^+ ions, and MAL molecules using the Langevin thermostat with a 1 ps^{-1} collision coefficient and the Berendsen barostat[32] with a 2 ps coupling constant. The first MD simulation and the two remaining simulations used time steps of integration of 0.25 and 2 fs, respectively. During quenching, the atomic coordinates of U1A were restrained by the harmonic potential around the initial atomic coordinates with a force coefficient of 100 kcal mol^{-1} $Å^{-1}$.

After the quenched MD simulation, six sequential simulations were executed as follows: (i) *NVT* (0.1 K, 10 ps) → (ii) *NVT* (1–300 K, 20 ps) → (iii) *NVT* (300 K, 20 ps) → (iv) *NPT* (300 K, 11 bar, 50 ns) → (v) *NVT* (300 K, 10 ps) → (vi) *NVT* (300 K, 20 ns). The initial atomic velocities were randomly assigned from the Maxwellian distribution. The first three *NVT* simulations are the same as those used for the $U1A^{closed}$ system, performed for equilibration of temperature. The pressure of the system was equilibrated through the fourth simulation using the Berendsen barostat,[32] where the coupling constant was switched from 0.5 to 2 ps at 35 ns. The pressure value averaged over the time period between 42 and 50 ns was 11.4 bar over the 25 MD trajectories (95% confidence interval of 10.6–12.2 bar). Then, in the fifth simulation, the *NPT* condition was switched to the *NVT* condition. The size of the simulation box at 50 ns was replaced with the average over 8000 snapshot structures, obtained during the period between

Table 2. Molecular Components and Simulation Times

system	no. of water molecules	no. of counterions	no. of malonate acids	simulation duration (ns)	no. of trajectories
$U1A^{closed}$	6011	7 Cl^-	none	20	25
$U1A^{closed}$–MLA	6011	40 Na^+, 7 Cl^-	20	20	25

dx.doi.org/10.1021/bi401708q | *Biochemistry* 2014, 53, 3646–3657

Table 3. Ratios of Residual SASAs of Open-like Forms (open-folded form-1, of1; open-folded form-2, of2; open-unfolded form-1, ou1; open-unfolded form-2, ou2) and a Vertical Form (v) to Those of the Closed (c) or Open (o) Form[a]

	SASA (Å^2)		ratio of residual SASAs									
			open-like form									
	closed form	open form	open-folded form-1		open-folded form-2		open-unfolded form-1		open-unfolded form-2		vertical form	
residue	S_c	S_o	S_{of1}/S_c	S_{of1}/S_o	S_{of2}/S_c	S_{of2}/S_o	S_{ou1}/S_c	S_{ou1}/S_o	S_{ou2}/S_c	S_{ou2}/S_o	S_v/S_c	S_v/S_o
His10	58.9	33.3	1.2	2.2	0.9	1.7	1.4	2.5	1.1	1.9	0.8	1.5
Tyr13	19.9	42.4	2.5	**1.2**	1.9	**0.9**	2.2	**1.0**	1.9	**0.9**	1.9	**0.9**
Leu44	10.0	89.5	8.6	**1.0**	9.6	**1.1**	5.0	**0.6**	5.3	**0.6**	8.2	**0.9**
Val45	6.9	15.5	1.5	**0.7**	4.8	**2.1**	2.8	**1.3**	2.5	**1.1**	4.0	**1.8**
Phe56	0.9	52.1	34.7	**0.6**	65.3	**1.2**	11.5	**0.2**	29.4	**0.5**	45.5	**0.8**
Asp92	122.9	125.6	0.1	0.1	0.7	0.7	0.7	0.7	0.1	0.1	0.5	0.5
Ile93	46.8	80.3	1.7	**1.0**	2.3	**1.3**	2.9	**1.7**	1.1	**0.6**	1.1	0.7
Ile94	56.7	18.0	1.5	4.9	0.7	2.1	0.6	1.8	2.1	6.6	0.7	2.2
Met97	44.6	127.5	1.6	**0.6**	0.9	0.3	3.4	**1.2**	2.2	**0.8**	2.2	**0.8**

[a]Bold entries indicate that the local structure is more similar to the open form than to the closed form.

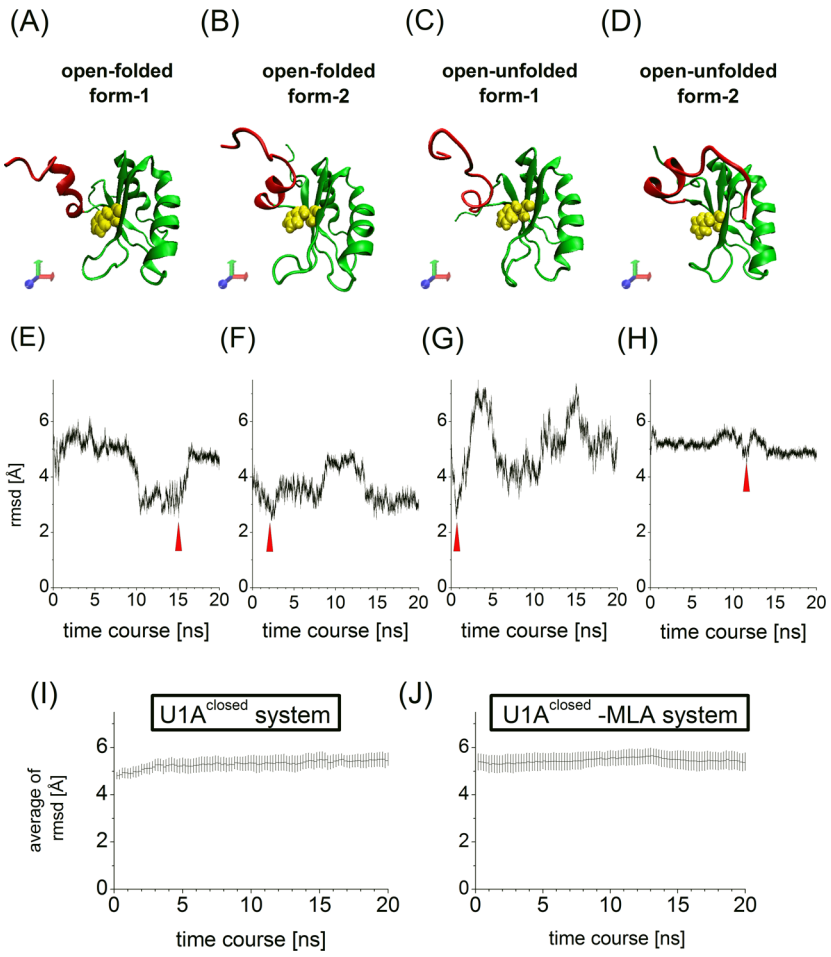


Figure 3. Representative open-like structures determined from MD simulations. (A and C) Open-folded form-1 and open-unfolded form-1 detected in an MD trajectory of the U1A^{closed} system, respectively. (B and D) Open-folded form-2 and open-unfolded form-2 detected in an MD trajectory of the U1A^{closed}–MAL system, respectively. (E–H) Time course analyses of rmsds from the open form, derived for each of the MD trajectories assuming open-like structures: (E) open-folded form-1, (F) open-folded form-2, (G) open-unfolded form-1, and (H) open-unfolded form-2. (I and J) Time course analyses of rmsds from the open form, where the values were averaged over the 25 MD trajectories at each time point: (I) U1A^{closed} system and (J) U1A^{closed}–MAL system. In panels A–D, U1A, the C-terminal region, and Phe56 are colored as in Figure 1. In panels E–H, each time point when the representative conformation appears is assigned by a red arrow. In panels I and J, the time course is divided by each 200 ps time interval. A set of 25 average values within each 200 ps time interval is averaged, where error bars indicate the 95% confidence interval.

42 and 50 ns of the *NPT* simulation. The energy of water molecules was quenched using the Berendsen thermostat[32] with a 0.001 ps coupling constant and a time step of integration of 0.25 fs, where atomic coordinates of solute molecules (U1A, MAL, and Na^+ and Cl^- ions) were restrained by the harmonic potential around the initial atomic coordinates with a force

dx.doi.org/10.1021/bi401708q | *Biochemistry* 2014, 53, 3646–3657

Table 4. Ratios of Interatomic Distances of Open-like Forms (open-folded form-1, of1; open-folded form-2, of2; open-unfolded form-1, ou1; open-unfolded form-2, ou2) and the Vertical Form (v) to Those of the Closed (c) or Open (o) Form[a]

	interatomic distance between Cα atoms (Å)		ratio of interatomic distances between Cα atoms									
			open-like forms									
	closed form	open form	open-folded form-1		open-folded form-2		open-unfolded form-1		open-unfolded form-2		vertical form	
residues	d_c	d_o	d_{of1}/d_c	d_{of1}/d_o	d_{of2}/d_c	d_{of2}/d_o	d_{ou1}/d_c	d_{ou1}/d_o	d_{ou2}/d_c	d_{ou2}/d_o	d_v/d_c	d_v/d_o
His10−Ile94	11.8	7.4	1.2	1.9	0.8	1.2	0.8	1.2	1.0	1.6	1.6	2.5
Leu44−Asp92	12.3	11.2	1.0	1.1	1.2	1.4	1.1	1.2	0.9	0.9	1.1	1.3
Tyr13−Ile93	11.8	15.0	1.2	**1.0**	1.3	**1.0**	1.2	**1.0**	1.1	0.9	1.3	**1.0**
Leu44−Met97	7.0	18.2	2.7	**1.0**	1.7	**0.7**	2.2	**0.8**	2.0	**0.8**	2.8	**1.1**
Val45−Met97	8.1	21.9	2.7	**1.0**	2.0	**0.7**	2.2	**0.8**	2.1	**0.8**	2.5	**0.9**
Phe56−Ile93	9.1	11.9	1.5	**1.2**	1.4	**1.1**	1.2	**0.9**	1.3	**1.0**	1.7	**1.3**
Phe56−Ile94	8.3	12.3	1.7	**1.1**	1.5	**1.0**	1.1	0.7	1.3	**0.9**	1.8	**1.2**
Phe56−Met97	10.8	18.0	1.7	**1.0**	1.3	**0.8**	1.5	**0.9**	1.4	**0.8**	1.9	**1.1**

[a]Bold entries indicate that the local structure is more similar to the open form than to the closed form.

coefficient of 100 kcal mol^{-1} Å^{-1}. The fourth and sixth simulations were executed using Berendsen thermostat[32] with a 5 ps coupling constant and a time step of integration of 2 fs. The final 20 ns *NVT* simulations were used for the following analyses. The pressure value averaged over the time period between 12 and 20 ns was 22.8 bar over the 25 MD trajectories (95% confidence interval of 19.6−26.0 bar). Then, we assumed that the system was equilibrated under a pressure of 22.8 bar during this time period. All MD simulations were performed using the sander and PMEMD modules in the molecular dynamics package of Amber 11.[33]

Analysis of MD Trajectories. Interatomic distances for the residue pairs, which significantly change upon reorientation of Helix-C (see Table 3 for details of residue pairs), hydrogen bond (H-bond) formation, the vector from the Cα atom of Asp90 to that of Lys98, and the root-mean-square deviation (rmsd) were calculated using the ptraj and cpptraj modules of the Amber Tools 1.5 package.[23] The secondary structure annotation and solvent-accessible surface area (SASA) were calculated using STRIDE[34] and NACCESS.[35] A H-bond was defined such that the X−Y distance was <3.3 Å and the X−H−Y angle was >120°, where X, Y, and H stand for acceptor, donor, and hydrogen atoms, respectively. rmsd fitting was performed for residues 10−102 according to backbone atoms (Cα, N, C, and O), and rmsd values were calculated using these atoms. The apo-U1A structure in the open form was prepared from chain B of the X-ray structure (PDB entry 1URN),[12] where the apo-U1A structure was converted to the wild-type structure by introducing two point mutations (H31Y and R36Q) and was compensated for the missing residues (Met1 and Lys97−Val102). Similar to the closed form of apo-U1A, the open form of apo-U1A was energetically minimized using the molecular mechanics simulation. We define contact between an MAL molecule and an amino acid residue when the minimal interatomic distance between them was <3 Å. We determined the number of MAL molecules that made contact with Arg52 for each snapshot of 25 MD trajectories of the U1A^{closed}−MAL system. Molecular structures were illustrated using Visual Molecular Dynamics (VMD),[36] and the density distribution of MAL molecules was calculated using the Volmap plugin of VMD.

RESULTS AND DISCUSSION

Existence of Open-like Forms: Open-Folded and Open-Unfolded Forms. Among the 50 MD trajectories obtained, we identified four trajectories undergoing reorientation of the C-terminal region toward the open form: two derived from the U1A^{closed} system and the other two from the U1A^{closed}−MAL system. For each of the four trajectories, the representative conformation analyzed (Figure 3A−D) was the conformation with the minimal rmsd from the open form, i.e., U1A in the complex with RNA (Figure 1B). The conformations shown in panels A and C of Figure 3 were obtained from the U1A^{closed} system, while those shown in panels B and D of Figure 3 were obtained from the U1A^{closed}−MAL system.

Panels E−H of Figure 3 show time course analyses of rmsds for the MD trajectories that assume the conformations discussed in panels A−D of Figure 3, respectively. As for each of the three MD trajectories discussed in panels E−G of Figure 3, the rmsd can be taken to represent relatively small values compared with the averaged rmsd values (Figure 3I,J). It is, therefore, possible to say that some of the U1A conformations in these three MD trajectories come to the open form. On the other hand, the other trajectory discussed in Figure 3H shows quite stable rmsd values similar to the averaged rmsd values during the simulation, although Helix-C moves away from the RNA-binding surface (see Figure 3B). This could be attributed to the uptake of the residues connecting Helix-C and the other part of U1A, inducing larger rmsds from the open form.

As shown in Figure 3A−D, these representative conformations are similar to the open form in that their C-terminal regions move away from the RNA-binding surface. Hereafter, these conformations are called "open-like forms". On the other hand, they differ in their conformational characteristics, i.e., have a folded or unfolded C-terminal region. Then, the folding condition of the C-terminal region was characterized using the helix formation index (HFI). HFI is defined as the number of residues assigned to a helix, i.e., α-helix, 3^{10}-helix, or π-helix, within Helix-C (i.e., residues 91−98). The helix assignment for each residue is given on the basis of STRIDE secondary structure annotation.[34] We defined that a conformation assumes a folding condition if the value of HFI is >5.

According to HFI, the conformations shown in panels A and B of Figure 3 possess the folded C-terminal region because of

dx.doi.org/10.1021/bi401708q | Biochemistry 2014, 53, 3646−3657

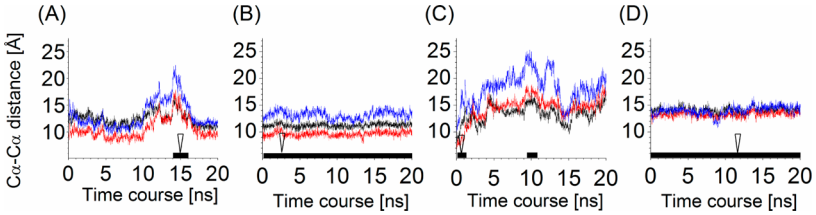


Figure 4. Time course analyses of interatomic distances. (A–D) MD trajectories assuming open-folded form-1, open-folded form-2, open-unfolded form-1, and open-unfolded form-2, respectively. The distances between Phe56 and Ile93, Phe56 and Ile94, and Phe56 and Met97 are shown by black, red, and blue curves, respectively. An open arrowhead on the horizontal axis indicates the time at which each representative open-like form was detected. The time interval when the open-like state appears is indicated by the black rectangle on the horizontal axis at the bottom. U1A, the C-terminal region, and Phe56 are colored as in Figure 1.

the HFI values of 7 and 5, respectively, and are termed open-folded form-1 and open-folded form-2, respectively. In addition, conformations shown in panels C and D of Figure 3 possess a partially (or even fully) unfolded C-terminal region because of the HFI values of 3 and 4, respectively, and are termed open-unfolded form-1 and open-unfolded form-2, respectively. Although the numbering is used to identify the MD trajectories, a U1A conformation should be simply termed the open-folded or open-unfolded form if it possesses the folded or unfolded C-terminal region, respectively.

To compare their structural characteristics with those of the open and closed forms, the residual SASA [S (square angstroms)] and the interatomic distance [d (angstroms)] were calculated for each representative open-like form (Tables 3 and 4), with a focus on the residues that changed significantly relative to the reorientation of the C-terminal region.[5] Structural similarity was estimated according to the S_x/S_y and d_x/d_y ratios, where x denotes open-folded form-1 (of-1), open-folded form-2 (of-2), open-unfolded form-1 (ou1), or open-unfolded form-2 (ou2) and y denotes the open form (o) or the closed form (c). Thus, a conformation is locally similar to the open form if S_x/S_o (d_x/d_o) is closer to 1 than S_x/S_c (d_x/d_c). For the ratios of residual SASA (Table 3), more than five of nine residues showed local similarity to the open form, indicating that those residues in the RNA-binding surface (e.g., Phe56) are exposed to permit RNA access. On the other hand, for the ratios of interatomic distances (Table 4), more than five of eight atom pairs were more locally similar to the open form than to the closed form, indicating that their C-terminal region moves away from the RNA-binding surface. According to these analyses, we concluded that these open-folded and open-unfolded conformations are similar to the open form, as a whole.

Stability of the Open-like State. The open-like state conformation of U1A is defined as that in which the C-terminal region is farther from Phe56, similar to the open form conformation. Further, the open-like state can assume an open-folded or open-unfolded state according to the conformation of the C-terminal region. We performed time course analyses of interatomic distances by focusing on pairing between Phe56 and X (where X is Ile93, Ile94, or Met97) to characterize the effects of the removal of Helix-C from Phe56 (Figure 4A–D). Furthermore, time course analyses of secondary structure formation were performed by focusing on the residues in Helix-C, namely residues 91–98 (Figure 6), to characterize the folding condition of the C-terminal region. To determine the structural stability of the open-like state, the four MD trajectories mentioned above were examined according to these quantities.

The open-folded state appearing in the MD trajectory of the U1A^{closed} system is unstable and returns to the closed form after reorientation (Figures 4A and 5A). The distances started to

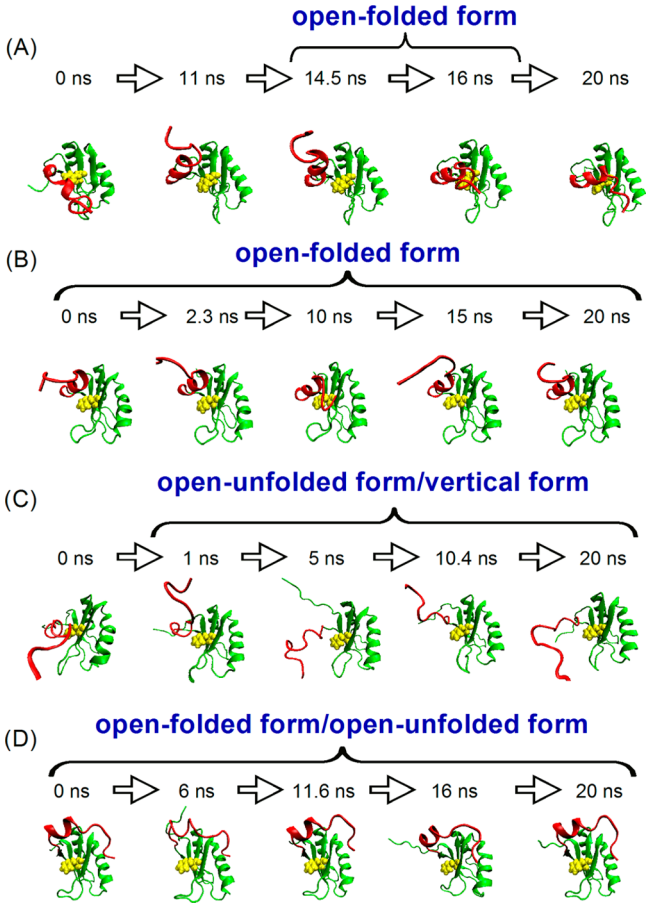


Figure 5. Helix-C reorientation over time. (A–D) MD trajectories assuming open-folded form-1, open-folded form-2, open-unfolded form-1, and open-unfolded form-2, respectively.

increase at approximately 10 ns, reached their maxima at approximately 14.5 ns, and immediately decreased to reach the plateau after 17.5 ns. In this trajectory, Met97 is relatively distant from Phe56 compared with Ile93 and Ile94, being similar to the open form (see Table 4). Therefore, the open-folded state, where Helix-C is almost held (Figure 6A,E), appeared between 14 and 16 ns.

The open-unfolded state in the MD trajectory of the U1A^{closed} system is also structurally unstable and instantaneously appears at approximately 0.6 or 10 ns in the MD

dx.doi.org/10.1021/bi401708q | *Biochemistry* 2014, 53, 3646–3657

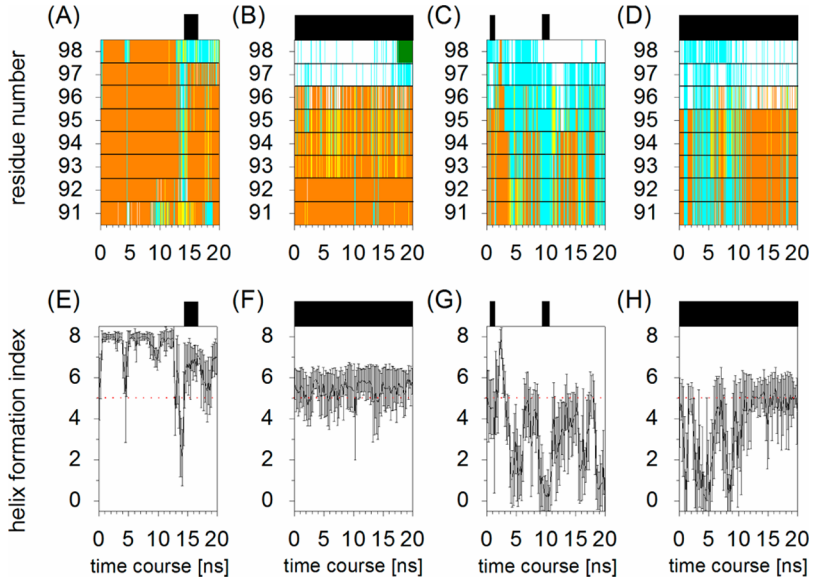


Figure 6. Time course analyses of the secondary structure of U1A. Panels A–D and E–H correspond to the MD trajectories assuming open-folded form-1, open-folded form-2, open-unfolded form-1, and open-unfolded form-2, respectively. (A–D) Time course analyses of STRIDE-based secondary structure annotation. Orange, yellow, green, blue, and white represent α-helix, 3_{10}-helix, β-strand, turn, and random coil, respectively. (E–H) Time course analyses of the helix formation index (HFI). The red dotted line indicates a HFI of 5. The time course is divided by 200 ps time intervals, and the values in each time interval are averaged, where an error bar indicates the standard deviation for the values in each 200 ps time interval. For each graph, the time interval where the open-like state appears is indicated by the black rectangle on the horizontal axis at the top.

trajectory (Figures 4C and 5C). Although the rmsd values are around 4 Å near 10 ns (see Figure 3C), their C-terminal regions move away from the RNA-binding surface. Helix-C assumes a partially folded or fully unfolded conformation (Figure 6C,G). It should be noted that values of HFI fluctuate around 5 during between 0 and 2 ns. This observation suggests that the open-folded state transiently converts into the open-unfolded state, and vice versa. After the initial conversion to the open-unfolded state, the C-terminal region orients vertically toward the RNA-binding surface. The conformation at 5 ns in Figure 4C is termed the vertical form and was analyzed as a representative example. The data listed in Tables 3 and 4 indicate a tendency for the vertical form to move away from the closed form. U1A is in the vertical state if its C-terminal region is in the same orientation as that in the vertical form. In this MD trajectory, the vertical state appears at approximately 5 ns, and then U1A almost assumes the open-unfolded or vertical state.

It is worth noting that the open-unfolded state in the U1A^{closed} system appears with the vertical state in the MD trajectory (Figure 5C). Then, it can be assumed that the vertical state may convert more frequently to the open-unfolded state than to the closed form. To test this assumption, we performed another set of MD simulations starting with the vertical form in the absence of MAL and found that such a conversion from the vertical state to the open-unfolded state occurs irrespective of the presence of MAL (see section SI-1 of the Supporting Information for details). Therefore, it is plausible that the vertical state may be on the edge of the basin of the open-unfolded form in the conformational repertoire of U1A in aqueous solution.

The open-like states in the U1A^{closed}–MAL system are structurally stable throughout the simulation (Figure 4B,D and Figure 5B,D). In the MD trajectory assuming open-folded form-2, averaged HFI values fluctuate around 6 (Figure 6B,F). Meanwhile, the error bars for HFI values indicate transient conversion between the open-folded and open-unfolded states. This might occur regardless of the orientation of Helix-C (see section SI-2 of the Supporting Information for details). In this trajectory, Met97 is relatively distant from Phe56 compared with Ile94 and Ile93 (Figure 4D), similar to the case in the open form (see Table 4). Then, it is possible to say that the orientation of Helix-C is similar to those of the open form and the open-folded state in the MD trajectory of the U1A^{closed} system.

The open-unfolded state in the MD trajectory of the U1A^{closed} system was mainly found during the three time periods between 1 and 2 ns, between 3 and 6 ns, and between 8 and 10 ns. Within the other time periods, the open-unfolded state transiently converts into the open-folded state, and vice versa. Ile93 and Ile94 are as distant from Phe56 as Met97. This could be attributed to the uptake of the residues connecting Helix-C and the other part of U1A (see Figure 3D). Meanwhile, in the open form and in the MD trajectories assuming the open-like state in the U1A^{closed} system, Met97 is relatively distant from Phe56 compared with the other two (see Table 4 and Figure 4A–C). Then, the open-like conformations in this trajectory are slightly different from those appearing in the other three trajectories, but they still permit RNA to access the RNA-binding surface. Accordingly, it is possible to say that Helix-C in the open-like state assumes a variety of conformations with respect to both the orientation and the folding condition.

These two open-like states in the U1A^{closed}–MAL system were maintained relatively longer (20 ns) than those in the U1A^{closed} system. This should be partially attributed to the influence of MAL because MAL molecules were distributed at a part of the RNA-binding surface of the open-unfolded form (Figure 7A,B), and they could prevent the return to the closed form. In Figure 7B, MAL molecules seem to contribute to the maintenance of the orientation of Helix-C through U1A–MLA interaction. On the other hand, MAL was not present at the site

dx.doi.org/10.1021/bi401708q | *Biochemistry* 2014, 53, 3646–3657

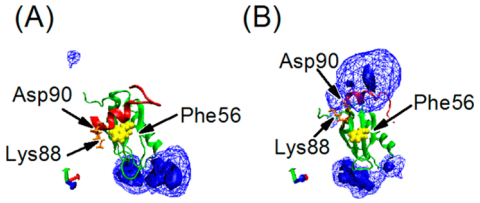


Figure 7. Density distribution of MAL around U1A in the open-unfolded state. (A) MD trajectory that assumes open-unfolded form-1. (B) MD trajectory that assumes open-unfolded form-2. The distributions were calculated using the MD trajectories from 0 to 20 ns. Locations of MAL with densities of >0.01 and >0.05 Å^{-3} are drawn as blue wire frames and solid surfaces, respectively. U1A, the C-terminal region, Phe56, Lys88, and Asp90 are colored as in Figures 1 and 2.

occupied by Cyt10 in the U1A−RNA complex (compare each panel in Figure 7 with panels A and B of Figure 2). Such coordination of MAL may occur for a longer time but is not essential for the reorientation of Helix-C. Therefore, it is reasonable to conclude that MAL molecules contribute, in general, to the stabilization of the open-like state. Moreover, despite the presence of MAL, the C-terminal region remained partially unfolded and should refold into Helix-C as shown in the alternative open form[5] (Figure 2A). Such a refolding process would take longer time than that in the simulations presented here. Thus, the closed−open conversion may differ mechanistically from the refolding of the C-terminal region. Specifically, they are relaxed independently on two different time scales. Furthermore, refolding requires U1A−RNA interactions. In terms of the formation of the U1A−RNA complex, it is therefore possible to say that induced fit works only for the refolding of the C-terminal region and the local interactions between U1A and RNA. Taken together, these analyses suggest that the open-like states, both the open-folded and open-unfolded states, appear spontaneously because they are part of the conformational repertoire of U1A in aqueous solution. This observation suggests that the reorientation of the C-terminal region occurs through conformational selection.

Coexistence of the Open-like and Closed States. Annuciado et al. employed time-resolved fluorescence anisotropy to examine the segmental motion of the C-terminal region[3] in the absence of RNA on a nanosecond time scale. They represented the motion of the C-terminal region with a cone whose vertex is the N-terminus of Helix-C. From their point of view, the cone angle should be close to 90° if closed−open conversion occurs. However, the observed value was 20°. This means that under the circumstances, U1A is usually in the "closed state". Therefore, its C-terminal region may assume the same orientation as that in the closed form, which possesses the C-terminal region covering Phe56. Therefore, their observation disagrees with ours because the C-terminal region in the closed state spontaneously reorients toward that in the open-like state in aqueous solution on the same time scale.

To determine the reason for this disagreement, we used our MD trajectories to examine the segmental motion of the C-terminal region by calculating the cone angle. The cone angle was defined as the ensemble average of the librational angle of motion of the C-terminal region, Ψ, i.e., ⟨Ψ⟩. The broken brackets denote the ensemble average of a variable within them over MD trajectories. Ψ is defined as follows:

$$\Psi = \cos^{-1}(\vec{\mu}\cdot\langle\vec{\mu}\rangle/|\vec{\mu}|\cdot|\langle\vec{\mu}\rangle|) \quad (1)$$

where $\vec{\mu}$ is a vector from the Cα atom of Asp90 to that of Lys98 [it should be noted that Asp90 is a linker region connecting the C-terminal region (residues 90−102) and the other (residues 1−89) (see Figure 2)]. Before calculating $\vec{\mu}$, we fit each MD snapshot structure to the initial snapshot in each MD trajectory on the basis of the rmsd, using the Cα atoms of the residues in the stable β-sheet of U1A (Leu41, Ile43, and Gln54). This procedure removes the entire rotational motion of U1A, clarifying the segmental motion of the C-terminal region in U1A.

We assumed that the U1A^{closed} system was in equilibrium after 12 ns (Figure 8A) because the ensemble average of the

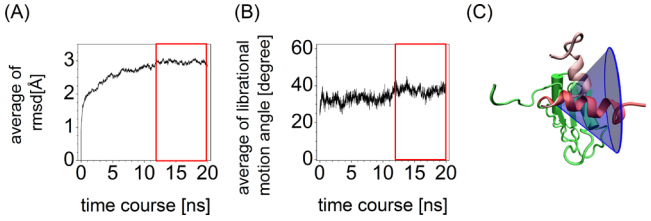


Figure 8. Cone angle analysis of the motion of the C-terminal region. (A) Average rmsd (Cα atoms of residues 10−102). (B) Average of the librational motion angle of the C-terminal region. (C) Mobile field of the C-terminal region in the conformational repertoire of U1A, represented by the blue circular cone with a cone angle of 37°. In panel A and B, the average values are calculated at each time point over 25 MD trajectories of the U1A^{closed} system. The assumed time interval for equilibrium is boxed in red. In panel C, the closed and open forms of the C-terminal region (residues 89−102) are colored dark and light red, respectively.

rmsd with respect to residues 10−102 over 25 trajectories was convergent. Although the average of Ψ over 25 trajectories fluctuated, we assumed that it was in equilibrium after 12 ns (Figure 8B). We then employed a set of 25 partial MD trajectories consisting of each last 8 ns interval to evaluate the cone angle in the equilibrium of the rmsd. An average of the cone angle was taken over the time interval of 8 ns, and then the average over the 25 partial MD trajectories was taken.

The cone angle calculated from these partial MD trajectories was 37.2° (95% confidence interval of 32.2−42.2°), which is slightly larger than that of the earlier experimental study but smaller than 90°. In our simulations of U1A^{closed} system, the pressure of the system is 13 bar (see Materials and Methods), thus being slightly higher than that of the normal experimental condition, 1 bar. According to the volume theorem, "partial molar volume of a protein decreases with the decrease of its conformational order",[37] it is possible that the segmental motion of the C-terminal region becomes flexible because of the higher pressure. Additionally, the C-terminal region might be more stably folded under the normal experimental condition.

There was one partial MD trajectory that existed in the open-folded state for approximately 2 ns (1%) of the entire duration (refer to Figure 4D and the related discussion). However, even if we excluded this partial MD trajectory to calculate the cone angle, the result was 37.3° as an average of the 24 trajectories (95% confidence interval of 32.1−42.5°). Therefore, we conclude that the disagreement between the observations of Annuciado et al.[3] and ours can be attributed to the inability of time-resolved fluorescence anisotropy to detect such a minor state as the open-folded state, which could be detected by high-pressure NMR. In fact, Helix-C in the open-folded state is

dx.doi.org/10.1021/bi401708q | Biochemistry 2014, 53, 3646−3657

partially unfolded (Figure 6B). The volume theorem tells us that partially unfolded conformations should become dominant under higher pressures. This should apply to the open-unfolded state because of complete Helix-C unfolding (Figure 6C,D).

On the basis of the discussion presented above, we concluded that the open-like state coexists with the closed state in aqueous solution. However, in MD simulation studies starting from the closed form,[3,7,16] no open-like form was reported. These results may be observed because of an insufficient number of samplings that would not reveal such a minor state. Moreover, only a single MD simulation of <10 ns was executed. In contrast, in our study, we performed MD simulations for a few hundred nanoseconds to sample the conformations and succeeded in detecting these open-like forms.

Stabilization of the Encounter Complex by Exchanging Inter- and Intramolecular Interactions. U1A binds to RNA through the concerted "stacking of RNA bases with aromatic side chain of U1A and many direct and water-mediated hydrogen bonds",[12] where the RNA bound to U1A is folded into a hairpin form. On the other hand, an NMR study indicates that free RNA assumes a conformation different from such a hairpin form,[38] suggesting that the RNA undergoes a conformational change during the formation of the complex with U1A. Therefore, in the initial step of the formation of the U1A–RNA complex, such concerted stacking and hydrogen bonding may not form between U1A and RNA.

When U1A initially approaches an RNA molecule (the first step of the TN conjecture), the U1A–RNA encounter complex is formed using "the edge of a positively charged region of the protein"[14] (the second step of the TN conjecture). This allows U1A to anchor to RNA rigidly enough to wait for the closed–open conversion. A positive charge on the protein would be attracted to negative charges on the RNA (the phosphate backbone[10,39]). Using the two negative charges and the size of an MAL molecule, an MAL molecule must be located on a positively charged residue in U1A as RNA phosphate groups interact with U1A. Therefore, the U1A–MAL interaction should be as stable as the U1A–RNA interaction during the formation of the U1A–RNA encounter complex.

In the second step of the TN conjecture, Glu19 and Arg52 play important roles in the sequence-specific recognition of RNA. Thus, according to the X-ray crystallographic structure of the U1A–RNA complex[12] and the results of the MD simulation by Tang and Nilsson,[14] Glu19 and Arg52 form specific close contacts with RNA bases (i.e., Ade6, Ura7, and Gua9) (see the right inset of Figure 1B). Arg52 forms a salt bridge with the phosphate group of Gua16, contributing to higher-affinity binding to RNA.[13,40] In the closed form of U1A, Arg52 is close enough to Glu19 to form H-bonds between them (see the inset of Figure 1A). When U1A makes contact with RNA, the H-bonds should break to form H-bonds between Arg52 and RNA (see the left inset of Figure 1B). Our MD simulations revealed that H-bonds between Glu19 and Arg52 were formed and broken (Figure 9A,B). For the $U1A^{closed}$ system, the average number of H-bonds was 0.82 in equilibrium after 12 ns for the 25 trajectories (95% confidentce interval of 0.65–0.99), while for the $U1A^{closed}$–MAL system, the average number of H-bonds was 0.35 in equilibrium after 12 ns for the 25 trajectories (95% confidence interval of 0.11–0.61). MAL molecules gradually made contact with Arg52 during this time interval (Figure 9C), and the average number of MAL molecules was 1.24 in equilibrium after 12 ns for the

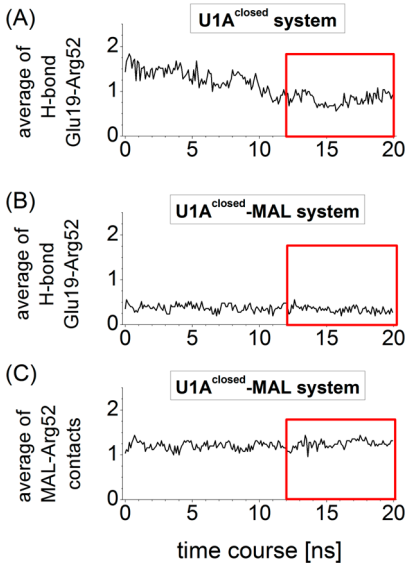


Figure 9. Formation of hydrogen bonds between Glu19 and Arg52 and coordination of MAL with Arg52. (A and B) Numbers of hydrogen bonds between Glu19 and Arg52 for (A) the $U1A^{closed}$ system and (B) the $U1A^{closed}$–MAL system. (C) Number of MAL molecules making contact with Arg52 in U1A. In panels A–C, the average value of 25 MD trajectories at each time point is shown. The assumed time interval of equilibrium is boxed in red.

25 trajectories (95% confidence interval of 0.88–1.60). The pressure values of the $U1A^{closed}$ and $U1A^{closed}$–MAL systems are 13.0 and 22.8 bar, respectively (see Materials and Methods). However, we can assume that the difference in the pressure should be not so significant if we consider the circumstances of high-pressure NMR studies, where the effect of pressure is seriously discussed only after the pressure in one experiment becomes 10-fold higher than that of the other.[18] Therefore, we conclude that the intramolecular interaction between Glu19 and Arg52 was partially disrupted and was exchanged for the intermolecular interaction between MAL and Arg52.

According to these observations, we deduced the following mechanism for stable anchoring of U1A to RNA. The intramolecular interaction is exchanged with intermolecular interactions upon the formation of the encounter complex. Thus, RNA induces the disruption of the H-bonds between Glu19 and Arg52, and Arg52 makes contact with phosphate groups in RNA to anchor U1A to the RNA; on the other hand, Glu19 searches for and makes contact with the recognition bases. Such an exchange between intra- and intermolecular interactions would favor the anchoring of U1A to the RNA until it undergoes the closed–open conversion. Considering that the disruption of H-bond formation occurs under the influence of MAL molecules, it is reasonable to conclude that such an exchange between intra- and intermolecular interactions should be promoted by the formation of the U1A–RNA encounter complex.

According to the discussion presented above, we speculate about how U1A searches the target sequence in RNA. Anchoring of RNA to U1A may involve several collisions between them through librational motion at the anchoring point. Under these circumstances, U1A incidentally makes contact with RNA on the RNA-binding surface only when it is in the open-like state. Subsequently, if U1A finds the target

sequence, the formation of a stable complex is mediated by specific interactions between U1A and RNA. If this does not occur, U1A is released from RNA and would wander until it succeeds in finding the target sequence.

Combined Mechanism of Conformational Selection and Induced Fit in the Formation of the U1A−RNA Complex: Modified TN Conjecture. We suggest that prior to the correct recognition of RNA, U1A makes contact with RNA on the RNA-binding surface by conformational selection because the open-like state spontaneously appears in the conformational repertoire of U1A in aqueous solution. This assumption is essentially different from the third step of the TN conjecture[14] because Helix-C reorientation occurs under the coupling to RNA folding "in response to the correct recognition of the RNA".[14] Thus, Tang and Nilsson[14] discussed RNA recognition consistently in the context of induced fit. Although the open-like state spontaneously appears, we here found that conformations in the open-like state differ from the open form because they are structurally unstable, incomplete, or both with respect to Helix-C formation in the C-terminal region.

We can speculate that in the open-like state, U1A incidentally makes contact with RNA on the RNA-binding surface and then "the correct recognition of the RNA"[14] induces the stabilization of the U1A−RNA complex, entailing the refolding into the α-helix. This mechanism is similar to that proposed in the W conjecture,[15] which maintains that in mutually induced fit, U1A would fold together with RNA to form the complex. Furthermore, we suggest that not only induced fit but also conformational selection plays an essential role in the formation of the U1A−RNA complex. As pointed out above, the open-like forms are a minority of the U1A conformations in aqueous solution. However, from the viewpoint of the chemical equilibrium, once the open-like forms turn to the open forms for the U1A−RNA recognition, some of the remaining U1A molecules in the closed form should convert into open forms. Such conformational conversion would continue until the U1A−RNA recognition process becomes equilibrated.

This insight reminds us of an earlier lower-resolution MD simulation study by Okazaki and Takada, which proposed a combined mechanism of conformational selection and induced fit according to the physicochemical mechanism underlying the formation of the protein−ligand complex.[41] Using Markov state models, Silva et al. recently reported an example of such combined mechanisms on the basis on all-atom simulations.[42] The results of these studies seem reasonable because molecular interactions in a cell are much more complex than we previously imagined such that the concepts of conformational selection and induced fit may occur independently. Thus, both were originally proposed to interpret the kinetic mechanism in which molecules are often represented by characteristic states, e.g., the relaxed−tense form in hemoglobin allostery.[43] Therefore, it may be impossible to classify a molecular recognition mechanism as conformational selection or induced fit.

The preceding demonstration of the reorientation of the C-terminal region in U1A and its subsequent contact with RNA, i.e., conformational selection-driven events, should be regarded as an independent step in U1A−RNA molecular recognition. We here propose a four-step modification of the three-step reaction scenario of the TN conjecture. The first and second steps are identical to those of the original TN conjecture, i.e., the approach of U1A to RNA through electrostatic interactions and the formation of the U1A−RNA encounter complex, respectively. However, the third step involves the spontaneous appearance of the open-like state and the subsequent contact with RNA on the RNA-binding interface. In the fourth step, the stabilization of the U1A−RNA complex is mediated by specific interactions between U1A and RNA.

■ CONCLUDING REMARKS

In this study, we performed all-atom MD simulations and found that the open-like states spontaneously appear in the conformational repertoire of U1A in aqueous solution. According to our MD simulations, the open-like states represent a minority of the U1A conformations in aqueous solution. However, high-pressure NMR[18,37] should detect these states because the Helix-C structure is less ordered.

The presence of the open-like states suggests that irrespective of "the correct recognition of the RNA",[14] U1A undergoes a change in its conformation to permit RNA access and subsequently makes contact with RNA on the RNA-binding surface. Therefore, "the correct recognition of the RNA"[14] would work only for the stabilization of the U1A−RNA complex. Finally, we suggest that U1A−RNA molecular recognition is mediated by a combined mechanism of conformational selection and induced fit.

According to the data presented here, we propose a further modification of the TN conjecture that explains the mechanism underlying U1A−RNA molecular recognition. This knowledge may serve as a basis for understanding all protein−RNA interactions.

■ ASSOCIATED CONTENT

Supporting Information

Details of another MD simulation starting from the vertical form (section SI-1, Figure S1, and Tables S1 and S2) and time course analyses of interatomic distances and secondary structure formation on average over the 25 MD trajectories (section SI-2 and Figures S2 and S3). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Graduate School of Information Science, Nagoya University, Furo-cho Chikusa-ku, Nagoya 464-8601, Japan. E-mail: mnagaoka@is.nagoya-u.ac.jp. Telephone and fax: +81-52-789-5623.

Funding

This work was partially supported by the Core Research for Evolutional Science and Technology (CREST), "High Performance Computing for Multi-scale and Multi-physics Phenomena" and "Establishment of Molecular Technology towards the Creation of New Functions", of the Japan Science Technology Agency and by a Grant-in-Aid for Science Research from the Ministry of Education, Culture, Sport, Science and Technology of Japan. I.K. is thankful for the support from the Japan Society for the Promotion of Science (JSPS) via a Research Fellowship for Young Scientists.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The calculations were partially performed using several computing systems at the Information Technology Center of Nagoya University.

■ ABBREVIATIONS

MAL, malonate; MD, molecular dynamics; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; GAFF, general Amber force field; SASA, solvent-accessible surface area; rmsd, root-mean-square deviation; VMD, Visual Molecular Dynamics; o, open form; c, closed form; v, vertical form; of1, open-folded form-1; of2, open-folded form-2; ou1, open-unfolded form-1; ou2, open-unfolded form-2.

■ REFERENCES

(1) Hieronymus, H., and Silver, P. A. (2004) A systems view of mRNP biology. *Genes Dev. 18*, 2845−2860.
(2) Cooper, T. A., Wan, L. L., and Dreyfuss, G. (2009) RNA and Disease. *Cell 136*, 777−793.
(3) Anunciado, D., Agumeh, M., Kormos, B. L., Beveridge, D. L., Knee, J. L., and Baranger, A. M. (2008) Characterization of the dynamics of an essential helix in the U1A protein by time-resolved fluorescence measurements. *J. Phys. Chem. B 112*, 6122−6130.
(4) Anunciado, D., Dhar, A., Gruebele, M., and Baranger, A. M. (2011) Multistep Kinetics of the U1A-SL2 RNA Complex Dissociation. *J. Mol. Biol. 408*, 896−908.
(5) Avis, J. M., Allain, F. H. T., Howe, P. W. A., Varani, G., Nagai, K., and Neuhaus, D. (1996) Solution structure of the N-terminal RNP domain of U1A protein: The role of C-terminal residues in structure stability and RNA binding. *J. Mol. Biol. 257*, 398−411.
(6) Kormos, B. L., Baranger, A. M., and Beveridge, D. L. (2007) A study of collective atomic fluctuations and cooperativity in the U1A-RNA complex based on molecular dynamics simulations. *J. Struct. Biol. 157*, 500−513.
(7) Kormos, B. L., Benitex, Y., Baranger, A. M., and Beveridge, D. L. (2007) Affinity and specificity of protein U1A-RNA complex formation based on an additive component free energy model. *J. Mol. Biol. 371*, 1405−1419.
(8) Kormos, B. L., Pieniazek, S. N., Beveridge, D. L., and Baranger, A. M. (2011) U1A Protein-Stem Loop 2 RNA Recognition: Prediction of Structural Differences from Protein Mutations. *Biopolymers 95*, 591−606.
(9) Law, M. J., Chambers, E. J., Katsamba, P. S., Haworth, I. S., and Laird-Offringa, I. A. (2005) Kinetic analysis of the role of the tyrosine 13, phenylalanine 56 and glutamine 54 network in the U1A/U1 hairpin II interaction. *Nucleic Acids Res. 33*, 2917−2928.
(10) Law, M. J., Linde, M. E., Chambers, E. J., Oubridge, C., Katsamba, P. S., Nilsson, L., Haworth, I. S., and Laird-Offringa, I. A. (2006) The role of positively charged amino acids and electrostatic interactions in the complex of U1A protein and U1 hairpin II RNA. *Nucleic Acids Res. 34*, 275−285.
(11) Law, M. J., Rice, A. J., Lin, P., and Laird-Offringa, I. A. (2006) The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA. *RNA 12*, 1168−1178.
(12) Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., and Nagai, K. (1994) Crystal-structure at 1.92 angstrom resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature 372*, 432−438.
(13) Qin, F., Chen, Y., Wu, M. Y., Li, Y. X., Zhang, J., and Chen, H. F. (2010) Induced fit or conformational selection for RNA/U1A folding. *RNA 16*, 1053−1061.
(14) Tang, Y., and Nilsson, L. (1999) Molecular dynamics simulations of the complex between human U1A protein and hairpin II of U1 small nuclear RNA and of free RNA in solution. *Biophys. J. 77*, 1284−1305.
(15) Williamson, J. R. (2000) Induced fit in RNA-protein recognition. *Nat. Struct. Biol. 7*, 834−837.
(16) Pitici, F., Beveridge, D. L., and Baranger, A. M. (2002) Molecular dynamics simulation studies of induced fit and conformational capture in U1A-RNA binding: Do molecular substates code for specificity? *Biopolymers 65*, 424−435.
(17) Rupert, P. B., Xiao, H., and Ferre-D'Amare, A. R. (2003) U1A RNA-binding domain at 1.8 angstrom resolution. *Acta Crystallogr. D59*, 1521−1524.
(18) Kitahara, R., Yokoyama, S., and Akasaka, K. (2005) NMR snapshots of a fluctuating protein structure: Ubiquitin at 30 bar-3 kbar. *J. Mol. Biol. 347*, 277−285.
(19) Sawaya, M. R., and Kraut, J. (1997) Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: Crystallographic evidence. *Biochemistry 36*, 586−603.
(20) Luo, X. L., Tang, Z. Y., Xia, G. H., Wassmann, K., Matsumoto, T., Rizo, J., and Yu, H. T. (2004) The Mad2 spindle checkpoint protein has two distinct natively folded states. *Nat. Struct. Mol. Biol. 11*, 338−345.
(21) Yang, F., and Phillips, G. N. (1996) Crystal structures of CO-, deoxy- and met-myoglobins at various pH values. *J. Mol. Biol. 256*, 762−774.
(22) Berman, H. M. (2008) The Protein Data Bank: A historical perspective. *Acta Crystallogr. A64*, 88−95.
(23) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B. P., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) *AmberTools 1.5*, University of California, San Francisco.
(24) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf. 65*, 712−725.
(25) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(26) Joung, I. S., and Cheatham, T. E. (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B 112*, 9020−9041.
(27) Joung, I. S., and Cheatham, T. E., III (2009) Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. *J. Phys. Chem. B 113*, 13279−13290.
(28) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605−1612.
(29) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Rob, M. A., Cheeseman, J. R., Jr., Montgomery, J. A., Jr., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2003) *Gaussian 03*, Gaussian, Inc., Wallingford, CT.
(30) Wang, J. M., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell. 25*, 247−260.

(31) Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general Amber force field. *J. Comput. Chem. 25*, 1157−1174.
(32) Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A., and Haak, J. R. (1984) Molecular-dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(33) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B. P., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) *Amber 11*, University of California, San Francisco.
(34) Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment. *Proteins: Struct., Funct., Genet. 23*, 566−579.
(35) Hubbard, S. J., and Thornton, J. M. (1993) *NACCESS*, Department of Biochemistry and Molecular Biology, University College London, London.
(36) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics Modell. 14*, 33−38.
(37) Akasaka, K. (2003) Exploring the entire conformational space of proteins by high-pressure NMR. *Pure Appl. Chem. 75*, 927−936.
(38) Hall, K. B. (1994) Interaction of RNA Hairpins with the Human U1A N-Terminal RNA-Binding Domain. *Biochemistry 33*, 10076−10088.
(39) Katsamba, P. S., Myszka, D. G., and Laird-Offringa, I. A. (2001) Two functionally distinct steps mediate high affinity binding of U1A protein to U1 hairpin II RNA. *J. Biol. Chem. 276*, 21476−21481.
(40) Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) Crystal-Structure of the RNA-Binding Domain of the U1 Small Nuclear Ribonucleoprotein-A. *Nature 348*, 515−520.
(41) Okazaki, K. I., and Takada, S. (2008) Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. *Proc. Natl. Acad. Sci. U.S.A. 105*, 11182−11187.
(42) Silva, D. A., Bowman, G. R., Sosa-Peinado, A., and Huang, X. H. (2011) A Role for Both Conformational Selection and Induced Fit in Ligand Binding by the LAO Protein. *PLoS Comput. Biol. 7*, 1−11.
(43) Bellelli, A. (2010) Hemoglobin and Cooperativity: Experiments and Theories. *Curr. Protein Pept. Sci. 11*, 2−36.