JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Structural Aspects for the Recognition of ATP by Ribonucleopeptide Receptors

Shun Nakano,[†] Tsukasa Mashima,[†] Akimasa Matsugami,[$] Masafumi Inoue,[†] Masato Katahira,[†] and Takashi Morii*[,†,‡]

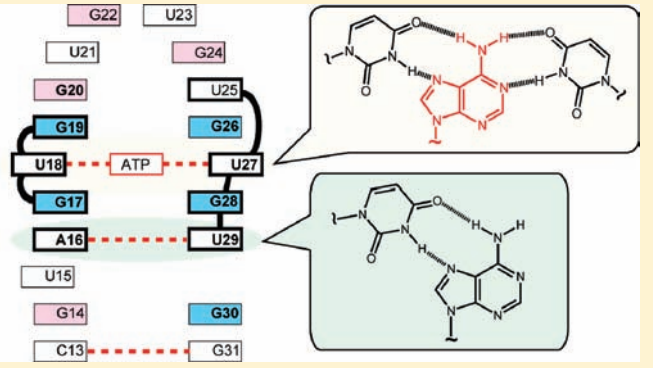
[†]Institute of Advanced Energy, Kyoto University

[‡]CREST, JST, Uji, Kyoto 611-0011, Japan

[$]Department of Supramolecular Biology, Graduate School of Nanobioscience, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Supporting Information

ABSTRACT: A modular structure of ribonucleopeptide (RNP) affords a framework to construct macromolecular receptors and fluorescent sensors. We have isolated ATP-binding RNP with the minimum of nucleotides for ATP binding, in which the RNA consensus sequence is different from those reported for RNA aptamers against the ATP analogues. The three-dimensional structure of the substrate-binding complex of RNP was studied to understand the ATP-binding mechanism of RNP. A combination of NMR measurements, enzymatic and chemical mapping, and nucleotide mutation studies of the RNP-adenosine complex show that RNP interacts with the adenine ring of adenosine by forming a U:A:U triple with two invariant U nucleotides. The observed recognition mode for the adenine ring is different from those of RNA aptamers for ATP derivatives reported previously. The RNP-adenosine complex is folded into a particular structure by formation of the U:A:U triple and a Hoogsteen type A:U base pair. This recognition mechanism was successfully utilized to convert the substrate-binding specificity of RNP from ATP- to GTP-binding with a C^+:G:C triple recognition mode.



INTRODUCTION

In vitro selection or SELEX (systematic evolution of ligands by exponential enrichment) provides powerful strategies for obtaining functional RNA molecules.[1−5] The selection and evolution technique produces RNA aptamers that specifically bind a wide range of targets[6−8] including small molecules, proteins or even cell membranes, and ribozymes that catalyze a variety of chemical reactions[9,10] from randomized pools of oligoribonucleotides. Recent advances in the in vitro selection method have produced signaling nucleic acid aptamers[11−16] applicable for biosensors and allosteric ribozymes regulated by both ligands and proteins.[17−21] The key feature of RNA to perform such a variety of functions has been attributed to the fact that a random pool of RNA provides a large number of different shapes, which increase the chance of selecting RNA molecules with the desired functionality.[22]

Application of the selection and evolution technique is not limited to obtaining functional macromolecules solely composed of RNA or DNA. It can also be used on ribonucleopeptides (RNP), such as the structurally well characterized complex of the Rev peptide and RRE (Rev Responsive Element) RNA.[23−28] A randomized nucleotide sequence was introduced into the RNA subunit of RNP to construct an RNP library onto which the in vitro selection method was applied (Figure 1a).[23−25] In the second step, the Rev peptide was modified with a fluorophore without greatly disturbing the affinity and specificity of the RNP receptor.[26] In the absence of a ligand for RNP, fluorescence emission was effectively quenched in the RNP complex, but it recovered upon ligand binding. RNP sensors thus created allow monitoring of the ligand-binding event by measuring the fluorescent signals. RNP receptors targeting ATP,[23,24] histamine,[27] or phosphotyrosine in the specific amino acid sequence[28] have been isolated and were successfully converted to the corresponding fluorescent sensors.[26−28]

Comparison of the characteristics of the RNA aptamers and RNPs targeting the same ligand provides an opportunity to understand the structural features that influence the fundamental molecular recognition mechanisms. To establish a proof of principle for the RNP-based selection strategy, we initially chose ATP as a target for selection from the RNP library[23,26] and compared the characteristics of this ATP-binding RNP with

Received: November 30, 2010
Published: March 03, 2011

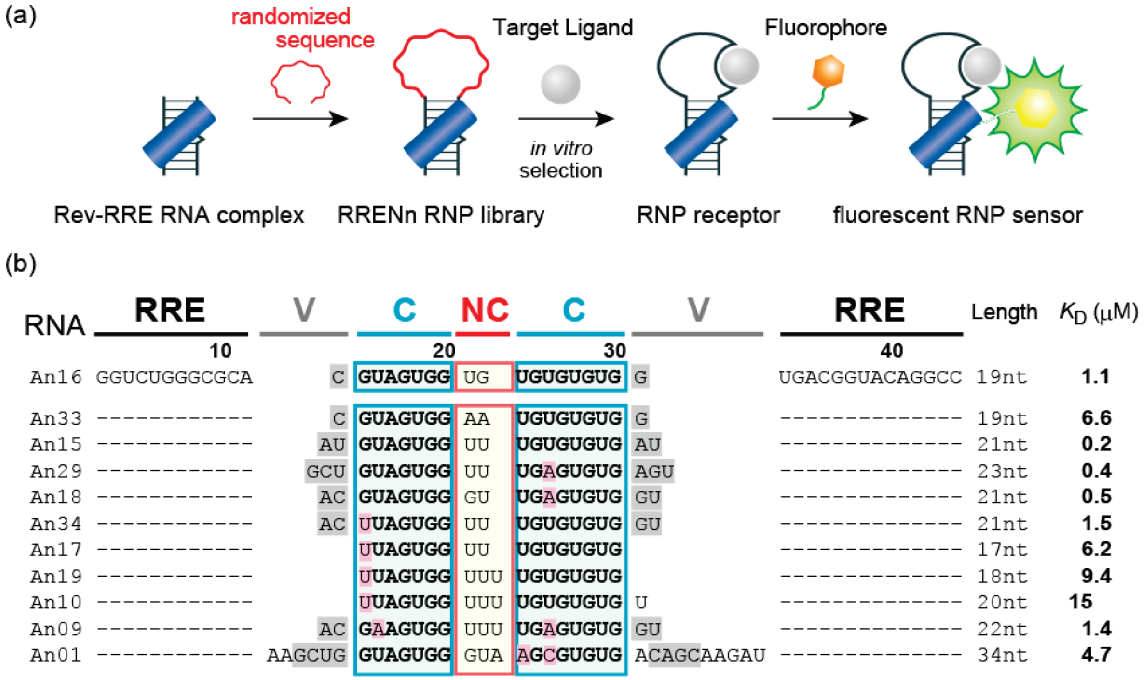


Figure 1. (a) A scheme shows the stepwise construction of RNP receptors and fluorescent RNP sensors. The Rev peptide and RRE RNA form a stable, structurally well-characterized RNP complex.[37] Randomized RNA sequences are introduced as a loop into the stable RNP complex to afford the RRENn RNP library.[28] By applying the in vitro selection method on this library, RNP receptors that bind specifically to various target ligands are obtained. RNP receptors are subsequently converted to RNP sensors by introducing a fluorophore at the N-terminus of the Rev peptide.[26,28] (b) Nucleotide sequences of ATP-binding RNP receptors isolated from the RRENn RNP library. The consensus region (C) is indicated in bold with the irregular nucleotides marked in pink. The nonconsensus region (NC) is boxed in red. Nucleotides expected to form base pairs in the variable region (V) are shaded in gray. The equilibrium dissociation constant (K_D) for the ATP complex of each RNP was obtained from the fluorescence spectral changes of each RNP complexed with 7-methoxy coumarin-labeled Rev (7mC-Rev) by titration with ATP.[26]

those of the reported ATP-binding RNA aptamer[6] and other RNA aptamers targeting the ATP analogues cAMP,[29] FAD,[30] NAD^+,[30] and S-adenosyl methionine (SAM).[31] These aptamers are all selected against the ATP analogues by using the target ligand tethered through the C8 position of the adenine ring and share quite similar consensus sequences for recognizing the adenosine moiety. The isolated ATP-binding RNPs from the RREN30 RNP library,[26] in contrast, reveal a consensus sequence that is different from the ones already reported for the ATP-binding aptamers. Because in vitro selection would afford diverse binding structures that vary significantly in accordance with the nature of ligands,[5,32,33] ATP-binding RNPs obtained by the selection against ATP immobilized through the 2′- or 3′-OH group of ribose may reveal a different consensus sequence, as suggested by Burke and Gold.[8] It has been postulated that RNA aptamers obtained by in vitro selection from random sequence pools tend to be the simplest, most abundant structures for a given ligand.[34] Then, RNA aptamers can specifically bind to the adenine ring of ATP by the naturally occurring A:U Watson−Crick or Hoogsteen base pairing, or the triplex type U:A:U formation by both Watson−Crick and Hoogsteen base pairing. However, structural analyses of the AMP complex of the ATP-binding RNA aptamer[35,36] revealed that the aptamer did not realize the base selectivity by simple Watson−Crick or Hoogsteen base pairing.

In this study, we have isolated ATP-binding RNP from the RRENn RNP library, in which the RNA consensus sequence is different from those previously reported for the ATP-binding RNA aptamers. NMR measurements, enzymatic and chemical mapping, and nucleotide mutation studies of the RNP-adenosine (Ado) complex show that RNP binds to the adenine base of adenosine by forming a U:Ado:U triple. The substrate-binding specificity of ATP-binding RNP was successfully converted to GTP-binding RNP by mutating from a U:A:U triple into a C^+:G:C triple recognition mode.

RESULTS

In Vitro Selection of ATP-Binding RNP Receptors. ATP-binding RNP receptors have been previously isolated from the RREN30 RNP library,[26] in which RNA has a randomized region with a fixed length of 30 nucleotides (Figure S1). Because only a limited number of nucleotides would participate in ATP-binding, these RNPs contain nucleotides that are not directly involved in the formation of the ATP-binding pocket. Therefore, selection of an RRENn RNP library[28] that possessed 7 to 40 nucleotides in the randomized region was carried out to obtain ATP-binding RNP with a minimum of nucleotides for ATP binding. After the selection from the RRENn RNP library by using ATP attached through the ribose 2′- or 3′-hydroxyl groups via a 22 atom spacer to agarose beads, RNAs with the randomized region of various lengths from 17 to 34 nucleotides were found (Figure 1b). Comparison of the nucleotide sequence of each RNA revealed a consensus sequence 5′-GUAGUGG---UGUGUGUG-3′ that was the same as the one found in ATP-binding RNP selected from the RREN30 RNP library[26] (Figure S1). This consensus sequence differs from those reported for the ATP aptamer,[6] cAMP aptamer,[29] adenine aptamer,[38] adenine riboswitch,[39] and S-adenosyl methionine responsive riboswitch.[40] The nucleotides in the nonconsensus region between the 5′-half and the 3′-half of the

dx.doi.org/10.1021/ja110725d | J. Am. Chem. Soc. 2011, 133, 4567–4579

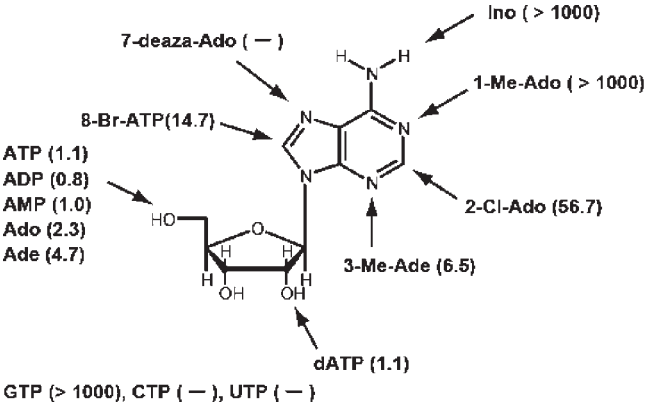


Figure 2. A schematic illustration summarizes the affinity and the selectivity of the ATP-binding An16RNP for ATP analogues. Equilibrium dissociation constants of An16RNP for ATP, ADP, AMP, Ado, adenine, 1-Me-Ado, 2-Cl-Ado, 3-Me-Ade, 7-deaza-Ado, Ino, 8-Br-ATP, dATP, GTP, CTP, and UTP obtained by the fluorescence titration of An16RNP complexed with 7mC-Rev[26] with the indicated ligand (Figure S2) are shown in parentheses (μM).

consensus sequence are different for each RNP. Furthermore, there is a variable region connecting the RRE and the consensus region, which often contains complementary nucleotides expected to form base pairs.

The affinity of RNP selected from the RRENn RNP library for the substrate ATP was analyzed by converting the RNP receptor to the fluorescent RNP sensor, for which 7-methoxy coumarin-labeled Rev (7mC-Rev) was utilized for the RNP complex formation.[26,28] Titration of the fluorescent RNP with an increasing concentration of ATP gave the titration curve for the complex formation of RNP and ATP (Figure S2), from which the equilibrium dissociation constants (K_D) for the complex of ATP and RNP could be deduced. The K_D values for the ATP binding ranged from 0.2 to 15 μM and are listed in Figure 1b. In general, RNP receptors selected from the RRENn RNP library showed a higher affinity to ATP than RNPs selected from the RREN30 RNP library (Figures 1b and S1).

For further structural study, we have focused on An16RNP from the RRENn RNP library because its RNA subunit An16 (Figure 1b) had a relatively small ATP-binding site of 19 nucleotides with the 5′-GUAGUGG---UGUGUGUG-3′ consensus sequence and it showed a relatively high affinity to ATP (K_D = 1.1 μM). Formation of the RNP complex by An16RNA and the Rev peptide was verified by titration of An16RNA with the Rev peptide. A gel mobility shift assay showed that An16RNA (1 μM) formed a 1:1 RNP complex (87%) in the presence of 1.2 equiv of Rev (Figure S3). In addition, An16RNP showed a higher affinity for ATP than An16RNA alone[23] (Figure S4). These characteristics made An16RNP as an ideal model for studying the structure of the ATP-binding RNP receptor.

Functional Evaluation of the ATP-Binding RNP Receptor. An16RNP revealed low affinities for UTP, CTP, or GTP but high affinities for AMP, ADP, dATP, adenosine (Ado), and adenine that were similar to that for ATP (Figures 2 and S2). This indicates that the adenine ring dominantly contributes in the complex formation. As An16RNP barely showed affinity to the adenosine analogues 1-methyl-adenosine (1-Me-Ado), inosine (Ino), and 7-deaza-adenosine (7-deaza-Ado), N1 and N7 atoms and the 6-amino group are essential for An16RNP to bind the adenine ring (Figure 2). The affinity for 3-methyl-adenine

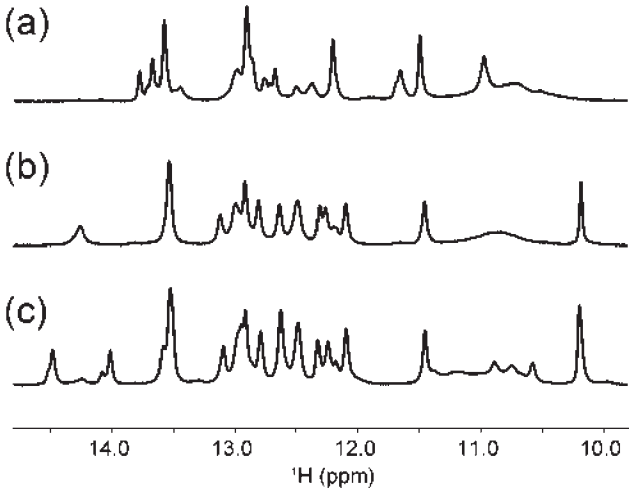


Figure 3. Imino proton spectra of An16RNA in a free state (a), the binary complex state with Rev peptide (b), and the ternary complex state with Rev peptide and Ado (c) at 20 °C.

(3-Me-Ade), on the other hand, is almost similar to that for adenine, indicating that An16RNP does not contact the adenine N3 position. The lower affinity for 2-chloro-adenosine (2-Cl-Ado) is possibly caused by a steric hindrance or an electrostatic repulsion. The affinity of An16RNP for the 8-bromo-ATP (8-Br-ATP) was found to be lower than that for ATP. Because 8-Br-ATP preferentially takes the *syn* conformation when free in solution,[41] the observed lower affinity is possibly caused by the unfavorable energetic cost to change the conformation of the substrate from *syn* to *anti*. Alternatively, An16RNP might bind 8-Br-ATP in its *syn* conformation with reduced complex stability due to the steric hindrance at the ribose moiety. These results indicated that An16RNP recognizes mainly the adenine base moiety and interacts less with the phosphate group and the ribose moiety (Figure 2).

NMR Spectral Change and Stoichiometry on Binary and Ternary Complex Formation. Figure 3a shows the imino proton spectrum of free An16RNA. When the Rev peptide was added step by step to the An16RNA solution, several new peaks appeared and grew. At the same time, several original peaks began to decrease and finally disappeared. This result indicates that a free state of An16RNA and its binary complex state with the Rev peptide are in the slow exchange regime in an NMR time scale. The spectral change of An16RNA ended when an equimolar amount of the Rev peptide was added, giving the spectrum shown in Figure 3b. This reveals that the An16RNA:Rev stoichiometry is 1:1 in An16RNP.

To obtain the ternary complex, Ado was added step by step to the An16RNA:Rev binary complex. Although the affinity of An16RNA:Rev for Ado is slightly lower than that for ATP, the amount of Mg^{2+} required for binding Ado, ca. 3 mM, is less than for ATP, ca. 5 mM (Figure S5). As the high concentration of Mg^{2+} tends to make NMR signals broad, Ado is a more suitable for the NMR experiment and was thus used as a substrate in all the following experiments. Again, the spectral change characteristic to the slow exchange was observed in the course of the addition of Ado. This result indicates that the binary complex and the An16RNA:Rev:Ado ternary complex are in the slow exchange regime. The spectral change ended when an equimolar amount of Ado was added, exhibiting the spectrum shown in Figure 3c. Thus, it was concluded that the An16RNA:Rev:Ado

stoichiometry in the ternary complex is 1:1:1. It is apparent from the comparison of Figure 3b and 3c that the number of imino proton resonances increased upon binding of Ado. This suggests the formation of direct hydrogen bonds between An16RNA and Ado. It is also likely that the binding pocket is formed with the binding of Ado. It is also likely that the binding pocket is formed with Ado and that the access of water to some imino protons of An16RNA is hindered, resulting in the observation of those resonances due to the restriction of the exchange with water.

Structure of the Stem Region of An16RNA in the Ternary Complex. The top and second panels of Figure 4a show, respectively, the intraresidue GH1-GN1 and UH3-UN3 correlation peaks of the HNN-COSY spectrum of the ternary complex.[42,43] From these spectra, imino protons of G and U residues are easily discriminated. With the help of this discrimination, sequential assignments of the imino protons of the ternary complex were carried out with the NOESY spectrum (Figure 4b). As expected from the predicted secondary structure derived from the RRE:Rev peptide study,[37] a very strong G imino-U imino cross peak was observed for the U3:G43 base pair (Figures 4d and 5a). Thus, this cross peak formed a good starting point. The connectivity from the G2 residue to the G7 residue of the lower part of the stem was accomplished as shown in green as described previously for the other oligonucleotides.[44−46] The reach of the connectivity to the G7 residue demonstrates the formation of the noncanonical G7:A39 base pair (Figure 5b). The imino proton of the terminal G1 residue was not observed probably due to partial end fraying. Another characteristic point of Figure 4b is the very strong G imino-G imino cross peak, as expected for the noncanonical G8:G37 base pair of the predicted secondary structure (Figures 4d and 5c). Therefore, this cross peak was used as another good starting point to obtain the connectivity from the G8 residue to the G31 residue in the upper part of the stem shown in purple. As expected, the found secondary structure of the upper and lower stem region of An16RNA in the ternary complex with Rev peptide and Ado is identical to the previously described structure of RRE in complex with the Rev peptide. The stem in both the binary and ternary complex contains noncanonical G:U, G:A, and G:G base pairs, together with canonical G:C and A:U base pairs. The C13:G31 base pair is the end of the stem.

Binding of Ado to An16RNP through the Formation of Intermolecular U:A:U Base Triple in the Ternary Complex. Correlation peaks in the third to fifth panel of Figure 4a originate from the presence of the two bond $^{15}N-^{15}N$ scalar coupling across the hydrogen bonds, $^{2h}J_{NN}$, in base pairs. Observation of seven GH1-CN3 correlation peaks in the third panel (Figure 4a) is in good agreement with the concluded seven G:C base pairs in the stem of the RRE region, excluding the terminal G1:C35 base pair (Figure 4d). The two UH3-AN1 correlation peaks in the fourth panel are also consistent with two A:U base pairs in the stem of the RRE region. Notably, observation of the G7H1-AN1 correlation peak in the fourth panel confirms the noncanonical G7:A39 base pair described above.

The unassigned imino proton of uracil U#1 gives a correlation peak to AN7 (fifth panel, Figure 4a), which indicates the formation of a Hoogsteen (or reversed Hoogsteen) A:U base pair in the loop region. From the five A residues of An16RNA, A12, A39 ,and A41 are involved in either a Watson−Crick A:U base pair or a noncanonical G:A base pair in the stem of the RRE region and A34 is located in the stem of the RRE region and therefore should be far away from the loop. A16 as the only remaining A residue is likely to be involved in the Hoogsteen A:U base pair with U#1. A possible assignment of U#1 will be discussed later. Noncanonical G:U and G:G base pairs in the stem do not result in correlation peaks for these base pairs, because the acceptor of the hydrogen bond is an oxygen atom, not a nitrogen atom. This way, all correlation peaks originating from the $^{15}N-^{15}N$ scalar couplings across the hydrogen bonds have been rationally interpreted.

When Ado in the ternary complex was replaced by ^{13}C, ^{15}N-labeled Ado, two new correlation peaks appeared in the HNN-COSY spectra, as indicated by arrows in Figure 4c. One unassigned imino proton of uracil U#3 gives a correlation peak with AN1, and the other unassigned imino proton of uracil U#2 does so with AN7. It is clear that the AN1 and AN7 resonances belong to Ado, because the correlation peaks were not observed with nonlabeled Ado (the fourth and fifth panel of Figure 4a). Thus, the newly observed correlation peaks in Figure 4c indicate the formation of intermolecular hydrogen bonds between An16RNA and Ado. The U#3H3-AN1 hydrogen bond reveals that Ado is trapped by An16RNA through a Watson−Crick A:U base pair, while the U#2H3-AN7 hydrogen bond reveals that Ado is also trapped through a Hoogsteen A:U base pair. As a whole, Ado is trapped by An16RNA in the ternary complex through the formation of a U:A:U base triple (Figure 5d).

Structure of the Loop Region of An16RNA in the Ternary Complex. The imino protons of the five unassigned G residues are labeled G#1 to G#5. Those of four unassigned U residues are labeled U#1 to U#4. GH1-GN1 and UH3-UN3 intraresidue correlation peaks in the HNN-COSY spectrum are strong for G#1, G#2, G#3, U#1, U#2, and U#3 and weak for G#4, G#5, and U#4 (Figure 4a, top and second panel). In the NOESY spectrum, imino−imino connectivities are found for G#1-U#3-U#2-G#2-U#1 and for U#2-U#4 (Figure 4e). All imino protons giving strong intraresidue correlation peaks are involved in the connectivities, except that of G#3. No connectivity was found for G#3.

Except for A16, the loop is composed of exclusively G and U residues. As described above, A16 is supposed to be involved in a Hoogsteen A:U base pair. Therefore, the observation of imino protons of either G or U residues from the loop implies the formation of G:G, U:U, or G:U base pairs in the loop (see Figure 6-1 of ref 47 for the 28 possible base pairs involving A, G, T(U), and C residues). As will be shown below, however, these base pairs are not found in the loop. When a G:U base pair is formed, a strong G imino-U imino NOE should be observed (Figure 5a, and XXVII and XXVIII of Figure 6-1 of ref 47). In fact, a strong NOE is observed for the G43:U3 base pair of the stem of the RRE region. The absence of this strong NOE for the residues of the loop excludes the formation of G:U base pairs. When a U:U base pair is formed, a strong U imino-U imino NOE should be observed (XII, XIII, and XVI of Figure 6-1 of ref 47). The absence of this strong NOE for U residues of the loop excludes also the formation of U:U base pairs. There are four kinds of G:G base pairs (III, IV, VI, and VII of Figure 6-1 of ref 47). Each G:G base pair has one of the following characteristics: the strong G imino-G imino NOE (III), the strong G amino-G amino NOE (IV), the interresidue GN2-GN7 correlation in HNN-COSY (VI), or the interresidue GN1-GN7 correlation in HNN-COSY (VII). As none of these features is found, no G:G base pairs are formed in the loop. Furthermore, the formation of both a G-tetrad and a G-triple that utilize the G:G base pair as a building block is also excluded for the same reason.[48] Thus, experimental

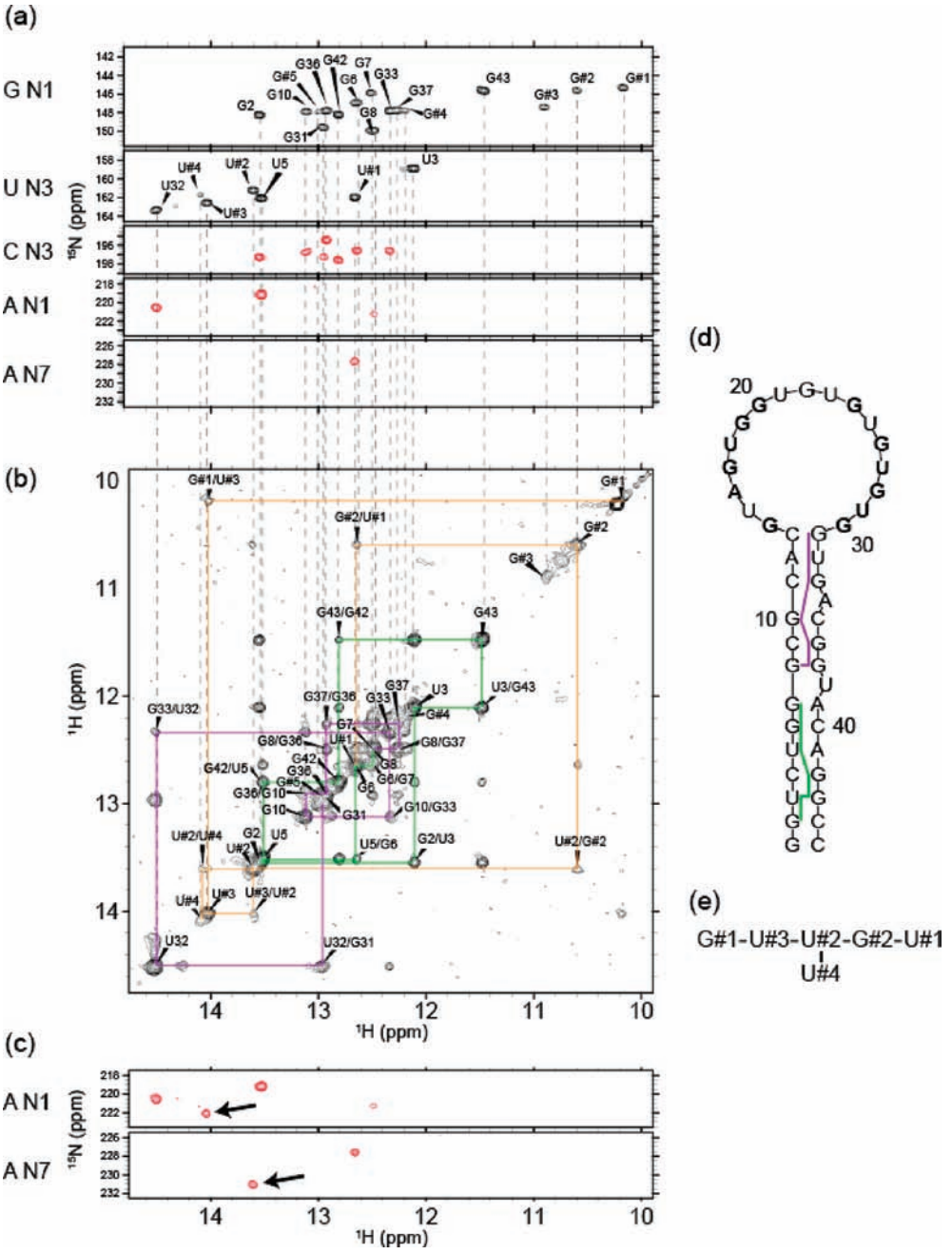


Figure 4. (a) HNN-COSY spectra of the ^{13}C, ^{15}N-labeled An16RNA:nonlabeled Rev:nonlabeled Ado ternary complex at 20 °C. Positive and negative peaks are shown in black and red, respectively. Intraresidue GH1-GN1 and UH3-UN3 correlation peaks are shown in the top and the second panels, respectively, labeled with their residue numbers. Unassigned intraresidue correlation peaks are also labeled with either G#1-G#5 or U#1-U#4. Interresidue GH1-CN3 correlation peaks for G:C base pairs, UH3-AN1 correlation peaks for Watson−Crick A:U base pairs, and UH3-AN7 correlation peaks for Hoogsteen A:U base pairs are shown in the third, fourth, and fifth panels, respectively. (b) The imino proton region of the NOESY spectrum of the ternary complex. The sequential assignments are indicated for the two parts of the stem region, with green and purple lines, respectively. Another connectivity of the cross peaks for the loop is shown with yellow lines. (c) HNN-COSY spectra of the ^{13}C, ^{15}N-labeled An16RNA:nonlabeled Rev:^{13}C, ^{15}N-labeled Ado ternary complex. Interresidue UH3-AN1 correlation peaks for Watson−Crick A:U base pairs and UH3-AN7 correlation peaks for Hoogsteen A:U base pairs are shown in the top and the second panels, respectively. Arrows indicate the peaks appearing on the replacement of nonlabeled Ado by ^{13}C, ^{15}N-labeled Ado. (d) Stem-loop illustration of An16RNA with the identified two NOESY connectivities in green and purple, respectively. (e) The found imino−imino NOESY connectivities for unassigned G and U residues of the loop.

results suggest that in addition to the U:Ado:U base triple, the Hoogsteen A:U base pair, and the C13:G31 base pair at the end of the stem, no other base pair is formed in the loop region.

It was confirmed that neither the imino protons of G#1-G#5 nor those of U#1-U#4 gave an NOE to aliphatic resonances of the Rev peptide (data not shown). The Rev peptide thus does not make close contact with the residues of the loop region. It is not likely that the Rev peptide interacts directly with Ado either, although it is difficult to completely rule out this possibility. In spite of ^{13}C, ^{15}N-labeling of An16RNA, the Rev peptide, and

dx.doi.org/10.1021/ja110725d |*J. Am. Chem. Soc.* 2011, 133, 4567–4579

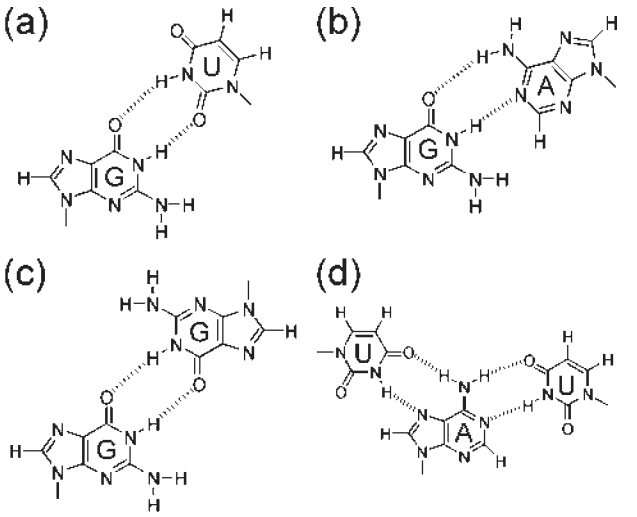


Figure 5. G:U (a), G:A (b), and G:G (c) base pairs, and an intermolecular U:A:U base triple (d) found in the An16RNA:Rev:Ado ternary complex.

Ado, assignment of the resonances could not be achieved due to heavy overlapping, in particular for the sugars. The structural determination of the ternary complex at atomic resolution is therefore not feasible at this moment.

Analyses of the Solution Structure of An16RNP by Enzymatic and Chemical Mapping. We next studied the solution structure of An16RNP by enzymatic digestion (Figure 6a), in-line probing (Figure 6b),[49] and chemical probing with dimethyl sulfate (DMS)[50] (Figure 6c) of An16RNP and its Ado complex. Ado was also utilized here as the substrate for An16RNP as ATP was known to interfere with the reaction of RNases.[51,52] In general, mapping of the ternary structure revealed that the intensity of scission bands decreased with increasing concentrations of Ado. An16RNP formed a folded structure upon binding to Ado that was not easily accessed by the hydrolytic enzymes. These results correlated well with the increase in the number of imino proton resonances found in the NMR measurement of the binding of Ado (Figure 3b and 3c). Within the RRE region, strong cleavage bands were observed at C11, A12, and G42 by RNase V1, which cleaves at the duplex region, as shown in the previously reported results of enzyme mapping for the Rev-RRE complex.[53] Outside the RRE region, a distinct scission band at C13 was observed, indicating that C13 base-paired with G31. The nucleotides from U18 to U25 including the nonconsensus region were efficiently cleaved by Nuclease S1, which cleaves single-stranded nucleotides, and by RNase A and RNase T1, which detect the single-stranded pyrimidine and G nucleotides, respectively (Figure 6a). The appearance of an intense scission band at U15 by RNase A in the absence and presence of Ado indicated that the 5′-site of the consensus sequence near the RRE region did not form the duplex structure even in the An16RNP-Ado complex. The consensus region is not an efficient substrate for RNase V1 or Nuclease S1. The structure at the consensus region is unlikely to be in the duplex form, which paralleled with the finding that none of G:U, G:G, and U:U base pairs was observed by NMR for this region.

In-line probing of the An16RNP-Ado complex was also carried out to estimate the structure of the phosphate backbone. The nucleotides that are mobile to take on the in-line conformation or held in the in-line conformation are prone to be hydrolyzed by the phosphodiester linkage.[49] Upon incubation of An16RNP with an increasing concentration of Ado, a progressive decrease in the intensity of the cleavage band within the regions of G14-G21 and U23-G31 was observed (Figures 6b and S6), which also suggested the structural change of An16RNP from a rather flexible structure to a well-defined folded structure. Interestingly, the in-line probing showed that U21 and G22 in the nonconsensus region were cleaved efficiently in free An16RNP. The cleavage bands of G22 became more intense upon formation of the An16RNP-Ado complex, and that of U21 also showed a relatively strong band in the binding state. U21 and G22 would take the in-line structure that was stabilized by the substrate binding, as has been demonstrated in the ATP aptamer constructed by Sassanfar et al.,[6] in which the backbone curvature forced the nucleotide to take an "ideal" in-line structural site.[49] If this is the case, An16RNP in the Ado bound form would take a curved backbone structure at U21 and G22 of the nonconsensus region. Additionally, A34 and U38 in the RRE region taking the bulge structure in the Rev-RRE complex[37] consistently showed the intense scission bands both in the free and in the substrate-bound forms of An16RNP in the in-line probing (Figure 6b).

In the chemical probing with DMS (Figure 6c), the band intensities of G14, G22, and G24 were strong even in the presence of the substrate Ado, indicating that the N7 positions of G14, G22, and G24 were exposed to the solvent regardless of the formation of the An16RNP-Ado complex. G20 was less exposed to the solvent than G14, G22, and G24. The intensities of the bands at G17, G19, G26, G28, and G30 were quite weak in the presence of Ado. This shows that G17, G19, G26, G28, and G30 in the consensus region were masked from the solvent and were likely to form the structural core of the An16RNP-Ado complex. These structural characteristics of the An16RNP-Ado complex were also found by enzymatic hydrolysis mapping.

Identification of the U Nucleotides that Directly Bind to the Substrate. There are three invariant U nucleotides, U18, U27, and U29, and three less preserved ones, U15, U23, and U25, in the consensus sequence (Figure 1b). In order to identify the U nucleotides that participate in the U:Ado:U triple formation, mutants of An16RNP at U18, U27 or U29 were evaluated for their binding to Ado. Each An16RNP mutant was complexed with 7mC-Rev to form a fluorescent RNP to easily monitor the Ado binding affinity.[26] The RNP mutated at U18 (Figure S7b) or U27 (Figure S7c) showed no fluorescence response to the increasing concentration of Ado, suggesting that these mutants completely lost either the substrate binding ability or the functionality to change fluorescence intensity upon substrate binding. Therefore the binding affinity to the ATP-immobilized resin was tested using 5′-^{32}P-labeled An16RNP and mutants. The ATP-binding assay of the 5′-^{32}P-labeled mutants of RNP at U18 or U27 indicated a complete loss of the ATP-binding ability (Figure S8). Titration of the fluorescent U29 mutant showed Ado-binding, but with significantly decreased affinity (Figure S7d) compared to the original An16RNP (Figure S7a). In contrast, mutation at U15, U23, or U25 barely affected the fluorescence responses of the mutants to Ado (Figure S9). These results indicate that the three invariant U nucleotides U18, U27, and U29 are essential for substrate binding. Especially U18 and U27 are likely to be involved in the formation of the U:Ado:U triple, because their mutation in An16RNP completely abolished the ATP-binding ability. In addition, the invariant U18 and U27 locate in the region where the intensities of the cleavage bands

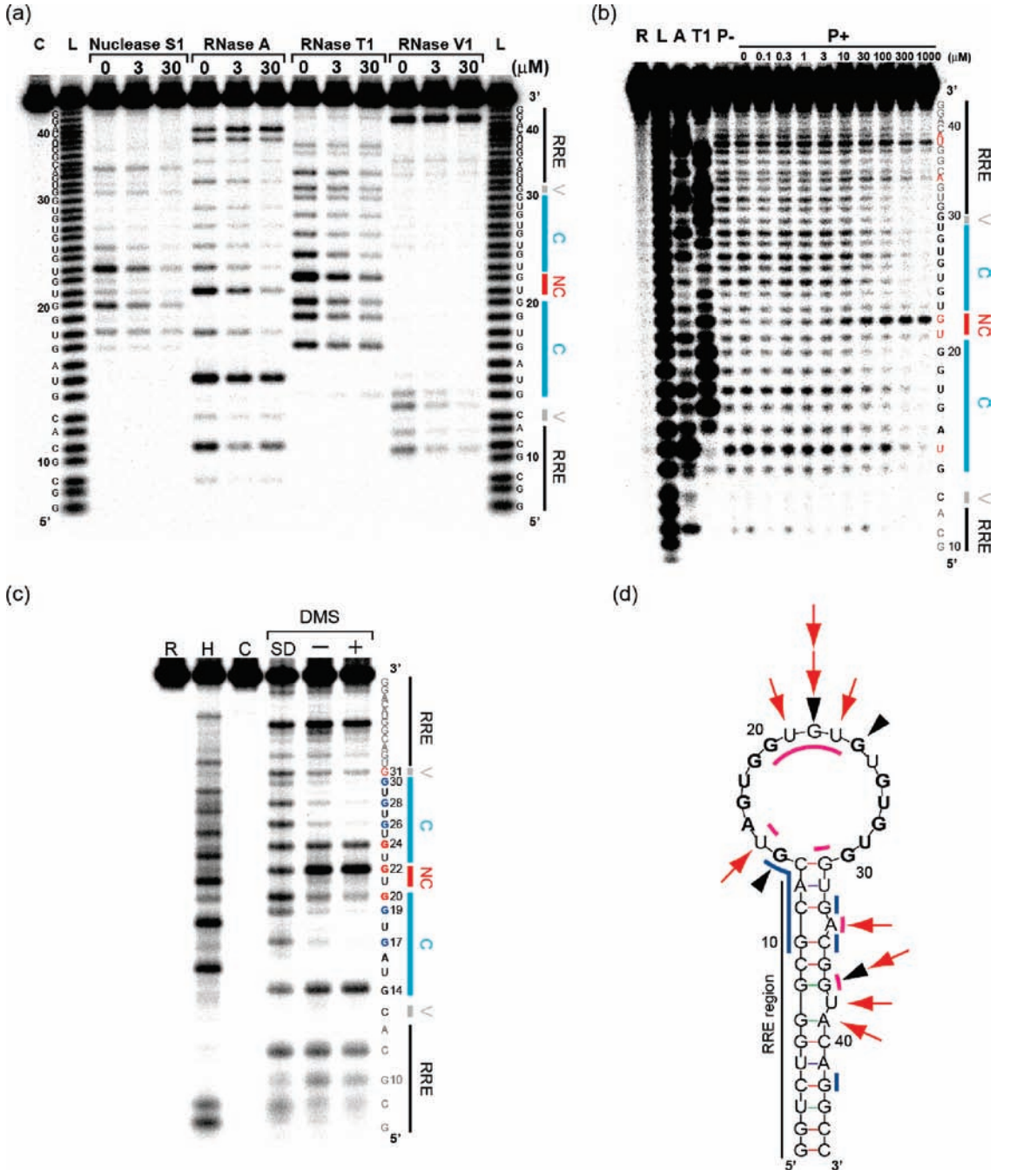


Figure 6. (a) An autoradiogram shows the structural mapping of the An16RNP-Ado complex by using hydrolytic enzymes in a buffer containing 10 mM Tris-HCl (pH 7.6), 100 mM NaCl, 10 mM $MgCl_2$, and 0.005% Tween 20 at 4 °C. C: intact RNA, L: OH^- marker. The enzyme and the concentration of Ado used for the hydrolytic reaction are shown above each lane. The consensus region (C) is indicated by blue lines, the nonconsensus region (NC) by a red line, and the variable region (V) by gray lines. (b) An autoradiogram shows the in-line probing of the An16RNP-Ado complex incubated for 100 h in a buffer containing 50 mM sodium-cacodylate (pH 7.0), 100 mM NaCl, and 10 mM $MgCl_2$ at 25 °C. The nucleotides that marked strong band intensities in the presence of Ado are shown in red. R: intact RNA. L: RNA hydrolyzed by heating with Mg^{2+}. A: pyrimidine marker digested by RNase A. T1: G marker digested by RNase T1. P—: RNA only. P+: RNP. The concentration of Ado (μM) is indicated above each lane. (c) An autoradiogram shows the DMS chemical probing of the An16RNP-Ado complex in a buffer containing 50 mM sodium-cacodylate (pH 7.0), 100 mM NaCl, and 10 mM $MgCl_2$ at 4 °C. The nucleotides that revealed high and low solvent accessibility are shown in red and blue, respectively. R: intact RNA. H: pyrimidine marker treated by hydrazine. C: RNA without the DMS reaction, but with $NaBH_4$ and aniline treatments. SD: RNP under semidenature condition (without Mg^{2+}, in the presence of 10 mM EDTA), (—): RNP in the absence of Ado, (+): RNP in the presence of 300 μM Ado. (d) A schematic illustration summarizes the characteristic RNA structures deduced from the results of the secondary structure analyses of the An16RNP-Ado complex. Invariant nucleotides in the consensus region are indicated in bold. Nucleotides that revealed a single-strand-like digestion pattern in the enzymatic digestion are indicated by a magenta line, and those that showed a double-strand-like pattern by a blue line. Nucleotides that show strong band intensities in the Ado-bound state by in-line probing are indicated with red arrows. G residues that show strong band intensities in the DMS chemical probing in the Ado-bound state are marked with filled triangles.

dx.doi.org/10.1021/ja110725d |*J. Am. Chem. Soc.* 2011, 133, 4567–4579

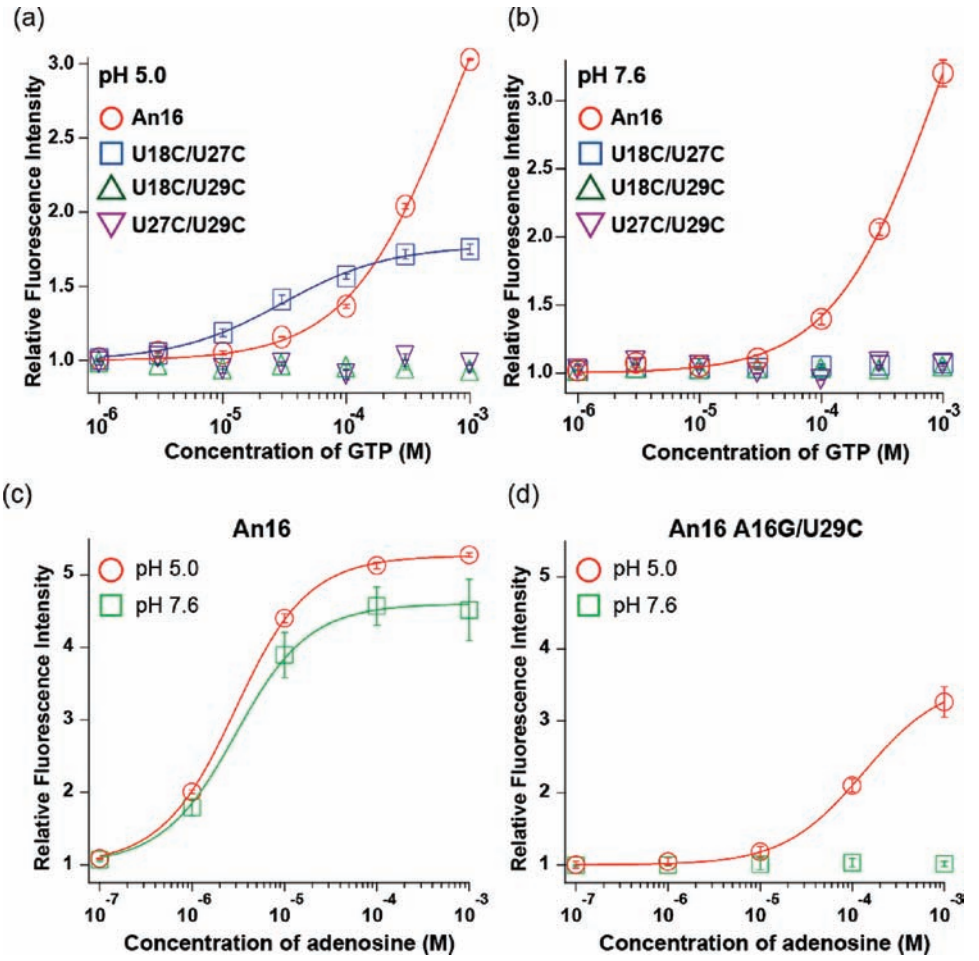


Figure 7. Titration curves for the relative fluorescence intensity changes of 7mC-Rev complexed An16RNP (red circles), U18C/U27C mutant (blue squares), U18C/U29C (green triangles) and U27C/U29C (purple reversed triangles) with increasing concentrations of GTP (a and b) in a buffer containing 10 mM Tris-acetate (pH 5.0) (a) or 10 mM Tris-HCl (pH 7.6) (b) with 100 mM NaCl, 10 mM $MgCl_2$, and 0.005% Tween 20 at 4 °C. Titration curves of relative fluorescence intensity changes for the resulting fluorescent RNP by Ado in neutral and acidic conditions are shown for An16 (c) and A16G/U29C (d). The neutral buffer (green squares) contained 10 mM Tris-HCl (pH 7.6), 100 mM NaCl, 10 mM $MgCl_2$, and 0.005% Tween 20, and the acidic buffer (red circles) contained 10 mM Tris-acetate (pH 5.0), 100 mM NaCl, 10 mM $MgCl_2$, and 0.005% Tween 20.

decreased upon addition of Ado both in the DMS chemical reaction and in the in-line probing.

U18 and U27 were mutated to a pair of C nucleotides (U18C/U27C) to test whether U18 and U27 actually participated in the formation of the U:Ado:U triple. In the triplex structure, the overall shape and the dimension of the U:A:U are comparable to the C^+:G:C triple. If the C mutations were introduced to the appropriate combination of U nucleotides, the mutant RNP is expected to bind to GTP through a C^+:G:C triple formation under the acidic conditions that are necessary for the protonation of C to stabilize the C^+:G:C triple. An appropriate set of mutations from U to C would convert ATP-binding An16RNP thus into GTP-binding RNP. In addition, the other possible combinations within the three invariant U nucleotides, C mutation of U18 and U29 (U18C/U29C) and C mutation of U27 and U29 (U27C/U29C), were also investigated.

The affinities of An16RNP and the three mutants for GTP and Ado were evaluated by fluorescence measurements. The mutants U18C/U27C, U27C/U29C, and U18C/U29C showed no fluorescence response to Ado in both acidic (Figure S10a) and neutral conditions (Figure S10b). The mutants U18C/U29C and U27C/U29C showed no detectable change in fluorescence intensity upon the addition of GTP, indicating that these mutants do not bind to GTP under acidic conditions and are thus not the correct combination. In contrast, the U18C/U27C mutant revealed distinct fluorescence intensity changes with increasing the concentration of GTP with a K_D value of 30 μM (Figure 7a). The fluorescence intensity change of the U18C/U27C mutant RNP upon the addition of GTP was not observed under neutral conditions (Figure 7b). The affinity of An16RNP for Ado in the acidic conditions (pH 5.0) (K_D = 2.4 μM) was similar to that at pH 7.6 (Figures 7c). Also the stability of An16RNP at pH 5.0 was similar to that at pH 7.6 (Figure S11). In acidic and neutral conditions, An16RNP also showed changes in the fluorescence intensity upon addition of GTP but bound GTP with a quite low affinity (K_D > 1000 μM) (Figure 7a and 7b, red circles). These fluorescence titration data show that under acidic conditions the U18C/U27C mutant bound GTP as the only mutant and with a much higher affinity than An16RNP. The results suggest that the U nucleotides participating in the U:Ado:U triple are U18 and U27.

Identification of the U Nucleotide that Forms the Hoogsteen A:U Base Pair. The consensus sequence contains only one invariant A nucleotide A16, suggesting that A16 participates in

dx.doi.org/10.1021/ja110725d |*J. Am. Chem. Soc.* 2011, 133, 4567–4579

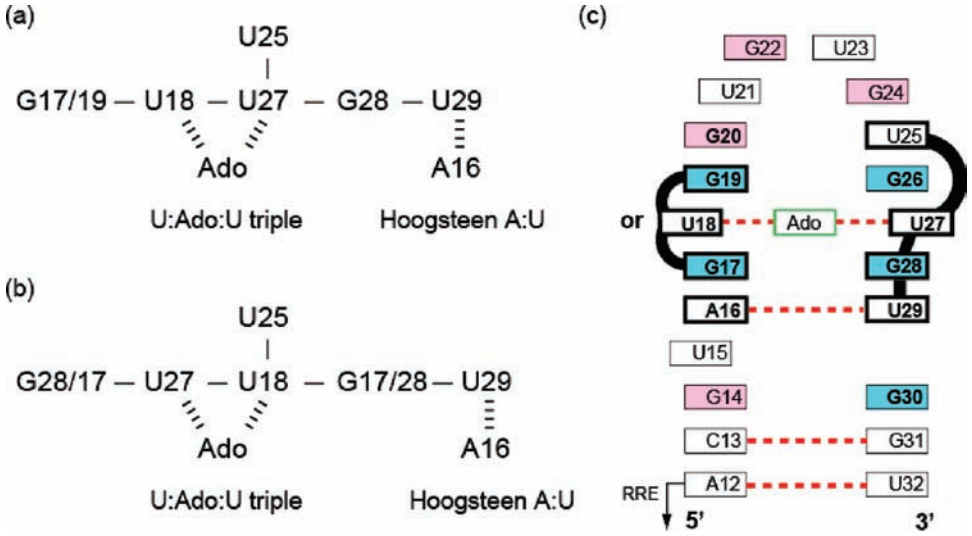


Figure 8. The information derived from the NMR method for the Ado-binding pocket. The imino−imino NOESY connectivities (solid lines) together with the U:Ado:U triple and the intramolecular Hoogsteen A:U base pair are shown (dotted lines). The possible assignments if U#2/U#3 is U27/U18 (a). The possible assignments if U#2/U#3 is U18/U27 (b). (c) A schematic illustration represents the interactions within the An16RNP-Ado complex. The nucleotides in the consensus sequence are indicated in bold face. The nucleotides that reveal high and low solvent accessibility are boxed in red and blue, respectively. The nucleotides assigned as in (a) are boxed with thick lines. Thick black lines represent the observed imino−imino NOE between the nucleotides, and the red dotted lines mark the hydrogen bond deduced from the NMR experiments.

the formation of a Hoogsteen A:U base pair. The single mutation of the remaining invariant nucleotide U29 to C29 greatly reduced the Ado binding activity of the resulting mutant RNP (Figure S7d); thus U29 could be the U involved in the Hoogsteen A:U base pair. A16 and U29 were mutated to G16 and C29, respectively, to obtain the mutant A16G/U29C. If A16 and U29 actually formed the Hoogsteen base pair, a mutation that replaces the A:U base pair with a G:C base pair would give under acidic conditions a parent An16RNP-like structure with a Hoogsteen G:C^+ base pair. The mutant RNP A16G/U29C displayed changes in fluorescence intensity upon addition of increasing concentrations of Ado at pH 5.0 (K_D = 131 μM), while no change in fluorescence intensity was observed upon addition of Ado at pH 7.6 (Figure 7d). That indicated an Ado-binding ability of the mutant RNP A16G/U29C only in acidic conditions due to G:C^+ base pair formation, which proves that U29 forms the intramolecular Hoogsteen A:U base pair with A16. Other mutant RNPs, such as A16G/U27C and A16G/U18C, showed no fluorescence response upon addition of Ado at pH 5.0 or 7.6 (Figure S12).

DISCUSSION

Binding Mode of ATP-Binding RNP Receptor. The RRENn RNP library that consists of the Rev peptide and RNA with an Rev peptide binding RRE region and various lengths of an additional randomized region was used for in vitro selection in this study with ATP immobilized at the 2′- or 3′-hydroxyl group to the resin. The obtained ATP-binding RNP receptors mostly utilized 17 to 23 nucleotides to construct the ATP-binding pocket and had the same consensus sequence as those obtained previously from the RREN30 RNP library with a randomized region of fixed 30 nucleotides (Figure S1).[26]

The results of NMR measurements revealed that the ATP-binding RNP composed of An16RNA and the Rev peptide binds to its substrate Ado by the formation of a U:Ado:U triple. This substrate recognition mechanism of An16RNP, where N1, N7 and the 6-amino group, but not N3, of the adenine ring are involved in the hydrogen bonding interactions, was also supported by measuring the binding affinities for ATP analogues (Figure 2). An16RNP has little specificity for the triphosphate and the ribose moieties. Modifications to the nucleobase moiety at N1, N7 or the 6-amino group almost abolished the substrate binding of An16RNP, while those at N3 or C8 showed little effect for the binding. In contrast to An16RNP, all the ATP aptamers[34−36] and the GTP aptamer[48] characterized by three-dimensional structural analyses have been shown to utilize the non-Watson−Crick type of hydrogen bonding for substrate binding. The ATP aptamer selected by Sassanfar et al.[6] mainly recognizes the N1, N3 and the 6-amino group by the direct interaction of a guanine base of the aptamer with the adenine base of the substrate ATP.[34−36] The cyclic-AMP aptamer selected by Koizumi et al.[29] interacts with N1, N7 and the 6-amino group of the adenine ring, which are also the determinants for the ATP complex formation of An16RNP. Additionally, the cAMP aptamer is sensitive to the conformation of the nucleotide and prefers to bind cAMP in the *anti* conformation of the adenine ring. Interestingly, naturally occurring RNA binding pockets utilize Watson−Crick or the Hoogsteen type of base pairing to recognize the adenine ring. In the SAM riboswitch-SAM complex, the adenine ring of SAM formed a Hoogsteen U:A base pair with U and hydrogen bonds with N3 and the 2′-OH group of A of the SAM riboswitch.[40] The adenine-binding riboswitch interacts with four U nucleotides by forming a Watson−Crick A:U base pair with the first U and hydrogen bonds between N3 and N9 and the second U, N9 and the third U, and N7 and the 2′-OH group of the fourth U nucleotide.[39]

The U:A:U triple interaction found in triplex RNA has not been reported so far as the binding mode of the adenine ring in an RNA aptamer. The observed difference in recognition mode of the adenine ring depends probably on the chemical nature of the ATP analogues utilized for in vitro selection. The ATP-binding

dx.doi.org/10.1021/ja110725d |*J. Am. Chem. Soc.* 2011, 133, 4567–4579

RNP receptor studied here has been selected with the target ATP immobilized with the 2′- or 3′-hydroxyl group onto the resin. This displays the adenine ring of ATP to the solvent and thus allows the RNA binding pocket to access it for both Watson−Crick and Hoogsteen interactions. The ATP aptamer[6] selected by using an ATP-immobilized resin via the C8 position of adenine mainly recognizes the Watson−Crick face possibly due to the steric constraints. Actually, RNP receptors sharing the same consensus sequence as found in the Sassanfar ATP aptamer were isolated from the in vitro selection from the RREN30 RNP library by using the ATP-immobilized resin linked via the C8 position (Figure S13). These observations clearly indicate that the linkage of the substrate to the resin greatly influences the substrate interaction mode and the structure of the resulting binding pocket in RNA. Therefore, the chemistry for the immobilization of substrate to the resin is an important parameter to control the selectivity and the affinity of aptamers. The immobilization of nucleotide derivatives at the 2′- or 3′-hydroxyl group to the resin would be an advantageous way to select aptamers that are sensitive to the recognition of nucleic acid bases.

From the secondary structural analyses of ATP-binding RNP by mfold,[54,55] several ATP-binding RNP with a relatively long stem in the variable region that connected the RRE region and the consensus region were observed (Figure 1b). Such RNP with a seemingly stable structure did not always show higher affinity to ATP than An16RNP. In the temperature range studied here, the Rev-RRE motif likely offers enough stability for the ATP-RNP complex. Although no direct contact of the Rev peptide to the substrate is observed in the NMR measurement of the An16RNP-Ado complex, the affinity of An16RNP to Ado is higher than that of An16RNA alone to Ado (Figure S4), indicating that the Rev peptide stabilizes the structure of the RNP-Ado complex. The RRE/Rev region in ATP-binding RNP contributes to the increase in the stability of the substrate-RNP complex by acting as a structural anchor for the ligand-binding site and to defining the boundary of the substrate-binding region. In this respect, the RNP library containing the RRE RNA-Rev complex[23,26,28] is conceptionally similar to the partially structured RNA libraries.[32]

Insight of the Ado-Binding Pocket Derived from the Integration of the Information with NMR and Biochemical Methods. The NMR method revealed that Ado is trapped by An16RNP through the formation of the U:Ado:U triple and that the unassigned U#3 of An16RNP makes a Watson−Crick A:U base pair with Ado while U#2 makes a Hoogsteen A:U base pair. The biochemical method strongly suggested that U18 and U27 are involved in the U:Ado:U triple. The NMR method also revealed that An16RNP forms an intramolecular Hoogsteen A:U base pair and that A16 is most likely to be involved in this base pair with the unassigned U#1. The mutation study supported that A16 and U29 form this base pair. The NMR method gave the imino−imino NOESY connectivities for the Ado-binding pocket of An16RNP (Figure 4e). Now, U#1 is assigned to U29, and U#2/U#3 is assigned to either U18/U27 or U27/U18. Figure 8a shows the possible assignments of the connectivities if U#2/U#3 is U27/U18, while Figure 8b does if U#2/U#3 is U18/U27. In the case of Figure 8a, the assignments are given to make the connectivities sequential for G17/G19-U18 and U27-G28-U29 segments. A similar consideration is also applied for the construction of the assignments for Figure 8b, although it is not feasible to make the connectivities completely sequential. The assignments shown in Figure 8a are most probable, although the other possibilities of the assignments cannot be completely excluded.

Finally, U21, U23, or U25 can be assigned to U#4 by considering the nucleotide sequence of An16RNA. The in-line probing experiments reveal that the cleavage intensity at U25 decreases on addition of Ado, while those at U21 and U23 are not significantly changed (Figures 6b and S6). These results imply that U25 undergoes a larger conformational change than U21 and U23 upon An16RNP-Ado complex formation. The U#4 resonance appeared only in the An16RNP-Ado complex (Figure 4a). Therefore, it is likely that U#4 is assignable to U25, as shown in Figure 8a and 8b.

A Possible Structural Model for the Complex of ATP and An16RNP. The overall structure of the An16RNP-ATP complex is considered based on combining the information obtained from the structural analysis of the An16RNP-Ado complex with the NMR measurements, the chemical and the enzymatic mapping of the An16RNP-Ado complex, and the structure−activity relationships evaluated by using the mutants of An16RNP (Figure 8c). The NMR measurement confirms that the structure of the RRE region is similar to the previously reported Rev-RRE complex[37] and that C13 and G31 in the variable region form a base pair. There is no G:G, G:U, or U:U base pair in the nonconsensus and the consensus regions, while a U:Ado:U triple and Hoogsteen base pair are formed as U18:Ado:U27 and A16:U29, respectively. The imino−imino NOESY connectivities considered above are also applied for the plausible structural model for the An16RNP-Ado complex.

It is natural to consider that the N7 positions of G nucleotides with the low solvent accessibility (G17, G19, G26, G28, and G30) are buried in the folded An16RNP-Ado complex. The connectivity of nucleotides (Figure 8a) indicates that G17 (or G19) and G28 likely stack to U18 and U27, respectively, of the U18:Ado:U27 triple in the An16RNP-Ado complex. Both the in-line probing and the DMS chemical probing experiments reveal that G22 shows enhanced cleavage upon formation of the An16RNP-Ado complex. The results indicate that G22 is solvent-exposed and takes the in-line structure in the substrate-bound structure. In the ATP aptamer,[6] nucleotides with such characteristics have been identified to take the ideal in-line structure in the conformation, which causes backbone curvature.[34−36,49] It is thus possible that G22 locates at the curved backbone in the ideal in-line structure. Moderate cleavage intensities are observed at U21 and U23 in the in-line probing of the An16RNP-Ado complex. The cleavage intensities at U21 and U23 by RNase A digestions are weakened upon formation of the Ado complex. These results indicate that the backbone at U21 and U23 could form tight, in-line-like conformations in the An16RNP-Ado complex.

In the possible structure of the ATP-An16RNP complex (Figure 8c), the G nucleotides with low solvent accessibility at the N7 position, such as G17 (or G19) and G28, sandwich the U18:Ado:U27 triple, which is consistent with the connectivity deduced from the NOE observation. The NOE observation also indicates an interaction of U27 in the triple with U25. G28 not only stacks to U27 but also locates in close proximity to U29 that forms the Hoogsteen base pair with A16. U21 and G22 in the nonconsensus region are located in the curved backbone in the in-line structure, and G22 and G24 are exposed to solvent. G14 is identified to exist in an asymmetric loop without forming a base pair. This asymmetric loop formation is important for the

stabilization of the substrate-bound complex of An16RNP because a mutation of G14 to A14 considerably reduces the ATP-binding affinity of RNP (K_D > 1000 μM).

The results of NMR experiments indicate no base pair formation except the Hoogsteen A:U base pair and no NOEs between the successive nucleotides within G19 to G26. The ATP-binding pocket of An16RNP thus folds into a novel structure that contains no normal base pair but a Hoogsteen-type A:U base pair and the U:A:U triple. The An16RNP-ATP complex probably contains a network of interactions between nucleotides and Mg^{2+} ions because the ATP-binding ability of An16RNP disappears in the absence of the Mg^{2+} ion (Figure S5).

Conversion of the Substrate-Binding Specificity of ATP-Binding RNP to GTP-Binding RNP. The U18C/U27C mutant of An16RNP reveals fluorescence intensity changes with increasing the concentration of GTP at pH 5.0 (Figure 7a) with a dissociation constant of 30 μM, but not under neutral conditions (Figure 7b). The results are well explained by considering the formation of the C^+:G:C triple at C18 and C27 in the complex of the U18C/U27C mutant with GTP, which is stabilized under acidic conditions, instead of the U:A:U triple in the An16RNP-Ado complex. These data clearly indicate that a structure-based conversion of the ligand selectivity of An16RNP is possible owing to its rather simple ligand-binding mechanism. As has been demonstrated for the structure-based conversion of adenine-riboswitch to azacytosine-riboswitch,[56] the ligand-binding mechanism utilizing the Watson−Crick and/or the Hoogsteen base pairing offers an opportunity to explore the structure-based design of parent receptors.

CONCLUSIONS

The structural information on the ATP-binding receptor constructed by using an RNP complex as scaffold was obtained for the first time from the combination of NMR measurements, enzymatic and chemical mapping analyses, and nucleotide mutation studies of the An16RNP-Ado complex. The Rev peptide binds to An16RNA to form a 1:1 RNP complex as expected and increases the ATP-binding affinity of RNP through the stabilization of the ternary complex. An16RNP and the substrate Ado formed a 1:1 complex by trapping Ado, and possibly ATP, in the U:Ado:U triple with two invariant U nucleotides. The RNP structure was maintained by a few base pairs. The recognition mechanism of Ado by the U:Ado:U triple formation differed from those of RNA aptamers targeting ATP analogues reported previously. The ATP recognition mechanism by U:Ado:U triple formation is possibly caused by the way substrate ATP was displayed on the resin used for the selection of ATP-binding RNP receptors. The attaching position of the substrate to the resin would control the substrate-binding mode of RNP, thereby tuning the affinity and the selectivity of RNP receptors obtained from the in vitro selection method. The ATP recognition mechanism of An16RNP disclosed here was successfully utilized to rationally convert the ATP-binding specificity of An16RNP to a GTP-binding RNP with the C^+:G:C triple recognition mode.

MATERIALS AND METHODS

Materials. ATP-agarose (immobilized on cross-linked 4% beaded agarose) and nucleotides (ATP, UTP, CTP, GTP, inosine, 1-methyl-adenosine, 2-chloro-adenosine, 3-methyl-adenine, 7-deaza-adenosine, and 8-bromo-ATP) were purchased from Sigma-Aldrich. *N*-Fmoc-protected amino acids, HBTU (2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate), 1-hydroxybenzotriazole (HOBt), DIEA (*N*,*N*-diisopropylethylamine), TFA (trifluoroacetic acid), and distilled DMF (*N*,*N*-dimethylformamide) were obtained from Watanabe Chemical Industry. Fmoc-PAL-PEG resin (0.38 mmol/g) was purchased from Applied Biosystems. Nuclease S1, RNase A, RNase T1, and RNase V1 were purchased from Ambion. Gel electrophoresis grade acrylamide and bisacrylamide were obtained from Wako Chemicals. All other chemicals were reagent grade and used without further purification. The Rev peptide modified with 7-methoxycoumarin-3-carboxylic acid (7mC-Rev) was synthesized as described previously.[26]

Nucleic Acids Preparations. The original double-stranded DNA pools were constructed by Klenow polymerase (New England Biolabs) reaction from a synthesized oligonucleotide containing 40 random nucleotides [5′-GGAATAGGTCTGGGCGCA-N40-P(OH)-3′] and a synthesized oligonucleotide containing 7 random nucleotides (5′-GGAATAGGCCTGTACCGTCA-N7-OH-3′) followed by PCR amplification to add the promoter for T7 RNA polymerase using Pyrobest DNA polymerase (TaKaRa) with 3′-DNA (5′-GGAATAGGCCTGTACCGTC-3′) and 5′-DNA primer (5′-TCTAATACGACTCACTATAGGAATAGGTCTGGGCGCA-3′: T7 RNA promoter is underlined). RNA transcription was performed using an AmpliScribeT7 kit (Epicenter) for 3 h at 37 °C, according to the supplier's recommended protocols. The resulting RNA was phenol/chloroform extracted, precipitated with ethanol, and pelleted by centrifugation. The RNA was purified by denaturing polyacrylamide gel electrophoresis and eluted. Concentrations of RNAs were determined by UV spectroscopy.

Preparation of Ribonucleopeptide Receptor. RNPs that bound ATP were selected as follows: RNA was heated at 80 °C for 3 min and cooled to room temperature for 2 h for a proper secondary structure. A binding buffer (100 μL, 10 mM Tris-HCl (pH 7.6), 100 mM NaCl, 10 mM $MgCl_2$) containing 1.0 μM RNA, 1.5 μM Rev, and a 50 μL volume of ligand-conjugated agarose was incubated to allow formation of a specific ribonucleopeptide complex for 30 min on ice. RNA-peptide-resin complexes were washed three times with 300 μL (6-volume of resin) of binding buffer to remove unbound RNA-peptide complexes and eluted three times with 150 μL (3-volume of resin) of binding buffer containing 4 mM ATP. Recovered ribonucleopeptide complexes were precipitated with ethanol and resuspended in TE buffer. After reverse transcription with AMV reverse transcriptase (Promega) of the selected RNA using the 3′-DNA primer used in PCR amplification and a successive PCR amplification (RT-PCR) using the 5′- and 3′-DNA primer, DNA templates were transcribed, and the resulting RNAs were subjected to the next round of selection. Selected RNA pools were converted to DNA and PCR-amplified to introduce *Bam*HI, *Eco*RI restriction sites by using primers 5′-GCGGGATCCTTTCGGCCTGTACCGTCA-3′ and 5′-CGGAATTCTAATACGACTCACTATAGG-3′. After enzymatic digestion (New England Biolabs), DNAs were cloned into the pUC19 vector using Ligation Kit Version 2 (TaKaRa) and sequenced using a BigDyeTerminator Cycle Sequencing Kit (Applied Biosystems) with a model 377 DNA sequencer and 3130/3130 × l genetic analyzer (Applied Biosystems).

Fluorescence Measurements on the Microplate. The 96-well fluorescence measurements were performed on a Wallac ARVOsx 1420 multilabel counter. A binding solution (100 μL) containing 1 μM of fluorescent RNP in 10 mM Tris-HCl (pH 7.6), 100 mM NaCl, and 10 mM $MgCl_2$ with an indicated concentration of ligand was gently swirled for a few minutes and allowed to sit for 30 min at 20 °C. Emission spectra were measured with an appropriate filter set for each fluorophore. Excitation and emission wavelengths for 7mC-Rev were 355 and 390 nm, respectively. Images of the fluorescence intensity of wells were obtained by using the Wallac 1420 software version 2.00. Color-coded

images of fluorescent intensity of each well were obtained by using the Wallac 1420 software version 2.00.

Determination of Binding Affinity by Fluorescence Titrations. The binding affinities of fluorescent RNP to ATP, GTP, and other adenosine analogues were obtained by fitting the substrate titration data using the following equation:

$$F_{obs} = A((([FRNP]_T + [substrate]_T + K_D)\text{-}((([FRNP]_T + [substrate]_T + K_D)^2\text{-}4[FRNP]_T[substrate]_T)^{1/2})/2[FRNP]_T$$

where A is the increase in fluorescence at saturating substrate concentrations ($F_{max} - F_{min}$), K_D is the equilibrium dissociation constant, and $[FRNP]_T$ and $[substrate]_T$ are the total concentrations of fluorescent RNP and the substrate, respectively.

NMR Spectroscopy. Nonlabeled and ^{13}C, ^{15}N-labeled An16 aptamers were synthesized *in vitro* by transcription as described previously.[44] Nonlabeled Rev peptides were chemically synthesized as described previously.[23−28] ^{13}C, ^{15}N-labeled Rev peptides were synthesized with *E. coli.* as described previously.[37,13]C, ^{15}N-labeled adenosine was purchased from Sigma. The lyophilized An16 was dissolved in 10 mM sodium phosphate buffer (pH 6.5) containing 10 mM NaCl, 3 mM $MgCl_2$, and 3 mM NaN_3. The concentration of An 16 was 0.1−0.6 mM. The An16:Rev peptide binary complex was prepared by the addition of a concentrated Rev peptide solution to An16 step by step. Then, either nonlabeled or labeled Ado was added step by step to prepare the An16:Rev:Ado ternary complex.

NMR spectra were recorded at 5−30 °C with Bruker DRX600 and DRX800 spectrometers equipped with a cryoprobe with a Z-gradient. DSS was used as an internal chemical shift reference. Spectra were processed and analyzed with XWIN-NMR/TopSpin (Bruker), NMRPipe,[57] Capp/Pipp/Stapp,[58] and Sparky.[59] The HNN-COSY spectra were shifted by $^1J_{HN}/2$ along the ^1H axis for adjustment with the NOESY spectrum.

Enzymatic Digestion Procedures. A solution containing 5′-^{32}P-labeled An16RNA (100000 cpm) and nonlabeled An16RNA (1.0 μM) was heated at 80 °C for 3 min and then quickly chilled on ice. A sample of RNA, Rev peptide (1.2 μM), and Ado (3 or 30 μM) were incubated in a buffer containing 10 mM Tris-HCl (pH 7.6), 100 mM NaCl, 10 mM $MgCl_2$, and 0.005% Tween 20 at 4 °C for 30 min. RNase was added to the RNP solution and incubated at 4 °C for 30 min (RNase A, T1, and V1) or for 1 h (Nuclease S1). Samples were extracted with phenol/chloroform, precipitated with ethanol, and then pelleted by centrifugation. The pellet was suspended in 80% formamide loading dye, run on a 12% denaturing polyacrylamide/8 M urea gel, and analyzed by a STORM phosphor imager (GE).

DMS Reaction Procedures. Chemical probing of the RNP structure was carried out as described.[50] A solution containing 5′-^{32}P-labeled An16RNA (∼100 000 cpm) and nonlabeled An16RNA (1.0 μM) was heated at 80 °C for 3 min and then chilled on ice in a buffer containing 50 mM sodium cacodylate (pH 7.0), 100 mM NaCl, and 10 mM $MgCl_2$. A solution (50 μL) of RNA, 1.2 μM Ac-Rev peptide, and 30 μM Ado was incubated at 4 °C for 30 min, and then 5 μL of a dimethyl sulfate solution (diluted to 1/10) were added and allowed to react at 4 °C for 50 min. The reaction was quenched by addition of a solution (12.5 μL) containing 1 M Tris acetate (pH 7.5), 1 M 2-mercaptoethanol, and 1.5 M sodium acetate and then precipitated by addition of 125 μL of cold ethanol. The RNA pellet was precipitated again from 50 μL of 0.3 M sodium acetate and 125 μL of ethanol and was rinsed with 70% ethanol and then lyophilized. The RNA dissolved in 100 μL of 10 mM Tris-HCl (pH 8.0) was added to 10 μL of freshly prepared 0.2 M $NaBH_4$ and then incubated on ice for 30 min. The reaction was quenched by an addition of cold 3 M sodium acetate (5 μL) and 75 μL of cold ethanol. The pellet was rinsed with 70% ethanol and then lyophilized. The sample was then treated with aniline.

Treatments of RNPs with hydrazine were carried out as previously described.[50] A solution containing 5′-^{32}P-labeled An16RNA (∼100 000 cpm) and nonlabeled An16RNA (1.0 μM) was heated at 80 °C for 3 min, then chilled on ice, and incubated with 30% hydrazine/H_2O (v/v) at 4 °C for 30 min. The reaction was quenched by adding 5 μL of cold 3 M sodium acetate and 125 μL of cold ethanol. The sample was chilled and pelleted, reprecipitated from 50 μL of 0.3 M sodium acetate and 125 μL of ethanol, rinsed with 70% ethanol, and then lyophilized. The sample was treated with aniline.

The chemically modified RNA was dissolved in 20 μL of 1.0 M aniline/acetate buffer (pH 4.5) and incubated in the dark at 60 °C for 20 min. The reaction was quenched by freezing with liquid nitrogen and lyophilized. The sample was resuspended in 20 μL of water and then lyophilized. The sample was reprecipitated from 50 μL of 0.3 M sodium acetate and 125 μL of ethanol. The lyophilized sample was dissolved in 90% formamide loading dye, heated at 80 °C for 3 min, chilled quickly, run on a 12% denaturing polyacrylamide/8 M urea gel, and analyzed by a STORM phosphor imager (GE).

In-line Probing Procedures. The in-line probing reactions were carried out as described previously.[49] A solution containing 5′-^{32}P-labeled An16RNA (120000 cpm) and nonlabeled An16RNA (1.0 μM) was heated at 80 °C for 3 min and then chilled on ice in 50 mM sodium cacodylate (pH 7.0), 100 mM NaCl, and 10 mM $MgCl_2$, with Rev (1.2 μM) and the indicated concentration of substrate (1, 3, 10, 30, 100, 300 μM, 1, 3, and 10 mM). The solution was incubated at 25 °C for 4 or 5 days. An aliquot (5 μL) of the reaction solution was diluted with 5 μL of a loading solution containing 8 M urea and 10 mM EDTA, and the 4 μL of the loading solution were run on a 12% denaturing polyacrylamide/8 M urea gel and then analyzed by a STORM phosphor imager (GE).

■ ASSOCIATED CONTENT

Supporting Information. Nucleotide sequences of ATP-binding RNPs isolated from the RREN30 RNP library (Figure S1), titration curves for the relative fluorescence intensity changes of An16RNP by ATP analogues (Figure S2), an autoradiogram of the gel shift assay for the An16RNA/Rev complex formation (Figure S3), titration curves for the binding of ATP to 5′-^{32}P-labeled An16RNA in the absence or presence of Rev (Figure S4), effects of Mg^{2+} on the affinity of An16RNP to Ado (Figure S5), titration of the relative cleavage band intensities by in-line probing of An16RNP with increasing concentration of Ado at U15, U21, and G22 (Figure S6), titration curves for the relative fluorescence intensity changes of the mutants of An16RNP at the nonpreserved U nucleotide by Ado (Figure S7), titration curves for the Ado-binding of the 5′-^{32}P-labeled U mutants of An16RNP (Figure S8), saturation curves for the relative fluorescence intensity changes of the mutants of An16RNP at the nonpreserved U nucleotide (Figure S9), titration curves for the binding of U18C/U27C, U18C/U29C, and U27C/U29C mutants with Ado (Figure S10), stability of An16RNP in the neutral and acidic conditions (Figure S11), titration curves of relative fluorescence intensity changes for the fluorescent RNP by Ado in neutral and acidic conditions for An16 A16G/U27C and An16 A16G/U18C mutants (Figure S12), and nucleotide sequences of ATP-binding RNPs isolated from the RREN30 RNP library by using the ATP-immobilized resin that anchored from the C-8 position of ATP (Figure S13). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
t-morii@iae.kyoto-u.ac.jp

■ ACKNOWLEDGMENT

This work was supported in part by the Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan to T.M. (Nos. 20241051 and 22121510). M.K. was supported by grants from MEXT (21370047 and 22121517), JST-SENTAN, Naito Foundation, YCU-AMRC, and CPLCEP-IAE-KU. The authors thank Dr. Eiji Nakata and Lieselot Carrette for comments.

■ REFERENCES

(1) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
(2) Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427–429.
(3) Gold, L.; Polisky, B.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763–797.
(4) Osborne, S. E.; Ellington, A. D. *Chem. Rev.* **1997**, *97*, 349–370.
(5) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611–647.
(6) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550–553.
(7) Burgstaller, P.; Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1084–1087.
(8) Burke, D, H.; Gold, L. *Nucleic Acids Res.* **1997**, *25*, 2020–2024.
(9) Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411–1418.
(10) Prudent, J. R.; Uno, T.; Schultz, P. G. *Science* **1994**, *264*, 1924–1927.
(11) Potyrailo, R. A.; Conrad, R. C.; Ellington, A. D.; Hieftje, G. M. *Anal. Chem.* **1998**, *70*, 3419–3425.
(12) Jhaveri, S.; Rajendran, M.; Ellington, A. D. *Nat. Biotechnol.* **2000**, *18*, 1293–1297.
(13) Brody, E. N.; Gold, L. *J. Biotechnol.* **2000**, *74*, 5–13.
(14) Yamamoto, R.; Baba, T.; Kumar, P. K. *Genes Cells* **2000**, *5*, 389–96.
(15) Lee, M.; Walt, D. R. *Anal. Biochem.* **2000**, *282*, 142–146.
(16) Hamaguchi, N.; Ellington, A. D.; Stanton, M. *Anal. Biochem.* **2001**, *294*, 126–131.
(17) Tang, J.; Breaker, R. R. *Chem. Biol.* **1997**, *4*, 453–459.
(18) Tang, J.; Breaker, R. R. *Nucleic Acids Res.* **1998**, *26*, 4214–4221.
(19) Araki, M.; Okuno, Y.; Hara, Y.; Sugiura, Y. *Nucleic Acids Res.* **1998**, *26*, 3379–3384.
(20) Koizumi, M.; Soukup, G. A.; Kerr, J. N. Q.; Breaker, R. R. *Nat. Struct. Biol.* **1999**, *6*, 1062–1071.
(21) Robertson, M. P.; Ellington, A. D. *Nucleic Acids Res.* **2000**, *28*, 1751–1759.
(22) Fontana, W.; Konings, D. A. M.; Stadler, P. F.; Schuster, P. *Biopolymer* **1993**, *33*, 1389–1404.
(23) Morii, T.; Hagihara, M.; Sato, S.; Makino, K. *J. Am. Chem. Soc.* **2002**, *124*, 4617–4622.
(24) Sato, S.; Fukuda, M.; Hagihara, M.; Tanabe, Y.; Ohkubo, K.; Morii, T. *J. Am. Chem. Soc.* **2005**, *127*, 30–31.
(25) Hasegawa, T.; Ohkubo, K.; Yoshikawa, S.; Morii, T. *e-J. Surf. Sci. Nanotech.* **2005**, *3*, 33–37.
(26) Hagihara, M.; Fukuda, M.; Hasegawa, T.; Morii, T. *J. Am. Chem. Soc.* **2006**, *128*, 12932–12940.
(27) Hasegawa, T.; Hagihara, M.; Fukuda, M.; Morii, T. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1277–1281.
(28) Hasegawa, T.; Hagihara, M.; Fukuda, M.; Nakano, S.; Fujieda, N.; Morii, T. *J. Am. Chem. Soc.* **2008**, *130*, 8804–8812.
(29) Koizumi, M.; Breaker, R. R. *Biochemistry* **2000**, *39*, 8983–8992.
(30) Burgstaller, P.; Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1084–1087.
(31) Burke, D. H.; Gold, L. *Nucleic Acids Res.* **1997**, *25*, 2020–2024.
(32) Davis, J.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11616–11621.
(33) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 5130–5135.
(34) Lorsch, J. R.; Szostak, J. W. *Acc. Chem. Res.* **1996**, *29*, 103–110.
(35) Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. *RNA* **1996**, *2*, 628–640.
(36) Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature* **1996**, *382*, 183–186.
(37) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, *273*, 1547–1551.
(38) Meli, M.; Vergne, J.; De′cout, J. L.; Maurel, M. C. *J. Biol. Chem.* **2002**, *277*, 2104–2111.
(39) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. *Chem. Biol.* **2004**, *11*, 1729–1741.
(40) Montange, R. K.; Batey, R. T. *Nature* **2006**, *441*, 1172–1175.
(41) Uesugi, S.; Ikehara, M. *J. Am. Chem. Soc.* **1977**, *99*, 3250–3253.
(42) Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147–14151.
(43) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293–8297.
(44) Matsugami, A.; Kobayashi, S.; Ouhashi, K.; Uesugi, S.; Yamamoto, R.; Taira, K.; Nishikawa, S.; Kumar, P. K. R.; Katahira, M. *Structure* **2003**, *11*, 533–545.
(45) Matsugami, A.; Ohyama, T.; Inada, M.; Inoue, N.; Minakawa, N.; Matsuda, A.; Katahira, M. *Nucleic Acids Res.* **2008**, *36*, 1805–1812.
(46) Mashima, T.; Matsugami, A.; Nishikawa, F.; Nishikawa, S.; Katahira, M. *Nucleic Acids Res.* **2009**, *18*, 6249–6258.
(47) Saenger, W. *Principles of Nucleic Acids Structure*; Springer-Verlag: New York, 1984.
(48) Carothers, J. M.; Davis, J. H.; Chou, J. J.; Szostak, J. W. *RNA* **2006**, *12*, 567–579.
(49) Soukup, G. A.; Breaker, R. R. *RNA* **1999**, *5*, 1308–1325.
(50) Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4679–4682.
(51) Oleson, A. E.; Hoganson, E. D. *Arch. Biochem. Biophys.* **1981**, *211*, 478–484.
(52) Kumar, K.; Jenkins, J. L.; Jardine, A. M.; Shapiro, R. *Biochem. Biophys. Res. Commun.* **2003**, *300*, 81–86.
(53) Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A. *EMBO J.* **1992**, *11*, 1119–1129.
(54) Zuker, M. *Science* **1989**, *244*, 48–52.
(55) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406–3415.
(56) Dixon, N.; Duncan, J. N.; Geerlings, T.; Dunstan, M. S.; McCarthy, J. E. G.; Leys, D.; Micklefield, J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2830–2835.
(57) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
(58) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214–220.
(59) Goddard, T. D.; Kneller, D. G. *SPARKY 3*; University of California: San Francisco, 2006.