RNA (2002), *8*:440–451. Cambridge University Press. Printed in the USA.
Copyright © 2002 RNA Society.
DOI: 10.1017.S1355838202026043

Solution structure of an RNA fragment with the P7/P9.0 region and the 3′-terminal guanosine of the *Tetrahymena* group I intron

AYA KITAMURA,[1] YUTAKA MUTO,[1] SATORU WATANABE,[1,6] INSIL KIM,[1] TAKUHIRO ITO,[1] YOICHI NISHIYA,[1,7] KENSAKU SAKAMOTO,[1] TAKASHI OHTSUKI,[2] GOTA KAWAI,[3] KIMITSUNA WATANABE,[2] KAZUMI HOSONO,[3,8] HIROSHI TAKAKU,[3] ETSUKO KATOH,[4] TOSHIMASA YAMAZAKI,[4] TAN INOUE,[5] and SHIGEYUKI YOKOYAMA[1]

[1]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo, Japan
[2]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo, Japan
[3]Department of Industrial Chemistry, Chiba Institute of Technology, Narashino, Chiba, Japan
[4]Biochemistry Department, National Institute of Agrobiological Sciences, Kannondai, Tsukuba, Ibaraki, Japan
[5]Graduate School of Biostudies, Kyoto University, Kyoto, Japan

ABSTRACT

In the second step of the two consecutive transesterifications of the self-splicing reaction of the group I intron, the conserved guanosine at the 3′ terminus of the intron (ωG) binds to the guanosine-binding site (GBS) in the intron. In the present study, we designed a 22-nt model RNA (GBS/ωG) including the GBS and ωG from the *Tetrahymena* group I intron, and determined the solution structure by NMR methods. In this structure, ωG is recognized by the formation of a base triple with the G264•C311 base pair, and this recognition is stabilized by the stacking interaction between ωG and C262. The bulged structure at A263 causes a large helical twist angle (40 ± 8°) between the G264•C311 and C262•G312 base pairs. We named this type of binding pocket with a bulge and a large twist, formed on the major groove, a "Bulge-and-Twist" (BT) pocket. With another twist angle between the C262•G312 and G413•C313 base pairs (45 ± 10°), the axis of GBS/ωG is kinked at the GBS region. This kinked axis superimposes well on that of the corresponding region in the structure model built on a 5.0 Å resolution electron density map (Golden et al., *Science*, 1998, *282*:345–358). This compact structure of the GBS is also consistent with previous biochemical studies on group I introns. The BT pockets are also found in the arginine-binding site of the HIV-TAR RNA, and within the 16S rRNA and the 23S rRNA.

Keywords: NMR; ribozyme; RNA structure; self-splicing

INTRODUCTION

The group I intron is a self-splicing RNA. The group I intron excises itself from a precursor RNA, connects the flanking exons, and generates the mature RNA efficiently and accurately. This self-splicing reaction consists of two consecutive steps involving transesterifications (Cech et al., 1992). In the first step, an exogenous guanosine is bound to the guanosine-binding site (GBS) that is conserved in the core region of the intron. Then, the 5′ splice site is cut by the transesterification. In the second step, the guanosine residue that is conserved at the 3′ terminus of the intron (ωG) replaces the exogenous guanosine (Fig. 1A). Then, the intron releases the 3′ exon, and the two exons are connected together by the second transesterification. The two transesterification reactions both require the guanosine-binding interaction, and are analogous to the reverse of each other (Inoue et al., 1986; Been & Perrotta, 1991). Therefore, it is considered that the guanosine-binding mechanism for the first step, which has been investigated by biochemical experiments, also describes the reaction in the second step.

Reprint requests to: Shigeyuki Yokoyama, Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1Hongo, Bunkyo-ku, Tokyo, Japan; e-mail: yokoyama@biochem.s.u-tokyo.ac.jp.
[6]Present address: Biophysics Division, National Cancer Center Research Institute, Tsukiji, Chuou-ku, Tokyo, Japan.
[7]Present address: Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Ltd., Asahi-machi, Machida-shi, Tokyo, Japan.
[8]Present address: Nippon Sanso Co., Nishi-shin bashi, Minato-ku, Tokyo, Japan.

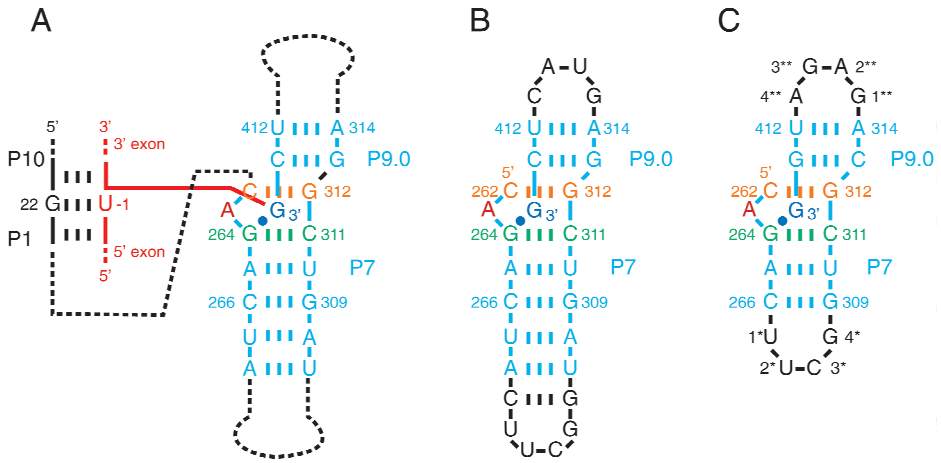


FIGURE 1. **A**: Sequence of the guanosine binding site (GBS) in the second transesterification step in the *Tetrahymena* intron. **B**: Sequence of P7/P9.0/G. **C**: Sequence of GBS/ωG. The helices corresponding to the P9.0 and P7 helices are colored light blue. The C262•G312 base pair is colored orange, A263 is colored red, the G264•C311 base pair is colored green, and ωG is colored blue. Broken lines represent base pairs. A blue-filled circle represents the interaction between ωG and the GBS predicted by the substitution experiment (Michel et al., 1989). Strands of the exons are colored red.

Substitution experiments revealed that a conserved G•C base pair in the P7 helix (hereafter G264•G311, according to the numbering for the *Tetrahymena* group I intron; Fig. 1A) was essential for the guanosine binding, and the G264•C311 base pair was proposed to form a base triple with the guanosine substrate (Michel et al., 1989). On a side of this G•C base pair (G264•C311), there are a bulged nucleotide (A263) and a Watson–Crick base pair (C262•G312; Fig. 1A). The guanosine substrate was proposed to interact with the C262 base through stacking (Yarus & Majerfeld, 1992). On the other hand, it was also proposed that the substrate guanosine interacts with the A•U base pair (A265•U310) on the other side (Fig. 1A), probably through hydrogen bonding (Yarus et al., 1991).

Moreover, some structural studies of group I introns have been reported. Three-dimensional models of group I introns, including the GBS region, were presented on the basis of phylogenetic and experimental analyses (Michel & Westhof, 1990; Lehnert et al., 1996). The P4-P6 region of the *Tetrahymena* group I intron was determined by X-ray crystallography at 2.8 Å resolution (Cate et al., 1996). The structure of the P5 helix, the internal bulge at J5a/5b, and the P1 helix were each determined by NMR (Allain & Varani, 1995; Luebke et al., 1997; Colmenarejo & Tinoco, 1999). However, these structures did not include the GBS, and thus provide no information about the guanosine recognition. Recently, the global architecture of a region containing the P3-P9 region was determined at 5 Å resolution (Golden et al., 1998). This P3-P9 fragment includes the GBS and ωG, but the elucidation of the guanosine recognition mechanism at an atomic resolution will require a higher resolution structure.

We previously designed a 31-nt model RNA of the GBS, termed P7/P9.0/G (Fig. 1B), from the *Tetrahymena* group I intron, and indicated that P7/P9.0/G retains the same specific interaction between the GBS and ωG as in the group I intron by substitution experiments (Watanabe et al., 1996). In the present study, we revamped the model RNA, named it GBS/ωG (Fig. 1C), and determined its solution structure by NMR.

RESULTS AND DISCUSSION

Design of GBS/ωG

The GBS/ωG RNA (Fig. 1C) was newly designed on the basis of P7/P9.0/G (Fig. 1B). In GBS/ωG, the U267•A308 and A268•U307 base pairs in the P7 stem are eliminated. The P9.0 stem of GBS/ωG is capped with a GAGA tetraloop, in order to stabilize the structure of the P9.0 stem in GBS/ωG. A model RNA with the wild-type sequence might form undesirable base pairs, such as C413•G312 and G313•C262. Therefore, to avoid the formation of alternative secondary structures, the C413•G313 base pair in P7/P9.0/G was replaced by the G413•C313 base pair in GBS/ωG. Because the same substitution in the group I intron has no effect on the splicing reaction (Russell & Herschlag, 1999), this corresponding substitution in GBS/ωG may not interfere with the guanosine binding by the GBS. As we expected, the GBS/ωG RNA yielded high-quality NMR spectra for the structural analysis.

We previously showed that the melting temperatures of some variants of P7/P9.0/G quantitatively indicate the interaction between the GBS and ωG, and found that P7/P9.0/G retains the same specificity as the group I intron (Watanabe et al., 1996). To confirm the same specificity in GBS/ωG, we measured the melting temperatures of two derivatives of GBS/ωG. One is the GBS RNA, which lacks the nucleotide corresponding to

ωG. The other is the GBS/ωA RNA, in which ωG is replaced by an adenosine. The melting temperature of the GBS/ωG RNA (70.5 °C) is appreciably higher than those of the GBS (67.5 °C) and GBS/ωA (67.0 °C) RNAs. This order is the same as that of the melting temperatures of the corresponding variants of P7/P9.0/G. Furthermore, we measured the NMR spectra of a variant of GBS/ωG, GBS/ωU, in which ωG is replaced by a uridine. From the spectra, we found that C262 and G312 in GBS/ωU do not form a Watson–Crick base pair (data not shown). Previous data indicated that C262 and G312 form a Watson–Crick base pair in the group I intron (Couture et al., 1990). The present results indicate that GBS/ωG distinguishes a guanosine from an adenosine or a uridine, as the group I intron does. Thus, the structure of GBS/ωG is expected to represent that of the guanosine-binding site at the second transesterification step in the group I intron.

Resonance assignments

The residues in GBS/ωG are numbered according to that in the *Tetrahymena* group I intron, except for the residues in the two tetraloops. The residues in the UUCG tetraloop are numbered as 1*–4*, and those in the GAGA tetraloop are numbered as 1**–4** (Fig. 1C). The assignment of the proton resonances was performed using the well-established NOE pathways (Wüthrich, 1986). Due to the excellent resonance dispersion in the spectra measured at 750 MHz (Fig. 2A), the ^{1}H NMR spectra were sufficient to assign all of the resonances in GBS/ωG, except for some of the H5′/H5″ and amino proton resonances (Table 1). Furthermore, to identify the resonances derived from ωG unambiguously, we synthesized a sample that is ^{13}C/^{15}N-labeled only at ωG, and measured its ^{13}C-^{1}H HSQC and ^{15}N-^{1}H HSQC spectra. On the basis of this assignment, we obtained unambiguous structural information about ωG.

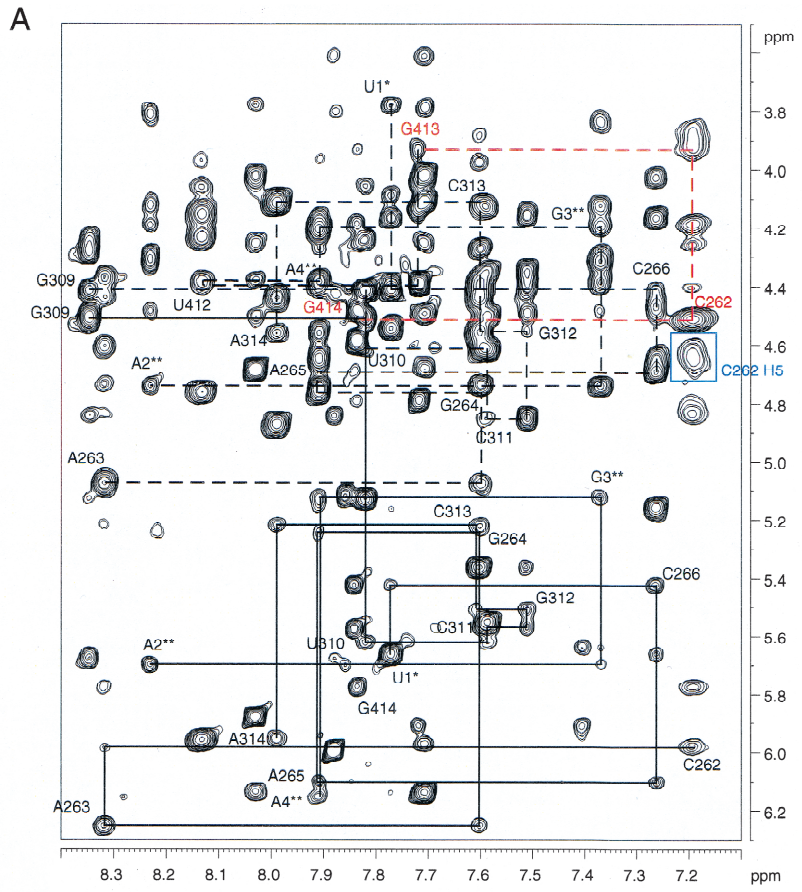
The chemical shifts and the NOE patterns of the UUCG and GAGA tetraloops in GBS/ωG are similar to those of the tetraloops determined previously, and indicate the same helix-loop conformation (Orita et al., 1993; Allain & Varani, 1995; Jucker et al., 1996). For example, the observations of the 2′-OH resonance of U1* (6.7 ppm), and the upfield-shifted resonance (10.0 ppm) of the imino proton of G1** are distinctive features of the UUCG and GAGA tetraloops, respectively (Orita et al., 1993; Allain & Varani, 1995).

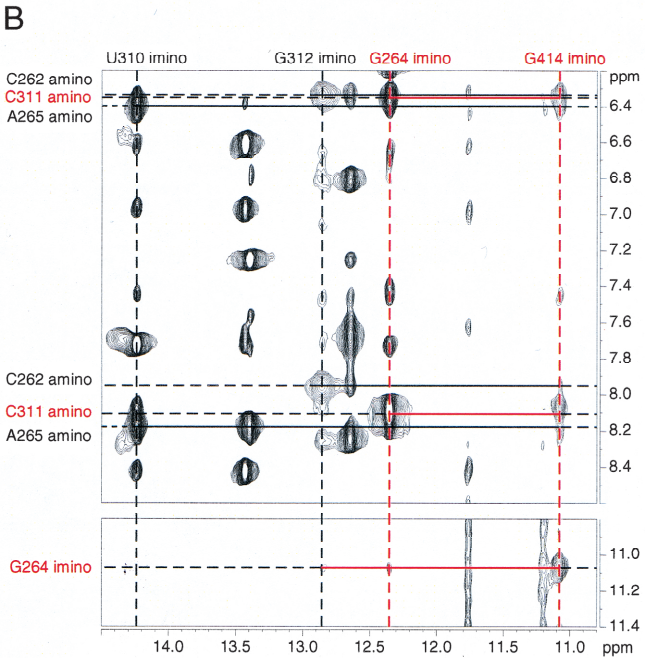
In the NOESY spectra in $^{2}H_2O$, a sequential walk could be performed from G264 to C266, from G309 to A314, and from U412 to G413 (Fig. 2A). These NOE connectivities of the P7 stem (residues G264 to C266, and G309 to C311) and of the P9.0 stem (residues C313 to A314, and U412 to G413) agree with the regular helical conformation of these regions (Varani et al., 1996). “Sequential NOEs” are also observed between the P7 and P9.0 stems. NOEs were observed between

TABLE 1. Proton and phosphorus chemical shifts (in parts per million) for GBS/ωG.[a]

	H8/H6	H2/H5	H1′	H2′	H3′	H4′	H5′/H5″	amino	imino	2′OH	P
C262	7.19	4.63	5.97	4.52	4.82	4.40	3.86/3.89	7.97/6.34	na		
A263	8.32		6.24	5.07	4.60	4.73			na		
G264	7.60	na	5.24	4.74	4.24	4.55	4.27/4.38		12.35		
A265	7.91	7.84	6.09	4.69	4.55	4.65	4.16/4.64		na		
C266	7.26	5.15	5.42	4.41	4.16	4.45	4.03/4.47	8.43/6.98	na		−3.94
U1*	7.77	5.65	5.66	3.78	4.53	4.38	4.52/4.09	na	11.77	6.65	−4.32
U2*	8.03	5.87	6.13	4.68	4.02	4.49	4.25/4.04	na			
C3*	7.71	6.13	5.96	4.11	4.50	3.80	3.62/2.70		na		−4.94
G4*	7.88	na	5.98	4.84	5.66	4.41	4.19/4.41		9.80		−4.92
G309	8.35	na	4.51	4.50	4.24	4.41	4.28/4.48	8.79/6.62	13.43		−2.22
U310	7.82	5.12	5.62	4.61	4.51	4.41	4.52/4.06	na	14.23		−4.77
C311	7.59	5.54	5.57	4.85	4.35	4.50	4.05/4.13	8.09/6.36	na		
G312	7.51	na	5.50	4.55	4.38	4.48	4.16/4.40		12.85		
C313	7.61	5.36	5.21	4.11	4.44	3.88	4.40/3.98	8.26/6.81	na		
A314	7.99	7.41	5.94	4.56	4.87	4.41			na		
G1**	9.89	na	5.63	4.12	4.48	4.29	4.08/4.41		9.99		−2.61
A2**	8.23	7.86	5.69	4.74	4.31	4.19	4.29/3.81		na		−1.17
G3**	7.37	na	5.12	4.21	4.38	4.27	3.84/4.19				−2.30
A4**	7.91	8.28	6.14	4.37	4.76	4.41	4.64/4.16		na		−4.97
U412	8.13	5.95	4.06	4.38	4.14	3.97		na	13.39		−1.78
G413	7.72	na	5.90	3.93	4.78	4.38	4.39/4.06		12.65		
G414	7.84	na	5.77	4.58	4.48	4.253	4.18/4.18	8.18/6.40	11.09		

[a]Imino, amino, and 2′-OH resonances were measured on a DRX600 spectrometer. Phosphorus resonances were measured on a DMX500 spectrometer. All other protons were measured on a DMX750 spectrometer.
na: not applicable.





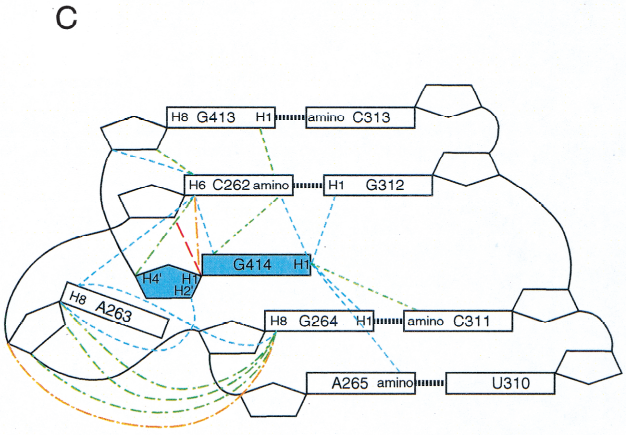


FIGURE 2. A: The base H8/H6 to H1′/H5/H2′ region is shown in a 150-ms two-dimensional NOESY spectrum in ^2H_2O at 283 K. Solid lines indicate H8/H6-H1′ correlations, and broken lines indicate H8/H6-H2′ correlations. Red broken lines indicate H8/H6-H2′ correlations between the P9.0 and P7 stems. The boxed peak corresponds to the H6-H5 correlation of C262. **B**: The imino–imino/amino region in the 200-ms two-dimensional NOESY spectrum in 90% H_2O/10% ^2H_2O at 283 K. Red lines indicate the correlation among the imino and amino protons in G414 (ωG) and the G264•C311 base pair. **C**: Schematic representation of observed NOEs about ωG, C262, and A263. Broken lines represent observed NOEs. Red, orange, green, and blue lines represent NOEs in the category of strong, medium, weak, and very weak, respectively. Intraresidue and intra-base-pair NOEs are not represented.

the H2′/H3′ of G413 and the H6 of C262, and between the H2′/H3′ of G312 and the H6 of C313 (Fig. 2A,C), although G413 and C262 are not covalently connected. These NOEs suggest that the G413•C313 base pair is stacked with the C262•G312 base pair.

Some NOEs were observed between the bases of ωG and the G264•C311 base pair (Fig. 2B,C). These NOEs indicate the formation of a base triple between these nucleotides, which is the same as the model proposed previously (Michel et al., 1989).

Structure determination

The structure of GBS/ωG (Fig. 3A) was determined by structure calculations with the information from NMR measurements (Table 2). Hydrogen bond constraints were used for the six Watson–Crick base pairs in the two stems. No hydrogen bonds were presumed for ωG in the structural calculation.

The subdivided parts of GBS/ωG are very well converged, although the overall structures are relatively diverged (Table 2). The r.m.s.d. for the P7 stem region (two Watson–Crick base pairs and the UUCG tetraloop) is 0.97 Å (Fig. 3B; Table 2). The r.m.s.d. for the P9.0 stem region (two Watson–Crick base pairs and the GAGA tetraloop) is 1.20 Å (Fig. 3C; Table 2). The structures of these two tetraloops are identical to those of the corresponding tetraloops in other RNA molecules (Allain & Varani, 1995; Jucker et al., 1996). The structure of the GBS region, whose solution is the purpose of this study, is also well defined, with an r.m.s.d. of 1.48 Å (Fig. 3D; Table 2).

TABLE 2. Structure statistics.

Number of restraints	
NOE distance restraints	334
Dihedral angle restraints	107
Hydrogen bond restraints	40
Total number of restraints	481
R.m.s.d. for all heavy atoms relative to the mean structure (Å)	
P7 stem and the UUCG tetraloop	0.97
P9.0 stem and the GAGA tetraloop	1.20
GBS region (residues 262–264, 311–312, 414)	1.48
All nucleotides	3.26
R.m.s.d. from idealized geometry	
Bonds (Å)	0.003
Angles (°)	0.61
Impropers (°)	0.47
R.m.s.d. for restraints	
Distance restraints (Å)	0.011
Dihedral restraints (°)	0.39

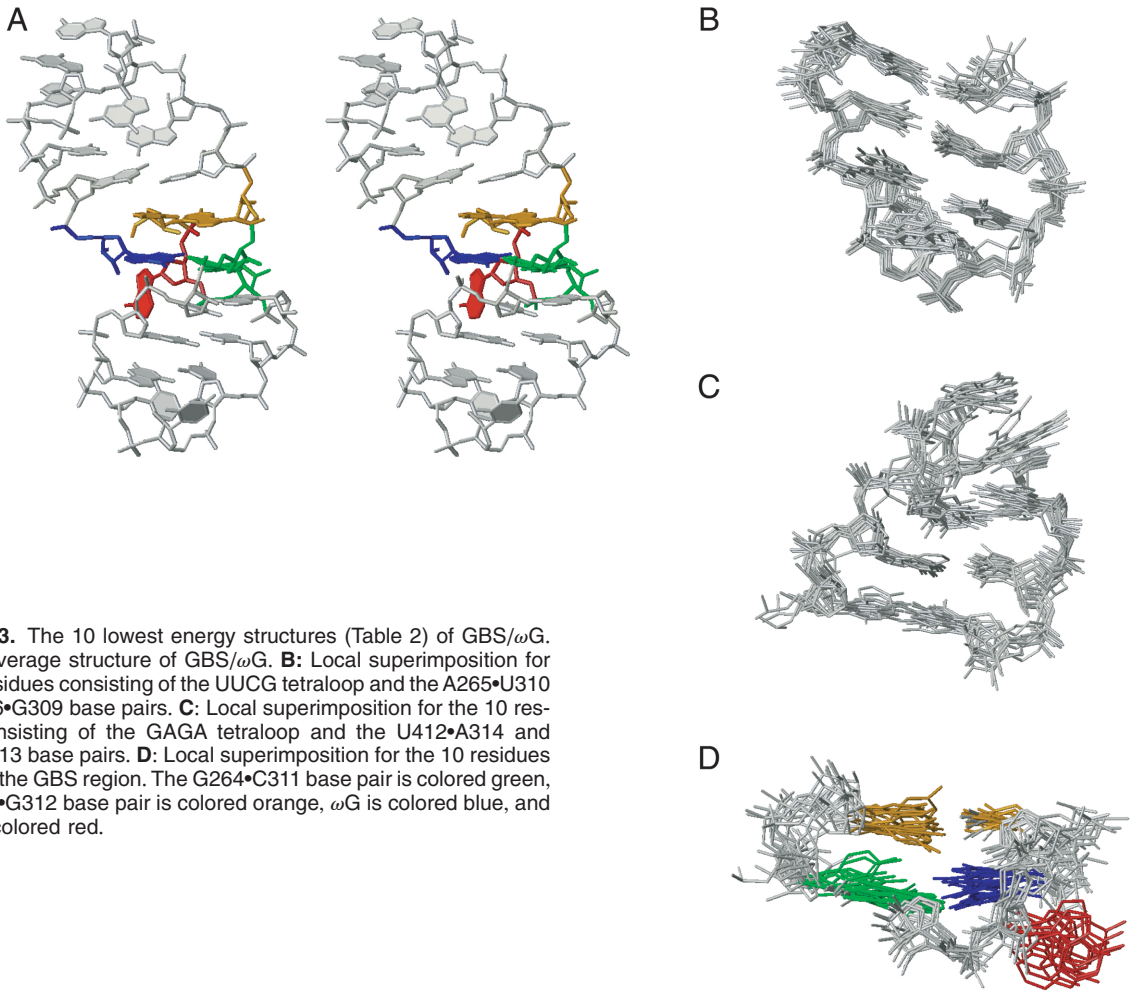


FIGURE 3. The 10 lowest energy structures (Table 2) of GBS/ωG. **A**: The average structure of GBS/ωG. **B:** Local superimposition for the 10 residues consisting of the UUCG tetraloop and the A265•U310 and C266•G309 base pairs. **C**: Local superimposition for the 10 residues consisting of the GAGA tetraloop and the U412•A314 and G413•C313 base pairs. **D**: Local superimposition for the 10 residues that form the GBS region. The G264•C311 base pair is colored green, the C262•G312 base pair is colored orange, ωG is colored blue, and A263 is colored red.

Base triple interaction

ωG forms a base triple with the G264•C311 base pair in GBS/ωG (Fig. 4A). Two hydrogen bonds were identified between ωG and G264 in this structure. One is between the amino proton of ωG and the N7 of G264, and the other is between the imino proton of ωG and the O6 of G264. Michel et al. (1989) showed that a variant with the substitution of an A•U base pair for the G264•C311 base pair captures a 2-aminopurine more tightly than a guanosine. It was also shown that the specificity of the GBS in the second step was changed from ωG to an adenosine by the substitution of a C•G base pair for the G264•C311 base pair (Been & Perrotta, 1991). The hydrogen bonds observed in GBS/ωG are consistent with these substitution experiments, and agree with the base triple model that was proposed by these substitution experiments.

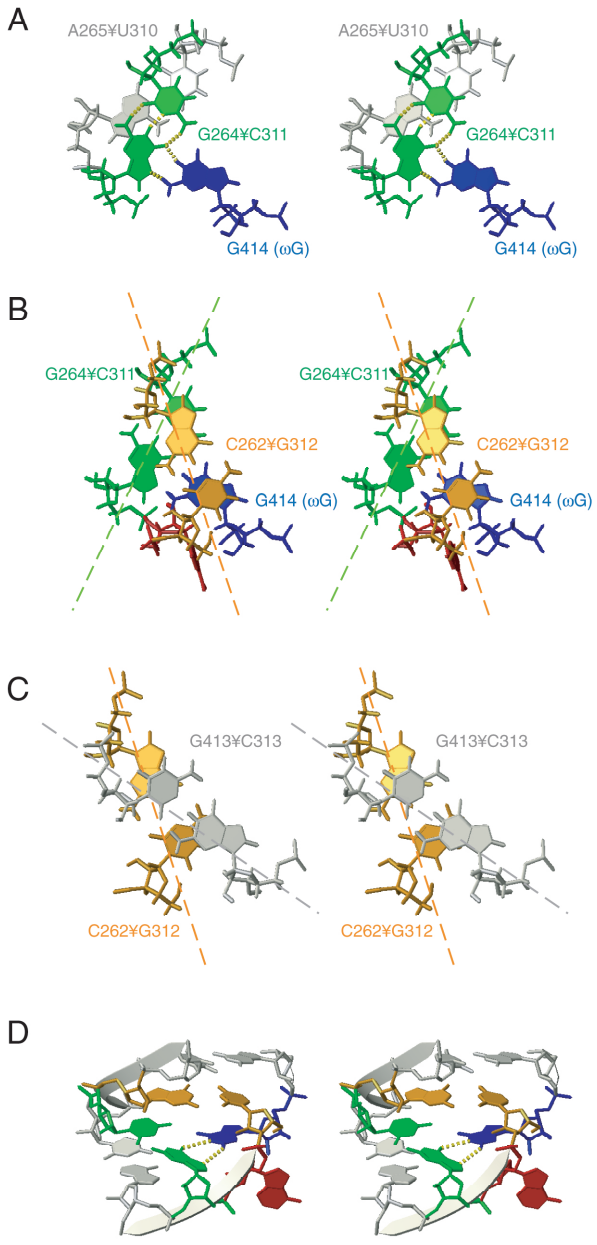


FIGURE 4. Stereoview of the GBS region of the average structure. Twist angles were derived from the 50 lowest energy structures. **A**: ωG forms a base triple with the G264•C311 base pair. **B**: C262 is stacked on ωG. The twist angle between the G264•C311 and C262•G312 base pairs is large (39.8 ± 7.6°). **C**: The twist angle between the G413•C313 and C262•G312 base pairs is also large (45.1 ± 10.4°). **D**: The GBS region with the backbones represented by ribbons. The color scheme is the same as in Figure 3.

Stacking interaction of C262 on ωG

The base of C262 is stacked on the base of G414 (ωG; Fig. 4B). Due to this stacking interaction, H5 of C262 is placed just above the purine ring of G414. The upfield-shifted resonance (4.63 ppm) originating from H5 of C262 (Fig. 2A) can be attributed to the deshielding effect by the purine ring of G414, and confirms the noncanonical base-stacking structure between C262 and G414. This stacking interaction would stabilize the recognition of G414 (ωG) by the GBS in GBS/ωG. The base of C262 intercalates well between the two bases of G413–G414 (Fig. 4D).

Yarus and Majerfeld (1992) exchanged C262 for other types of nucleotides, and found that the splicing activity is almost proportional to the stacking properties of the nucleotide at this position. From this result, they suggested that C262 is stacked on the guanosine substrate and stabilizes the binding interaction. The stacking interaction between C262 and ωG in GBS/ωG indicates that C262 is stacked not only on the guanosine substrate but also on ωG.

A263 bulge

It was suggested that C262 and G264 form Watson–Crick base pairs with G312 and C311, respectively, and A263 does not form a base pair, but instead exists in a bulge structure (Burke, 1988; Couture et al., 1990). In the present structure, A263 is not close enough to hydrogen bond with any of the bases in the major groove of the P7 stem (Figs. 3A and 4D). The A263 base is located near the back of the ωG ribose moiety, whereas its front side is involved in the reaction (Fig. 4D).

The nucleotide corresponding to A263 is semiconserved, as either an adenosine or a cytidine, among all of the sequenced group I introns, and covaries with the base pair corresponding to C262•G312. A263 corresponds to the base pair of G262•C312 or C262•G312, whereas C263 corresponds to the base pair of A262•U312 or U262•A312. However, the replacement of A263 with other nucleotides affects neither the self-splicing nor the substrate specificity in the *Tetrahymena* group I intron (Yarus et al., 1991). Furthermore, there was no interference from any of the base modified analogs of A263 (Ortoleva-Donnelly et al., 1998). These

observations are consistent with the present structure: No specific interaction around A263 was observed in GBS/ωG. We found that this bulge structure at A263 causes a large helical twist angle (40 ± 8°) between the C262•G312 and G264•C311 base pairs. This twist angle is greater than that in an A-form helix (32.7°). This large twist angle causes a pocket in the major groove, and enables G414 (ωG) to stack on C262 as well as to form the base triple with the G264•C311 base pair. We named this type of binding pocket, formed on the major groove with a bulge and a large twist, a "Bulge-and-Twist" (BT) pocket.

Couture et al. (1990) showed that the splicing reaction was impaired in the variant, in which the nucleotide at position 263 (N263) could form a base pair with that at position 312 (N312). They indicated that N263 covaries with the N262•N312 base pair so as to avoid base pairing between N263 and N312 (Couture et al., 1990). Their variants with the improper N263•N312 base pair may have a normal twist angle between the N263•N312 and G264•C311 base pairs, as in the canonical A-form helix. Then, the stacking interaction of ωG (or the substrate guanosine) with N262 and the base triple formation with the G264•C311 base pair would be prevented in these variants.

Kink between the two helices

A large twist angle between the G413•C313 and C262•G312 base pairs (45 ± 10°; Fig. 4C) is present at the stem–stem junction (Fig. 1C). This angle is as large as that between the C262•G312 and G264•C311 base pairs (40 ± 8°; Fig. 4B). These large twist angles may have naturally caused the kink of the axis in the GBS region (Fig. 5).

A kink around the GBS in GBS/ωG is also observed within the corresponding region in the P3-P9 structure model built on a 5.0-Å resolution electron density map (Golden et al., 1998; Fig. 5). When the stem region in GBS/ωG is superimposed on the corresponding region in the P3-P9 structure, the bases in the P9.0 stem of GBS/ωG are located at almost at the same positions as the corresponding bases in the crystal structure (Fig. 5). However, the positions of G264-C266 in GBS/ωG are shifted by one base as compared with those in the 5.0 Å crystal structure (Fig. 5).

The X-ray crystallographic study (Golden et al., 1998) and the biochemical studies (Pyle et al., 1992; Wang & Cech, 1992; Szewczak et al., 1998, 1999; Strobel & Ortoleva-Donnelly, 1999) of the group I introns indicated that many segments are brought together to form the catalytic core of the intron, such as P7, P9.0, P1, J4/5, and J8/7. It is likely that the kink between the P7 and P9.0 helices observed in GBS/ωG makes space for these segments for the proper packing of the catalytic core.

Contribution of the A265•U310 base pair to the guanosine binding

Mutations of the A265•U310 base pair, adjacent to the G264•C311 base pair, affected the splicing reaction of the group I intron (Yarus et al., 1991). Variants with the G265•U310 or U265•U310 base pair were inferior to the wild-type group I intron in their reactivity toward the guanosine. Nevertheless, the G265•U310 variant preferred a 2,6-diaminopurine (2,6-DAP) to the guanosine as the substrate, although this reactivity of the variant toward the 2,6-DAP was still lower than that of the wild-type intron toward the guanosine. The specificity of the G265•U310 variant was interpreted as the formation of a hydrogen bond between the 6-NH_2 of the 2,6-DAP and the O6 of G265. This hydrogen-bonding scheme has been called the axial model, in which the 2,6-DAP tilts and interacts with the two adjacent base pairs, G264•C311 and G265•U310 (Yarus et al., 1991).

On the basis of the structure of GBS/ωG, we examined the axial hydrogen-bonding scheme for the G265•U310 variant with the 2,6-DAP. The replacement of ωG by the 2,6-DAP prevents the formation of the hydrogen bond with the O6 of G264. However, the replacement of the A265•U310 by a G•U wobble base pair enables the 2,6-DAP to form two hydrogen bonds with G265 and U310, because the U in a G•U wobble base pair extrudes into the major groove (Fig. 6A). Therefore, the 2,6-DAP in this model forms more hydrogen bonds than ωG in the base triple conformation, while it does not stack on C262 (Fig. 6B). The hydrogen-bonding scheme in this model can account for the affinity of the variant for the 2,6-DAP.

On the basis of the reactivity of the variant toward the 2,6-DAP, it was also proposed that a similar hydrogen-bonding scheme may be possible for the wild-type group I intron with the guanosine substrate (Yarus et al., 1991). The proposed hydrogen bond is between the O6 of the guanosine and the 6-NH_2 of A265, and corresponds to that between the 6-NH_2 of the 2,6-DAP and the O6 of G265 in the variant. Due to the tilt of the ωG base, this hydrogen-bonding scheme may contradict the stacking interaction between the substrate and C262 (Yarus & Majerfeld, 1992).

Thus, we examined the probability of this hydrogen-bonding scheme in the intact group I intron. In GBS/ωG, ωG is stacked on C262, and forms the base triple (Fig. 6C). If the base of ωG in GBS/ωG is tilted, as in the model described above, then ωG could form only two hydrogen bonds, one with G264 and the other with A265, and could not stack on C262. Therefore, the axial conformation does not seem preferable when the substrate is a guanosine. However, in the full-length group I intron, it is possible that some interactions that are not included in the structure of GBS/ωG may allow the axial conformation in conformational changes during the splicing reaction.

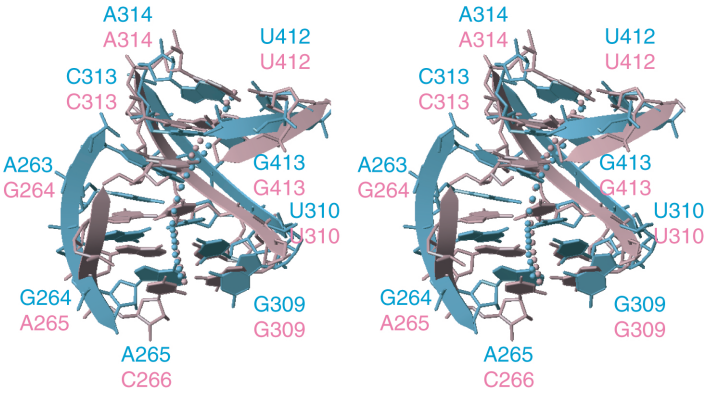


FIGURE 5. Stereoview superimposition of the stem region in the average structure of GBS/ωG, with the corresponding region in the crystal structure (Golden et al., 1998). Axes are represented by lines of spheres. GBS/ωG is colored pink. The crystal structure is colored sky blue.

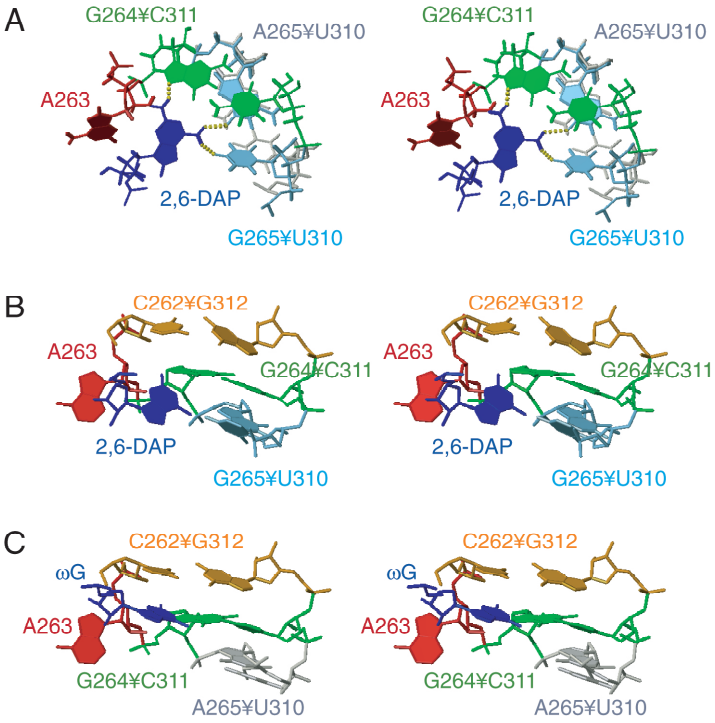


FIGURE 6. A,B: Stereo depiction of the model structure with a 2,6-diaminopurine (2,6-DAP) and the G265•U310 wobble base pair. The A265•U310 base pair in GBS/ωG is superimposed on the G265•U310 base pair. The 2,6-DAP is colored blue, the G264•C311 base pair is colored green, the G265•U310 base pair is colored sky blue, and the A265•U310 base pair in GBS/ωG is colored beige. The putative hydrogen bonds are represented by yellow broken lines. **C:** Stereoview of the GBS region in GBS/ωG.

Connection to the P10 helix

How compatible is the present GBS/ωG structure with the connection to the P10 helix? We constructed a model, by substituting the GBS/ωG structure for the corresponding part of the full-length model structure of the group I intron (Michel & Westhof, 1990), as shown in Figure 7. On the front side of ωG, there are the 3′ exon that is connected to the ωG 3′-OH group, and the phosphate group following ωG that is to react with U(−1) at the 3′ end of the 5′ exon. It has been reported that the 3′ and 2′-OH groups of ωG, and the 3′-OH group of U(−1) are bound with magnesium ions (Piccirilli et al., 1993; Weinstein et al., 1997; Shan et al., 1999). Whereas

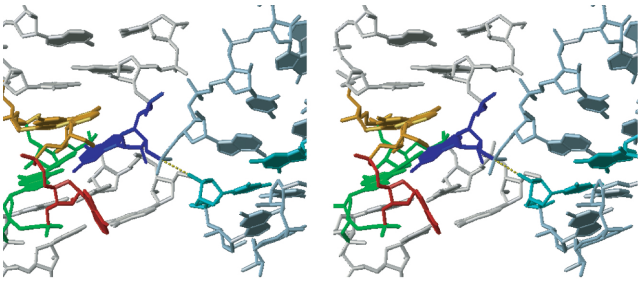


FIGURE 7. Stereoview of the model structure derived from the Michel–Westhof model (Michel & Westhof, 1990) and GBS/ωG. The P10 helix is colored light blue, and the G22•U(−1) base pair is colored royal blue. The broken line connects the phosphate of A415 and the O3′ of U(−1).

the back of the ωG ribose faces the base of A263, the location of ωG relative to the bulged base is compatible with the connection to the P10 helix and the magnesium ion bindings. U412 and G413 of the P9.0 helix are close to the P10 helix, but with no steric hindrance. Therefore, the present GBS/ωG structure appears to be compatible with the P10 helix connection and the splicing reaction.

Comparison with other RNA molecules

The GBS in the group I intron is also known to bind an arginine, because of the structural similarity between the guanidino group of the arginine side chain and the isosteric set of atoms on the guanosine base (Fig. 8A,B; Yarus, 1988, 1989). The HIV TAR RNA also recognizes an arginine residue in the arginine-rich domain of the Tat protein (Calnan et al., 1991) by a mechanism similar to that elucidated for the GBS in GBS/ωG (Fig. 8C,D; Brodsky & Williamson, 1997). The HIV TAR RNA forms a binding pocket that is very similar to the BT pocket in GBS/ωG, although this binding pocket recognizes an arginine instead of a guanosine. First, the BT pocket in the TAR RNA consists of two base pairs and a bulge between them. Second, the substrate arginine forms two hydrogen bonds with G26 in the G•C base pair, using a guanidino group (Fig. 8B). Third, the twist angle between the A22•U40 and G26•C39 base pairs (47°) is larger than that in the canonical A-form. With this large twist angle, A22 is stacked with the arginine (Fig. 8C,D). Beside these features that are common to the BT pocket, the TAR RNA has an additional stacking interaction with the arginine. In the TAR RNA, one of the bulged residues, U23, forms a base triple with the A27•U38 base pair (Puglisi et al., 1992), and is stacked with the substrate arginine. On the other hand, the bulge A263 in GBS/ωG forms no hydrogen bond, and is not stacked with the substrate.

The BT pocket is also found in other RNA molecules, and has a role in binding either a free nucleotide or a nucleotide residue from another RNA segment. For example, a part of the 16S rRNA (Fig. 8E,F; Wimberly et al., 2000) contains the distinct structural features of the BT pocket, which recognizes a cytidine residue (C1200). The target C1200 forms a base triple with the A1055•U1205 base pair. The twist angle between the A1055•U1205 and U1052•A1206 base pairs is also large (43°) with the two residues in a bulge. U1052 in the U•A base pair is stacked with C1200. Furthermore, one of the bulged residues, G1053, is stacked with the target C1200, like U23 in the TAR RNA. In this BT pocket, the target nucleotide is recognized by forming a (U•A)•C base triple, whereas in the GBS BT pocket, the target guanosine is recognized by forming a (C•G)•G base triple. Another example of the BT pocket is within the 23S rRNA (Fig. 8G,H; Ban et al., 2000). In this BT pocket, the target residue, A2612, forms an A2612•(A2095•U2650) base triple, and is stacked with G2097 in the G2097•C2647 base pair. The twist angle between the G2097•C2647 and A2095•U2650 base pairs is large (46°), with the bulge of A2648-A2649. In these two rRNA BT pockets, the substrates are specifically recognized by forming base triples that are different from each other, and from the base triple in the GBS BT pocket. This indicates that the specificity of the BT pocket for a ligand is variable, according to the type of base triple.

MATERIALS AND METHODS

RNA preparation

RNA samples were synthesized on a DNA/RNA synthesizer and were deprotected as described (Kim et al., 1997). The [^{13}C,^{15}N] 3′,5′-GDP was obtained as described previously (Ohtsuki et al., 1998). The RNA sample that was specifically labeled at G414 (ωG) was prepared enzymatically by the ligation of [^{13}C,^{15}N] 3′,5′-GDP to the RNA fragment that lacks a nucleotide at position 414. These RNA samples were purified as described (Kim et al., 1997) and were dissolved in the desired buffers. For the NMR experiments involving exchangeable protons, the 3.0 mM of unlabeled sample and the 0.5 mM of labeled sample were prepared in 90% H_2O/10% ^2H_2O containing 10 mM sodium phosphate, pH 5.5, and 50 mM sodium chloride. All other experiments for nonexchangeable protons were performed in 99.96% ^2H_2O containing 10 mM sodium phosphate, pH 6.5, and 50 mM sodium chloride.

Melting profile analysis

Melting profiles of the RNAs were measured with a Gilford Response II UV spectrometer at a heating rate of 1 °C/min (Watanabe et al., 1996). The UV absorbance at 260 nm of the sample, in 10 mM sodium cacodylate buffer (pH 6.5) containing 100 mM sodium chloride, was measured at 0.5 °C increments from 25 to 100 °C.

NMR spectroscopy

The two- and three-dimensional NMR experiments were acquired on Bruker DMX500, DRX600, and DMX750 spectrometers. All spectra were acquired at 283 K except for the two-dimensional ^{1}H-^{15}N HSQC spectra, which were acquired at 278 K. For spectra in the H_2O NMR buffer, pulse sequences with jump and return water suppression (Plateau & Guéron, 1982) were used. For spectra in the ^2H_2O NMR buffer, low power irradiation of the residual HOD resonance was applied during the relaxation delay. The following spectra were acquired using an unlabeled RNA sample: NOESY spectra (50-, 150-, 200-, 400-ms mixing times) in the ^2H_2O NMR buffer; a NOESY spectrum (200-ms mixing time) in the H_2O NMR buffer; a DQF-COSY spectrum in the ^2H_2O NMR buffer; a TOCSY experiment (150-ms mixing time) with the MLEV-17 mixing sequence (Bax & Davis, 1985) in the ^2H_2O NMR buffer; and a ^{1}H, ^{31}P HetCor experiment (50- and 100-ms mixing times) with the WALTZ-16 isotropic mixing sequence

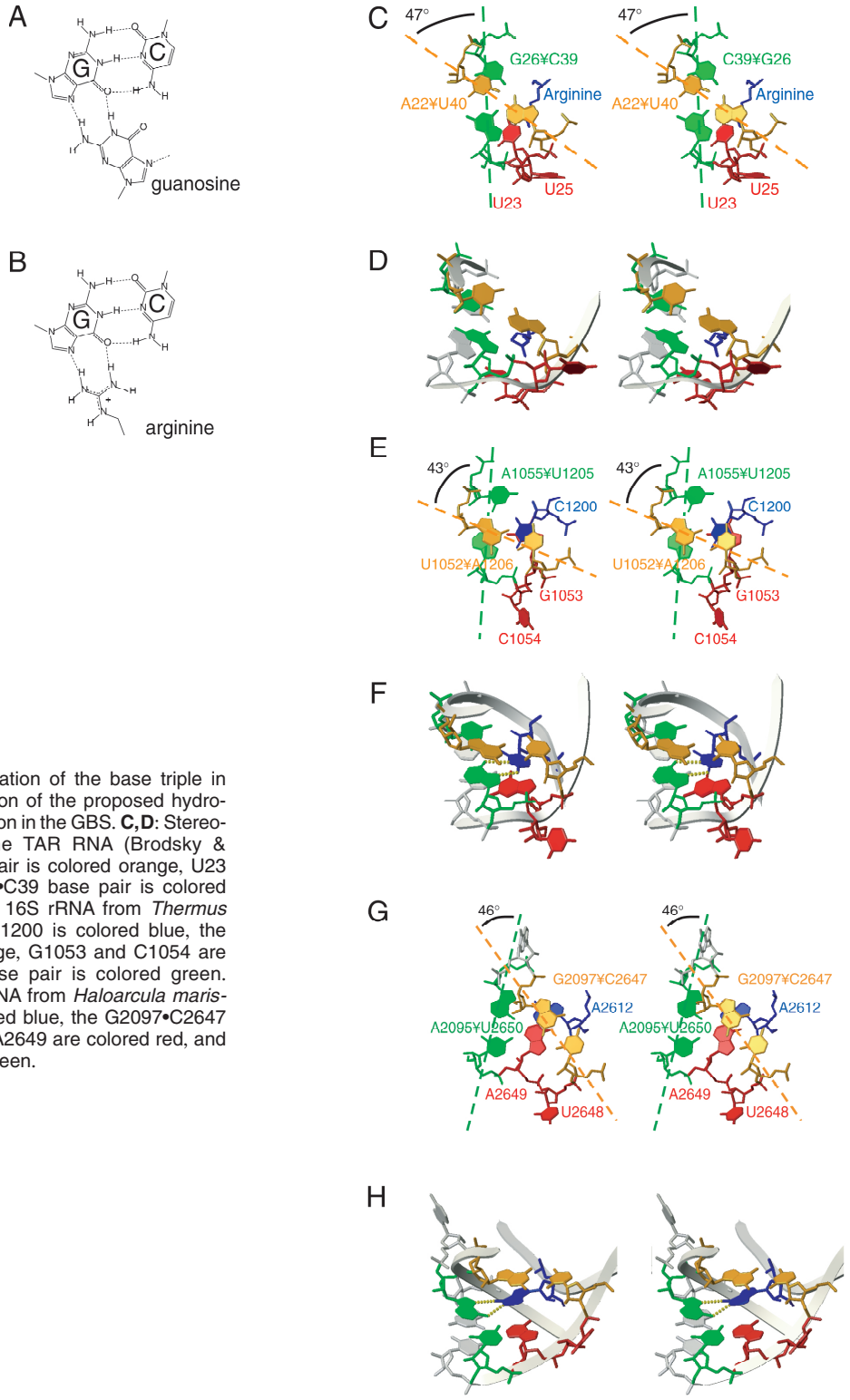


FIGURE 8. A: The schematic representation of the base triple in GBS/ωG. **B**: The schematic representation of the proposed hydrogen bonds in the arginine-binding interaction in the GBS. **C,D**: Stereoview of the arginine-binding site in the TAR RNA (Brodsky & Williamson, 1997). The A22•U40 base pair is colored orange, U23 and U25 are colored red, and the G26•C39 base pair is colored green. **E,F**: Stereoview of a part of the 16S rRNA from *Thermus thermophilus* (Wimberly et al., 2000). C1200 is colored blue, the U1052•A1206 base pair is colored orange, G1053 and C1054 are colored red, and the A1055•U1205 base pair is colored green. **G,H**: Stereoview of a part of the 23S rRNA from *Haloarcula marismortui* (Ban et al., 2000). A2612 is colored blue, the G2097•C2647 base pair is colored orange, U2648 and A2649 are colored red, and the A2095•U2650 base pair is colored green.

(Kellogg & Schweitzer, 1993) in the $^{2}H_2O$ NMR buffer. The RNA sample selectively labeled at G414 was used to acquire the following spectra: a two-dimensional ^{1}H-^{13}C HMQC spectrum in the $^{2}H_2O$ buffer, and a two-dimensional ^{1}H-^{15}N HSQC spectrum in the H_2O buffer.

NOE distance constraints

Semiquantitative distance constraints between nonexchangeable protons were estimated from the cross-peak intensities in the NOESY spectrum with a 50-ms mixing time. Using the

covalently fixed pyrimidine H5-H6 distance (2.43 Å) as a reference, the peak intensities were classified as strong, medium, weak, or very weak. The upper bound distance constraints of the corresponding proton pairs were set to 3.0, 4.0, 5.0, or 6.0 Å, respectively. Distance constraints involving exchangeable protons were estimated from the two-dimensional 200-ms mixing time NOESY spectrum, and were classified as either weak, very weak, or extremely weak. For the cross-peaks that were classified as extremely weak, the upper bound distance constraints were set to 7.0 Å in order to account for the possibility of spin diffusion. In the case of the intra-base-pair distances for A•U H2-NH, G•C NH-NH2, and G•U NH-NH, they were classified as strong constraints. Intraresidue sugar-to-sugar constraints were not included in the calculations.

Dihedral angle constraints

Sugar pucker conformations were determined from $^3J_{H1'-H2'}$ couplings that were semiquantitatively estimated from ^{31}P-decoupled two-dimensional DQF-COSY and TOCSY spectra. Residues with couplings >7 Hz (C262, U2*, C3*, and G414) were constrained to the C2′-*endo* conformation (Varani et al., 1996) through four of the torsion angles in the ribose sugar ring (Wimberly, 1992). Residues with weak $^3J_{H1'-H2'}$ couplings and large $^3J_{H3'-H4'}$ couplings (A265-U1*, G4*-A2**, A4**-G413) were constrained to the C3′-*endo* conformation. Residues with intermediate $^3J_{H1-H2'}$ couplings (A263, G264, and G3**) were left unconstrained. For residues clearly lacking H4′-H5′ and H4′-H5″ peaks in the DQF-COSY spectra, the γ-torsion angles were constrained to the *gauche*$^+$ conformation (60 ± 20°; Smith & Nikonowicz, 1998). For residues clearly lacking P-H5′ and P-H5″ peaks in the HetCor spectra, the β torsion angles were constrained to the *trans* conformation (180 ± 20°; Smith & Nikonowicz, 1998). We used dihedral constraints to keep the amino groups planar.

Hydrogen bond constraints

Watson–Crick base pairs were identified using two criteria: the observation of a significantly downfield shifted NH or NH_2 proton resonance and the observation of strong G•C NH-NH_2 or A•U H2-NH NOEs. The G•A base pair in the GAGA tetraloop and the U•G base pair in the UUCG tetraloop were identified by the typical chemical shifts and the NOE patterns reported previously (Varani et al., 1991; Orita et al., 1993; Allain & Varani, 1995; Jucker et al., 1996). Hydrogen bonds were introduced as distance restraints of 2.9 ± 0.3 Å between donor and acceptor heavy atoms and 2.0 ± 0.2 Å between acceptor and hydrogen atoms (Smith & Nikonowicz, 1998).

Structure calculation

The structures were calculated using X-PLOR 3.8 (Brünger, 1992) with restrained molecular dynamics (rMD). One hundred starting structures were generated with randomized torsion angles. The protocol for structure calculation was divided into four stages. For all of the stages, the force constants for bond length, bond angle, and NOE distance were set to 1,000 kcal mol^{-1} Å^{-2}, 500 kcal mol^{-1} rad^{-2}, and 50 kcal mol^{-1} Å^{-2}, respectively. At the first stage, each of the starting structures was subjected to 70 ps of rMD at 1,000 K using hydrogen bond and NOE-derived distance constraints. In the second step, each of the structures was subjected to 15 ps of rMD at 1,000 K with the shake protocol. The weight of the van der Waals force constant was increased during the last 10 ps. Then, each structure was cooled to 300 K over 14 ps and was minimized (1,000 steps). At the third stage, 1 ps of rMD at 1,000 K with hydrogen bond and NOE-derived distance constraints was performed with an increasing force constant for the torsion angle (5 to 50 kcal mol^{-1}). Each of the structures was cooled to 300 K over 3.5 ps and was minimized (1,000 steps). At the final stage, each of the structures was subjected to 1,000 steps of restrained energy minimization using the Lennard–Jones potential and all of the NMR-derived constraints.

Model buildings

The model structure that explains the result of the substitution experiment by Yarus et al. (1991) was constructed by modifying the coordinates of GBS/ωG. The A265•U310 base pair and ωG were replaced with a wobble G•U base pair and a 2,6-DAP, respectively. The modified coordinates were subjected to energy minimization using the Lennard–Jones potential with three hypothetical hydrogen bond constraints (2,6-DAP•G264 2-NH_2-N7, 2,6-DAP•G264 6-NH_2-O6, 2,6-DAP•U310 6-NH_2-O4).

The model structure of the GBS region with the P10 and P1 helices was constructed from the full-length model of the group I intron (Michel & Westhof, 1990) by substituting the helical region of GBS/ωG for the corresponding region. The P7 and P9.0 helices within GBS/ωG and the P1, and P10 helices were subjected to energy minimization.

Coordinates

Structure coordinates for the averaged structure have been deposited in the Protein Data Bank (PDB ID code 1K2G).

ACKNOWLEDGMENTS

We are grateful to Michael Yarus and François Michel for valuable critical reading of the manuscript. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas (No. 04272103) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the "Research for the Future" Program (JSPS-RFTF97L00503) from the Japan Society for the Promotion of Science.

Received October 1, 2001; returned for revision November 6, 2001; revised manuscript received January 14, 2002

REFERENCES

Allain FH, Varani G. 1995. Structure of the P1 helix from group I self-splicing introns. *J Mol Biol 297*:877–893.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science 289*:905–920.

Bax A, Davis DG. 1985. Practical aspects of two-dimensional transverse NOE spectroscopy. *J Magn Reson 63*:207–213.
Been MD, Perrotta AT. 1991. Group I intron self-splicing with adenosine: Evidence for a single nucleoside-binding site. *Science 252*:434–437.
Brodsky AS, Williamson JR. 1997. Solution structure of the HIV-2 TAR-argininamide complex. *J Mol Biol 267*:624–639.
Brünger AT. 1992. *X-PLOR Version 3.1 Manual.* New Haven, Connecticut: Yale University.
Burke JM. 1988. Molecular genetics of group I introns: RNA structures and protein factors required for splicing—A review. *Gene 73*:273–294.
Calnan BJ, Tidor B, Biancalana S, Hudson D, Frankel AD. 1991. Arginine-mediated RNA recognition: The arginine fork. *Science 252*:1167–1171.
Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science 273*:1678–1685.
Cech TR, Herschlag D, Piccirilli JA, Pyle AM. 1992. RNA catalysis by a group I ribozyme. Developing a model for transition state stabilization. *J Biol Chem 267*:17479–17482.
Colmenarejo G, Tinoco I Jr. 1999. Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme. *J Mol Biol 290*:119–135.
Couture S, Ellington AD, Gerber AS, Cherry JM, Doudna JA, Green R, Hanna M, Pace U, Rajagopal J, Szostak JW. 1990. Mutational analysis of conserved nucleotides in a self-splicing group I intron. *J Mol Biol 215*:345–358.
Golden BL, Gooding AR, Podell ER, Cech TR. 1998. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science 282*:259–264.
Inoue T, Sullivan FX, Cech TR. 1986. New reactions of the ribosomal RNA precursor of *Tetrahymena* and the mechanism of self-splicing. *J Mol Biol 189*:143–165.
Jucker FM, Heus HA, Yip PF, Moors EHM, Pardi A. 1996. A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J Mol Biol 264*:968–980.
Kellogg GW, Schweitzer BI. 1993. Two- and three-dimensional ^{31}P-driven NMR procedures for complete assignment of backbone resonances in oligodeoxynucleotides. *J Biomol NMR 3*:577–595.
Kim I, Muto Y, Inoue M, Watanabe S, Kitamura A, Yokoyama S, Hosono K, Takaku H, Ono A, Kainosho M, Sakamoto H, Shimura Y. 1997. NMR analysis of the hydrogen bonding interactions of the RNA-binding domains of the *Drosophila* sex-lethal protein with target RNA fragments with site-specific [3-^{15}N] uridine substitutions. *Nucleic Acids Res 25*:1565–1569.
Lehnert V, Jaeger L, Michel F, Westhof E. 1996. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: A complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem Biol 3*:993–1009.
Luebke KJ, Landry SM, Tinoco I Jr. 1997. Solution conformation of a five-nucleotide RNA bulge loop from a group I Intron. *Biochemistry 36*:10246–10255.
Michel F, Hanna M, Green R, Bartel DP, Szostak JW. 1989. The guanosine binding site of the *Tetrahymena* ribozyme. *Nature 342*:391–395.
Michel F, Westhof E. 1990. Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol 216*:585–610.
Ohtsuki T, Kawai G, Watanabe K. 1998. Stable isotope-edited NMR analysis of *Ascaris suum* mitochondrial tRNA^{Met} having a TV-replacement loop. *J Biochem 124*:28–34.
Orita M, Nishikawa F, Shimayama T, Taira K, Endo Y, Nishikawa S. 1993. High-resolution NMR study of a synthetic oligoribonucleotide with a tetranucleotide GAGA loop that is a substrate for the cytotoxic protein, ricin. *Nucleic Acids Res 21*:5670–5678.
Ortoleva-Donnelly L, Szewczak AA, Gutell RR, Strobel SA. 1998. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA 4*:498–519.
Piccirilli JA, Vyle JS, Caruthers MH, Cech TR. 1993. Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature 361*:85–88.
Plateau P, Guéron M. 1982. Exchangeable proton NMR without baseline distortion, using new strong-pulse sequences. *J Am Chem Soc 104*:7310–7311.
Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. 1992. Conformation of the TAR RNA-Arginine complex by NMR spectroscopy. *Science 257*:76–80.
Pyle AM, Murphy FL, Cech TR. 1992. RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature 358*:123–128.
Russell R, Herschlag D. 1999. Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme. *RNA 5*:158–166.
Shan S-O, Yoshida A, Sun S, Piccirilli JA, Herschlag D. 1999. Three metal ions at the active site of the *Tetrahymena* group I ribozyme. *Proc Natl Acad Sci USA 96*:12299–12304.
Smith JS, Nikonowicz EP. 1998. NMR structure and dynamics of an RNA motif common to the spliceosome branch-point helix and the RNA-binding site for phage GA coat protein. *Biochemistry 37*:13486–13498.
Strobel SA, Ortoleva-Donnelly L. 1999. A hydrogen-bonding triad stabilizes the chemical transition state of a group I ribozyme. *Chem & Biol 6*:153–165.
Szewczak AA, Ortoleva-Donnelly L, Ryder SP, Moncoeur E, Strobel SA. 1998. A minor groove RNA triple helix within the catalytic core of a group I intron. *Nat Struct Biol 5*:1037–1042.
Szewczak AA, Ortoleva-Donnelly L, Zivarts MV, Oyelere AK, Kazantsev AV, Strobel SA. 1999. An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc Natl Acad Sci USA 96*:11183–11188.
Varani G, Aboul-ela F, Allain FH. 1996. NMR investigation of RNA structure. *Prog NMR Spectr 29*:51–127.
Varani G, Cheong C, Tinoco I Jr. 1991. Structure of an unusually stable RNA hairpin. *Biochemistry 30*:3280–3289.
Wang JF, Cech TR. 1992. Tertiary structure around the guanosine-binding site of the *Tetrahymena* ribozyme. *Science 256*:526–529.
Watanabe S, Kawai G, Muto Y, Watanabe K, Inoue T, Yokoyama S. 1996. An RNA fragment consisting of the P7 and P9.0 stems and the 3′-terminal guanosine of the *Tetrahymena* group I intron. *Nucleic Acids Res 24*:1337–1344.
Weinstein LB, Jones BCNM, Cosstick R, Cech TR. 1997. A second catalytic metal ion in a group I ribozyme. *Nature 388*:805–808.
Wimberly BT. 1992. *NMR derived structures of RNA loops: The conformation of eukaryotic 5S ribosomal loop E*. Thesis, University of California, Berkeley.
Wimberly BT, Brodersen DE, Clemons WM Jr, Morgan-Warren RJ, Carted AP, Vonrhein C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. *Nature 407*:327–339.
Wüthrich K. 1986. *NMR of proteins and nucleic acids.* New York: John Wiley & Sons, Inc.
Yarus M. 1988. A specific amino acid binding site composed of RNA. *Science 240*:1751–1758.
Yarus M. 1989. Specificity of arginine binding by the *Tetrahymena* intron. *Biochemistry 28*:980–988.
Yarus M, Illangesekare M, Christian E. 1991. An axial binding site in the *Tetrahymena* precursor RNA. *J Mol Biol 222*:995–1012.
Yarus M, Majerfeld I. 1992. Co-optimization of ribozyme substrate stacking and L-arginine binding. *J Mol Biol 225*:945–949.