J|A|C|S COMMUNICATIONS

Published on Web 03/07/2006

Conformation of the Group II Intron Branch Site in Solution

Jörg C. Schlatterer, Samuel H. Crayton, and Nancy L. Greenbaum*

Department of Chemistry and Biochemistry, Florida State University, 214 DLC, Tallahassee, Florida 32306-4390

Received November 20, 2005; E-mail: nancyg@chem.fsu.edu

Group II introns are self-splicing RNAs found in certain bacteria and in organellar genes of plants, fungi, and yeast.[1−3] These ribozymes employ a two-step transesterification mechanism essentially identical to that of the spliceosome, a dynamic ribonucleoprotein complex that splices eukaryotes pre-mRNA transcripts,[3,4] and there is limited secondary structural similarity between RNA components of the two. These findings fuel the speculation that the two systems share a common evolutionary ancestor.[5−7]

The secondary structure of Group II introns is organized into six domains.[8,9] Activity is a result of Mg^{2+}-dependent tertiary interdomain contacts.[3,10,11] Domain 6 (D6) contributes the nucleophile for the first transesterification in the form of a ribose 2′OH, typically that of a conserved adenosine, called the branch site (Figure 1A).[12] It is possible that, in the folded, active Group II intron, the branch site region is recognized and positioned for catalysis only after formation of long-range interactions facilitated by divalent metal ions. Therefore, the features by which the branch site region of the native D6 is recognized are important.

Biochemical studies have revealed features of the Group II intron branch site.[13,14] Terbium cleavage and phosphoro-thioate interference studies of D6 suggest metal ion binding within the GNRA tetraloop,[14] the internal loop,[14,15] and the branch site region[14] (Figure 2A). Also, mutational analysis of D6 indicates the importance of the internal loop in splicing activity.[16] Crystallographic studies of a highly modified D5−D6 construct depict both the branch site A and the 3′ adjacent U in an extrahelical position.[17] However, it is possible that this conformation is buttressed by crystal contacts with an RNA strand cut as one of the modifications.

Here, we report spectroscopic evidence that, in solution, the branch site adenosine of the ai5γ Group II intron is in an extrahelical position, and the two flanking nucleotides stack intrahelically, analogous to the conformation observed in the branch site helix of the spliceosome.[18] We also find that the structure of the branch site motif depends on the presence of the internal loop of D6.

To investigate the position of the branch site residue, we utilized 2-aminopurine (2ap), a fluorescent analogue of adenine,[19] as a structural probe because it fluoresces when exposed to solvent and emission is quenched upon stacking. When 2ap was substituted for the branch site adenine in the full sequence, $D6_{2ap}$ (Figure 1B), we observed ∼9-fold greater fluorescence than for a complementary helix, $D6_{comp}$ (Figure 1C), a result similar to that found for a duplex representing the spliceosomal branch site helix[18] (Supporting Information). In the latter helix, an NMR-derived structure shows the branch site adenosine extruded from the helix, stabilized in the minor groove by formation of a base triple. Addition of 5 mM Mg^{2+} or Ca^{2+} to $D6_{2ap}$ resulted in ∼10% increase in fluorescence (Figure 1C). Fluorescence of 2ap in an analogous hairpin in which the internal loop was deleted ($D6_{2ap,del}$) was only ∼6 times that of the complementary helix, and, in contrast with the enhancement seen in $D6_{2ap}$, addition of 5 mM Mg^{2+} to $D6_{2ap,del}$ resulted in an ∼8% decrease in fluorescence. These data imply that the branch site nucleoside of D6 is exposed to solvent and offers its 2′OH for branching. The internal loop and physiological concentrations of divalent metal ions facilitate this position.

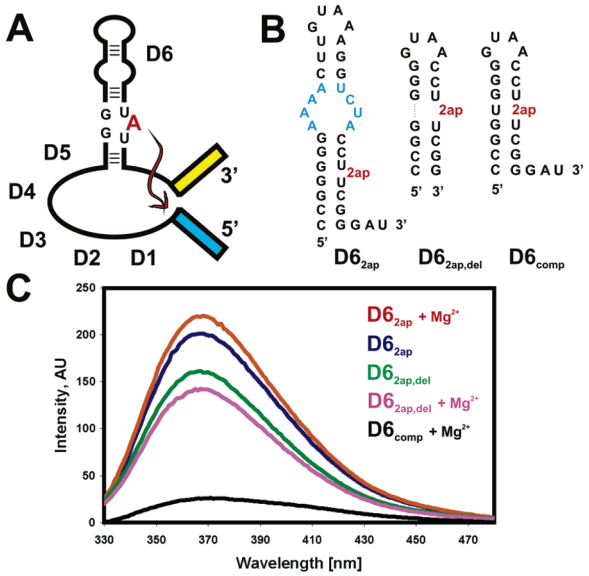

Figure 1. (A) Scheme of an ai5γ Group II intron (branch site A in red). (B) Constructs for fluorescence assays from left to right: full-length D6 ($D6_{2ap}$), deletion of the internal loop ($D6_{2ap,del}$), and the complementary sequence ($D6_{comp}$). (C) Fluorescence spectra of $D6_{2ap}$, $D6_{2ap,del}$, and $D6_{comp}$ ± 5 mM Mg^{2+}.

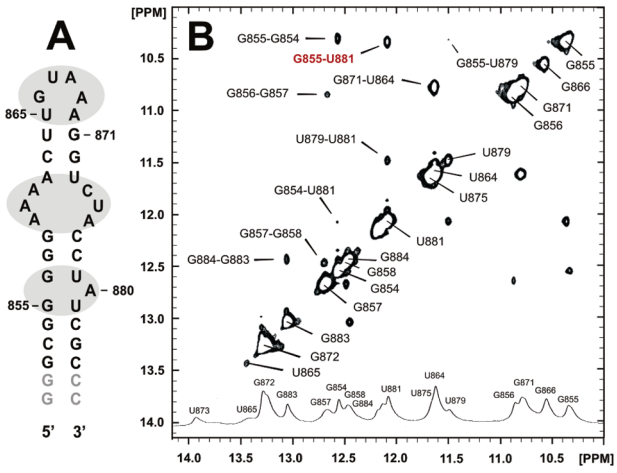

Figure 2. (A) D6 constructs for NMR investigations: 33nt sequence in black was used for homonuclear experiments; 37nt RNA sequence (with gray GC base pairs) included ^{15}N-labeled G's and U's. The shading indicates the regions of resonances impacted by Mg^{2+} and Ca^{2+}. Residues are numbered according to their position in the intact intron. (B) Region of a 2D NOESY spectrum of exchangeable protons showing imino−imino cross-peaks.

To evaluate local features of the branch site motif, we chose a 33nt D6 construct for solution NMR studies. Some transcribed samples had two additional GC base pairs at the termini to improve

10.1021/ja0578754 CCC: $33.50

transcriptional yield (Figure 2A). The presence of a single fold was verified by nondenaturing PAGE, and the hairpin structure was identified via NOE cross-peaks of the tetraloop.[20] Exchangeable protons were assigned from 2D ^{1}H,^{1}H NOESY spectra of 33nt D6, and from ^{1}H−^{15}N HSQC spectra of 37nt D6 transcribed with ^{15}N-GTP and ^{15}N-UTP. Base pair assignments confirmed six GC, two AU, and two GU base pairs.

1D ^{1}H NMR spectra indicated no major conformational change upon addition of Mg^{2+} and Ca^{2+} up to 5 mM, although the resonances of several imino protons attributed to bases within the loops and branch site region shifted upon alkaline earth metal ion addition. In particular, we observed upfield shifts for G855 (by 0.12 ppm), U881 (0.06 ppm), and U871 (0.01 ppm) and downfield shifts for G856 (by 0.18 ppm), U879 (0.03 ppm), and G866 (0.07 ppm) (see Supporting Information). Mg^{2+} and Ca^{2+} induced almost identical resonance shifts (±0.002 ppm), except for the imino protons of G854 and G856, the bases flanking the branch site, for which Mg^{2+} induced larger shifts (by 0.04 and 0.07 ppm, respectively) than the shifts observed with Ca^{2+}. These NMR data support binding of metal ions in the proposed regions;[14] enhancement of $D6_{2ap}$ fluorescence in the presence of Mg^{2+} and Ca^{2+} further supports the conclusion of a metal ion-binding site within the branch site region.

We then asked about intra- vs extrahelical conformation of the bases neighboring the branch site. We identified NOEs corresponding to two GU pairs (Figure 2B) in the imino region of the D6 spectrum. The strong NOE (T_{mix} = 80−300 ms) corresponding to the G871−U864 pair near the hairpin loop was unequivocally assigned from resonance shifts seen when the adjacent U−A base pair was switched to A−U. An imino−imino cross-peak between G855 and U881 indicates base pair formation between these bases. There was no NOE between G856 and U879, but the upfield location of each broad imino resonance peak implies protection from rapid exchange, with or without hydrogen bond formation, consistent with an intrahelical position. A strong $n \rightarrow n + 2$ NOE between imino protons of U879 and U881 indicates stacking of the bases flanking the branch site adenosine.

By comparison with those of D6, 1D ^{1}H spectra of $D6_{del}$ exhibit an upfield shift for G856h1 at 10.18 ppm (vs 10.86 ppm in D6) and a further upfield shift of 0.05 ppm in the presence of Mg^{2+}. We see a strong NOE between $D6_{del}$ G856h1 and a nonexchangeable resonance at 6.21 ppm. Assignment of this resonance to A880h2 was based on (1) systematic exclusion of all H1′, H5, H6, and H8 protons of bases in that vicinity by previous assignment; (2) appearance of the resonance peak in D_2O; and (3) perturbation of this resonance location upon substitution of the branch site adenosine by 2ap. The NOE suggests a base pair between G856 and the branch site adenosine in $D6_{del}$. No $n \rightarrow n + 2$ NOEs between U879 and U881 were observed, as when the internal loop was present. These results agree strongly with the 2ap fluorescence data and imply that the branch site adenosine of $D6_{del}$ is at least partly sandwiched between the neighboring nucleosides and less exposed to solvent than in D6. Mutational analysis indicated that the internal loop enhances the efficiency and fidelity of branching.[16] Both NMR and fluorescence results shown here confirm that the internal loop contributes directly to the conformation and divalent ion binding properties and, thereby, recognition of the ai5γ Group II intron branch site prior to splicing.

The conformation of the branch site motif as modeled in the crystallographic structure[17] contrasts with structural models for both the spliceosomal[18] and Group I self-splicing introns:[21,22] in both of these cases, positioning of the nucleophilic nucleoside is achieved by formation of a base triple. Our current results support a model in which the branch site adenosine of the ai5γ Group II intron is in an extrahelical position, flanked by intrahelically stacked uridines. This conformation is analogous to that observed in the spliceosomal helix,[15] lending further support to the conjecture of common ancestry. The internal loop and divalent metal ions play an important role in its structure and may assist in positioning of the nucleophilic adenosine with its 2′OH.

Acknowledgment. We thank the NMR facilities at the National High Magnetic Field Laboratory (Tallahassee, FL) and the Department of Chemistry and Biochemistry, Florida State University, and the Biochemical Analysis Sequencing & Synthesis laboratory. This research was funded by NIH grant RO1 GM54008 and NSF grant MCB 0316494 (to N.L.G.).

Supporting Information Available: Experimental procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Michel, F.; Umesono, K.; Ozeki, H. *Gene* **1989**, *82*, 5−30.
(2) Bonen, L.; Vogel, J. *Trends Genet.* **2001**, *17*, 322−331.
(3) Michel, F.; Ferat, J. L. *Annu. Rev. Biochem.* **1995**, *64*, 435−461.
(4) Sontheimer, E. J.; Gordon, P. M.; Piccirilli, J. A. *Genes Dev.* **1999**, *13*, 1729−1741.
(5) Sharp, P. A. *Cell* **1985**, *42*, 397−400.
(6) Cech, T. R. *Cell* **1986**, *44*, 207−210.
(7) Jacquier, A. *Trends Biochem. Sci.* **1990**, *15*, 351−354.
(8) Michel, F.; Jacquier, A.; Dujon, B. *Biochimie* **1982**, *64*, 867−881.
(9) Schmelzer, C.; Schmidt, C.; Schweyen, R. J. *Nucleic Acids Res.* **1982**, *10*, 6797−6808.
(10) Costa, M.; Michel, F. *Embo J.* **1995**, *14*, 1276−1285.
(11) Costa, M.; Deme, E.; Jacquier, A.; Michel, F. *J. Mol. Biol.* **1997**, *267*, 520−536.
(12) Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H. L. *Cell* **1986**, *44*, 213−223.
(13) Boudvillain, M.; Pyle, A. M. *Embo J.* **1998**, *17*, 7091−7104.
(14) Sigel, R. K.; Vaidya, A.; Pyle, A. M. *Nat. Struct. Biol.* **2000**, *7*, 1111−1116.
(15) Chanfreau, G.; Jacquier, A. *Science* **1994**, *266*, 1383−1387.
(16) Chu, V. T.; Adamidi, C.; Liu, Q.; Perlman, P. S.; Pyle, A. M. *Embo J.* **2001**, *20*, 6866−6876.
(17) Zhang, L.; Doudna, J. A. *Science* **2002**, *295*, 2084−2088.
(18) Newby, M. I.; Greenbaum, N. L. *Nat. Struct. Biol.* **2002**, *9*, 958−965.
(19) Menger, M.; Tuschl, T.; Eckstein, F.; Porschke, D. *Biochemistry* **1996**, *35*, 14710−14716.
(20) Jucker, F. M.; Heus, H. A.; Yip, P. F.; Moors, E. H.; Pardi, A. *J. Mol. Biol.* **1996**, *264*, 968−980.
(21) Golden, B. L.; Gooding, A. R.; Podell, E. R.; Cech, T. R. *Science* **1998**, *282*, 259−264.
(22) Guo, F.; Gooding, A. R.; Cech, T. R. *Mol. Cell* **2004**, *16*, 351−362.

JA0578754