Structural Insights into Group II Intron Catalysis and Branch-Site Selection

Lan Zhang[1] and Jennifer A. Doudna[1,2]*

Group II self-splicing introns catalyze autoexcision from precursor RNA transcripts by a mechanism strikingly similar to that of the spliceosome, an RNA-protein assembly responsible for splicing together the protein-coding parts of most eukaryotic pre-mRNAs. Splicing in both cases initiates via nucleophilic attack at the 5′ splice site by the 2′ OH of a conserved intron adenosine residue, creating a branched (lariat) intermediate. Here, we describe the crystal structure at 3.0 Å resolution of a 70-nucleotide RNA containing the catalytically essential domains 5 and 6 of the yeast ai5γ group II self-splicing intron, revealing an unexpected two-nucleotide bulged structure around the branch-point adenosine in domain 6.

Removal of intron RNA from coding sequences is an essential step in the maturation of eukaryotic pre-messenger RNAs. Group II self-splicing introns, found in organellar genes of yeast, plants, and fungi and in eubacteria, catalyze intron excision by a two-step mechanism analogous to that of the spliceosome, a dynamic RNA-protein complex that splices nuclear pre-mRNA transcripts (*1, 2*). In the first step, the 2′ hydroxyl group of a conserved intron adenosine residue attacks the phosphodiester bond at the 5′ splice site, producing a branched (lariat) form of the intron that contains a 2′-5′ linkage and releasing the 5′ exon (*3*). The second step involves exon ligation and removal of the intron lariat. Although the branch-point adenosine in both group II introns and pre-mRNAs is predicted to be unpaired, a lack of high-resolution structures for group II introns has precluded an understanding of branch-point selection or the possible structural similarity between self-splicing and spliceosomal RNAs. Here, we describe a crystal structure and chemical probing and functional assays that suggest that the branch point of the ai5γ group II intron consists of a two-nucleotide bulged conformation, providing a structural explanation for branch-site choice. Revealing the molecular structure of a group II intron–related RNA, this work provides a basis for detailed comparison of self-splicing and spliceosomal catalytic mechanisms.

Group II introns share a common secondary structure consisting of six domains (*1*) (Fig. 1A). Domain 5 binds domain 1 extensively to form the catalytic core (*4–6*) and contains a universally conserved AGC sequence essential for splicing activity (*7*). Domain 6 contains the branch-point adenosine whose 2′ hydroxyl serves as the nucleophile during the first step of self-splicing. Domains 5 and 6 are able to catalyze 5′ junction cleavage and lariat formation when added to domain 1 in trans, with reaction kinetics similar to those determined for splicing by the intact intron (*3, 8, 9*). Moreover, these two domains share sequence and secondary structural features with U6 small nuclear RNA (snRNA) and base-paired U2 snRNA/pre-mRNAs, respectively (*2*). Thus, domains 5 and 6, spanning an ~70-nucleotide region of the intron, are attractive initial targets of structural investigation because of their importance in forming the intron active site and presenting the branch-point adenosine for its role in nucleophilic attack.

Although constructs of the wild-type domains 5 and 6 (d56) from the ai5γ intron failed to crystallize, perhaps because of inherent flexibility, crystals were obtained of a modified construct (d56-TR), in which domains 5 and 6 are coaxially stacked and the loop nucleotides of domain 6 are deleted (*10*) (Fig. 1B). A tetraloop receptor sequence, inserted between the two domains, mimics the tetraloop receptor in domain 1 that docks with the tetraloop of domain 5—the ζ-ζ′ interaction (Fig. 1A)—potentiating intermolecular interactions in the crystal (*4, 10*). This construct was shown to be active in the hydrolytic pathway for group II intron splicing in trans (*10*). Furthermore, addition of the domain 6 loop nucleotides into a similar construct restored full branching activity in trans, although it is not clear whether domain 5 and 6 in this construct are coaxially stacked (*11, 12*). D56-TR RNA prepared by in vitro transcription (*13*) was crystallized by modification of previous conditions to improve the diffraction limit to 3.0 Å resolution (*13*). To obtain phase information, bromine atoms were introduced into the RNA by incorporation of 5-bromo UTP during in vitro transcription (*13, 14*). Phases were determined by MAD using a four-wavelength x-ray diffraction dataset collected from a bromine-containing crystal; a native dataset was used for refinement (*13, 15*) (Table 1). In the

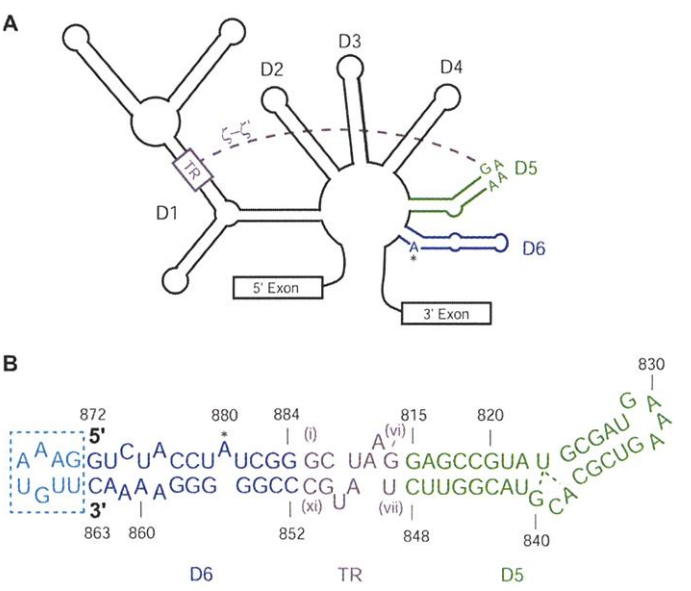

Fig. 1. Secondary structure of the group II intron and the crystallization construct. **(A)** Schematic diagram of the ai5γ group II intron from yeast mitochondria. The six helical domains emanating from the central wheel are called D1 to D6; 5′ and 3′ exons are represented by boxes. The long-range interaction ζ-ζ′ between the terminal GAAA tetraloop in domain 5 and a tetraloop receptor (TR, gray box) in domain 1, which is mimicked by the crystal contact, is denoted by a gray dashed line. The branch-point adenosine is marked with an asterisk (*). The color coding (domain 5, green; tetraloop receptor, gray; domain 6, blue) and the branch-point asterisk mark are the same for all other figures. **(B)** Sequence and predicted secondary structure of crystallization construct d56-TR. Domains 5 and 6 are coaxially stacked with the tetraloop receptor in between. Residues in domains 5 and 6 are numbered according to their positions in the intact intron—domain 5 from residue 815 to 848 and domain 6 from residue 852 to 884. Note that the terminal loop of domain 6 is removed (residues 864 to 871, shown in light blue), and the new 5′ end (G872) and 3′ end (C863) are indicated. Residues in the tetraloop receptor are indicated in roman numerals.

[1]Department of Molecular Biophysics and Biochemistry and [2]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA.

*To whom correspondence should be addressed. E-mail: jennifer.doudna@yale.edu

structure, d56-TR adopts a long-helical conformation (Fig. 2A). Domain 5 is largely A form, with a two-nucleotide A838/C839 bulge in the middle. This internal bulge, predicted from both biochemical and phylogenetic covariation data (*16, 17*), does not affect the overall helical axis. To our surprise, the branch site of domain 6 also contains a two-nucleotide bulge: an A880/U881 bulge is observed instead of the proposed single-nucleotide A880 bulge (Fig. 2A). Unlike the bulge in domain 5, the bases of the two nucleotides at the branch site are stacked. Furthermore, the two strands of domain 6 separate at the end closest to the deleted loop nucleotides: one strand folds back, and its penultimate residue (A862) stacks with the branch-point bases; the other strand makes a crystal contact by base-pairing with its symmetry mate. As intended, the tetraloop receptor docks with the domain 5 tetraloop of an adjacent molecule in the crystal lattice (*18*) and has the same conformation as seen in the group I intron P4-P6 domain crystal structure (*19*).

As do most ribozymes and RNA-protein catalysts, group II introns require divalent metal ions for both folding and catalysis (*20, 21*). Several structurally or catalytically important metal ion–binding sites have been mapped in domains 5 and 6 by different methods (*22–24*). In the experimental electron density map for d56-TR, two magnesium ions were modeled on the basis of coordination geometry, distance from RNA ligands, and temperature-factor refinement. One magnesium ion binds the phosphate oxygen of G833 near the GAAA tetraloop of domain 5; the other one binds at the phosphate oxygen of U(vii) in the tetraloop receptor (Fig. 2, A and B). Because the tetraloop receptor does not exist in the wild-type d56, the latter binding site is probably unrelated to intron catalysis. To identify additional metal ion–binding sites, native crystals were soaked with manganese and cobalt hexammine-containing solutions (*13*), which have been used as probes for nonhydrated magnesium and hydrated magnesium binding sites, respectively [for example (*19, 25*)]. Difference Fourier syntheses calculated from diffraction data from these crystals revealed a single manganese ion coordinated to N7 of adenosine 838 at the bulge in domain 5, whereas a cobalt hexammine ion bound the major groove O4 of guanosine 879 of a G-U wobble pair flanking the branch-point A (Fig. 2, A and B). Three of the metal ion–binding sites observed in the native and difference electron density maps—at the d5 terminal loop, d5 bulge, and d6 branch point—correlate well with previous Tb^{3+} cleavage and solution nuclear magnetic resonance (NMR) data (*22, 24*), suggesting that these represent magnesium ions that bind, and may stabilize, the ground-state intron RNA (Fig. 2, A and B). The latter two sites may not be occupied by magnesium ions in the native crystal because of the high salt concentration used for crystallization and the weaker binding affinities measured by solution NMR for these sites (*24*). Recent phosphorothioate substitution of pro-S_p oxygens in domain 5 revealed a structurally important metal ion–binding site in the conserved AGC triad (residue 816 to 818, Fig. 2B) (*23*). Its possible requirement for additional interactions with domain 1 of the intron may explain the absence of this metal ion in the d56-TR structure.

The two-nucleotide bulge in domain 5 is of particular interest because of its possible

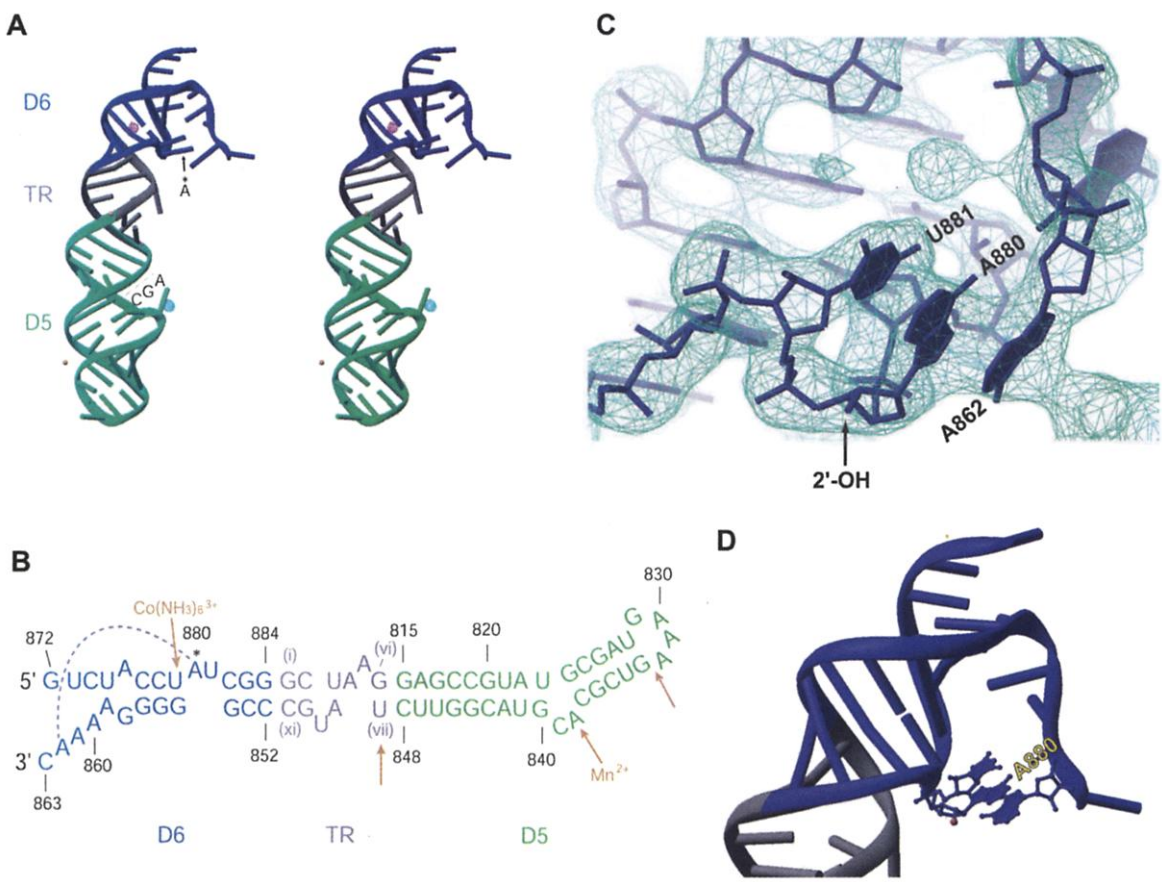

Fig. 2. Crystal structure of d56-TR. (**A**) Stereo ribbon diagram representation of the crystal structure of d56-TR. Green helix, domain 5; blue helix, domain 6; gray helix, tetraloop receptor; gold spheres, magnesium ions. Difference Fourier maps contoured at 5 σ are superimposed with the structure to show cobalt hexammine (in pink) and manganese ion (in cyan) binding sites. The branch-point A880 is marked with an asterisk (*) and the conserved AGC sequence (816 to 818) in domain 5 is also labeled. (**B**) Observed secondary structure of d56-TR. Color coding and sequence numbering are the same as in Fig. 1B. Predicted metal ion–binding sites, orange arrows; base stacking, gray dashed line. (**C**) Part of the $2F_o - F_c$ map, contoured at 1.5 σ, around the domain 6 branch-point region; the structural model of domain 6 is in blue. The stacking residues U881, A880 (and its 2′-OH group), and A862 are labeled. (**D**) Close-up view of domain 6 structure around the branch point. The structural model of domain 6 is represented by blue ribbons except for residues A880, U881, and A862. A880 and U881 form the double-bulged branch-point structure and stack with A862. The branch-point adenosine (A880) is labeled and its 2′-OH group is shown in red. Fig. 2, A, C, and D are created using RIBBONS (*37*).

roles both in binding to the rest of the group II intron domains and in catalysis (*6, 26, 27*). A C839/G840 bulge had been predicted until recently, when an alternative secondary structure containing an A838 and C839 bulge was proposed based on chemical probing and sequence covariation analysis (*16, 17*). In the d56-TR crystal structure, we observe the newly predicted U823-G840 wobble pair and the A838/C839 bulge, confirming the previous studies. The internal AC bulge bears increased similarity to the predicted secondary structure of base-paired U2/U6 snRNAs, which varies according to the distance between this bulge and the universally conserved AGC triad (*16, 17*). In the crystal structure, this particular spacing positions the AC bulge and the AGC sequence closer in space and aligns critical functional groups on the same "catalytic face" of domain 5 RNA (*17, 26, 27*) (Fig. 2A). Furthermore, a metal ion binds to the bulge in the crystal structure, and phosphorothioate substitution of the pro-Rp oxygen of C839 or of the equivalent residue in U6 snRNA blocks splicing activity (*27, 28*). One notable feature of the electron density map is the weak density and high temperature factors observed for the conserved AGC sequence compared with the rest of the structure (*18*). This observation is puzzling, because the sequence is embedded within the domain 5 stem and forms standard base pairs with the opposite GUU sequence. The unusual dynamics in this region suggests the potential for alternate conformations or interactions (perhaps with domain 1) that may facilitate catalysis. Moreover, the disorder of this region might also explain the lack of an observed metal ion–binding site (*23*).

The most unexpected feature in the crystal structure is the two-nucleotide bulge at the branch point in domain 6. The highly conserved stretch of purine residues opposite the branch site in domain 6 supports this double bulge by enabling alternative base pairings (Fig. 2B). Stacking of the branch-point adenosine, A880, with the 3′ nucleotide (U881) creates a local backbone geometry that exposes the 2′ hydroxyl group of the branch-point A to solvent, possibly facilitating its role as a nucleophile to attack the 5′ splice site (Fig. 2, C and D). We wondered whether the same conformation exists in the wild-type domain 6, and if so, whether this structure is required for splicing activity.

To test whether the two-nucleotide bulge around the branch point in domain 6 is a property of the RNA sequence or an artifact due to crystal packing forces or the design of the crystallization construct, an RNA construct containing the wild-type domains 5 and 6 (d56) (Fig. 3A) was probed by using 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluene sulfonate (CMCT). CMCT modifies unpaired uridines by reacting with the N3 group of the base, forming a covalent adduct that can be detected by primer extension using reverse transcriptase (*13*). To facilitate the primer extension assay, a d56 RNA construct was prepared containing an additional 15 nucleotides 3′ to the wild-type intron sequence to enable primer annealing (*13*). Activity assays (*13*) confirmed that this construct has wild-type branching activity in the presence of domain 1 (*18*). CMCT modification followed by primer extension using this construct revealed that U881, the predicted bulged residue in the wild-type d56, has a striking modification level of 10 times that of background (Fig. 3). The only other uridine to be significantly modified in the sequence is

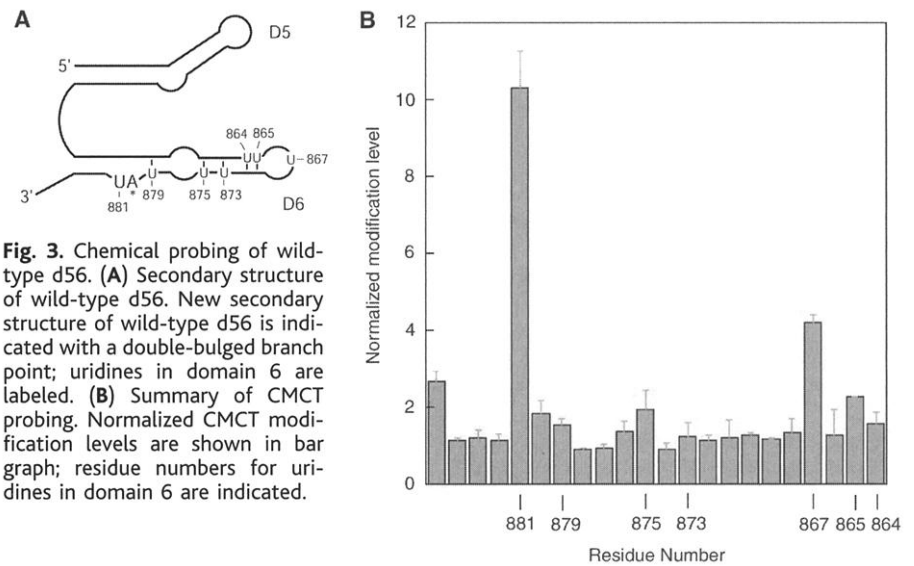

Fig. 3. Chemical probing of wild-type d56. (**A**) Secondary structure of wild-type d56. New secondary structure of wild-type d56 is indicated with a double-bulged branch point; uridines in domain 6 are labeled. (**B**) Summary of CMCT probing. Normalized CMCT modification levels are shown in bar graph; residue numbers for uridines in domain 6 are indicated.

Table 1. X-ray data collection, MAD phasing, and structure refinement. Space group $P6_122$, unit cell $a = b = 90.99$ Å, $c = 235.89$ Å (native). $a = b = 91.68$ Å, $c = 241.65$ Å (bromo). The mean figure of merit (FOM) after density modification was 0.9438 (solvent content = 83.3%). FOM is calculated from the cosine of the error in phase angle for a specific reflection. Refinement of a native data set, 1497 nonhydrogen atoms and two magnesium ions. Numbers after slash refer to the highest resolution bin.

Parameter	Native	Bromo at wavelength			
		λ1	λ2	λ3	λ4
		Crystallographic data			
Wavelength (Å)	1.1	0.9200	0.9194	0.9273	0.9020
Resolution (Å)	30–3.0	30–3.6	30–3.6	30–3.6	30-3.6
Completeness (%)	85.8/51.7	89.8/70.2	87.7/62.4	86.5/57.3	83.5/47.5
R_{sym}* (%)	5.4/43.0	4.9/28.7	4.4/30.8	3.7/32.4	4.0/34.0
I/δ	23.4/3.0	22.5/2.7	22.8/2.5	25.7/2.3	23.1/2.0
Redundancy	15.5	28.7	28.7	28.7	28.7
		MAD phase determination			
		λ1 (inflection)	λ2 (peak)	λ4 (hard remote)	
Resolution (Å)		30–3.6	30–3.6	30–3.6	
R_{cullis}†		0.7573	0.7225	0.7754	
Phasing power‡		0.8398	1.3290	1.0340	
		Refinement			
Resolution (Å)		30–3.0			
R_{cryst}§ (%)		24.8			
R_{free}§ (%)		26.8			
rmsd bonds (Å)		0.006591			
rmsd angles (°)		1.15807			
Average B-factor‖ (Å²)		99.55			

*$R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the statistically weighted absolute intensity of multiple measurements of symmetry related reflections. † $R_{cullis} = \Sigma\|F_{PH} \pm F_P| - F_H(calc)|/\Sigma|F_{PH}|$ reported for all centric reflections. ‡Phasing power = $\langle|F_H|\rangle/\langle\|F_P + F_H| - |F_{PH}\|\rangle$ reported for all reflections. §$R = \Sigma\| F_o - k|F_c\|/\Sigma|F_o|$, R_{cryst} from the working set and R_{free} from the test set. ‖Despite the high average B-factor, electron density for the RNA is of good quality (Fig. 2C). Similar values of average B-factor were obtained previously for structures of RNAs of similar size and diffraction resolution [e.g. (*39*)].

U867 (Fig. 3), which is expected to be exposed in the terminal GUAA tetraloop. Uridines in the sequence that are base-paired in the predicted structure are modified only slightly compared with the control (Fig. 3). This result supports a two-nucleotide bulge conformation in wild-type d56 RNA, including the branch-point A and the adjacent 3′ U, as observed in the d56-TR crystal structure. In similar CMCT modification experiments with the crystallization construct d56-TR, the uridine 3′ of the branch-point adenosine is strongly modified compared with the base-paired uridines in the RNA (*18*). These results suggest that a two-nucleotide bulge does exist in domain 6 in solution.

To determine whether the observed domain 6 structure is required for splicing activity, a set of domain 6 mutants were constructed in the context of the d56 RNA to test whether U879 base pairs with G856, as in the previously predicted model, or with G855 as in the crystal structure (Fig. 4A) (*13*). It has been shown that G-U wobble pairs are likely to flank the branch-point adenosine and to be important for splicing (*29, 30*), so we constructed a U879C point mutant to create a G-C base pair at this position (Fig. 4A). Two additional mutants of the d56 RNA were designed to restore wobble base-pairing geometry with the U879C substitution by changing additionally either G856 or G855 to A (Fig. 4A). According to the original model, the G856A change should restore the wobble geometry adjacent to the branch-point A by creating an A-C pair; however, the new model predicts that the G855A change will restore the wobble pair. The U879C mutation diminished branching in an in trans splicing assay (*13*), consistent with previous results suggesting that the wobble base-pair geometry at this position is important for lariat formation during the first step of splicing (Fig. 4) (*30*). The U879C/G856A double mutant likewise showed very low branching activity, comparable to that of the U879C mutant. However, the U879C/G855A double mutant showed significant branching activity compared with the other mutants, although not as efficient as the wild-type d56 RNA (Fig. 4). These data suggest that when domains 5 and 6 dock into the rest of intron, U879 is base-paired with G855, and U881 is instead likely to be exposed and to perform some other function. One explanation for the weak rescue effect is that the G-U base pair flanking the branch point, although important for its wobble geometry, makes sequence-specific contacts within the group II intron active site that are disrupted in the A-C mutant. In support of this idea, G855 is the most conserved residue in domain 6 other than the branch-point A (*31*). In addition, the resulting A-C wobble base pair is less stable, forming only one hydrogen bond at neutral pH at which we performed the splicing assays.

It is noteworthy that the nucleotide 3′ to the branch-point A is nearly always a pyrimidine in both group II introns and pre-mRNAs (*2, 31*). There is no rationale for this conservation based on the crystal structure of d56-TR, suggesting that this residue makes additional contacts during splicing. The conserved purine-rich 5′ strand of domain 6 makes alternative base pairings in this region possible, explaining the difficulty in predicting the secondary structure based on covariation analysis (*31*). Might a similar structure occur in the spliceosome? U2 snRNA lacks a purine-rich sequence analogous to that in domain 6, and thus does not support alternative base pairings that can occur in the group II intron domain 6. In the crystal structure of a self-complementary oligonucleotide designed to mimic a yeast U2 snRNA/pre-mRNA complex, the adenosine 5′ to the branch-point A was bulged out of the helix and contacted an adjacent helix in the crystal lattice (*32*). It is unclear how this structure relates to the branch-point conformation in the spliceosome. However, the fact that the base-pairing between U2 snRNA and pre-mRNA is weak downstream of the branch-point adenosine suggests that the two-nucleotide bulge model is not inconsistent with the currently available sequences of spliceosomal and pre-mRNAs (*2, 33–36*).

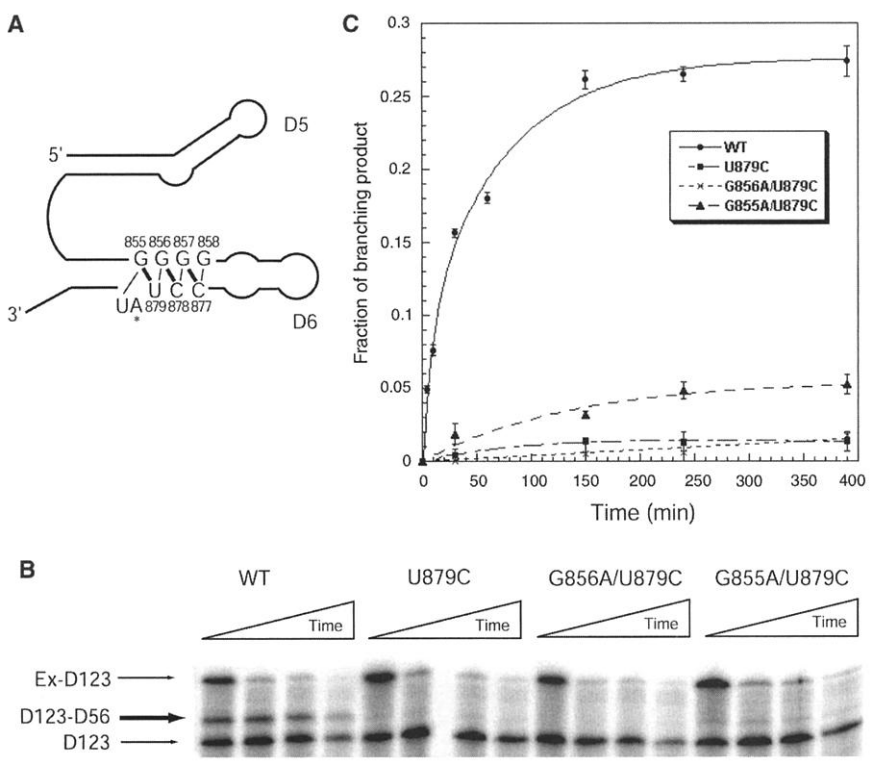

Fig. 4. Branching activity of wild-type and mutant d56 constructs assayed using an in trans splicing system. (**A**) Secondary structure of wild-type d56. New secondary structure of wild-type d56 is indicated with a double-bulged branch point and new base-pairing scheme. Base-pairing according to the new prediction is indicated by thick black lines and the old prediction, by thin lines. (**B**) Comparison of splicing activities of wild-type d56 and d56 mutants. Reactions were incubated under branching conditions for the times indicated. Precursor (Ex-D123) and hydrolysis product (D123) are indicated by thin arrows. Branching product (D123-D56) is indicated by a thick arrow. (**C**) Summary of the splicing activities. Fraction of branching product was plotted as a function of time, and curves were fit as a biphasic reaction described previously (*38*). Wild-type d56 activity is shown with solid circle, U879C mutant with solid square, G855A/U879C double mutant with solid triangle and G856A/U879C double mutant with x.

References and Notes

1. F. Michel, J. L. Ferat, *Annu. Rev. Biochem.* **64**, 435 (1995).
2. C. A. Collins, C. Guthrie, *Nature Struct. Biol.* **7**, 850 (2000).
3. C. L. Peebles *et al.*, *Cell* **44**, 213 (1986).
4. M. Costa, F. Michel, *EMBO J.* **14**, 1276 (1995).
5. M. Boudvillain, A. M. Pyle, *EMBO J.* **17**, 7091 (1998).
6. M. Boudvillain, A. de Lencastre, A. M. Pyle, *Nature* **406**, 315 (2000).
7. C. L. Peebles, M. Zhang, P. S. Perlman, J. S. Franzen, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4422 (1995).
8. K. Chin, A. M. Pyle, *RNA* **1**, 391 (1995).
9. A. M. Pyle, J. B. Green, *Biochemistry* **33**, 2716 (1994).
10. A. R. Ferre-D'Amare, K. Zhou, J. A. Doudna, *J. Mol. Biol.* **279**, 621 (1998).
11. M. Podar, P. S. Perlman, *RNA* **5**, 318 (1999).
12. In a construct d56-TRB, nucleotides 864 to 871 of domain 6 were included and the new 5′ and 3′ ends were at residue 815 at the beginning of domain 5 and at residue G(vi) of the tetraloop receptor, respectively (see Fig. 1B). Branching efficiency was comparable to that of wild-type d56 (*13*).
13. Supplementary material is available on *Science* On-

line at www.sciencemag.org/cgi/content/full/1069268/DC1.
14. C. Baugh, D. Grate, C. Wilson, *J. Mol. Biol.* **301**, 117 (2000).
15. W. A. Hendrickson, *Science* **254**, 51 (1991).
16. M. Costa, E. L. Christian, F. Michel, *RNA* **4**, 1055 (1998).
17. B. B. Konforti *et al.*, *Mol. Cell.* **1**, 433 (1998).
18. L. Zhang, J. Doudna, unpublished observations.
19. J. H. Cate *et al.*, *Science* **273**, 1678 (1996).
20. E. J. Sontheimer, P. M. Gordon, J. A. Piccirilli, *Genes Dev.* **13**, 1729 (1999).
21. A. M. Pyle, *Science* **261**, 709 (1993).
22. R. K. Sigel, A. Vaidya, A. M. Pyle, *Nature Struct. Biol.* **7**, 1111 (2000).
23. P. M. Gordon, J. A. Piccirilli, *Nature Struct. Biol.* **8**, 893 (2001).
24. R. K. Sigel, A. M. Pyle, personal communication.
25. J. H. Cate, R. L. Hanna, J. A. Doudna, *Nature Struct. Biol.* **4**, 553 (1997).
26. D. L. Abramovitz, R. A. Friedman, A. M. Pyle, *Science* **271**, 1410 (1996).
27. G. Chanfreau, A. Jacquier, *Science* **266**, 1383 (1994).
28. S. L. Yean, G. Wuenschell, J. Termini, R. J. Lin, *Nature* **408**, 881 (2000).
29. P. Z. Qin, A. M. Pyle, *Curr. Opin. Struct. Biol.* **8**, 301 (1998).
30. V. T. Chu, Q. Liu, M. Podar, P. S. Perlman, A. M. Pyle, *RNA* **4**, 1186 (1998).
31. V. T. Chu, C. Adamidi, Q. Liu, P. S. Perlman, A. M. Pyle, *EMBO J.* **20**, 6866 (2001).
32. J. A. Berglund, M. Rosbash, S. C. Schultz, *RNA* **7**, 682 (2001).
33. J. Wu, J. L. Manley, *Genes Dev.* **3**, 1553 (1989).
34. Y. Zhuang, A. M. Weiner, *Genes Dev.* **3**, 1545 (1989).
35. C. C. Query, M. J. Moore, P. A. Sharp, *Genes Dev.* **8**, 587 (1994).
36. R. Parker, P. G. Siliciano, C. Guthrie, *Cell* **49**, 229 (1987).
37. M. Carson, *J. Appl. Crystallogr.* **24**, 958 (1991).
38. D. L. Daniels, W. J. Michels, A. M. Pyle, *J. Mol. Biol.* **256**, 31 (1996).
39. A. R. Ferre-D'Amare, K. Zhou, J. A. Doudna, *Nature* **395**, 567 (1998).

40. The authors wish to thank A. Ferre-D'Amare for providing plasmid pD5627 and initiating the work, A. M. Pyle and R. Sigel for providing plasmid pJD13′-673 and sharing NMR data, the Yale Center for Structural Biology for crystallographic and computer support, J. Kieft, D. Battle, J. Murray, S. Szep, and the staff at the National Synchrotron Light Source for help with data collection. We also thank A. M. Pyle, J. Steitz, D. Rio, D. Battle, P. Adams, and B. Sagar for critical reading of this manuscript and insightful discussions, as well as C. Guthrie, R. Batey, and members of the Doudna laboratory for helpful discussions. This work was funded in part by a grant from the NIH. Coordinates of d56-TR have been deposited in the Protein Data Bank (RCSB ID code RCSB015439 and PDB ID code 1KXK).

21 December 2001; accepted 4 February 2002
Published online 21 February 2002;
10.1126/science.1069268
Include this information when citing this paper.

Living with Lethal PIP3 Levels: Viability of Flies Lacking PTEN Restored by a PH Domain Mutation in Akt/PKB

Hugo Stocker,[1] Mirjana Andjelkovic,[2]* Sean Oldham,[1] Muriel Laffargue,[3] Matthias P. Wymann,[3] Brian A. Hemmings,[2] Ernst Hafen[1]†

The phosphoinositide phosphatase PTEN is mutated in many human cancers. Although the role of PTEN has been studied extensively, the relative contributions of its numerous potential downstream effectors to deregulated growth and tumorigenesis remain uncertain. We provide genetic evidence in *Drosophila melanogaster* for the paramount importance of the protein kinase Akt [also called protein kinase B (PKB)] in mediating the effects of increased phosphatidylinositol 3,4,5-trisphosphate (PIP3) concentrations that are caused by the loss of PTEN function. A mutation in the pleckstrin homology (PH) domain of Akt that reduces its affinity for PIP3 sufficed to rescue the lethality of flies devoid of PTEN activity. Thus, Akt appears to be the only critical target activated by increased PIP3 concentrations in *Drosophila*.

Mutations in the tumor suppressor gene *PTEN* (the phosphatase and tensin homolog on chromosome 10) are frequent in glioblastomas, endometrial carcinoma, melanomas, and prostate cancer (*1*). Furthermore, two dominant hamartoma syndromes, Cowden disease and Bannayan-Zonana syndrome, are linked to germ line mutations in *PTEN* (*1*). The PTEN protein carries a phosphatase domain resembling those of dual-specificity protein phosphatases (*2–4*). Although it can dephosphorylate protein substrates such as focal adhesion kinase (*5*) and the adapter protein Shc (*6*), PTEN's predominant enzymatic activity appears to be the dephosphorylation of phosphoinositides at the D3 position. Because PTEN uses the second messenger PIP3 as a substrate, PTEN antagonizes the function of phosphatidylinositol-3 kinase (PI3K) (*7*, *8*). Immortalized mouse embryonic fibroblasts or embryonic stem cells lacking PTEN function show an approximately twofold increase in PIP3 concentrations (*9*, *10*). PIP3 interacts with a wide variety of PH domain-containing proteins, including the serine-threonine kinases Akt (also called PKB) and phosphoinositide-dependent kinase 1 (PDK1), Btk family tyrosine kinases, guanine nucleotide exchange factors for the Rho and Arf families of small guanosine triphosphatases, and phospholipase Cγ (*11*, *12*). The plethora of proteins that are potentially regulated by PIP3 provides widespread signaling potential for this lipid second messenger.

Genetic analyses in model organisms have implicated PTEN as a negative regulator of insulin receptor signaling. In the nematode *Caenorhabditis elegans*, PTEN antagonizes the activity of the PI3K AGE-1 in the regulation of metabolism, development, and life span (*13–16*). In the fruit fly *Drosophila melanogaster*, PTEN counteracts signaling downstream of the insulin receptor to control cellular growth (*17–19*). There are, however,

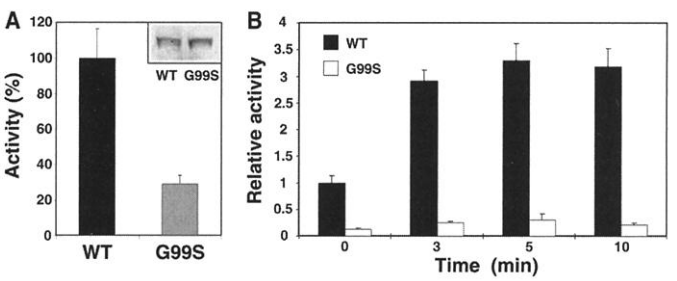

Fig. 1. Reduced kinase activity caused by an amino acid substitution in the PH domain of dAkt. (**A**) Effect of the G99S substitution in the PH domain on dAkt kinase activity from larval extracts (*42*). Activity from wild-type larvae was considered to be 100%. Inset, dAkt protein was detected in 40 μg of larval extracts using the same antiserum. (**B**) Reduced insulin-induced activation of the G99S mutant dAkt. The dAkt constructs were expressed in HEK 293 cells (*43*). Transfected cells were starved for 24 hours before stimulation with insulin for the indicated time periods, and dAkt kinase activity was determined (*44*). The activity of wild-type dAkt from unstimulated cells was considered to be relative activity = 1.

[1]Zoologisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. [2]Friedrich Miescher Institute, Maulbeerstrasse 66, CH-4058 Basel, Switzerland. [3]Université de Fribourg, Rue du Musée 5, CH-1700 Fribourg, Switzerland.

*Present address: Department of Vascular and Metabolic Diseases, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland.

†To whom correspondence should be addressed. E-mail: hafen@zool.unizh.ch