Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 11105–11109, December 1991
Biochemistry

Folding of group I introns from bacteriophage T4 involves internalization of the catalytic core

[RNA catalysis/ribozyme/RNA self-splicing/RNA structure/Fe(II)-EDTA]

TIMOTHY S. HEUER*, P. SCOTT CHANDRY†‡§, MARLENE BELFORT†, DANIEL W. CELANDER*, AND THOMAS R. CECH*¶

*Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215; †Molecular Genetics Program, Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, NY 12201-0509; and ‡Department of Microbiology and Immunology, Albany Medical College, Albany, NY 12208

Contributed by Thomas R. Cech, September 13, 1991

ABSTRACT **Fe(II)-EDTA, a solvent-based cleavage reagent that distinguishes between the inside and outside surfaces of a folded RNA molecule, has revealed some of the higher-order folding of the group IB intron from *Tetrahymena thermophila* pre-rRNA. This reagent has now been used to analyze the bacteriophage T4 *sunY* and *td* introns, both of which are members of the group IA subclass. Significant portions of the phylogenetically conserved secondary structure are protected from Fe(II)-EDTA cleavage. However, the P4 secondary structure element, which is substantially protected in the *Tetrahymena* intron, is available for cleavage in the two T4 introns. We conclude that a family of catalytic RNAs (ribozymes) that possess similar secondary structures and have similar activities fold into similar but nonidentical tertiary structures that nevertheless serve to internalize portions of the catalytic center. Furthermore, comparison of cleavage patterns of the *sunY* and *td* intron RNAs indicates that conserved nucleotides outside as well as within the catalytic core participate in the tertiary structure.**

Group I introns must fold into a catalytically active structure in order to mediate RNA self-splicing (1, 2). While the secondary structures of group I introns have been established by comparative sequence analysis (3–5), their higher-order folding is only beginning to be understood. Fe(II)-EDTA combines with oxygen to generate free radicals that promote oxidative damage to ribose moieties and result in nucleic acid strand scission (6–9). Because it has little base-sequence specificity and similar reactivity toward single- and double-stranded forms of RNA (10), Fe(II)-EDTA provides a useful probe for the higher-order folding of RNA. In a previous study (6) this methodology was used to analyze the native structure of the L−21 *Sca* I RNA, a shortened form of the *Tetrahymena* group IB intron that has enzymatic activity as a sequence-specific endoribonuclease (11). The RNA, in its Mg^{2+}-stabilized active conformation, was found to possess significant cleavage-resistant regions, including portions of elements P3, P4, and P7 that reside in the catalytic center of the ribozyme. Parallel studies with yeast $tRNA^{Phe}$, a molecule of known structure, supported the interpretation that protection from cleavage was related primarily to low solvent accessibility (6).

To assess the generality of the inside-outside folding pattern among group I introns, the bacteriophage T4 *sunY* (a split gene of unknown function) and *td* (thymidylate synthase) mRNA introns were selected for study. The *sunY* and *td* introns are self-splicing group IA introns and share common secondary structures as well as considerable primary sequence similarity (ref. 5; Fig. 1). Although the general secondary structure of group IA and IB introns is conserved in the core region consisting of P3, P4, P5, P6, P7, P8, and P9, the *sunY* and *td* intron RNAs differ significantly from the *Tetrahymena* intron RNA in primary sequence and in the number and arrangement of peripheral secondary structure elements that decorate the conserved core (ref. 5; Fig. 1).

METHODS

Plasmid Construction and RNA Synthesis. The wild-type *sunY* intron and exon coding sequences were subcloned into a pTZ19U plasmid, which places the *sunY* transcription unit under the transcriptional control of a phage T7 promoter. Exon I along with the first 13 base pairs at the 5′ end of the intron and 781 base pairs of the open reading frame extending from L9.1 were deleted by site-specific mutagenesis. The *td* L−7 ribozyme was derived from a pTZ*td*ΔP6-2 construct which is deleted for the intron open reading frame (13). Exon I along with the first 7 base pairs at the 5′ end of the intron were deleted to create a T7 transcription unit giving the 5′ end indicated. For both introns, the T7 run-off transcript underwent self-catalyzed hydrolysis at the 3′ splice site (14) to produce the ribozyme that was utilized for analysis.

Fe(II)-EDTA Cleavage. End-labeled RNA transcripts were gel-purified and ethanol-precipitated from 0.25 M NaCl/10 mM Tris·HCl, pH 8/1 mM EDTA. The RNA was then resuspended in 10 mM Tris·HCl, pH 8/1 mM EDTA. *sunY* RNA cleavage reaction mixtures contained 50 mM Tris·HCl (pH 7.4), 2 mM Na_2EDTA, 1 mM $(NH_4)_2Fe(SO_4)_2$, 5 mM dithiothreitol, and the indicated $MgCl_2$ concentration. The reaction mixtures were incubated for 90 min at 42°C. Cleavage of the *td* intron was performed similarly except that the reaction mixtures contained 50 mM Tris·HCl (pH 7.4), 4 mM Na_2EDTA, 2 mM $(NH_4)_2Fe(SO_4)_2$, and 10 mM dithiothreitol and were incubated at 42°C for 120 min. Control reactions were performed by incubation of the RNA in buffers containing 0 and 10 mM $MgCl_2$ at 42°C for 90 min (*sunY*) or 120 min (*td*) in the absence of Fe(II)-EDTA.

RESULTS AND DISCUSSION

DNA templates encoding shortened forms of the *sunY* and *td* introns that lack the 5′ splice site were constructed. The corresponding intron RNAs, whose 5′ ends comprise the 5′ exon-binding site (internal guide sequence; refs. 1 and 4), were transcribed using bacteriophage T7 RNA polymerase. Both RNAs were shown to be catalytically active in 10 mM $MgCl_2$ by two criteria. First, transcripts extending beyond the 3′-terminal guanosine of the intron underwent 3′ splice-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

§Present address: University of Melbourne, Russell Grimwade School of Biochemistry, Parkville, Victoria 3052, Australia.
¶To whom reprint requests should be addressed.

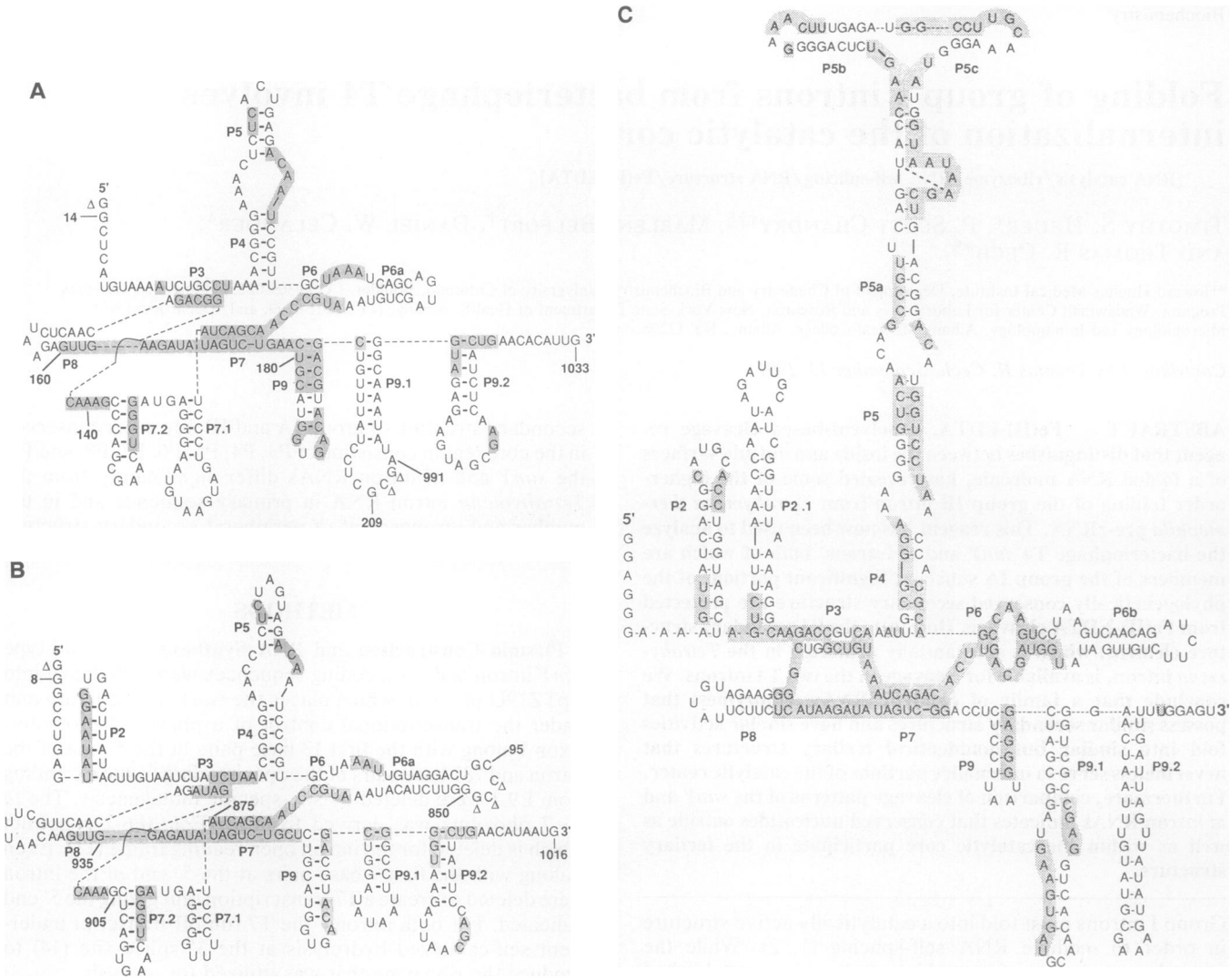

FIG. 1. Secondary structures and regions of protection from Fe(II)-EDTA cleavage for three ribozymes derived from group I introns. The *sunY* L−13 ribozyme (*A*), the *td* L−7 ribozyme (*B*), the *Tetrahymena* L−21 ribozyme (*C*) are shown with regions of protection from Fe(II)-EDTA cleavage shaded. In the L−*n* notation, *n* = number of nucleotides missing from the 5′ end of the intron. Triangles indicate foreign nucleotides at regions where open reading frames were deleted and a single guanosine at the 5′ end that was added for more efficient transcription. Paired regions are designated P, and joining regions (J) are numbered according to the paired elements they connect, 5′ to 3′ (e.g., J6/7 joins P6 and P7). Secondary structures for phage T4 introns are from ref. 5. For the *Tetrahymena* ribozyme, the secondary structure of P9 and the bulge in P5a have been redrawn (12) and the Fe(II)-EDTA protection was redetermined in the present study, based on analysis of both 5′- and 3′-end-labeled RNA (data not shown), to ensure that it was directly comparable to the phage introns. The map is similar but not identical to that published earlier (6). Protection from Fe(II)-EDTA cleavage was first assigned by visual criteria and then confirmed and quantified by phosphorimager or densitometric scanning (Figs. 2 and 3).

site hydrolysis (14). Second, the shortened introns had activity as sequence-specific endoribonucleases (11, 15) when supplied with exogenous guanosine and an oligoribonucleotide substrate complementary to the intron's 5′ exon-binding site (data not shown). These activities provide evidence that the Fe(II)-EDTA protection pattern observed for each intron corresponds to the active structure of the RNA. Although both introns are known to be highly efficient in self-splicing (13, 16–18), we have not measured the fraction of active molecules in the population.

Fe(II)-EDTA cleavage analyses were performed following pre equilibration of ^{32}P end-labeled RNAs in the presence and absence of 10 mM $MgCl_2$ (Figs. 2 and 3). For both RNAs, mapping was performed from the 5′ as well as the 3′ end and gave consistent results in the overlap region. A fairly uniform pattern of cleavage was observed when the *sunY* and *td* RNAs were reacted in the absence of $MgCl_2$. In the presence of 10 mM $MgCl_2$, however, significant regions of protection from Fe(II)-EDTA-induced cleavage were seen in both RNAs. Phosphorimager analysis of the radioactivity distribution for *sunY* showed that the regions of greatest protection (the valleys in the 10 mM $MgCl_2$ data in Fig. 2*B*) typically contained 3- to 4-fold less radioactivity than neighboring regions of strong cleavage (the peaks in the 10 mM $MgCl_2$ data), with a maximum difference of about 7-fold. This extent of protection is of the same order as that measured earlier for the *Tetrahymena* intron (average of 4-fold, maximum about 10-fold; ref. 6).

The location of each protected region was mapped by alignment of the products of Fe(II)-EDTA cleavage with RNA sequencing ladders on the same gel (Figs. 2*A* and 3*A*). Nucleotides found to be consistently protected in multiple experiments (three to five analyses of each position for both introns) are shaded in Fig. 1 *A* and *B*. Notably, most of the conserved secondary structural elements are similarly protected in both *sunY* and *td* intron RNAs. Within the conserved catalytic center, portions of P3 and all of P7 are protected from cleavage. The joining regions J5/4, J6/6a,

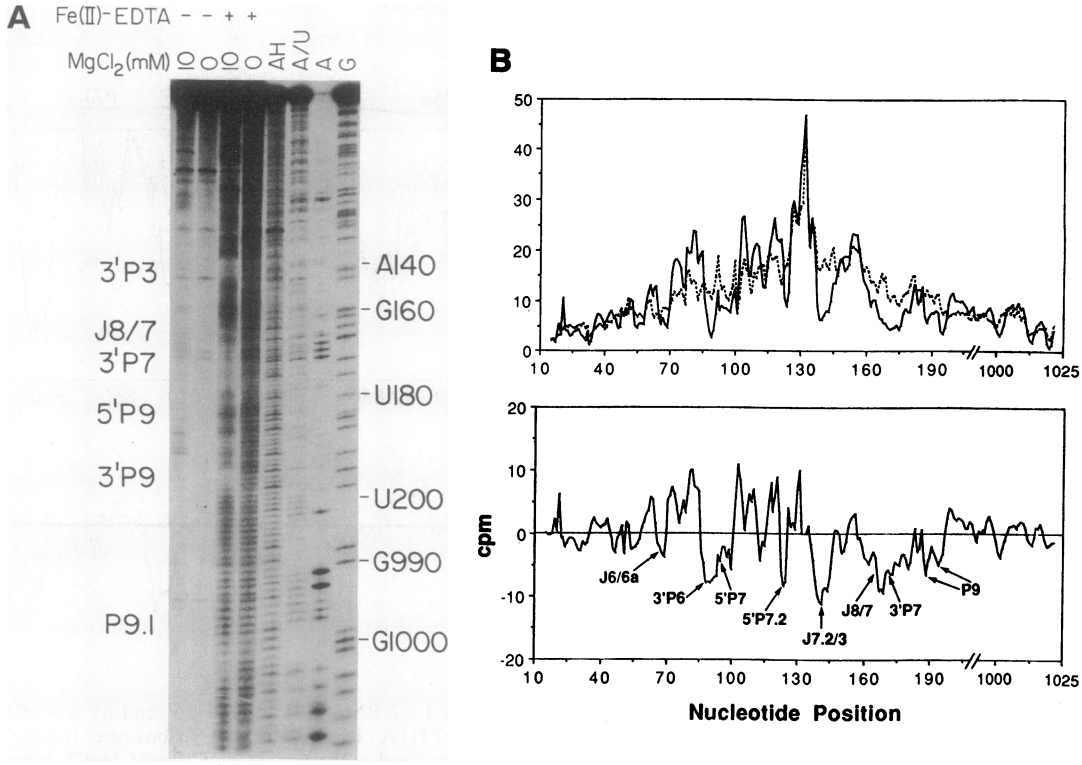

FIG. 2. Fe(II)-EDTA cleavage of the *sunY* ribozyme. (*A*) Autoradiogram of *sunY* L−13 RNA labeled at the 3′ end with 5′-[^{32}P]pCp. Lanes G, A, and A/U, sequencing standards generated by partial, base-specific cleavage of the RNA by endonucleases T_1, U_2, and PhyM, respectively (19). Lane AH, products of partial alkaline hydrolysis of the RNA. Nucleotide positions determined from these sequencing standards are indicated at right. (*B*) Quantitation of Fe(II)-EDTA cleavage. Radioactivity of polyacrylamide gels including that shown in *A* was quantitated using the Molecular Dynamics phosphorimager. All values have been corrected for background hydrolysis observed in the starting material and then normalized to adjust for loading differences. (*Upper*) An overlay plot of Fe(II)-EDTA cleavage in the absence of $MgCl_2$ (dashed line) and in the presence of 10 mM $MgCl_2$ (solid line). (*Lower*) The difference plot, cpm for 10 mM $MgCl_2$ data minus cpm for 0 mM $MgCl_2$ data. Negative peaks in the difference plot represent regions of protection.

J7.2/3, J8/7, and J7/9 are also protected in both RNAs.‖ The secondary structure elements along the periphery of the catalytic core, including P5, P6, P8, P9.1, and P9.2, show similar patterns of limited protection in both intron RNAs. The differences in protection from Fe(II)-EDTA cleavage observed between *sunY* and *td* intron RNAs are few and subtle.‖ In the *td* intron RNA an asymmetric portion of the P2 element, which is absent from the *sunY* intron, is protected from Fe(II)-EDTA cleavage. Additionally, the P9 element shows more limited protection in the *td* intron than in the *sunY* intron.

The group IA-specific region P7.2 is partially protected from cleavage in both phage introns, while P7.1 is not protected. We suggest that P7.2 has a similar packing environment in the two introns, perhaps adjacent to or within the catalytic core, whereas P7.1 protrudes into the solvent. Such an interpretation is consistent with other observations with the *td* intron: mutations in P7.2 disrupt function (16), whereas P7.1 can be substituted with the different P7.1 elements of *sunY* or *nrdB* without disrupting function (ref. 20; M. Bryk and M.B., unpublished work). The J7.2/3 sequence that connects P7.2 to the core is protected in both introns. It may reside within the core, perhaps being equivalent to J7/3 in the *Tetrahymena* intron.

‖The *sunY* and *td* Fe(II)-EDTA cleavage experiments were performed and the protection maps were constructed independently in two different laboratories. The *sunY* analysis was performed in Boulder, and the *td* analysis in Albany. This makes the similarities found in the two maps all the more noteworthy but also means that minor differences between them might result from small differences in methods.

The stem–loops interrupted by open reading frames (the end of P9.1 in *sunY* and P6a in *td*) are generally accessible to cleavage. This indicates that they are held away from the catalytic core, which could be advantageous in preventing them from interfering with folding of the active site.

Note that some tertiary structures might not result in protection from Fe(II)-EDTA cleavage. For example, the triple-helical scaffold proposed for the P4–P6 region (21) is probably not tested by our experiments. In DNA triplexes, at least two of the three strands can still be cleaved by free radicals produced by chelated Fe(II) (22), so it seems likely that an RNA triplex could remain cleavable by Fe(II)-EDTA if not involved in additional interactions that excluded it from the solvent.

The majority of the nucleotides protected from Fe(II)-EDTA cleavage in both the core and the peripheral regions of the *sunY* and *td* intron RNAs are conserved between the two sequences and indeed, in many cases also conserved in the third T4 intron, *nrdB* (Fig. 4). These nucleotides may well be conserved at least in part because of their contributions to tertiary structure. It remains possible that these nucleotides have some additional function *in vivo*, such as protein recognition or RNA stability.

The general pattern of protection from Fe(II)-EDTA cleavage described for the group IA *sunY* and *td* intron RNAs is consistent with that observed in the group IB *Tetrahymena* L−21 *Sca* I RNA (Fig. 1*C*). In all three RNAs, significant regions of protection are observed in the phylogenetically conserved core region of the ribozyme. In particular, all of the J6/7, P7, and J8/7 elements are protected in each RNA, as well as portions of P3, P6, J6/6a, and J6a/6. These results demonstrate the ability of these three molecules to fold into

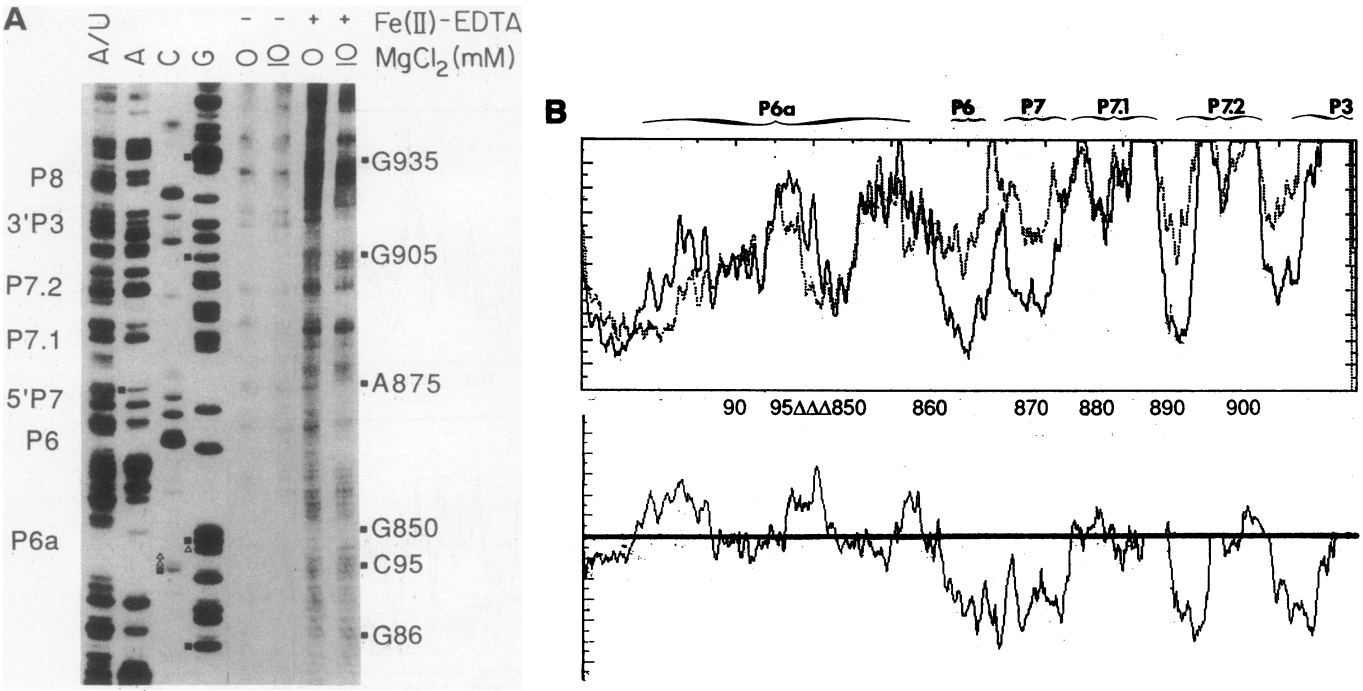

FIG. 3. Fe(II)-EDTA cleavage of the *td* ribozyme. (*A*) Autoradiogram of *td* L−7 RNA labeled at the 5′ end by use of [γ-^{32}P]ATP and polynucleotide kinase. Lanes are labeled as in Fig. 2. (*B*) Quantitation of Fe(II)-EDTA cleavage of *td* L−7 ribozyme. (*Upper*) An overlay of densitometric scans of Fe(II)-EDTA cleavage in the absence of $MgCl_2$ (dotted line) and in the presence of 10 mM $MgCl_2$ (solid line). (*Lower*) The difference plot as in Fig. 2*B*.

an inside-outside structure in which the catalytic center of the ribozyme is on the inside and suggest that this property may be conserved among self-splicing group I introns.

The most notable difference in protection from Fe(II)-EDTA cleavage observed in comparable elements of these three intron RNAs is in the P4 element of the core, which is substantially protected in the *Tetrahymena* L−21 *Sca* I RNA but accessible in the *sunY* and *td* intron RNAs. In the *Tetrahymena* intron, P4 is part of an independent folding domain that involves the P5abc extension (refs. 7 and 23; F. Murphy and T.R.C., unpublished work). In an RNA transcript that contains only this domain, protection of P4 from Fe(II)-EDTA cleavage requires the P5abc extension (F. Murphy and T.R.C., unpublished work). This extension

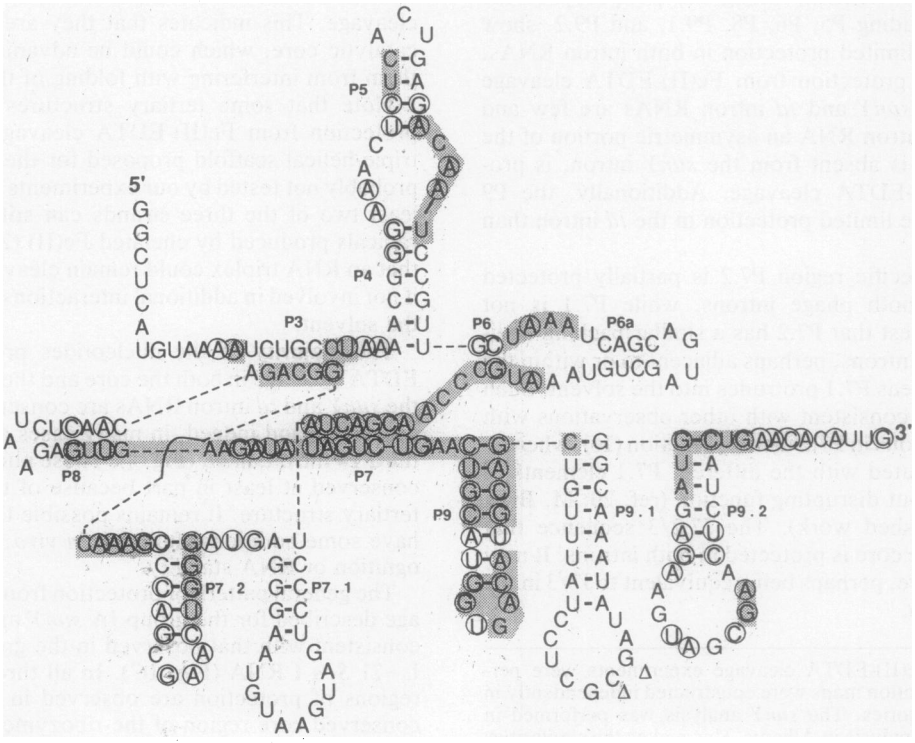

FIG. 4. Correlation between protected and conserved nucleotides in the phage T4 introns. Superimposed on the protection map of the *sunY* ribozyme are circles identifying those nucleotides conserved among the three phage introns in the *nrdB*, *sunY*, and *td* genes.

contributes to catalysis by stabilizing the core of the *Tetrahymena* intron (24, 25). Because P5abc is not found in the group IA introns, it had seemed possible that the group IA-specific P7.1 and P7.2 elements might also protect P4 to attain an equivalent structure. The accessibility of P4 in the phage introns provides physical evidence against this hypothesis. Instead, the differences in Fe(II)-EDTA protection between the two phage T4 introns and the *Tetrahymena* intron indicate that functionally equivalent tertiary structures can be formed in at least two different ways. Each of these two variations of a common inside-outside structure results in internalization of the catalytic center of the ribozyme.

We thank Alan Lambowitz and John Burke for helpful comments on the manuscript, Alice Sirimarco for manuscript preparation, and Tim Coetzee for help with Fig. 3. This research was supported by National Institutes of Health Grants GM28039 (T.R.C.) and GM39422 and GM44844 (M.B.) and by Damon Runyon–Walter Winchell Cancer Fund Postdoctoral Fellowship DRG-954 (D.W.C.). We also thank the W. M. Keck Foundation for generous support of RNA science on the Boulder campus.

1. Cech, T. R. (1990) *Annu. Rev. Biochem.* **59,** 543–568.
2. Belfort, M. (1990) *Annu. Rev. Genet.* **24,** 363–385.
3. Michel, F., Jacquier, A. & Dujon, B. (1982) *Biochimie* **64,** 867–881.
4. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. (1982) *Nature (London)* **300,** 719–724.
5. Shub, D. A., Gott, J. M., Xu, M.-Q., Lang, B. F., Michel, F., Tomaschewski, J., Pedersen-Lane, J. & Belfort, M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 1151–1155.
6. Latham, J. A. & Cech, T. R. (1989) *Science* **245,** 276–282.
7. Celander, D. W. & Cech, T. R. (1991) *Science* **251,** 401–407.
8. Tullius, T. D. & Dombroski, B. A. (1985) *Science* **230,** 679–681.
9. Hertzberg, R. & Dervan, P. B. (1984) *Biochemistry* **23,** 3934–3945.
10. Celander, D. W. & Cech, T. R. (1990) *Biochemistry* **29,** 1355–1361.
11. Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988) *Biochemistry* **27,** 8924–8931.
12. Michel, F. & Westhof, E. (1990) *J. Mol. Biol.* **216,** 585–610.
13. Salvo, J. L. G., Coetzee, T. & Belfort, M. (1990) *J. Mol. Biol.* **211,** 537–549.
14. Inoue, T., Sullivan, F. X. & Cech, T. R. (1986) *J. Mol. Biol.* **189,** 143–165.
15. Zaug, A. J., Been, M. D. & Cech, T. R. (1986) *Nature (London)* **324,** 429–433.
16. Belfort, M., Chandry, P. S. & Pedersen-Lane, J. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 181–192.
17. Xu, M. Q. & Shub, D. A. (1989) *Gene* **82,** 77–82.
18. Hicke, B. J., Christian, E. L. & Yarus, M. (1989) *EMBO J.* **8,** 3843–3851.
19. Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977) *Nucleic Acids Res.* **4,** 2527–2538.
20. Bryk, M. & Belfort, M. (1990) *Nature (London)* **346,** 394–396.
21. Michel, F., Ellington, A. D., Couture, S. & Szostak, J. W. (1990) *Nature (London)* **347,** 578–580.
22. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645–650.
23. Flor, P. J., Flanegan, J. B. & Cech, T. R. (1989) *EMBO J.* **8,** 3391–3399.
24. Joyce, G. F., Van der Horst, G. & Inoue, T. (1989) *Nucleic Acids Res.* **17,** 7879–7889.
25. Van der Horst, G., Christian, A. & Inoue, T. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 184–188.