Nucleic Acids Research, 1995, Vol. 23, No. 3 **491–499**

An RNA tertiary structure of the hepatitis delta agent contains UV-sensitive bases U-712 and U-865 and can form in a bimolecular complex

Andrea D. Branch[1,*], Bonnie J. Levine[1,2] and Jana A. Polaskova[2,+]

[1]Center for Studies of the Addictive Diseases, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA and [2]Department of Biochemistry, The Cornell University Medical College, 1300 York Avenue, New York, NY 10021, USA

Received September 14, 1994; Revised and Accepted December 2, 1994

ABSTRACT

Genomic RNA of the hepatitis delta agent has a highly conserved element of local tertiary structure. This element contains two nucleotides which become covalently crosslinked to each other upon irradiation with UV light. Using direct RNA analysis, we now identify the two nucleotides as U-712 and U-865 and show that the UV-induced crosslink can be broken by re-exposure to a 254 nm peak UV light source. In the rod-like secondary structural model of delta RNA, nucleotides U-712 and U-865 are off-set from each other by 5–6 bases, a distance too great to permit crosslinking. This model needs to be modified. Our data indicate that bases U-712 and U-865 closely approximate each other and suggest that the smooth helical contour proposed for delta RNA is interrupted by the UV-sensitive element. The nucleotide sequence shows that the UV-sensitive site does not have a particularly high density of conventional Watson–Crick base pairs compared to the rest of the genome. However, this element may have a number of non-Watson–Crick bonds which confer stability. Following UV-crosslinking and digestion with 1 mg/ml of RNase T1 at 37°C for 45 min in 10 mM Tris–HCl, 1 mM EDTA (conditions expected to give complete digestion), this element can be isolated as part of a 54 nucleotide long partial digestion product containing at least 16 internal G residues. UV-crosslinking analysis shows that this unusual tertiary structural element can form in a bimolecular complex.

INTRODUCTION

The human hepatitis delta agent is a novel subviral pathogen (1–4) with a circular genomic RNA (5–6) that has many features in common with viroid RNA and RNA of the circular satellites of certain plant viruses (7–10). Unlike its relatives which infect plants, delta encodes a protein, the delta antigen (11), which is required for replication (12–13). Circular genomic delta RNA, the delta antigen, and the surface antigen of hepatitis B virus combine to form infectious particles, which can be found at concentrations of 10^{11}/ml in blood (14). The delta agent is transmitted sexually, by contaminated blood, and through person-to-person contact (4). Delta hepatitis commonly persists in a chronic form. There is no generally effective therapy for the severe and rapidly progressive liver disease resulting from chronic delta hepatitis, which afflicts ~70 000 people in the USA (15–16).

Since its discovery in 1980 (17), much has been learned about the structure and molecular biology of the delta agent's viroid-like circular RNA genome and its rolling circle replication cycle (18–19). Genomic and antigenomic delta RNAs, like several other infectious circular RNAs (20–22), can self-cleave (23–30). The primary sequence of delta RNA, like that of viroid RNA, has an unusual segregation of purine and pyrimidine nucleotides, which may facilitate copying by cellular enzymes normally involved in the transcription of DNA (31). In its native structure, which is presumed to be the most thermodynamically favorable form, delta RNA exists as a collapsed circle, forming an extended helix, or rod-like structure (5–6) similar to that of viroids (32). This rod-like structure specifies an RNA editing event (33–35) and promotes binding of the delta antigen (36–38). In addition, several years ago we found that the native structure of delta RNA contains a novel element of local tertiary structure which lies in a highly conserved region of the genome (39). Its position in the genome suggests that one function of this element might be to protect the mature circular RNA from self-cleavage. We recently demonstrated that it functions as a ribozyme control element *in vitro*, inhibiting cleavage by a *trans*-active delta ribozyme (40). The local tertiary structure may also provide a protein binding site. This element was originally detected because it is susceptible to UV-induced crosslinking (39). Its presence in full-length genomic transcripts indicates that the UV-sensitive element is part of the native structure of delta RNA (39).

The analysis of UV-sensitive elements has provided information about a wide variety of RNAs. For example, UV-induced crosslinks helped to identify the native conformation of tRNA

* To whom correspondence should be addressed at +present address: The Division of Liver Diseases, The Mount Sinai Medical Center, 1 Gustave L. Levy Place, Atran 722, Box 1039, New York, NY 10029-6574, USA

(41–42), *E.coli* 16S rRNA (43), 7SL RNA of the signal recognition particle (44), and the self-splicing group I intron from *Tetrahymena* (45). In addition, UV-crosslinks identified points of interaction between tRNA and 16S rRNA (46) and between M1 RNA of RNase P and its substrate (47). UV-crosslinks provided the first indication of a common structural motif present in conserved sequences of viroid RNA and 5S rRNA (48).

The UV-sensitive site in 5S rRNA (loop E) is associated with the binding of at least two proteins, TFIIIA (49–50) and ribosomal protein L5 (51). Its nuclear magnetic resonance spectroscopic structure was recently determined by Wimberly *et al.* (52), who found that loop E contains four consecutive non-Watson–Crick bonds and has a sharp bend in its sugar–phosphate backbone. Unusual conformations of the sugar–phosphate backbone have been identified as a general feature of protein binding sites in RNA (53–54). UV-crosslinking provides one method for seeking such elements.

Although we have studied UV-sensitive elements in a number of RNAs (55), we have not found any others with properties like those of the UV-sensitive site in delta RNA. Following UV-induced crosslinking, this site is highly resistant to digestion by RNase T1 (39), suggesting that the non-crosslinked (native) form is also nuclease resistant. A high degree of nuclease resistance usually correlates with a high degree of thermal stability in RNA. Since the sequence of the UV-sensitive site indicates that it does not have a particularly high density of Watson–Crick base pairs compared to the rest of the delta genome, its nuclease resistance suggests that this element may contain non-Watson–Crick bonds which contribute significantly to its stability. While secondary structure is generally the major determinant of RNA stability (56), the Hoogsteen-paired quartets of G residues recently described by Cheong and Moore (57) provide an example of a very stable RNA structure resulting from unconventional non-Watson–Crick bonds.

Mapping the UV-induced crosslink is an important step in defining a UV-sensitive site. We have identified the two crosslinked bases in delta RNA, U-712 and U-865, by direct RNA analysis. We have also used UV-crosslinking and analysis of RNase T1 digestion products to show that the element of local tertiary structure in delta RNA can be produced in a bimolecular complex.

MATERIALS AND METHODS

The crosslink was mapped by direct RNA analysis. Several alternative methods could not be used due to peculiarities of the delta UV-sensitive site. For example, reverse transcriptase copying did not go to completion, despite the use of conditions which allowed full-length copies of viroid RNA to be synthesized (58); RNase H failed to produce useful cleavage products; and the initial transcript we studied was too long for the crosslink to be directly mapped from alkaline hydrolysis ladders. In addition, except for the one RNase T1 partial digestion product we describe in detail, we were not able to generate any other useful partial digestion products, either with RNase T1 or pancreatic RNase A.

The conditions we used for UV-crosslinking (55), RNA fingerprinting (59) and secondary analysis (60–62) have been described in detail previously. The subgenomic cDNA clones used in this work were kindly provided by Drs Bahige Baroudy (Gamble Institute of Medical Research, Cincinnati, OH) and John Gerin (Georgetown University School of Medicine, Rockville, MD). The DNA oligomers used as templates for the synthesis of short RNA hairpins by the Milligan approach (63) were generously prepared by Drs Alicia Buckler-White and John Gerin (Georgetown University School of Medicine, Rockville, MD).

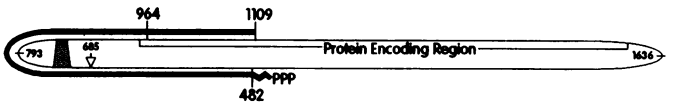

Figure 1. Landmarks of subgenomic delta transcripts. Two subgenomic cDNA clones (bases 482–964 and bases 482–1109) were transcribed under conditions described before (39). The RNA transcripts are depicted by the heavy line above the rod-like structure of circular genomic delta RNA. They contain a segment of vector sequences at their 5′ co-terminal ends, marked by a squiggle. The shaded segment of the rod-like structure identifies the position of the UV-sensitive tertiary element. The numbering of delta nucleotides is according to Wang *et al.* (5) and is presented in Arabic numerals. The protein encoding region is boxed (11). The self-cleavage site (23–30) is marked by an arrow.

Under the conditions used, RNase T1 cleaves after G residues; pancreatic RNase A cleaves after pyrimidine residues; RNase U2 cleaves after A residues; RNase T2 and nuclease P1 have no base specificity. RNase T1, RNase U2, pancreatic RNase A and RNase T2 yield 3′ phosphate termini. Nuclease P1, which has a 3′ phosphatase activity, cleaves to yield 5′ phosphate groups. Unless otherwise specified, an oligonucleotide terminates with a 5′ hydroxyl group and a 3′ phosphate group.

RESULTS

Mapping the UV-sensitive site in internally labeled delta transcripts

A combination of techniques was used to identify the two nucleotides in genomic delta RNA which become crosslinked to each other upon irradiation with UV light. The general approach we used was to find the approximate location of the UV-induced crosslink and then to use the process of elimination to identify the two crosslinked nucleotides. We previously reported (39) that the UV-sensitive element lies in a highly conserved portion of the genome, near the self-cleavage site (see Fig. 1). This UV-sensitive element of local RNA tertiary structure is made up of sequences which lie roughly opposite each other in the rod-like secondary structure of delta RNA proposed by Wang *et al.* (5).

In the first part of the current study, direct RNA analysis was carried out on subgenomic cDNA transcripts containing the UV-sensitive site. A diagram of these transcripts is shown in Figure 1. To permit nearest-neighbor analysis, delta transcripts were synthesized separately from each of the four [α-^{32}P]-labeled nucleoside triphosphates, yielding 'single-label' transcripts. Following UV-irradiation, gel electrophoresis was used to separate transcripts in which a crosslink had been induced from the remaining transcripts (39).

To prepare a large RNase T1 partial digestion product to be used in subsequent mapping studies, UV-crosslinked transcripts and control transcripts (which had been exposed to UV light but were not crosslinked) were incubated for 45 min at 37°C in 2 µl containing 1 mg/ml of RNase T1, 10 mM Tris–HCl (pH 7.4), 1 mM EDTA and 10 µg of tRNA. Under these conditions, RNase T1 does not cut in the vicinity of the two crosslinked bases in delta RNA, creating a large RNase T1 partial digestion product (see arrow in Fig. 2). As a result, the two crosslinked nucleotides and

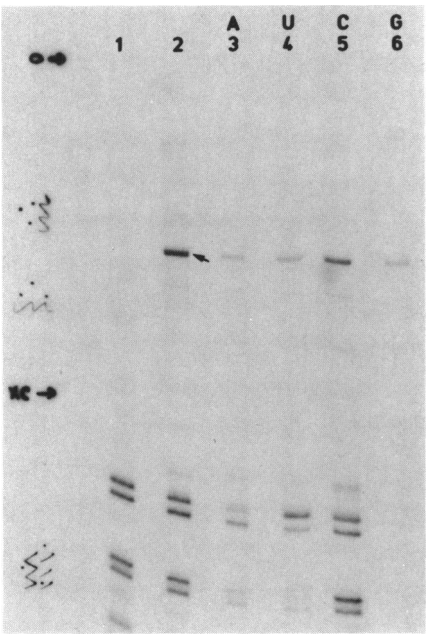

Figure 2. Gel electrophoresis of the RNase T1 partial digestion product of UV-crosslinked transcripts. *In vitro* transcripts were synthesized from the two subgenomic cDNA clones described in the legend to Figure 1 with each of the four [α-^{32}P]nucleoside triphosphates. The UV-crosslinked form of each transcript and its UV-irradiated but not crosslinked form were purified from gels and digested in 2 µl containing 10 mM Tris–HCl, pH 7.4, 1 mM EDTA, 1 mg/ml RNase T1 and 10 µg tRNA at 37°C for 45 min. Digestion products were separated by preparative scale electrophoresis. Products of crosslinked transcripts (lanes 2–6) included a partial digestion product (marked by an arrow) which was not present in control transcripts (lane 1). Products of the shorter transcript, which contains delta bases 482–964, appear in lanes 1 and 2; while those of the longer transcript, which contains delta bases 482–1109, appear in lanes 3–6. Letters above lanes 3–6 indicate the [α-^{32}P]nucleoside triphosphate used in the transcription reaction. 'O' indicates the origin of electrophoresis of this 20% polyacrylamide gel containing a Tris–borate buffer and 7 M urea; 'XC' the position of xylene cyanol blue. The idiosyncratic markings on the left-hand side were made by radioactive ink and were used to guide excision of the crosslink-specific partial digestion product.

Table 1. Secondary analysis of RNase T1 fingerprint spots

Spot	Possible Identity	Pancreatic RNase A	RNase U2	RNase T2	Nuclease P1
1.	CAACAUUCCG (726-735)	AAC, AU, G, C, U	+	C, A, G, U	pG, pA, pC, pU, [PO_4]
2.	CCUCCUCG (703-710)	G, C, U	–	C, G, U	pG, pC, pU, PO_4
3.	X - Oligo	X^1, G, [C]	–	C, G, [U], X^2	X^3, pG, PO_4
4.	CCCAG (859-863)	AG, C	+	C, A, G	pG, pA, pC, PO_4
5.	ACCG (869-872)	AC, G, C	+	C, A, G	pG, pC, PO_4
6.	AUG (856-858)	AU, G	+	A, G, U	pG, pU, PO_4
7.	CUG (711-713) (721-723)	G, C, U	–	C, G, U	pG, pU, PO_4
8.	CCG (717-719)	G, C	–	C, G	pG, pC, PO_4
9.	AG (875-876)	AG	+	A, G	pG, PO_4
10.	CG (715-716) (873-874)	G, C	–	C, G	pG, PO_4
11.	G (714) (720) (724) (725) (864) (868)	G	–	G	PO_4

Oligonucleotide designations refer to the RNase T1 fingerprint spots marked in Figure 3. the numbers in parentheses indicate the delta bases (5) potentially present in the spot. Oligonucleotides produced by RNase T1, pancreatic RNase A, RNase U2, and RNase T2 contain 5′ hydroxyl termini and 3′ phosphate termini; while those produced by nuclease P1 have 5′ phosphate termini and 3′ hydroxyl termini. Conditions for secondary enzymatic digestions have been described before (60–62). Products of pancreatic RNase A digestion were analyzed by high-voltage electrophoresis on DEAE paper, pH 3.5. Products of RNase U2 were analyzed on DEAE paper, pH 1.9; a '+' indicates that RNase U2 treatment caused a change in mobility, while a '–' indicates that it did not. Products of RNase T2 were analyzed on 3 MM paper, pH 3.5. Products of nuclease P1 were analyzed on DEAE paper, pH 3.5 and on 3 MM paper, pH 3.5. The products X^1, X^2, and X^3 did not co-migrate with any known markers, and thus could not be identified. Parentheses identify products which were either very faint, or were not consistently detected.

their flanking sequences can be separated from the remainder of the transcript (which is digested to completion) by fractionating the digests in 20% polyacrylamide gels containing 7 M urea (39). After elution from gels, single-label transcripts were either analyzed directly or were mixed together to provide uniformly labeled material for subsequent experiments.

Despite the unusual RNase T1 resistance of the crosslink-specific fragment, RNase T1 cleavage can be obtained by carrying out the digestion at 65°C for 1 h in a sealed capillary tube (39). These conditions were used to prepare the two-dimensional RNA fingerprint of the mixed-label sample shown in Figure 3. Oligonucleotides were recovered from fingerprint spots and analyzed further by treatment with a variety of nucleases in a process called 'secondary analysis'. Products released by the secondary enzymatic digestions are presented in Table 1. These data provide several pieces of information about the location of the UV-induced crosslink. For example, pancreatic RNase A digestion of spot 1 released AAC, AU, G, C, and U. Only one RNase T1 oligonucleotide in the sequence of genomic delta RNA, CAACAUUCCG (bases 726–735), would yield this set of products, indicating that one of the two segments joined by the covalent crosslink must contain bases 726–735 and neighboring sequences.

The segment containing the other crosslinked base was identified with the following data: spot 4 yielded pG_{OH}, pA_{OH}, pC_{OH}, and phosphate upon treatment with nuclease P1 and yielded AG and C upon digestion by pancreatic RNase A (Table 1). Considered together with the position of spot 4 in the fingerprint, these results indicate that spot 4 contains either CCAG or CCCAG; however, since CCAG does not occur in the partial cDNA clones used for the fingerprinting studies, spot 4 must contain CCCAG. This oligonucleotide occurs twice in the cDNA clones, at positions 628–632 and 859–863. However, only the CCCAG at position 859–863 is near other oligonucleotides also present in the fingerprint, such as spots 5 and 6. Spot 5 yielded AC, G and C upon digestion with pancreatic RNase A and pG_{OH}, pC_{OH} and free phosphate upon digestion with nuclease P1. Thus, spot 5 was identified as ACCG, an oligonucleotide which occurs at positions 538–541, 741–744 and 869–872. Spot 6 was identified as AUG based on its pancreatic RNase A products. AUG occurs many times in the sequence. However, the three oligonucleotides, AUG, CCCAG and ACCG, only occur in conjunction in the sequence GAUGCCCAGGUCGGACCG (bases 855–872), indicating that this segment comprises part of the UV-sensitive element in delta RNA.

A diagram of the crosslink-specific RNase T1 partial digestion product is shown in Figure 4. The oligonucleotides which could be positively and unambiguously identified by analysis of fingerprint spots are boxed by solid lines, with numbers corresponding to those

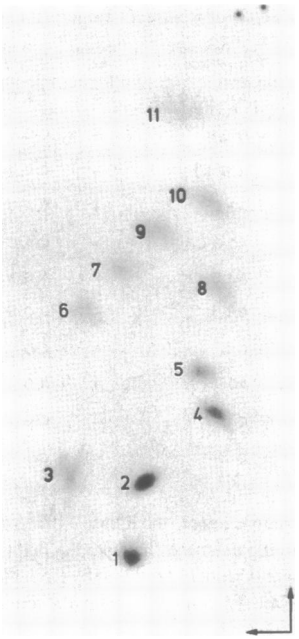

Figure 3. RNase T1 fingerprint of the UV-crosslinked RNase T1 partial digestion product. After elution from gels, the RNase T1 partial digestion product shown in Figure 2 was digested in 2 µl containing 1 mg/ml RNase T1 and 10 µg of tRNA at 65°C for 1 h in a sealed capillary tube and then the products were separated into a two-dimensional fingerprint (39,59). Oligonucleotides were eluted from the numbered positions and subjected to secondary analysis (60–62). The horizontal and vertical arrows indicate the directions of first and second dimension separation, respectively.

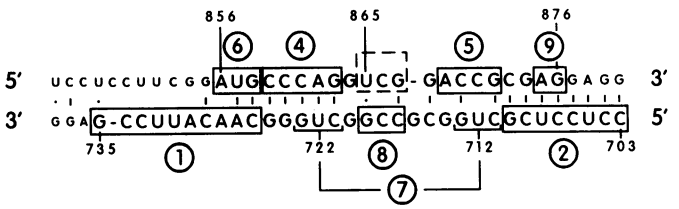

Figure 4. Diagram summarizing fingerprinting data about the UV-crosslinked RNase T1 partial digestion product. Each delta sequence which could be unambiguously identified from fingerprinting and conventional secondary analysis is boxed and bears the encircled number of its fingerprint spot (1, 2, 4–6, 8, 9). Spot 7 contained CUG, which occurs twice in the sequence and is underlined, rather than boxed, to indicate that the two CUG segments could not be distinguished. Spot 7 did not contain UCG, which is boxed with a broken line. The nucleotides flanking the region containing the crosslink are shown in small uppercase letters.

of the fingerprint spots of Figure 3. These oligonucleotides must consist of bases from flanking sequences. The oligonucleotides containing the two crosslinked bases would be missing from their usual places in the fingerprint. Thus, all the bases in (solid line) boxes can be eliminated from consideration as possible crosslinked bases. Taken together, these data establish the approximate boundaries of the RNase T1 partial digestion product and lead to a remarkable conclusion: this resistant region must contain a minimum of 16 internal G residues which escaped cleavage by 1 mg/ml RNase T1 (see Fig. 4). Such a high degree of nuclease resistance is usually associated with double-stranded RNA, and yet it is evident from the sequence that the UV-sensitive site does not

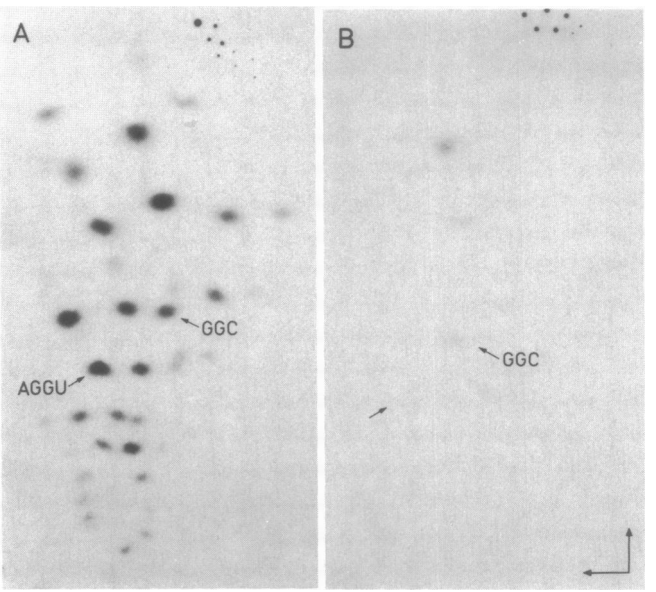

Figure 5. Pancreatic RNase A fingerprints of the UV-crosslinked RNase T1 partial digestion product from U-labeled RNA. From transcripts synthesized in the presence of [α-^{32}P]uridine triphosphate, the UV-irradiated but not crosslinked (control) form (**A**) and the RNase T1 partial digestion product of the crosslinked form (**B**) were prepared and digested at 37°C for 10 min in 2 µl containing 0.25 mg/ml pancreatic RNase A, 10 mM Tris–HCl, 1 mM EDTA and then the products were fingerprinted (59). An arrow marks the empty spot in the fingerprint of crosslinked RNA (B) where AGGU would appear if it were released from the crosslinked RNase T1 partial digestion product. The positions of GGC (identified in both panels) and AGGU/GAGU/GGAU (identified in A as AGGU) provide landmarks. The horizontal and vertical arrows in (B) indicate the directions of first and second dimension separation, respectively.

contain an exceptionally large number of Watson–Crick base pairs compared to delta RNA as a whole. While the crosslink itself might protect one or two G residues from cleavage, the extensive RNase T1 resistance of UV-crosslinked delta RNA has not been observed in studies of other UV-crosslinked RNAs (55) and thus the extensive resistance of crosslinked delta RNA probably reflects an unusual structural feature inherent in this site.

Once the approximate locations of the two segments present in the RNase T1 partial digestion product were determined, further experiments were carried out to map the crosslinked bases themselves. The first of the two crosslinked bases was identified in the following manner. In RNase T1 fingerprints, the two oligonucleotides CUG and UCG co-migrate, forming spot 7 in Figure 3. If both oligonucleotides were present in spot 7, nuclease P1 digestion would release $_pG_{OH}$, $_pC_{OH}$ and $_pU_{OH}$. However, $_pC_{OH}$ was not present in nuclease P1 digests (see Table 1), indicating that UCG was missing. Furthermore, the CUG/UCG spot did not appear in the fingerprint prepared from RNA synthesized in the presence of [α-^{32}P]-labeled CTP (data not shown). While CUG occurs twice in the RNase T1 partial digestion product, CUG would not be expected to appear in the fingerprint of C-labeled RNA (since neither copy of CUG is followed by a C nucleotide in the sequence; see Fig. 4). However, UCG would be expected in the C-label fingerprint of this segment of RNA. The total absence of the CUG/UCG spot from the C-labeled fingerprint indicates that one of the nucleotides in the trimer UCG (bases 865–867) is crosslinked to a second nucleo-

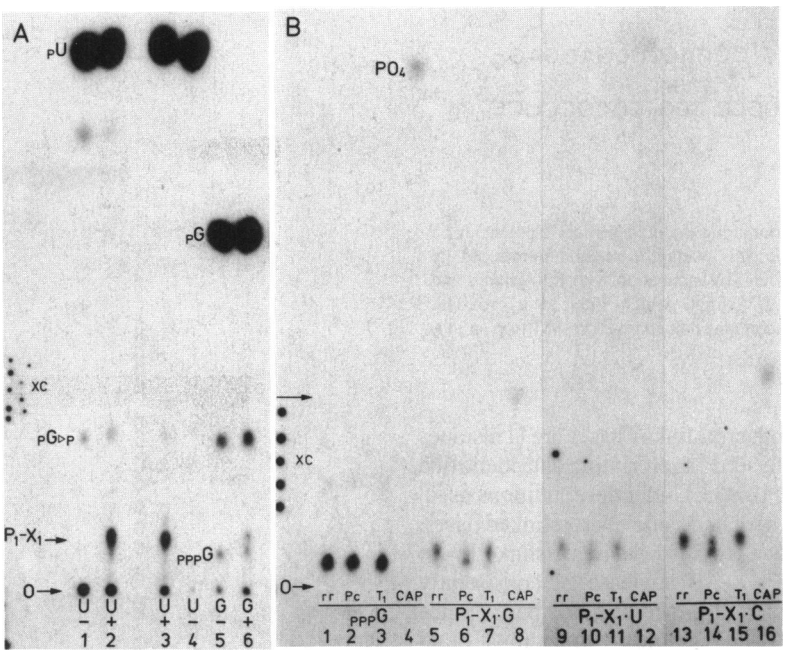

Figure 6. Further analysis of UV-crosslinked oligonucleotides. (**A**) Detection of P1-X1, a nuclease P1 digestion product. To provide material for this analysis, U-labeled transcripts (lanes 1–4) and G-labeled transcripts (lanes 5–6) were either directly digested by RNase T1 at 37°C or were first exposed to UV light and then digested. Products were fractionated by electrophoresis in 7 M urea–20% polyacrylamide gels. From the position of the crosslink-specific RNase T1 partial digestion product, RNAs were eluted from both control transcripts and from UV-irradiated transcripts, digested for 1 h at 37°C in 10 µl containing 0.33 mg/ml nuclease P1 in 10 mM sodium acetate, pH 6.0, applied to DEAE paper and fractionated by high voltage electrophoresis at pH 3.5 (60–62). Lanes 1, 4 and 5 depict products of control transcripts (–); lanes 2, 3 and 6 depict products of UV-irradiated transcripts (+). A UV-specific product is marked P_1-X_1. Several additional species are identified. 'O' indicates the origin of electrophoresis; 'XC' the position of xylene cyanol blue. (**B**) Analysis of the P_1-X_1 nuclease P1 digestion product. The control species, $_{PPP}G_{OH}$ (lanes 1–4, marked pppG), and the P_1-X_1 spot from G-labeled transcripts (lanes 5–8), U-labeled transcripts (lanes 9–12), and C-labeled transcripts (lanes 13–16) were individually eluted from DEAE paper and either resuspended in water (lanes 1, 5, 9 and 13) or digested by pancreatic RNase A (lanes 2, 6, 10 and 14) or digested by RNase T1 (lanes 3, 7, 11 and 15) or by calf alkaline phosphatase (lanes 4, 8, 12 and 16), under conditions described before (61) except that enzyme-treated samples also contained 10 µg of tRNA. Samples were applied to DEAE paper and fractionated by electrophoresis at pH 3.5. An arrow identifies the position of a phosphatase-resistant product present in samples of G-labeled and C-labeled transcripts, but not present in U-labeled transcripts. The position of free phosphate is also identified. 'O' indicates the origin of electrophoresis; 'XC' the position of xylene cyanol blue.

tide. Thus UCG must be part of fingerprint spot 3, which is identified as the crosslink-containing spot [because it releases unconventional products upon digestion with several enzymes (see Table 1)].

To identify which base in the UCG trimer was crosslinked, pancreatic RNase A fingerprints were prepared and analyzed. The UV-specific RNase T1 partial digestion product shown in Figures 2 and 4 was prepared. Crosslinked fragments were eluted from gels and then were incubated at 37°C for 10 min in 2 µl containing 0.25 mg/ml of pancreatic RNase A, 10 mM Tris–HCl, 1 mM EDTA and 10 µg of tRNA. Digestion products were fractionated into a two-dimensional fingerprint. The tetramer AGGU was missing from the fingerprint of U-labeled crosslinked RNA (Fig. 5B), indicating that AGGU contains one of the two crosslinked bases. The oligonucleotides AGGU and UCG overlap by 1 base, U-865 (in the sequence AGGUCG; see Fig. 4). Together, the results of RNase T1 and pancreatic RNase A fingerprinting studies identified U-865 as one of the two crosslinked bases.

Identification of the second crosslinked base was more difficult. Analysis of nuclease P1 digestion products provided a key piece of information, indicating that both crosslinked bases are uridine residues. As shown in Figure 6, nuclease P1 digestion released a slowly migrating product from crosslinked RNA which could be detected following high voltage electrophoresis on DEAE paper carried out at pH 3.5. This product, P_1-X_1, was missing from digests of control RNA (Fig. 6A, compare lane 1 to lanes 2 and 3). P_1-X_1 was released from crosslinked G-labeled, C-labeled and U-labeled RNA, but not from A-labeled RNA. It was much more prominent in products of U-labeled RNA than in those of G-labeled RNA (Fig. 6A, compare the intensity of P_1-X_1 in lanes 3 and 6) or C-labeled RNA (not shown), suggesting that P_1-X_1 contained two phosphate groups 5′ to U residues and one phosphate group 5′ to a G residue and one 5′ to a C residue. Following elution of P_1-X_1 from the DEAE paper, samples were treated with calf alkaline phosphatase. In the case of U-labeled P_1-X_1, the major product was free phosphate (Fig. 6B, lane 12). This result indicates that pU occurs only at the 5′ end of P_1-X_1. In the case of G-labeled P_1-X_1 and C-labeled P_1-X_1 (Fig. 6B, lanes 8 and 16), phosphatase treatment created a novel product, which migrated much more slowly than free phosphate. (A small amount of free phosphate was also released from the G-labeled P_1-X_1, probably due to contamination with $_{PPP}G_{OH}$, which migrates just below P_1-X_1 during electrophoresis on DEAE paper at pH 3.5; see Fig. 6A, lane 5.) Therefore P_1-X_1 contains at least one pG and one pC, neither at a 5′ end, supporting the conclusion that both 5′ ends are pU.

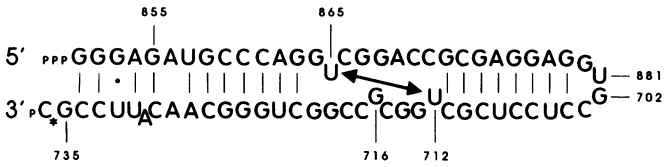

Figure 7. A UV-sensitive hairpin containing the delta tertiary element. A 65 base long RNA was transcribed *in vitro* using the method introduced by Milligan *et al.* (63), labeled at its 3′ end by incubation with RNA ligase and [^{32}P]$_{p}$C$_{p}$ (star), and exposed to UV light, which induced a crosslink (double-headed arrow). The hairpin contains delta bases 855–881 (top strand) and 702–735 (bottom strand).

These results suggest that both crosslinked bases are U residues according to the following analysis. Fingerprinting data identified one of the crosslinked bases as U-865. Under the conditions used, nuclease P1 cleaves on the immediate 5′ side of crosslinked bases, as proven by the evidence that U-865 is both a crosslinked base and is at the 5′ terminus of P_1-X_1. If nuclease P1 consistently cleaved on the immediate 5′ side of the crosslinked bases, since both 5′ ends are U residues, both crosslinked bases must be U residues. Given the constraints of the sequence, the second crosslinked U residue must be either U-712 or U-722 (see Fig. 4).

Crosslinking of either U-712 or U-722 would have escaped detection during the RNA fingerprinting analysis. The sequence in which they occur, GCUGG, occurs twice in the RNase T1 partial digestion product containing the crosslinking site (see Fig. 4, in which both copies of GCUGG are underlined). Thus the CUG spot would still appear in the RNase T1 fingerprint even if one copy of CUG were missing because it was crosslinked to another oligonucleotide. Because RNA fingerprinting is not a quantitative procedure, this loss would not be detectable. Furthermore, because the two copies of CUG occur in the same sequence context, nearest-neighbor analysis could not be used to distinguish them. Additional techniques were needed to complete the mapping of the crosslink.

Mapping the UV-sensitive site in short, end-labeled delta transcripts

In order to identify the second U residue in the crosslink, a number of short, end-labeled transcripts containing the UV-sensitive element were analyzed by partial alkaline hydrolysis, a technique we used previously to map a UV-sensitive site in the central conserved region of viroid RNA (48). Figure 7 presents a diagram of one of these transcripts, a 65 nucleotide long molecule, which was synthesized *in vitro* using the conditions described by Milligan *et al.* (63). It was then ligated to ^{32}P-labeled $_{p}$C$_{p}$. This end-labeled RNA was crosslinked by irradiation with UV light. Gel electrophoresis was used to separate the UV-crosslinked RNA from the remainder of the UV-exposed, but apparently non-crosslinked molecules (data not shown).

The crosslinked RNA hairpin was then partially cleaved by alkaline hydrolysis and the products were fractionated by electrophoresis in 20% polyacrylamide gels containing 7 M urea (Fig. 8, lanes c and d) to identify the position of the crosslink. For comparison, alkaline hydrolysis was also carried out on control transcripts, which had received no exposure to UV light (Fig. 8, lanes a and b), and on UV-exposed, but non-crosslinked transcripts (Fig. 8, lanes e and f). Partial digestion products of

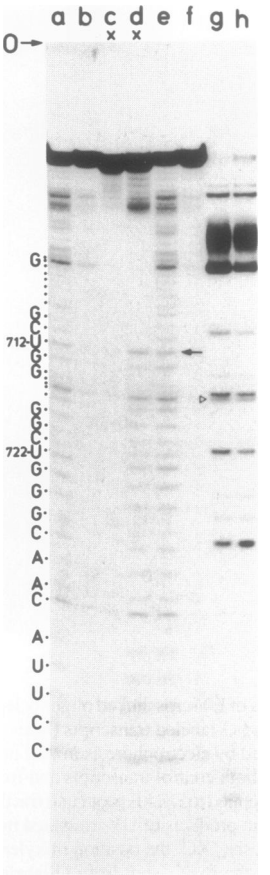

Figure 8. Gel electrophoretic analysis of partial digestion products of the UV-sensitive hairpin. Three forms of the 65 base long hairpin depicted in Figure 7 were prepared: hairpins that received no UV treatment (lanes a and b); hairpins that were UV-crosslinked (lanes c and d, marked x); hairpins that were UV-treated, but apparently non-crosslinked (lanes e–h). To produce alkaline hydrolysis ladders, samples were incubated at 90°C in 0.05 M sodium bicarbonate/carbonate at pH 9.2 for either 15 min (lanes a, d and e) or 5 min (lanes b, c and f). To produce a family of RNase T1 partial digestion products, non-crosslinked samples were incubated at 65°C in buffer containing RNase T1 and urea for 15 min (lane g) or 4 min (lane h), as described before (62). Each band in the alkaline ladder is marked by a dot, to assist in the counting of the bands. An arrow identifies the beginning of a gap in the alkaline ladder, which starts at U-712. An open arrowhead designates an under-represented RNase T1 partial digestion product associated with G-716. 'O' marks the origin of this 7 M urea–20% polyacrylamide gel.

RNase T1 were used to identify nucleotide positions in the gel (Fig. 8, lanes g and h). The partial digestion product associated with G-716 was fainter than expected, suggesting that this nucleotide may be protected or distorted (see Fig. 8, lanes g and h, open arrowhead).

Alkaline cleavage of RNAs containing no crosslinks produced uninterrupted ladders of digestion products (Fig. 8, lanes a and e). In contrast, the pattern of cleavage products from crosslinked RNA terminated at U-712 (see arrow in Fig. 8). From these, and supporting data from other short, end-labeled transcripts (data not shown), and from the fingerprinting data summarized in the preceding section, we conclude that U-712 and U-865 are the two nucleotides in delta genomic RNA joined by the UV-induced covalent bond. This bond is indicated by the double-headed arrow in Figure 7.

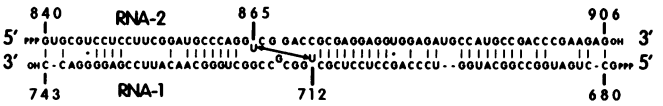

Figure 9. A UV-crosslinked bimolecular complex of RNA-1 (bases 680–743) and RNA-2 (bases 840–906) with a double-headed arrow identifying a UV-induced crosslink.

Generation of a UV-sensitive bimolecular complex containing the tertiary element

To learn the details of the tertiary structure in delta RNA, techniques such as NMR spectroscopy or X-ray crystallography will be needed. However, the delta hairpins we synthesized for mapping studies may not be optimal material for future structural studies because they have terminal loops, rather than free ends. Wimberly *et al.* (52) successfully prepared a model of the internal loop E of eukaryotic 5S rRNA by synthesizing two separate transcripts, which combined to form the desired internal loop. It has proven to be difficult to crystallize hairpin loop structures. In two separate studies, RNA oligomers containing tetraloop sequences formed double helices, rather than hairpin loops, when crystallized (64–65). Thus, we wished to determine whether the element of local tertiary structure in delta genomic RNA could be formed in a bimolecular complex in which two halves of the tertiary structure were present on separate RNA molecules.

For these experiments, two RNA molecules were synthesized *in vitro* under conditions introduced by Milligan *et al.* (63). RNA-1 was 65 nucleotides in length and represents sequences from the bottom strand of the rod-like secondary structure of delta genomic RNA (delta bases 680–743; see Fig. 9). RNA-2 was 67 nucleotides in length and represents sequences from the top strand of the delta rod-like structure (delta bases 840–906). RNA-1 and RNA-2 were combined and incubated in the presence of a magnesium-containing buffer to allow complexes to form. Products were fractionated by electrophoresis in pre-chilled 8% polyacrylamide gels containing a Tris–borate buffer, but no EDTA or urea. Only one new band, an RNA-1–RNA-2 complex, was generated. This new band (marked by the open arrowhead in Fig. 10) migrated much more slowly than free RNA-1 and free RNA-2 (marked by the solid arrowhead). A Molecular Dynamics PhosphorImager was used to determine the percentage of RNA in a given sample which was present in the RNA-1–RNA-2 complex. In samples incubated for 15 min at 60°C, the RNA-1–RNA-2 complex accounted for 64% of the RNA (Fig. 10, lane n), while it accounted for 80% in samples incubated for 60 min (Fig. 10, lane p). The slow rate of formation of the RNA-1–RNA-2 complex suggests that either RNA-1 or RNA-2 or both may have self-structure which inhibits complex formation. Palukaitis and Symons reported slow kinetics for hybridization to viroid RNA and attributed the slow rate to residual RNA secondary structure (66). The new band, representing the RNA-1–RNA-2 complex, was not generated when either RNA-1 alone (Fig. 10, lanes c and d) or RNA-2 alone (Fig. 10, lanes g and h) was incubated in the magnesium-containing buffer. Production of the RNA-1–RNA-2 complex required the addition of cations and did not take place when a mixture of RNA-1 and RNA-2 was incubated in water (Fig. 10, lane j); nor was it produced in samples kept on ice in the magnesium-containing buffer (Fig. 10, lane k).

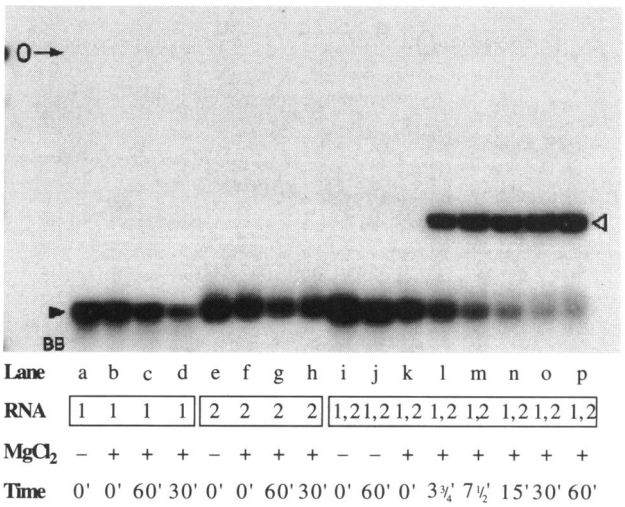

Figure 10. Detection of an RNA-1–RNA-2 complex by gel electrophoresis. RNA-1 (lanes a–d) and RNA-2 (lanes e–h) and an equimolar mixture of the two (lanes i–p) were heated to 90°C for 1 min in water and then were either maintained on ice in water (lanes a, e and i); incubated at 60°C for 60 min in water (lane j); or were maintained on ice in 2.5 mM $MgCl_2$, 5 mM HEPES, pH 6.5 (lanes b, f, k); or were incubated in sealed capillary tubes at 60°C for 60 min in 2.5 $MgCl_2$, 5 mM HEPES, pH 6.5 (lanes l–p). Samples were quenched on ice, mixed with a buffer containing HEPES, $MgCl_2$, sucrose and dye, and analyzed by electrophoresis in pre-chilled non-denaturing 8% polyacrylamide gels. 'O' indicates the origin of the gel; 'BB' marks the bromophenol blue dye; an open arrowhead identifies the RNA-1–RNA-2 complex, a solid arrowhead denotes free RNA-1 and RNA-2.

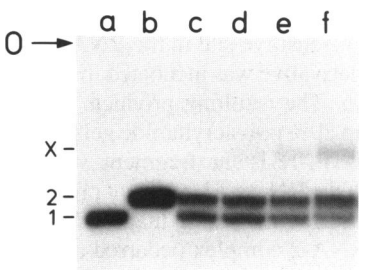

Figure 11. UV-Sensitivity of the RNA-1–RNA-2 complex. The RNA-1–RNA-2 complex, eluted from preparative scale non-denaturing gels, was resuspended in 0.5 mM $MgCl_2$, 5 mM HEPES and exposed to UV irradiation at room temperature for various times: 0 s (lane c); 7 s (lane d); 20 s (lane e); 60s (lane f). To release RNA-1 and RNA-2 from complexes which had not been crosslinked, samples were mixed with formamide dye containing 50 mM EDTA and heated at 90°C for 3 min. Samples were fractionated by electrophoresis to determine the extent of UV-crosslinking. An aliquot of RNA-1 is shown in lane a, RNA-2 in lane b; 'O' marks the origin of the 7 M urea–10% polyacrylamide gel.

Analysis of complexes eluted from gels revealed that RNA-1 and RNA-2 are present in a one-to-one ratio (data not shown) and showed that complexes are sensitive to UV-induced crosslinking. The crosslinked derivative constituted ~15% of the RNA in samples irradiated for 60 s (Fig. 11, lane f). We were not able to identify the factors that caused the crosslinking to plateau at 15%. Susceptibility of the crosslink to photolysis, non-specific UV damage and structural features of the complex may all contribute.

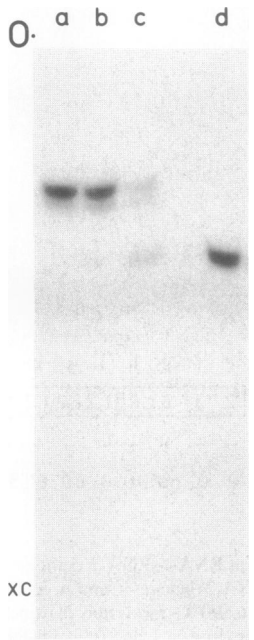

Figure 12. Photolysis of the UV-crosslink. The UV-crosslinked form of the subgenomic transcript, eluted from gels and purified as described before (39), was either analyzed directly (lane a) or was exposed to UV light (peak 245 nm) for 20 s (lane b) or for 2 min (lane c). Samples were fractionated by electrophoresis in a 7 M urea–4% polyacrylamide gel. An aliquot of the subgenomic transcript which received no UV treatment appears in lane d. 'O' indicates the origin of the gel; 'XC', the position of xylene cyanol blue.

To map the UV-sensitive site in the RNA-1–RNA-2 complex, the crosslinked derivative was incubated in 1 mg/ml RNase T1 at 37°C for 45 min. The resulting products were fractionated by electrophoresis in 20% polyacrylamide gels containing 7 M urea. A large, RNase T1-resistant fragment was generated which co-migrated with the RNase T1-resistant product of a crosslinked subgenomic transcript, indicating that the UV-crosslink induced in the RNA-1–RNA-2 complex occurred at the same site as the crosslink induced in the subgenomic transcript (data not shown).

Photolysis of the UV-induced crosslink

To determine whether the crosslink in delta RNA could be broken by re-exposure to the same UV light source which created it, aliquots of crosslinked subgenomic transcripts were exposed to UV light (254 nm peak output) and then were analyzed by electrophoresis. Following UV irradiation for 2 min, about 30% of the RNA co-migrated with non-crosslinked control transcripts, demonstrating that the delta crosslink is susceptible to photolysis (Fig. 12, lane c). Similar results were obtained in studies of crosslinked RNA-1–RNA-2 complexes (data not shown). These results are compatible with the delta crosslink being a pyrimidine–pyrimidine cyclobutane dimer, and thus are in accord with our other studies indicating that the crosslink connects two U residues. Additional photosensitivity tests are available (67–69) and have been used to define the chemistry of a UV-induced crosslink between tRNA and 16S rRNA (68) and a UV-induced crosslink in transcripts of tRNA (69).

DISCUSSION

Our results show that U-712 and U-865 are the two bases in delta RNA which become covalently crosslinked to each other upon irradiation with UV light, indicating that these two bases are very close to each other in the native conformation of delta RNA. In the rod-like secondary structure model of delta RNA (5) these two nucleotides are several bases away from each other. Thus, at least in this small region of the genome, the rod-like model needs to be modified. More work is needed to determine whether and in what way the smooth helical contour of the overall delta RNA molecule is interrupted by the tertiary structure. The NMR structure of loop E from *Xenopus laevis* 5S ribosomal RNA (52) might suggest one way in which the sugar–phosphate backbone could be locally folded in order to bring U-712 and U-865 into close proximity. Although the delta RNA does not have the same structural motif as that of 5S rRNA/viroid RNA, all three sites are flanked by Watson–Crick base paired helices and have crosslinked bases which are off-set from each other. In the case of loop E, 'loop' turns out to be a misnomer: the internal loop closes to form a G·A base pair and a reverse-Hoogsteen A·U base pair and perhaps two additional non-Watson–Crick base pairs, A·A and U·U. Extensive interstrand stacking leads to a structure susceptible to UV-induced crosslinking. While one strand of loop E has almost A-form geometry, the sugar–phosphate backbone of the other strand has a very irregular conformation and locally points in the direction opposite to the rest of the helix.

In our current studies, we used susceptibility to UV-induced crosslinking to show that the tertiary structure from delta RNA, like that of loop E of 5S rRNA, can be formed when its two halves are on separate RNA molecules. Susceptibility to UV-crosslinking was used previously by Wimberly *et al.* to verify that their 27 nucleotide complex had the proper conformation and thus could be used as a model of loop E in NMR studies (52). Downs and Cech have presented evidence that a UV-crosslinking reaction which induces an adenosine–adenosine bond in a shortened enzymatic version of the *Tetrahymena* self-splicing intervening sequence can be used as an assay for correctly folded structure and to test for the production of the native conformation under various conditions (45). As a cautionary note, data from Behlen *et al.* indicate that susceptibility to UV-crosslinking should be combined with additional tests of RNA structure and function whenever possible. Production of an intramolecular crosslink between C-48 (in the variable loop) and U-59 (in the T loop) of yeast $tRNA^{Phe}$ wild-type and mutant transcripts appeared 'in general to be somewhat less sensitive than lead cleavage in its ability to detect distant defects in the overall folding of $tRNA^{Phe}$' (69). However, Behlen *et al.* conclude that the C-48–U-59 crosslinking assay provides a useful means for determining whether a mutant sequence is folded into a structure similar to that of wild-type $tRNA^{Phe}$. Taken together, the studies of Wimberly *et al.* (52), Downs and Cech (45) and Behlen *et al.* (69), indicate that UV-crosslinking can be used to assess the fidelity of RNA folding, as we have done here.

Non-Watson–Crick bonds, such as those present in photosensitive elements, are important features of RNA structure and need to be identified and characterized using high resolution techniques. The two RNAs making up the RNA-1–RNA-2 complex are 65 and 67 bases in length, for a total of 132 nucleotides. This complex is about 30 nucleotides smaller than the P4–P6 domain from the *Tetrahymena* group I intron which was recently

crystallized using the sparse matrix approach (70). Thus, it appears that crystallographic studies of the delta tertiary structure may be possible in the near future.

ACKNOWLEDGEMENTS

We wish to thank Ms Sylvia Genus for outstanding technical support. This work was funded in part by NIH grants U01AI-31867 and R01AI31067 to Dr Hugh D. Robertson (Cornell University Medical College) and NIDA Research Center Grant P50DA-05130 directed by Dr Mary Jeanne Kreek (Rockefeller University), project 7 to ADB.

REFERENCES

1 Purcell,R.H. and Gerin,J.L. (1990) In Fields,B.N. (ed.), *Virology*. Raven Press, New York, Vol. 2, pp. 2275–2287.
2 Gerin,J.L., Purcell,R.H. and Rizzetto,M. (1991) *The Hepatitis Delta Virus*. Wiley-Liss, New York.
3 Taylor,J.M. (1992) *Annu. Rev. Microbiol.* **46**, 253–76.
4 Polish,L.B., Gallagher,M., Fields,H.A. and Hadler,S.C. (1993) *Clin. Microbiol. Rev.* **6**, 211–229.
5 Wang,K-.S., Choo,Q-.L., Weiner,A.J., Ou,J-.H., Najarian,R.C., Thayer,R.M., Mullenbach,G.T., Denniston,K.J., Gerin,J.L. and Houghton,M. (1986) *Nature* **323**, 508–514.
6 Kos,A., Dijkema,R., Arnberg,A.C., van der Meide,P.H. and Schellekens,H. (1986) *Nature* **323**, 558–560.
7 Diener,T.O. (1985) *The Viroids*. Plenum Press, New York.
8 Francki,R.I.B. (1985) *Annu. Rev. Microbiol.* **39**, 151–174.
9 Symons,R.H. (1991) *Mol. Plant Microb. Interact.* **4**, 111–121.
10 Branch,A.D., Levine,B.J. and Robertson,H.D. (1990) *Semin. Virol.* **1**, 143–152.
11 Weiner,A.J., Choo,Q.-L., Want,K.-S., Govindarajan S., Redeker, A.G., Gerin J.L. and Houghton,M. (1988) *J. Virol.* **62**, 594–599.
12 Chao,M., Hsieh,S.-Y. and Taylor,J. (1990) *J. Virol.* **64**, 5066–5069.
13 Lee,C.-Z., Lin,J.-H., Chao M., McKnight,K., and Lai,M.M.C. (1993) *J. Virol.* **67**, 2221–2227.
14 Ponzetto,A., Hoyer,B.H., Popper,H., Engle,R., Purcell,R.H. and Gerin,J.L. (1987) *J. Infect. Dis.* **155**, 72–78.
15 Farci,P., Mandas,A., Coiana,A., Lai,M.E., Desmet, V., Van Eyken,P., Gibo,Y., Caruso,L., Scaccabarozzi,S., Criscuolo, D., Ryff,J.-C. and Balestrieri,A. (1994) *N. Engl. J. Med.* **330**, 88–94.
16 Bisceglie,A.M. (1994) *N. Engl. J. Med.* **330**, 137–138.
17 Rizzetto,M., Canese,M.G., Arico,S., Crivelli,O., Trepo,C., Bonino,F. and Verme,G. (1977) *Gut* **18**, 997–1003.
18 Chen,P-.J., Kalpana,G., Goldberg,J., Mason,W., Werner,B., Gerin,J. and Taylor,J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8774–8778.
19 Branch,A.D. and Robertson,H.D. (1984) *Science* **223**, 450–454.
20 Prody,G.A., Bakos,J.T., Buzayan,J.M., Schneider,I.R. and Bruening,G. (1986) *Science* **231**, 1577–1580.
21 Hutchins,C.J., Rathjen,P.D., Forster,A.C.,Symons, R.H. (1986) *Nucleic Acids Res.* **14**, 3627–3640.
22 Feldstein,P.A.,Buzayan,J.M.,and Bruening,G. (1989) *Gene* **82**, 53–61.
23 Sharmeen,L., Kuo,M.Y.-P., Dinter-Gottlieb,G. and Taylor,J. (1988) *J. Virol.* **62**, 2674–2679.
24 Kuo,M.Y.-P., Sharmeen,L., Dinter-Gottlieb,G. and Taylor,J. (1988) *J. Virol.* **62**, 4439–4444.
25 Wu,H.-N., Lin,Y.-J., Lin,F.-P., Makino,S., Chang,M.-F. and Lai,M.C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1831–1835.
26 Perrotta,A.T. and Been,M.D. (1991) *Nature* **350**, 434–436.
27 Branch,A.D. and Robertson,H.D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10163–10167.
28 Suh,Y.-A., Kumar,P.K.R., Kawakami,J., Nishikawa,F., Taira,K. and Nishikawa,S. (1993) *FEBS Lett.* **326**, 158–162.
29 Gottlieb,P.A., Prasad,Y., Smith,J.B., Williams,A.P. and Dinter Gottlieb,G. (1994) *Biochemistry* **33**, 2802–2808.
30 Tanner,N.K., Schaff,S., Thill,G., Petit-Koskas,E., Crain-Denoyelle,A. and Westhof,E. (1994) *Curr. Biol.* **4**, 488–498.
31 Branch,A.D., Lee,S., Neel,O.D. and Robertson H.D. (1993) *Nucleic Acids Res.* **21**, 3529–3535.
32 Riesner,D., Henco,K., Rokohl,U., Klotz,G., Kleinschmidt,A.K., Domdey,H., Jank,P., Gross,H.J. and Sänger,H.L. (1979) *J. Mol. Biol.* **133**, 85–115.
33 Luo,G., Chao,M., Hsieh,S.-Y., Sureau,C., Nishikura,K. and Taylor,J. (1990) *J. Virol.* **64**, 1021–1027.
34 Zheng,H., Fu,T.-B., Lazinski,D. and Taylor,J. (1992) *J. Virol.* **66**, 4693–4697.
35 Casey,J.L., Bergmann,K.F., Brown,T.L. and Gerin.,J.L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7149–7153.
36 Chang.,M.-F., Baker,S.C., Soe,L.H., Kamahora,T., Keck,J.G., Makino,S., Govindarajan,S. and Lai,M.M.C. (1988) *J. Virol.* **62**, 2403–2410.
37 Lin,J.-H., Chang,M.-F., Baker,S.C., Govindarajan,S. and Lai,M.M.C. (1990) *J. Virol.* **64**, 4051–4058.
38 Chao,M., Hsieh,S.-Y. and Taylor,J. (1991) *J. Virol.* **65**, 4057–4062.
39 Branch,A.D., Benenfeld,B.J., Baroudy,B.M., Wells,F.V., Gerin,J.L., and Robertson,H.D. (1989) *Science* **243**, 649–652.
40 Branch,A.D., Polaskova,J.A., and Schreiber,D.R. (unpublished).
41 Yaniv,M., Favre,A. and Barrell,B.G. (1969) *Nature* **223**, 1331–1333.
42 Ninio,J., Favre,A. and Yaniv,M. (1969) *Nature* **223**, 1333–1335.
43 Atmajda,J., Brimacombe,R., Blöcker,H. and Frank,R. (1985) *Nucleic Acids Res.* **13**, 6919–6936.
44 Zwieb,C. and Schüler,D. (1989) *Biochem. Cell Biol.* **67**, 434–442.
45 Downs,W.D. and Cech,T.R. (1990) *Biochemistry* **29**, 5605–5613.
46 Prince,J.B., Taylor,B.H., Thurlow,D.L., Ofengand,J. and Zimmermann, R.A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5450–5454.
47 Guerrier-Takada,C., Lumelsky,N. and Altman,S. (1989) *Science* **246**, 1578–1584.
48 Branch,A.D., Benenfeld,B.J. and Robertson,H.D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6590–6594.
49 Romaniuk,P.J. (1989) *Biochemistry* **28**, 1388–1395.
50 Darsillo,P. and Huber,P.W. (1991) *J. Biol. Chemistry* **266**, 21075–21082.
51 Allison,L.A., Romaniuk,P.J. and Hayes-Bakken,A. (1991) *Devel. Biol.* **144**, 129–144.
52 Wimberly,B., Varani,G. and Tinoco,I.,Jr (1993) *Biochemistry* **32**, 1078–1087.
53 Green,M.R. (1991) *Curr. Biol.* **1**, 245–247.
54 Ellington,A.D. (1993) *Curr. Biol.* **3**, 375–377.
55 Branch,A.D., Benenfeld,B.J., Paul,C.P. and Robertson,H.D. (1989) *Methods Enzymol.* **180**, 418–442.
56 Jaeger,J.A., Santa Lucia,J.,Jr and Tinoco,I.,Jr (1993) *Annu. Rev. Biochem.* **62**, 255–287.
57 Cheong,C. and Moore,P.B. (1992) *Biochemistry* **31**, 8406–8414.
58 Tabler,M., Schnölzer,M. and Sänger,H.L. (1985) *Biosci. Rep.* **5**, 143–158.
59 Branch,A.D., Benenfeld,B.J. and Robertson,H.D. (1989) *Methods Enzymol.* **180**, 130–154.
60 Barrell,B.G. (1971) *Procedures Nucleic Acids Res.* **2**, 751–779.
61 Hannon,G.J., Maroney,P.A., Branch,A.D., Benenfeld,B.J., Robertson,H.D. and Nilsen,T.W. (1989) *Mol. Cell. Biol.* **9**, 4422–4431.
62 Paul,C.P., Levine,B.J., Robertson,H.D. and Branch,A.D. (1992) *FEBS Lett.* **305**, 9–14.
63 Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
64 Holbrook,S.R., Cheong,C., Tinoco,I.,Jr and Kim,S.-H. (1991) *Nature* **353**, 579–581.
65 Cruse,W.B.T., Saludjian,P., Biala,E., Strazewski,P., Prange,T. and Kennard,O. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4160–4164.
66 Palukaitis,P. and Symons,R.H. (1979) *Virology* **98**, 238–245.
67 Wang,S.Y. (1976) *Photochemistry and Photobiology of Nucleic Acids*. Academic Press, New York, Vols 1 and 2.
68 Ofengand,J. and Liou,R. (1980) *Biochemistry* **19**, 4814–4822.
69 Behlen,L.S., Sampson,J.R. and Uhlenbeck,O.C. (1992) *Nucleic Acids Res.* **20**, 4055–4059.
70 Doudna,J.A., Grosshans,C., Gooding,A. and Kundrot,C.E. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 7829–7833.