Volume 305, number 1, 9–14 FEBS 11180 June 1992

Transcripts of the viroid central conserved region contain the local tertiary structural element found in full-length viroid

C.P. Paul[a], B.J. Levine[a], H.D. Robertson[a,b] and A.D. Branch[b]

[a]*Department of Biochemistry, Cornell University Medical College, New York, NY 10021, USA* and [b]*Center for Studies of the Biological Correlates of Addiction, The Rockefeller University, New York, NY 10021, USA*

Received 8 January 1992; revised version received 27 March 1992

The viroid central conserved region (CCR) is highly conserved among different viroids and is thought to be involved in viroid replication. A novel tertiary structure occurs in the CCR of native circular potato spindle tuber RNAs. To permit more detailed studies of this structural element, a small RNA oligonucleotide containing the CCR of the viroid genome was synthesized. The tertiary structure of these CCR transcripts was examined by UV-crosslinking of the RNA, followed by mapping of the crosslink using limited alkaline digestion and classical RNA secondary analysis. The CCR transcript was found to undergo UV-crosslinking between the same two bases as in full-length viroid, indicating that the tertiary structure is the same and that the CCR transcript will be useful for the affinity purification of host components.

Viroid; RNA structure; UV-crosslinking; Viroid–host interaction

1. INTRODUCTION

Viroids are among the smallest plant pathogens known, consisting only of circular RNA molecules of 246 to 375 bases which exist in a rod-like structure stabilized by regions of conventional base pairing which alternate with regions devoid of Watson–Crick base pairs [1]. Viroids are not believed to code for any proteins; therefore, all of their interactions with their hosts are dependent on the sequence and structure of the RNA genome. They are dependent on host factors for RNA copying, ligation, cleavage and transport. One region of the potato spindle tuber viroid (PSTV) genome is of particular interest because it is very highly conserved among the different members of the potato spindle tuber viroid group and, for this reason, is thought to be inportant in viroid replication.

Interestingly, an element of local tertiary structure has been found in the viroid central conserved region (CCR) [2,3]. Like the element of local tertiary structure in the TψC arm, identified by both Robertus et al. [4] and Rich and RajBhandary [5], the tertiary element in the viroid CCR involves bases predicted to be near each other in the secondary structure model of PSTV [1]. The tertiary element in viroid RNA was identified because it encompasses two bases which become covalently linked upon exposure to ultraviolet (UV) light. These bases must be very close to one another to react when irradiated, suggesting that non-Watson-Crick bonds occur in this region of the viroid. This novel structural element is likely to be important in the interaction between the CCR and host cell factors. To permit physical studies of the CCR by techniques such as nuclear magnetic resonance spectroscopy, which are more readily applied to molecules much shorter than full-length viroid RNA, and to facilitate its use in affinity purification of host components with which it interacts, a small RNA transcript encompassing the UV-sensitive region of local tertiary structure in viroid RNA was synthesized. The UV-sensitivity was used to examine the tertiary structure of the small transcript and to compare it to that of full-length viroid. The tertiary structure of the 55-base transcript was found to be the same as that of the CCR when it is part of the complete viroid RNA molecule, indicating that this transcript contains all of the structural features necessary to produce the UV-sensitive element.

Correspondence address: C.P. Paul, Department of Plant Pathology, 351 Bessey Hall, Iowa State University, Ames, IA 50011, USA. Fax: (1) (515) 294-1337.

2. MATERIALS AND METHODS

2.1. *Synthesis of transcripts*

The 55-base transcripts representing the viroid CCR were synthesized according to the procedure of Milligan et al. [6] in 500 μl reactions. Synthetic DNA oligonucleotides used as transcription templates were obtained from the Memorial Sloan-Kettering Micro Core facility or were the kind gift of Drs. Bahige Baroudy and Alicia Buckler-White at Georgetown University. Transcripts were labeled at the 5′ end by the inclusion of [γ-^{32}P]GTP in the reaction mixture or were internally labeled using one of the four [α-^{32}P]-labeled nucleoside triphosphates. Transcripts to be 3′ end-labeled were synthesized in the absence of labeled nucleoside triphosphates and were subsequently labeled by the addition of [^{32}P]pCp to the 3′ terminus using T4 RNA ligase in 15 μl reactions [7]. Full-length transcripts were purified by electrophoresis in a 10% polyacrylamide gel containing 7 M urea using 0.5× Tris-

borate buffer (1×= 0.089 M Tris pH 8.2, 0.089 M boric acid and 0.0028 M EDTA).

Full-length viroid transcripts were synthesized from the linearized plasmid pHD1 using SP6 RNA polymerase in 20 μl reactions [8], yielding a plus strand monomeric PSTV sequence beginning with base 147 and ending with base 146 (in the numbering system of Gross et al [1]) flanked by vector sequences at the 5′ and 3′ ends (D. Rosen, unpublished data). The full-length viroid cDNA sequence was obtained from pAV401 (a gift of A. Van Kammen) [9].

2.2. *UV irradiation and purification of crosslinked RNAs*

Two-dimensional gel electrophoresis and UV-crosslinking were carried out as previously described in communications from this laboratory [2,3,10]. Separation in the first dimension was performed in 8% polyacrylamide gels for CCR transcripts and 5% polyacrylamide gels for HD1 transcripts. Urea gels used for second dimension separation were 10% polyacrylamide for CCR transcripts and 5% polyacrylamide for HD1 transcripts. The crosslinked (X-linked) and non-crosslinked (UV, not X-linked) forms of the transcripts were recovered from gels containing RNA exposed to UV light as were transcripts which were run on two-dimensional gels, but not exposed to UV light (No UV). Under the conditions used, an average of 5% of the sample was found to undergo crosslinking when CCR transcripts were exposed to UV light.

2.3. *Mapping the crosslink using alkaline hydrolysis of the RNA*

CCR transcripts were heated at 90°C in 10 μl of 0.05 M sodium bicarbonate/carbonate at pH 9.2 in the presence of tRNA and subjected to electrophoresis at 1000 V in 20% polyacrylamide gels containing 7 M urea. The 5′ end-labeled transcripts were separated on gels run in 1× Tris-borate buffer, while gels used for 3′ end-labeled transcripts were made with 0.5× Tris-borate buffer, with 0.5× buffer in the top reservoir and 1× buffer in the bottom reservoir. Non-crosslinked transcripts treated with alkali as described for crosslinked samples or digested with base-specific ribonucleases were run in neighboring lanes as size markers. Digestions with 1 mg/ml pancreatic RNase were performed in 2 microfilters of buffer (10 mM Tris pH 7.4, 1 mM EDTA) in the presence of 10 μg tRNA for 30 min at 37°C. Samples were treated with RNase U2 at a concentration of 10 U/ml in buffer containing 0.05 M sodium acetate pH 4.5 and 0.002 M EDTA for 2 h at 37°C in the presence of 10 μg tRNA. Incubations with RNase T1 were done under strong digestion conditions (1 mg/ml RNase T1 in 2 microliters of T1 buffer (10 mM Tris pH 7.4, 1 mM EDTA) in the presence of 10 μg tRNA for 65°C for one h) or under mild digestion conditions designed to result in partial digestion of the RNA (2 μg/ml RNase T1 in 5 μl of buffer containing 25 mM sodium citrate pH 5.0, 7 M urea, 1 mM EDTA, 0.035% xylene cyanol and 0.035% Bromophenol blue in the presence of 1.25 μg tRNA for 3 min at 65°C).

2.4. *Classical RNA secondary analysis; direct RNA sequence analysis by sequential enzymatic digestions*

In preparation for RNA secondary analysis, internally labeled crosslinked CCR and full-length transcripts were digested with 1 mg/ml RNase T1 in T1 buffer for 45 min at 37°C in the presence of 10 μg tRNA. The resulting RNase T1 digestion products were separated on 20% polyacrylamide, 7 M urea gels using 0.5× Tris-borate buffer. RNase T1-resistant oligonucleotides were recovered from gels and subjected to subsequent RNase digestion and one-dimensional high-voltage electrophoresis on Whatman 3MM or DEAE paper [11–13].

3. RESULTS

The CCR transcript contains the region of the viroid surrounding the local tertiary structural element found in the central conserved region of circular viroid [2] and full-length viroid transcripts (Branch, Levine, and Robertson, unpublished data) (Fig. 1). It is made up of two

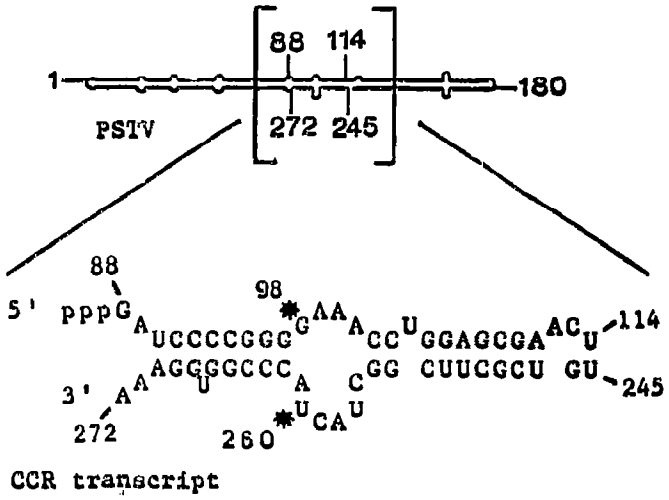


Fig. 1. The 55-base CCR transcript is made up of bases 88–114 and 245–272 from the central conserved region of PSTV (numbering system of Gross et al. [1]).

regions which are non-contiguous, but interacting in the viroid genome. In the rod-like structure of PSTV two helices lie on either side of a middle region devoid of Watson–Crick base pairs. In the CCR transcript, bases 88–111 and 247–272 of full-length viroid are joined by a loop made up of bases 112–114 and 245–246 (ACUUG), which were included to provide a linker allowing base-pairing between complementary regions. UV-crosslinking was used to determine whether the CCR transcript formed the local tertiary structure which is characteristic of this region in genomic PSTV.

Exposure of the CCR transcript to UV light produced a new species which could be resolved by two-dimensional gel electrophoresis (compare Fig. 2A and B). The crosslinked CCR transcript exhibited increased mobility relative to the non-crosslinked transcript in the second dimension. This change in mobility reflects the conformation of the crosslinked molecule after denaturation and allows purification and further analysis of the crosslinked RNA. When samples were loaded on to gels at high concentration, aggregation of transcripts could occur. Aggregates were unrelated to crosslinking as they separated in the first dimension, before exposure to UV light. Further analysis of CCR transcripts eluted from 2D gels on sequencing gels occasionally revealed the presence of species which did not comigrate with the crosslinked transcript. These species were not characterized, but may contain a crosslink at another site. These minor species did not affect the results in mapping of the major crosslink.

The positions of the bases involved in the UV-induced crosslink were mapped by alkaline digestion of end-labeled CCR transcripts. Digestion of 5′ end-labeled non-crosslinked transcripts under conditions which introduced an average of one cleavage per mole-

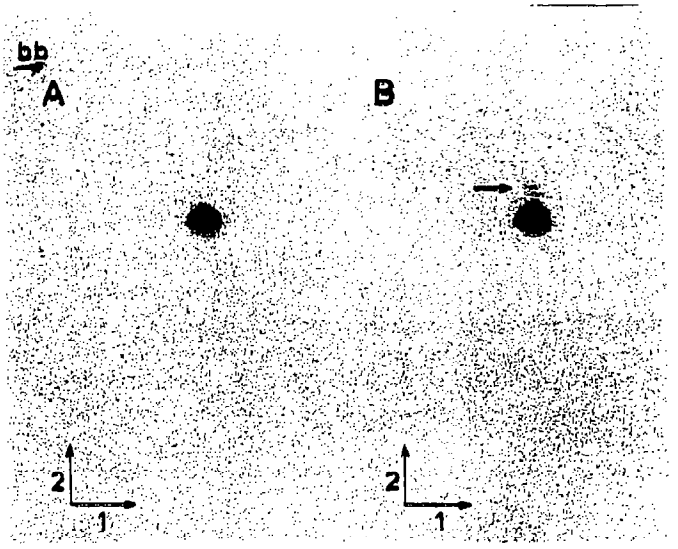
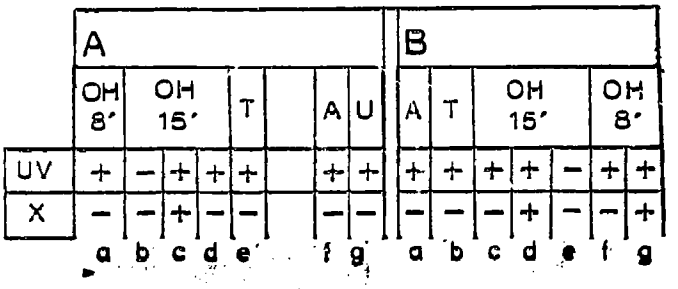


Fig. 2. Separation of crosslinked from non-crosslinked CCR transcripts by two-dimensional gel electrophoresis. (A) CCR transcripts which were not exposed to UV light. (B) CCR transcripts which were exposed to UV light. The position of the spot representing crosslinked RNA is indicated by an arrow. The pairs of arrows labeled 1 and 2 indicate the directions of electrophoresis in the first and second dimensions, respectively.

cule resulted in a ladder where each band differed in length by one base from the bands on either side of it when subjected to electrophoresis in sequencing gels (Fig. 3A, lanes a, b and d). When 5′ end-labeled crosslinked CCR transcripts were treated with alkali, a gap appeared in the ladder which began at the position of the crosslinked base (Fig. 3A, lane c). The crosslink, which itself is resistant to cleavage under alkaline conditions, connects pairs of fragments generated by the alkaline cleavage. These linked fragments migrated much more slowly in the gel. The identity of one of the crosslinked bases as G-98 (in the numbering system of full-length viroid) was determined by comparison with nearby lanes containing digests of non-crosslinked transcripts produced either by alkaline hydrolysis or by base-specific ribonucleases (Fig. 3A, lanes e, f and g).

A similar analysis using 3′ end-labeled transcripts suggested the identity of the other base involved in the crosslink. It was clear that either U-260 or A-261 participated in the crosslink with the weight of evidence pointing to U-260 as the crosslinked base (Fig. 3B, lane d). Interpretation of these results was often complicated due to band compressions and to the difference in mobility of the RNA fragments in the enzyme digestion buffer used for markers and the alkaline hydrolysis buffer.



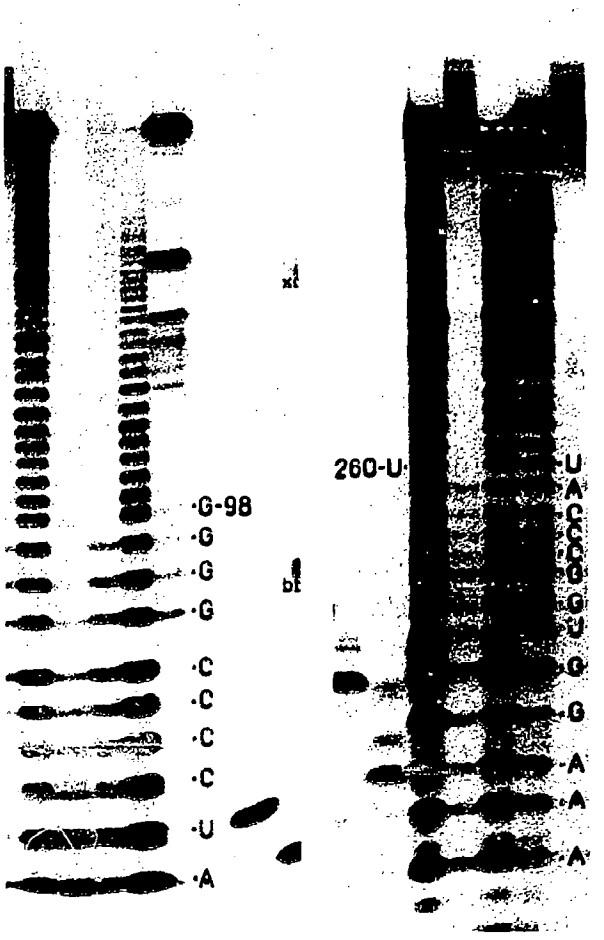


Fig. 3. Mapping of the UV-induced crosslink in the CCR transcript by limited alkaline hydrolysis. Digestion products were separated on a 20% polyacrylamide, 7 M urea gel. Lanes in which enzyme-digested RNA markers (T, RNase T1; A, pancreatic RNase, U, RNase U2) and alkali-treated samples (OH) were run are indicated, as is the time for the alkali treatment. Symbols indicate whether the RNA in each lane was exposed to UV light (+) or not (−) and whether it was crosslinked (+) or noncrosslinked (−). A. 5′ end-labeled CCR transcripts. Treatments: a, 8 min alkali; b–d, 15 min alkali; e, RNase T1 under mild digestion conditions; f, pancreatic RNase; g, RNase U2. Transcripts were recovered from two-dimensional gels: lane c, X-linked; lane b, no UV; lanes a and d–g, UV, not X-linked. The sequence is indicated, showing the position of G-98, the crosslinked base. An arrowhead indicates the origin of electrophoresis. The positions of the xylene cyanol (xc) and bromophenol blue (bb) are marked. B. 3′ end-labeled CCR transcripts. Treatments: a, pancreatic RNase; b, RNase T1 under strong digestion conditions; c–e, 15 min alkali; f and g, 8 min alkali. Transcripts were recovered from two-dimensional gels: lanes d and g, X-linked; lane e, no UV; lanes a–c and f, UV, not X-linked. The sequence is indicated in lane g, obscuring the bands. The position of U-260, the crosslinked base, is marked.

Fig. 4. One-dimensional gel separation of RNase T1-digested CCR transcripts. Transcripts were internally labeled by the addition of [α^{32}P]UTP to the transcription reaction mixture. The oligonucleotides generated by digestion of the transcripts with RNase T1 (1 mg/ml, 37°C) were subjected to electrophoresis in a 20% polyacrylamide gel containing 7 M urea. The symbols above the lanes indicate whether the RNA was exposed to UV light (+) and whether it was crosslinked (+) or non-crosslinked (−). C, CCR transcript; F, full-length transcript. Lanes: a, CCR transcript, X-linked; b, full-length transcript, X-linked; c, CCR transcript, UV, not X-linked; d, full-length transcript, UV, not X-linked. The arrow indicates the band which is present in crosslinked samples, but not in samples lacking the crosslink. Bands 1 and 2 are present in the lanes containing UV, not X-linked CCR transcripts, but not in the lanes containing crosslinked CCR transcripts. The positions of the xylene cyanol (xc) and bromophenol blue (bb) markers are indicated.

In order to ascertain which of these bases, U-260 or A-261, participated in the crosslink, further information was obtained through classical RNA secondary analysis. CCR and full-length transcripts, which were internally labeled with one of the four nucleoside triphosphates, were exposed to UV light and the crosslinked forms of the transcripts were recovered from two-dimensional gels. The transcripts were then digested with RNase T1 and the products separated by gel electrophoresis. RNase T1 digestion products of CCR transcripts which are not crosslinked are: CUACUACCCG, AAACCUG, AUCCCCG, AACUUG, CUUCG, UCG, AAA, AG, CG, UG, G, pppG. A novel RNase T1 product was found in crosslinked samples (Fig. 4, lanes a and b) that was not found in non-crosslinked samples (Fig. 4, lanes c and d). Furthermore, two prominent oligonucleotides were missing from the products generated from the crosslinked CCR transcript, suggesting that they form part of the crosslinked RNase T1 product (Fig. 4, lanes a and c, bands 1 and 2). Two-dimensional RNA fingerprinting, a technique used in mapping the crosslink in circular viroid [2], suggested the identity of these fragments as CUACUACCCG and AAACCUG (data not shown). For further characterization of the UV-induced crosslink in the CCR transcript, the RNA was eluted from bands representing the crosslinked RNase T1-generated oligonucleotides which were excised from one-dimensional gels similar to the one shown in Fig. 4. These species were then subjected to digestion with several base-specific ribonucleases followed by high voltage electrophoresis. Digestion products of the CCR transcripts were compared with those from the full-length transcripts.

Digestion of the crosslinked RNase T1-generated oligonucleotides with RNase U2 produced several species, including two which comigrated with CUA and CCCG. CUA oligonucleotides were found when transcripts were internally labeled with [α-^{32}P]ATP, [α-^{32}P]CTP, or [α-^{32}P]UTP, while CCCG oligonucleotides were found when the label was [α-^{32}P]GTP or [α-^{32}P]CTP. Examination of the CCR sequence shows that the RNase T1-generated fragment CUACUACCCG must be involved in crosslinking. This sequence contains U-260 and A-261, the candidates for crosslinking identified by alkaline hydrolysis of 3′ end-labeled transcripts. Ribonucleases often exhibit difficulty in cleaving the crosslinked base [2] and RNase U2 is not a particularly strong ribonuclease. However, RNase U2, which specifically recognizes adenine residues under these conditions, did not have any difficulty in cleaving after the adenine residues to give the CUA fragments. This suggests that the crosslink occurs at U-260. Thus, the combined evidence from the alkaline mapping and RNA secondary analysis show that U-260 is one of the crosslinked bases. Results of RNA secondary analysis also showed that, as in circular viroid [2], the partial RNase T1 digestion product GAAACCUG participated in the crosslink and confirmed the identity of the other base involved in the crosslink as G-98 (data not shown) as determined by the alkaline hydrolysis experiments.

4. DISCUSSION

The CCR transcripts undergo UV-crosslinking between the same two bases as do native circular viroid [2] and full-length transcripts (Branch, Levine, and Robertson, unpublished data) indicating that the same tertiary structural element found in the larger molecules is present in the 55-base transcripts. Upon irradiation with ultraviolet light, bases which are very near each other may become covalently linked, providing information on the three-dimensional arrangement of the bases. UV-crosslinking was useful in developing a model of tRNA structure which proved to be very nearly accurate when the crystal structure was determined [14,15] and has more recently been used to gather information on the tertiary structures of ribosomal RNAs [2,16,17], the hepatitis delta RNA [18], M1 RNA enzyme-substrate complex [19], the RNA from the signal recognition particle [20], and the *Tetrahymena* self-splicing intervening sequence RNA [21]. The two crosslinked bases in full-length viroids and CCR transcripts, G-98 and U-260, must be positioned very near each other in order to undergo crosslinking, but are in a region of the viroid that is depicted as an open loop between two base-paired regions in the secondary structure model [1]. Thus, additional tertiary structure is present in this region of the viroid. The CCR transcript was designed to allow for folding of the transcript so that base-pairing could occur between regions which form helices in the native rod-shaped structure of viroids. Thus, the small transcript has the sequence and secondary structure of the viroid CCR and these experiments have shown that it also folds into the same three-dimensional arrangement, indicating that these 55 bases are sufficient to allow formation of the local tertiary structural element.

The viroid CCR is a region in the middle of the rod-like secondary structure which is highly conserved among the different viroids and is thought to be involved in replication of viroid RNA [22]. The CCR transcript used here is taken from the sequence of PSTV; however, it shares extensive sequence and secondary structural similarity with the CCRs of the other viroids of the PSTV group [22–24]. The extent of conservation at both the sequence and structural levels suggests an important functional role for this region of the viroid. The CCR transcript, due to its conservation of tertiary structure, could provide a useful tool for elucidating this function. For example, scaled-up synthesis of the CCR transcript should allow for efficient isolation of host components which interact with it.

The role of RNA structure is particularly important in viroids as they are not believed to code for proteins; thus, host factors that are necessary for completion of their life cycle must interact directly with the RNA genome. The viroid central conserved region is likely to be one of a number of RNA domains which interact with host cell components involved in RNA copying, cleavage, ligation, or transport. Particular host protein–viroid interactions are suggested by the similarity in sequence and structure between the viroid CCR and 5S ribosomal RNA [2,25]. These similarities are in a region of 5S ribosomal RNA which contains the binding sites for transcription factor IIIA [26–30] and ribosomal protein L5 [31]; thus, these proteins may interact with viroid RNA as well. Riesner [32,33] and Forster and Symons [34] have discussed the role of possible alternative viroid RNA conformations. Furthermore, Wolff et al. [35] have published an assay for viroid RNA-host protein binding. The small CCR transcript provides an important tool for further experiments into viroid-host interactions and serves as a model for the design of small transcripts representing other viroid functional domains as well as domains of other biologically important RNAs.

Acknowledgements: This research was supported by a NIH postdoctoral fellowship to C.P.P. (1F32AI07940-01A2) and by a NIH Center Grant (DA-05130), Dr. M.J. Kreek, Program Director.

REFERENCES

[1] Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sänger, H.L. (1978) Nature 273, 203–208.
[2] Branch, A.D., Benenfeld, B.J. and Robertson, H.D. (1985) Proc. Natl. Acad. Sci. USA 82, 6590–6594.
[3] Branch, A.D., Benenfeld, B.J. and Robertson, H.D. (1985) Nucleic Acids Res. 13, 4889–4903.
[4] Robertus, J.D., Ladner, J.E., Finch, J.T., Rhodes, D., Brown, R.S., Clark, B.F.C. and Klug, A. (1974) Nature 250, 546–551.
[5] Rich, A. and RajBhandary, U.L. (1976) Annu. Rev. Biochemistry 45, 805–860.
[6] Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Nucleic Acids Res. 15, 8783–8798.
[7] England, T.E. and Uhlenbeck, O.C. (1978) Nature 275, 560–561.
[8] Robertson, H.D., Rosen, D.L. and Branch, A.D. (1985) Virology 142, 441–447.
[9] van Wezenbeek, P., Vos, P., van Boom, J. and van Kammen, A. (1982) Nucleic Acids Res. 10, 7947–7957.
[10] Branch, A.D., Benenfeld, B.J., Paul, C.P. and Robertson, H.D. (1989) Methods Enzymol. 180, 418–442.
[11] Barrell, B.G. (1971) in: Procedures in Nucleic Acids Research (G.L. Cantoni, and D.R. Davies, Eds.) vol. 2, Harper and Row, New York, pp. 751–779.
[12] Dickson, E., Pape, L.K. and Robertson, H.D. (1979) Nucleic Acids Res. 6, 91–110.
[13] Robertson, H.D., Dickson, E., Plotch, S.J. and Krug, R.M. (1980) Nucleic Acids Res. 8, 925–942.
[14] Yaniv, M., Favre, A. and Earrell, B.G. (1969) Nature 223, 1331–1333.
[15] Ladner, J.E., Jack, A., Robertus, J.D., Brown, R.S., Rhodes, D., Clark, B.F.C. and Klug, A. (1975) Proc. Natl. Acad. Sci. USA 72, 4414–4418.
[16] Zwieb, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2397–2411.
[17] Stiege, W., Glotz, C. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 1687–1706.
[18] Branch, A.D., Benenfeld, B.J., Baroudy, B.M., Wells, F.V., Gerin, J.L. and Robertson, H.D. (1989) Science 243, 649–652.
[19] Guerrier-Takada, C., Lumelsky, N. and Altman, S. (1989) Science 246, 1578–1584.

[20] Zwieb, C. and Schüller, D. (1989) Biochem. Cell Biol. 67, 434–442.
[21] Downs, W.D. and Cech, T.R. (1990) Biochemistry 29, 5605–5613.
[22] Keese, P. and Symons, R.H. (1985) Proc. Natl. Acad. Sci. USA 82, 4582–4586.
[23] Koltunow, A.M. and Rezaian, M.A. (1989) Intervirology 30, 194–201.
[24] Elena, S.F., Dopazo, J., Flores, R., Diener, T.O. and Moya, A. (1991) Proc. Natl. Acad. Sci. USA 88, 5631–5634.
[25] Branch, A.D., Levine, B.J. and Robertson, H.D. (1990) Seminars in Virology 1, 143–152.
[26] Pieler, T. and Erdmann, V.A. (1983) FEBS Lett. 157, 283–287.
[27] Andersen, J., Delihas, N., Hanas, J.S. and Wu, C.-W. (1984) Biochemistry 23, 5759–5766.
[28] Romaniuk, P.J. (1985) Nucleic Acids Res. 13, 5369–5387.
[29] Huber, P.W. and Wool, I.G. (1986) Proc. Natl. Acad. Sci. USA 83, 1593–1597.
[30] Christiansen, J., Brown, R.S., Sproat, B.S. and Garrett, R.A. (1987) EMBO J. 6, 453–460.
[31] Huber, P.W. and Wool, I.G. (1986) J. Biol. Chem. 261, 3002–3005.
[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] Riesner, D. (1991) Molecular Plant-Microbe Interactions 4, 122–131.
[34] Forster, A.C. and Symons, R.H. (1987) Cell 49, 211–220.
[35] Wolff, P., Gilz, R., Schumacher, J. and Riesner, D. (1985) Nucleic Acids Res. 13, 355–367.