Volume 14 Number 10 1986 Nucleic Acids Research

Two juxtaposed tyrosyl-OH groups participate in ϕX174 gene A protein catalysed cleavage and ligation of DNA

A.D.M.van Mansfeld*, H.A.A.M.van Teeffelen, P.D.Baas and H.S.Jansz

Institute of Molecular Biology and Laboratory for Physiological Chemistry, State University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received 25 March 1986; Accepted 23 April 1986

ABSTRACT

Bacteriophage ØX174 encoded gene A protein is an enzyme required for initiation and termination of successive rounds of rolling circle ØX DNA replication. This enzyme catalyses cleavage and ligation of a phosphodiester bond between nucleotide residues G and A at the ØX origin. The cleavage reaction which occurs during initiation involves formation of a free G_{OH} residue at one end and a covalent bond between tyrosine-OH of the gene A protein and 5' phosphate of the A residue, at the other end of the cleavage site. During termination the covalently bound gene A protein cleaves the phosphodiester bond between G and A at the regenerated origin and ligates the 3' and 5' ends of the displaced genome-length viral DNA to form a circle. Since tyrosyl-OH mediated rearrangements of phosphodiester bonds in DNA may also apply to other enzymes involved in replication or recombination such as topoisomerases we have studied this interesting mechanism in greater detail. Analysis of ^{32}P-labelled gene A protein-DNA complex by tryptic digestion followed by sequencing of ^{32}P-containing peptides showed that two tyrosyl residues in the repeating sequence tyr-val-ala-lys-tyr-val-asn-lys participate in phosphodiester bond cleavage. Either one of these tyrosyl residues can function as the acceptor of the DNA chain. In an α-helix the side chains of these tyrosyl residues are in juxtaposition. An enzymatic mechanism is proposed in which these two tyrosyl-OH groups participate in an alternating manner in successive cleavage and ligations which occur during phosphodiester bond rearrangements of DNA.

INTRODUCTION

The DNA replication of bacteriophage ØX174 proceeds by a rolling circle mechanism (for recent reviews see refs. 1 and 2). In order to start the process, the protein encoded by the viral gene A (gene A protein) cleaves the viral strand of the double-stranded replicative form DNA (RF DNA). Cleavage requires the presence of the 30 base pair origin region, a sequence which occurs in a highly conserved form in all isometric single-stranded DNA phages studied (3-6). Cleavage takes place at a specific site within this region (7), between nucleotides G and A at positions 4305 and 4306 of the ØX DNA sequence (8). The cleavage reaction involves the formation of a free 3'-hydroxyl group at the G residue (7,9,10) and a covalent bond (7,11,12) between the 5'-phos-

© IRL Press Limited, Oxford, England.

phate of the A residue and the 4'-hydroxyl group of one of the tyrosyl residues in gene A protein (13-15). The 3'-hydroxyl group is the primer for the DNA synthesis. The DNA at the 5' end of the cleavage site is displaced during DNA replication. After one round of DNA replication the covalently bound gene A protein cleaves the regenerated origin and binds covalently to the new 5' end. Simultaneously the 3' and 5' ends of the displaced genome-length viral DNA are ligated to form a circle, which is thrown off.

In order to study the mechanism by which this rearrangement of phosphodiester bonds proceeds, we have investigated which of the 20 tyrosyl residues in gene A protein bind covalently to DNA upon cleavage. Therefore, we have used the previous observation that gene A protein can cleave synthetic oligonucleotides which contain at least the first 10 nucleotides of the origin region of ØX DNA (16). Also in these conditions gene A protein forms a covalent bond with the phosphate at the 5' end of the cleavage site via the 4'-hydroxyl group of tyrosine (14). A second product of gene A is A* protein. A* protein is the result of an internal (in frame) translational start in the gene which encodes gene A protein (17). A* protein thus lacks the N-terminal third of the polypeptide chain of gene A protein. A* protein is present in ØX *am3* infected cells at 2 hours after infection in much larger amounts than gene A protein and is isolated simultaneously with gene A protein (unpublished observations). A* protein has retained a number of activities of gene A protein: A* protein cleaves oligonucleotides with the sequence of the origin region (16), binds, like gene A protein via a tyrosyl residue to the phosphate of the phosphodiester bond which is cleaved (16,18) and can ligate DNA (19-21). Analysis of the amino acid residues which surround this/these particular tyrosyl residue(s) in gene A protein or A* protein, in combination with the predicted amino acid sequence of gene A protein will then reveal which tyrosyl residue(s) is/are involved.

MATERIALS AND METHODS

Materials

Gene A protein and A* protein were purified as described by Langeveld *et al.* (19), with the modification that the ØX *am3* infected *E. coli* C cells were harvested at 2 hrs post infection. The internally ^{32}P-phosphate labelled oligonucleotide CAACTTG*ATATTAATAAC was constructed as described before (14). TPCK treated trypsin was from Worthington (Freehold, N.Y., U.S.A.) and proteinase K was from Merck (Darmstadt, F.R.G.).

Methods

Analysis of the gene A protein/A* protein-oligonucleotide complexes. The internally labelled oligonucleotide, 0.03 pmol, was incubated with approximately 0.1 pmol gene A protein or A* protein in 20 mM Tris/HCl pH 7.5, 1 mM EDTA, 5 mM $MgCl_2$, 5 mM dithiotreitol (DTT), 150 mM NaCl and 5% glycerol for 15 min at 30°C in a volume of 16 µl. Then 1.6 µl 50 mM $CaCl_2$ and 1 µl trypsin, 1 mg/ml, or 1 µl proteinase K, 5 mg/ml, which had been preincubated for 30 min at 37°C, were added and the incubations were continued for 3 hrs at 37°C. 15 µl 98% formamide containing 0.05% bromophenol blue and 0.05% xylene cyanol F as tracing dyes was added, the samples were heated for 3 min at 100°C and applied on a 20% polyacrylamide gel which contained 7 M urea and 100 mM Tris, 100 mM borate and 2 mM EDTA. After electrophoresis the radioactive products were detected by autoradiography.

Isolation of the peptide-oligonucleotide complexes which are formed by digestion of the A* protein-oligonucleotide complex with trypsin. Approximately 5 nmol A* protein were incubated with 15 nmol of the labelled oligonucleotide CAACTTG*ATATTAATAAC in 20 mM Tris/HCl pH 7.5, 1 mM EDTA, 5 mM $MgCl_2$, 5 mM dithiotreitol (DTT), 150 mM NaCl and 5% glycerol for 15 min at 30°C in a final volume of 60 ml, devided over 12 polyethylene tubes. The reaction was stopped by the addition of 0.3 ml 100 mM EDTA pH 7.5 to each tube. Then 5,3 ml ice-cold 20% trichloroacetic acid (TCA) was added to each tube and the protein-oligonucleotide complex was spun down after standing for 20 min at 0°C. The pellets were washed subsequently with ice-cold 10% TCA and ice-cold diethylether and dried. Each pellet was dissolved in 1 ml 50 mM Tris/HCl 8.0, 5 mM $CaCl_2$ containing 25 µg TPCK treated trypsin. The samples were incubated for 16 hrs at 37°C with continuous shaking, pooled and applied on a mono Q-anion exchange column (Pharmacia, Uppsala, Sweden). The column was eluted with a 0 to 2 M NaCl gradient in 10 mM Tris/HCl pH 7.5 using the Pharmacia FPLC equipment. The fractions of the broad peak which contained the ^{32}P radioactivity were pooled. The pool contained 3.05 nmol peptide-oligonucleotide complex, as could be calculated from the specific radioactivity of the oligonucleotide. 2 nmol of this pool were subjected to reversed phase chromatography on a pepRPC column (Pharmacia) using an acetonitrile gradient (buffer A: 10 mM NH_4Ac, buffer B: 10 mM NH_4Ac, 60% acetonitrile). The radioactivity of the fractions (0.5 ml) was determined and the U.V. absorption at 214 nm was recorded. The first peak, peak A, contained 0.45 nmol peptide-oligonucleotide complex and the second peak, peak B, contained 0.96 nmol peptide-oligonucleotide complex. The fractions were lyophilized, dissolved in 50 µl H_2O and lyophilized again.

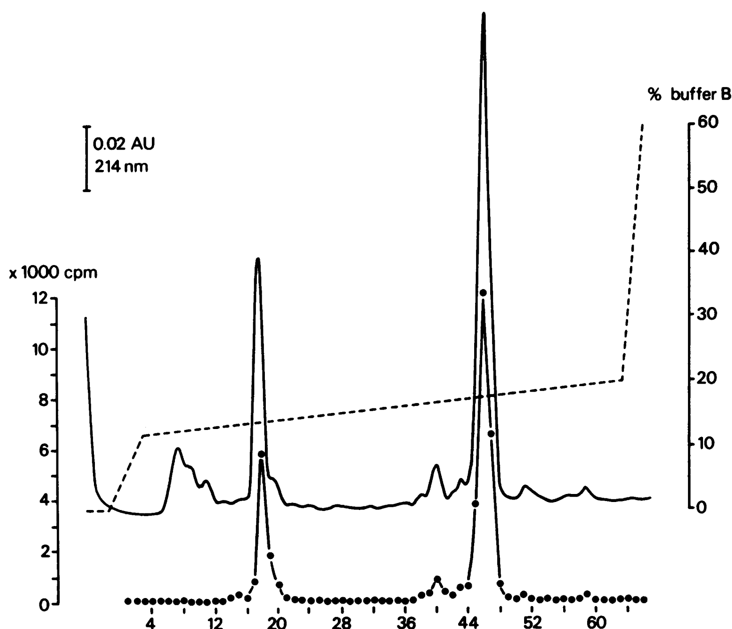

Figure 1. Reversed phase chromatography of the peptide-oligonucleotide complexes obtained after tryptic digestion of the A* protein-oligonucleotide complex. 2 nmol peptide-oligonucleotide complex, obtained after ion-exchange chromatography of the tryptic digest of A* protein-oligonucleotide complex, was submitted to reversed phase chromatography using an acetonitrile gradient (indicated by the broken line). The radioactivity-pattern is indicated by the line with closed circles and the UV absorption is indicated by the uninterrupted line.

Amino-acid sequencing of the peptide moieties of the two peptide-oligonucleotide complexes. 0.33 nmol of the peptide-oligonucleotide complex as present in the first peak after reversed phase chromatography was dissolved in 20 µl H_2O, applied on a polybrene-impregnated glass fiber filter and mounted in an Applied Biosystems 470 A gas phase sequencer (Applied Biosystems, Foster City, Ca., U.S.A.). The first degradation cycle was started after three precycles, using standard conditions. The PTH-amino acids were analysed on-line using an Applied Biosystems 120 A PTH Analyzer. The sequencing run was ended after 8 cycles. A considerable amount of diphenylthiourea (DPTU) was formed during cell cycles. Identification of the PTH-amino acids was unambiguous. 0.55 nmol of the peptide-oligonucleotide complex as present in the second peak after reversed phase chromatography was analysed as above. The sequencing run was ended after 11 cycles.

RESULTS

In order to label the tyrosyl residue(s) in gene A protein to which the DNA is bound covalently upon cleavage, gene A protein was incubated with the oligonucleotide CAACTTA*ATATTAATAAC. This oligonucleotide has a ^{32}P-phosphate (*) at the phosphodiester bond which is cleaved (14). The resulting ^{32}P-labelled gene A protein-oligonucleotide complex was digested with trypsin or proteinase K. In both cases two radioactive peptides were detected upon analysis of the digestion products by electrophoresis in polyacrylamide gel followed by autoradiography. Similar results were obtained by Röth *et al.* (13) and Sanhueza and Eisenberg (15). Not enough pure gene A protein was available for determination of the amino acid sequences of the peptide moieties of these peptide-oligonucleotide complexes. Therefore we performed these analyses using A* protein instead of gene A protein. A* protein was incubated with the above ^{32}P-labelled oligonucleotide and the A* protein-oligonucleotide complex was digested with trypsin or proteinase K. In both cases two radioactive peptides were detected upon analysis of the digestion products by electrophoresis in polyacrylamide gel followed by autoradiography. The products had exactly the same properties as the products obtained after proteolytic digestion of the gene A protein-oligonucleotide complex. Sanhueza and Eisenberg have also observed two radioactive products after tryptic digestion of ^{32}P-labelled A* protein-DNA complex following a different approach (18). To prepare larger quantities of the peptide-oligonucleotide complexes approximately 5 nmol A* protein were incubated with 15 nmol of the oligonucleotide which contained ^{32}P at the cleavage site. The protein-oligonucleotide complex was digested with trypsin and the peptide-oligonucleotide complexes were isolated by ion-exchange and reversed phase chromatography (Figure 1). The radioactivity patterns in the peaks coincide exactly with the UV absorption patterns. The yields were calculated from the known specific radioactivity of the labelled oligonucleotide. 0.33 nmol peptide-oligonucleotide complex as present in peak A and 0.55 nmol peptide-oligonucleotide complex as present in peak B were subjected to amino acid sequence analysis in a gas phase sequencer with on-line detection of the PTH-amino acid derivatives. The results of the PTH-amino acid analyses of 5 successive cycles of sample A and 10 successive cycles of sample B are shown in Figure 2. Analysis of the material of the first cycle of sample A does not yield a PTH-amino acid. The following cycles clearly show: val, asn, lys,-. The sequencing run was ended after 8 cycles. About 40% of the radioactivity which had been applied on the glass fiber filter in the sequencer was retained on the filter after the sequencing run. The PTH-amino acid analyses of succes-

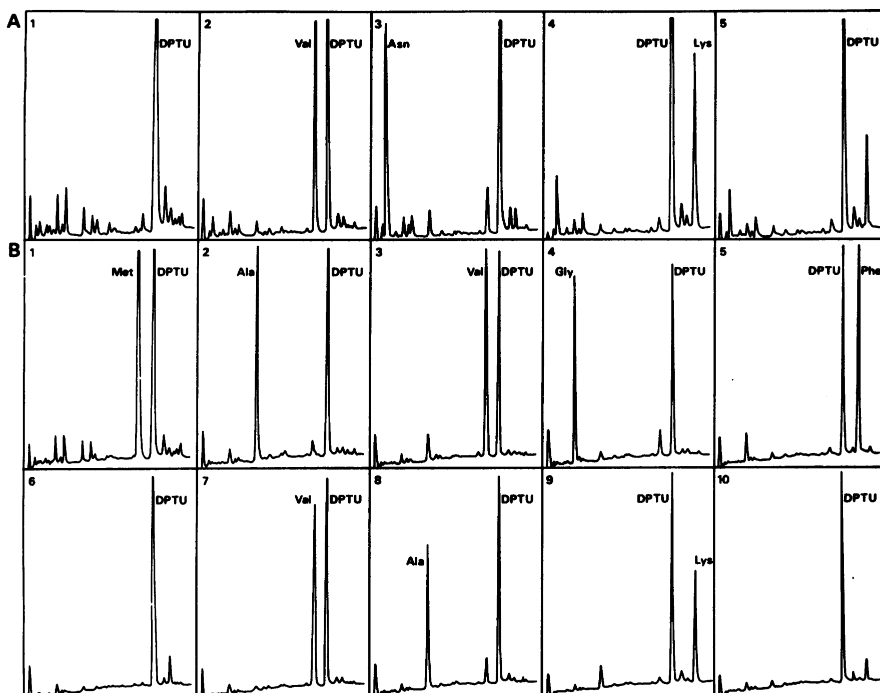

Figure 2. The amino acid sequences of the peptide moieties of the two peptide-oligonucleotide complexes. A. The results of PTH amino acid analyses of the first 5 gas phase sequencer cycles of peptide-oligonucleotide complex A. Panels A 1-5: -, val, asn, lys, -. B. The results of the PTH-amino acid analyses of the first 10 gas phase sequencer cycles of peptide-oligonucleotide complex B. Panels B 1-10: met, ala, val, gly, phe, - , val, ala, lys, - .

sive sequencing cycles of sample B show: met, ala, val, gly, phe,-, val, ala, lys,-. No PTH-amino acid was detected in cycle 6. The sequencing run was ended after 11 cycles. 70% of the radioactivity which had been applied on the glass fiber filter in the sequencer was retained on the filter after the sequencing run. Alignment of the two amino acid sequences with the predicted tyrosine containing tryptic peptides of A* protein (8) shows the following. The amino acid sequence of sample A corresponds to the tryptic peptide tyr, val, asn, lys which has a tyrosyl residue at position one. The amino acid sequence of sample B corresponds to the C-terminal part of the tryptic peptide ala, thr, ser, tyr, met, ala, val, gly, phe, tyr, val, ala, lys. This sequence has a tyrosyl residue at the position where no amino acid was detected in the sequencing run of sample B. These alignments suggest that tyrosyl residues were present in the

peptides at the positions where no amino acid was detected. These tyrosyl residues have probably been split off but, since they carry olignucleotides covalently bound to the hydroxyl groups of the tyrosyl residue, they produce polar tyrosyl-phenyl-thiazolinon derivatives that do not dissolve in the solvents and remain on the filters. The observation that in both cases a major part of the radioactivity which had been applied was still present on the filter after the sequencing run supports this explanation. The peptide moiety of sample B lacks the N-terminal part of the predicted tryptic peptide to which it corresponds. This is probably caused by a contaminating chymotryptic activity present during the digestion of the A* protein-oligonucleotide complex with trypsin. Chymotrypsin can easily cleave the peptide bond between the tyrosyl and the methionyl residues.

DISCUSSION

The results show that the peptide moieties of the two peptide-oligonucleotide complexes differ and thus show that two different tyrosyl residues in A* protein can bind covalently to DNA. This conclusion also holds for gene A protein since the same peptide-oligonucleotide complexes are observed when the gene A protein-oligonucleotide complex is digested with trypsin. The two peptide-oligonucleotide complexes are formed in about equal amounts so either one of the two tyrosyl residues can function as the acceptor of the DNA chain during cleavage of the phosphodiester bond in DNA. This indicates that the two tyrosyl-hydroxyl groups are functionally equivalent. Inspection of the amino acid sequence of gene A protein shows that the two tyrosyl residues lie only three amino acid residues apart in the repeating sequence: tyr, val, ala, lys, tyr, val, asn, lys. (The tyrosyl residues are residues 343 and 347 in the complete amino acid sequence of 513 residues of gene A protein which is deduced form the DNA sequence (8)). The tyrosyl side-chains protrude at the same side of the helix if this part of gene A protein occurs in an α-helix conformation. Computer-graphic modelling studies showed that the hydroxyl groups of the two tyrosyl side-chains in such a conformation can approach the phosphorus atom of the phosphodiester bond which is to be cleaved close enough for this reaction and further showed that the two hydroxyl groups occupy a symmetrical position towards the phosphorus atom. In Figure 3, we present a model for the successive gene A protein catalysed cleavage and cleavage-ligation reactions which take place during initiation and termination of the DNA replication. The model shows how the two juxtaposed tyrosyl residues play an equivalent role and how the two tyrosyl-hydroxyl groups cooperate as two active groups in one active

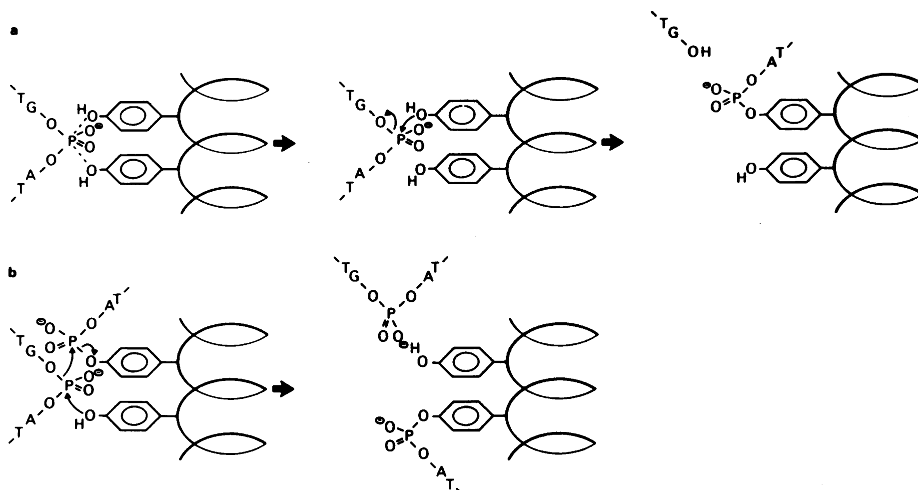

Figure 3. Model for the gene A protein catalysed cleavage and cleavage-ligation reactions which occur during initiation and termination of ØX rolling circle replication respectively. The origin sequence is shown schematically as: -TG-O-P-O-AT-.
O O⁻

The two tyrosyl residues at the active centre of gene A protein are supposed to be part of an α-helix which brings their side-chains in juxtaposition.
a. Cleavage. Binding of gene A protein with DNA places the P-atom at the origin in a symmetrical position towards the two tyrosyl-OH groups. Nucleophilic attack (→) leads to transesterification as shown. The results indicate that either one of the O-atoms of the two tyrosyl residues can function as the acceptor of the DNA chain in this reaction. b. Cleavage-ligation. After one round of replication gene A protein binds the regenerated origin sequence. A nucleophilic attack (→) of the hydroxyl group of the free tyrosyl residue initiates two successive transesterifications as indicated. This results in a phosphodiester bond between DNA and this tyrosyl residue, a free hydroxyl group at the other tyrosyl residue and formation of a phosphodiester bond which ligates the two DNA ends.

centre. The model also shows how the two tyrosyl residues act alternating during the successive cleavage and ligation reactions. The required positioning of the origing DNA sequence is supposed to be realised by an interaction of another part of the gene A protein with the DNA.

The results which are presented here make gene A protein the first enzyme that can cleave and ligate DNA for which the amino acids which are involved have been determined. Other enzymes that play a similar role during DNA replication are the gene II protein of bacteriophage fd (22) and the Rep c protein of plasmid pT181 (23). Rep c protein binds covalently to the 5' end of the DNA at the cleavage site in pT181 DNA. The nature of this bond is not known. Strikingly, two tyrosyl residues occur three amino acid residues apart in fd

gene II protein (24) in the sequence: arg, tyr, phe, gly, phe, tyr, arg... Although no covalent gene II protein-DNA complexes have been demonstrated (22) it is attractive to assume that these two tyrosyl residues are involved in the cleavage and ligation of the fd DNA. The cleavage and ligation reactions (nicking-closing reactions) which are catalysed by the topoisomerases and gyrases proceed via covalent DNA-enzyme complexes (25-28). The bond between protein and DNA in the intermediate is a tyrosyl-phosphate ester (28,29). In these properties the topoisomerases resemble gene A protein. It is not known whether one or two tyrosyl residues in the topoisomerases are involved in the cleavage and ligation reactions. Further research may show whether gene A protein can be regarded as a prototype of the enzymes in which tyrosyl residues are utilized to cleave and ligate DNA during DNA replication and recombination.

ACKNOWLEDGEMENTS

We thank dr. J.M. Vereyken (State University of Groningen, The Netherlands), and dr. J.P.H. Burbach (State University of Utrecht, The Netherlands) for advice, dr. R. van de Broek (Applied Biosystems, United Kingdom) for performing the amino acid sequencing during the FEBS course on "Micro-sequencing of proteins" in Berlin, F.R.G., to which one of us (A.D.M. van Mansfeld) participated, dr. J. de Vlieg (State University of Groningen, The Netherlands) for the computer grafic modelling, prof. J.H. van Boom and G.H. Veeneman (State University of Leiden, The Netherlands) for the synthetic oligonucleotides, and the Netherlands Organization for the Advancement of Pure Research (ZWO) for financial support.

*To whom correspondence should be addressed

REFERENCES

1. Baas, P.D. (1985) Biochim. Biophys. Acta 825, 111-139.
2. Marians, K.J. (1985) C.R.C. Crit. Rev. Biochem. 17, 153-215.
3. Fiddes, J.C., Barrell, B.G. and Godson, G.N. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 1081-1085.
4. Van Mansfeld, A.D.M., Langeveld, S.A., Weisbeek, P.J., Baas, P.D., van Arkel, G.A. and Jansz, H.S. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 331-334.
5. Heidekamp, F., Langeveld, S.A., Baas, P.D. and Jansz, H.S. (1980) Nucl. Acids Res. 8, 2009-2021.
6. Heidekamp, F., Baas, P.D. and Jansz, H.S. (1983) J. Virol. 42, 91-99.
7. Langeveld, S.A., van Mansfeld, A.D.M., Baas, P.D., Jansz, H.S., van Arkel, G.A. and Weisbeek, P.J. (1978) Nature 271, 417-420.
8. Sanger, F., Coulson, A.R., Friedmann, T., Air, G.M., Barrell, B.G., Brown, M.L., Fiddes, J.C., Hutchison, C.A. III, Slocombe, P.M. and Smith, M. (1978) J. Mol. Biol. 125, 225-246.
9. Eisenberg, S., Scott, J.F. and Kornberg, A. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 1594-1597.
10. Ikeda, J.-E., Yudelevich, A. and Hurwitz, J. (1977) Proc. Natl. Acad. Sci. U.S.A. 73, 2669-2673.

11. Eisenberg, S., Griffith, J. and Kornberg, A. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3198-3202.
12. Ikeda, J.-E., Yudelevich, A., Shimamoto, N. and Hurwitz, J. (1979) J. Biol. Chem. 254, 9416-9428.
13. Roth, M.J., Brown, D.R. and Hurwitz, J. (1984) J. Biol. Chem. 259, 10556-10567.
14. Van Mansfeld, A.D.M., van Teeffelen, H.A.A.M., Baas, P.D., Veeneman, G.H., van Boom, J.H. and Jansz, H.S. (1984) FEBS Lett. 173, 351-356.
15. Sanhueza, S. and Eisenberg, S. (1985) J. Virol. 53, 695-697.
16. Van Mansfeld, A.D.M., Langeveld, S.A., Baas, P.D., Jansz, H.S., van der Marel, G.A., Veeneman, G.H. and van Boom, J.H. (1980) Nature 288, 561-566.
17. Linney, E. and Hayashi, M. (1973) Nature New Biol. 245, 6-8.
18. Sanhueza, S. and Eisenberg, S. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 4285-4289.
19. Langeveld, S.A., van Arkel, G.A. and Weisbeek, P.J. (1980) FEBS Lett. 114, 269-272.
20. Eisenberg, S. and Finer, M. (1980) Nucl. Acids. Res. 8, 5305-5315.
21. Van Mansfeld, A.D.M., van Teeffelen, H.A.A.M., Zandberg, J., Baas, P.D., Jansz, H.S., Veeneman, G.H. and van Boom, J.H. (1982) FEBS Lett. 150, 103-108.
22. Meyer, T.F., Geider, K. (1982) Nature 296, 828-832.
23. Koepsel, R.R., Murray, R.W., Rosenblum, W.D. and Khan, S.A. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 6845-6849.
24. Beck, E., Sommer, R., Auerswald, E.A., Kurz, Ch., Zink, B., Osterburg, G., Schaller, H., Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1978) Nucl. Acids Res. 5, 4495-4504.
25. Champoux, J.J. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3800-3804.
26. Depew, R.E., Liu, L.F. and Wang, J.C. (1978) J. Biol. Chem. 253, 511-518.
27. Sugino, A., Peebles, C.L., Kreuger, K.N. and Cozzarelli, N.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4767-4771.
28. Tse, Y.-C., Kirkegaard, K. and Wang, J.C. (1980) J. Biol. Chem 255, 5560-5565.
29. Champoux, J.J. (1981) J. Biol. Chem. 256, 4805-4809.