Proc. Natl. Acad. Sci. USA
Vol. 77, No. 11, pp. 6425–6428, November 1980
Biochemistry

Protein p3 is linked to the DNA of phage ϕ29 through a phosphoester bond between serine and 5′-dAMP

(initiation of replication/*Bacillus subtilis*/nucleotidyl-peptide/acid and alkali hydrolysis)

JOSÉ M. HERMOSO AND MARGARITA SALAS

Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma, Canto Blanco, Madrid-34, Spain

Communicated by Severo Ochoa, July 28, 1980

ABSTRACT **To investigate the role of protein p3 in bacteriophage ϕ29 initiation of replication, we have studied the nature of the covalent linkage between protein p3 and ϕ29 DNA. The protein-DNA compound was digested with micrococcal nuclease and pronase resulting in a nucleotidyl-peptide that was further digested by alkaline phosphatase and snake venom phosphodiesterase yielding 5′-dAMP. The DNA-protein linkage is sensitive to alkali. Treatment of the nucleotidyl-peptide with 0.1 M NaOH at 37°C for 3 hr after phosphatase digestion released 5′-dAMP. Hydrolysis of the nucleotidyl-peptide with 5.8 M HCl at 110°C for 90 min yielded *O*-phosphoserine. These results, together with the sensitivity of the DNA-protein linkage to snake venom phosphodiesterase and its resistance to hydroxylamine, indicate that protein p3 is covalently linked to ϕ29 DNA through a phosphoester bond between L-serine and 5′-dAMP, namely a *O*,5′-deoxyadenylyl-L-serine bond.**

The genome of *Bacillus subtilis* phage ϕ29 consists of a linear, double-stranded DNA of M_r 11.8 × 10^6 (1) with a protein covalently attached to both 5′ termini (2–5). This protein was characterized as the product of cistron 3—p3 (2), an early protein of M_r 27,000 (6, 7) that is required for the initiation of DNA replication (8). A protein covalently linked to the 5′ termini of DNA has also been found in adenovirus (9, 10). Some RNA viruses also contain a protein bound to the 5′ end of their genome; this is the case for picornaviruses like polio (11, 12), encephalomyocarditis (13), and foot-and-mouth disease (14) viruses; vesicular exanthema virus (15); and several plant viruses (16–19).

One of the problems in the replication of linear DNAs is how the 5′ ends are primed (i.e., how the 3′-OH group required by the DNA polymerase is provided). Structural features of the terminal regions (cohesive ends, terminal repetitions, hairpins) may determine the way this problem is solved. The replication of both adenovirus and ϕ29 is initiated at either DNA end and proceeds by a strand displacement mechanism (20, 21). Models for the initiation of phage ϕ29 and adenovirus replication have been proposed in which a newly synthesized terminal protein plays the role of primer by the covalent attachment of the terminal nucleotide of the nascent DNA strand, providing the 3′-OH group required for chain elongation (9, 21, 22). This mechanism, if confirmed, would give a new answer to the problem of replication of the 5′ ends in linear DNAs. We have started the study of this mechanism focusing our attention on the linkage between protein p3 and ϕ29 DNA. In this paper, we show that protein p3 is linked to the 5′ termini of ϕ29 DNA through a phosphoester bond between serine and 5′-dAMP. A preliminary account of some of these findings has been presented (22).

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

MATERIALS AND METHODS

Preparation of ϕ29 p3-[^{32}P]DNA Compound: Enzymatic Digestions. *B. subtilis* 110NA *su*$^-$ was grown and infected with the ϕ29 delayed-lysis mutant *sus*14 (1242) in the presence of carrier-free ^{32}P [1 mCi/ml (1 Ci = 3.7 × 10^{10} becquerels); The Radiochemical Centre, Amersham] as described (23). After 150 min of incubation, protoplasts were prepared (24) and the lysates were treated with DNAse and RNase. After removal of bacterial debris, the phage was purified by centrifugation in a stepwise CsCl density gradient followed by a CsCl gradient to equilibrium (25). Protein p3-[^{32}P]DNA compound, prepared essentially as described (2), was incubated with 75 units/ml of micrococcal nuclease (Worthington) for 1 hr at 37°C in 50 mM Tris·HCl, pH 8.8/10 mM $CaCl_2$. After 1 hr of incubation, less than 0.05% of the initial ^{32}P radioactivity was recovered as acid-insoluble material. The nuclease-digested material was precipitated with trichloroacetic acid [final concentration, 10% (wt/vol)] in the presence of lysozyme as carrier (50 μg/ml; from egg white, grade I, Sigma). After 30 min at 0°C, the sample was centrifuged at 17,000 × *g* for 15 min at 4°C and the pellet was washed four times with an equal volume of cold 10% trichloroacetic acid, once with ethanol/ether, 1:1 (vol/vol), and once with ethanol/ether, 1:3 (vol/vol). The pellet was dried under vacuum, dissolved in 10 mM Tris·HCl, pH 7.8/0.1% sodium dodecyl sulfate, and digested with Pronase (1 mg/ml; B grade, Calbiochem) for 18 hr at 37°C. The resulting nucleotidyl-peptide was purified by adsorption to a DEAE-cellulose column and elution with 1 M ammonium acetate, pH 5.0. The salt was removed by drying under vacuum. If necessary, the nucleotidyl-peptide was digested again with micrococcal nuclease and Pronase and was purified by preparative high-voltage paper electrophoresis as described below.

Chemical and Enzymatic Treatments of the Nucleotidyl-Peptide. Alkaline hydrolysis was performed in 0.1 M NaOH for 3 hr at 37°C. Acid hydrolysis was in 5.8 M HCl in a vacuum-sealed tube at 110°C for the indicated times. Treatment with alkaline phosphatase (2.3 units/ml; calf intestine, grade I, Boehringer Mannheim GmBH) was carried out in 50 mM Tris·HCl, pH 8.8, for 1 hr at 37°C. The enzyme was inactivated by heating 1 hr at 65°C. Digestion with snake venom phosphodiesterase (2.2 units/ml; Worthington) was done in 10 mM Tris·HCl, pH 8.8/5 mM $MgCl_2$ for 1 hr at 37°C.

High-Voltage Paper Electrophoresis and Thin-Layer Chromatography. Electrophoresis was performed with Whatman 3 MM paper in 5% (vol/vol) acetic acid/0.5% pyridine, pH 3.5, at 3000 V in a Michl tank with xylene cyanol FF and acid fuchsin as visual markers.

The identification of nucleotides was also carried out by thin-layer chromatography on polyethyleneimine (PEI)-cellulose plates (6.5 × 6.5 cm; Polygram CEL 300 PEI, Ma-

cherey–Nagel (Düren, Germany)). Development in one direction, first with 0.5% formic acid and then with 0.15 M lithium formate, at pH 3.0 (26), resulted in a partial separation. Complete separation of the ribonucleotides and deoxyribonucleotides was achieved after a third run in the second direction in borate buffer (27).

Labeled compounds were located by autoradiography with Kodak X-Omat film using Ilford x-ray intensifying screens at −70°C. The following compounds were used as markers: *O*-phospho-L-serine, *O*-phospho-DL-threonine, L-serine, 5′-ribonucleotides, 5′-deoxyribonucleotides (all from Sigma), and *O*-phosphotyrosine (a gift from J. C. Wang, Harvard University).

RESULTS

Identification of the Nucleotide Linked to Protein p3. A major spot (Fig. 1, lane A, spot I) was obtained when the product of the protein p3-[^{32}P]DNA digest with micrococcal nuclease and Pronase was subjected to high-voltage paper electrophoresis. After alkaline phosphatase treatment, spot I released P_i and was converted into spot II (Fig. 1, lane B). The amount of ^{32}P radioactivity released as P_i was about the same as that of the phosphatase-resistant material, indicating that protein p3 is linked to the DNA through a single phosphate group. Because micrococcal nuclease digests DNA to form 3′ mono- and dinucleotides (28), spot I was identified as a nucleoside 3′,5′-bisphosphate linked to an amino acid or short peptide by a 5′-phosphate, with a terminal 3′-phosphate. Treatment of spot I with alkaline phosphatase followed by snake venom phosphodiesterase formed P_i and one other product with the same electrophoretic mobility as 5′-dAMP (Fig. 1, lane C).

To check further that 5′-dAMP was obtained from each 5′ end of φ29 DNA, we used restriction endonuclease fragments *Hha* I O and *Hin*dIII L (respectively corresponding to the left and right ends of φ29 DNA), which had been labeled by means of polynucleotide kinase and [γ-^{32}P]ATP at only the 5′ termini of the intact genome after removal of the protein by alkali and phosphatase treatments (unpublished data). After digestion with DNase I and snake venom phosphodiesterase, only one labeled product was obtained, with the same electrophoretic mobility as 5′-dAMP (Fig. 1, lanes D and E), in agreement with the results obtained by direct sequence determination of both 5′ ends of φ29 DNA (unpublished data). To test whether the protein is linked to the DNA through a ribonucleotide, we analyzed a sample corresponding to the treatment shown in Fig. 1, lane C by thin-layer chromatography in a system in which deoxyribo- and ribonucleotides are resolved. The result (Fig. 2) indicates that the protein is linked to a 5′-dAMP residue. Therefore, it can be concluded that the nucleotidyl moiety of spot I is pdAp.

Identification of the Amino Acid Linked to φ29 DNA. To characterize the amino acid involved in the linkage between protein p3 and φ29 DNA, spot I [obtained by digestion of the protein-DNA compound with micrococcal nuclease and Pronase (Fig. 3, lane A)] was hydrolyzed with 5.8 M HCl at 110°C for different times. As shown in Fig. 3, lanes B–D, the 3′-phosphate group was recovered as P_i, and three other spots were obtained after 45, 75, and 90 min of incubation. With longer incubation times there was a shift of radioactivity from the two slower mobility spots (one of them very faint) to the faster one that comigrates with the *O*-phospho-L-serine marker. The marker was only partially hydrolyzed to serine after incubation for 90 min under the conditions described above (Fig. 3, lane F). The two slower mobility spots probably represent partial hydrolysis products with different amino acid content (29), suggesting that spot I contains, at least, a tripeptide. Acid hydrolysis of spot I followed by alkaline phosphatase digestion

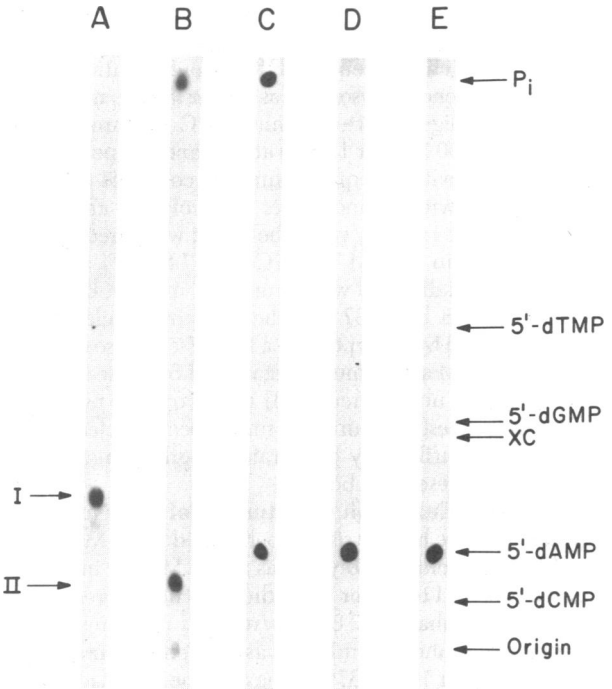

FIG. 1. High-voltage paper electrophoresis at pH 3.5 of enzymatic digestions of φ29 p3-DNA. Lanes: A, complex digested with micrococcal nuclease and Pronase purified by ion-exchange chromatography in DEAE-cellulose (spot I); B, sample A treated with alkaline phosphatase (spot II); C, sample B digested with snake venom phosphodiesterase; D and E, restriction endonuclease terminal fragments *Hha* I O and *Hin*dIII L, respectively, labeled with ^{32}P at the 5′ end of the intact genome and treated with DNase I and snake venom phosphodiesterase. 5′-Deoxyribonucleotides were used as internal standards and their positions were determined by UV absorbance. The locations of $^{32}P_i$ and of the visual marker xylene cyanol (XC) are also indicated.

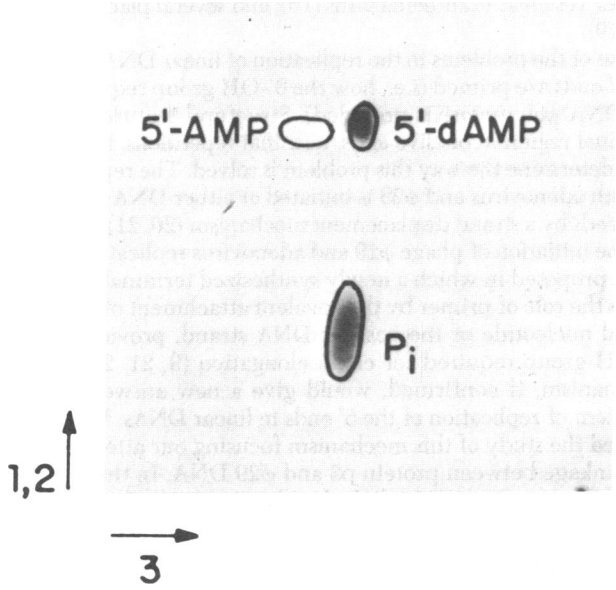

FIG. 2. Thin-layer chromatography in PEI-cellulose plates of sample C from Fig. 1. The chromatography was developed in the directions indicated. 5′-Ribonucleotides and 5′-deoxyribonucleotides were used as internal standards and their positions were determined by UV absorbance. $^{32}P_i$ was run in parallel.

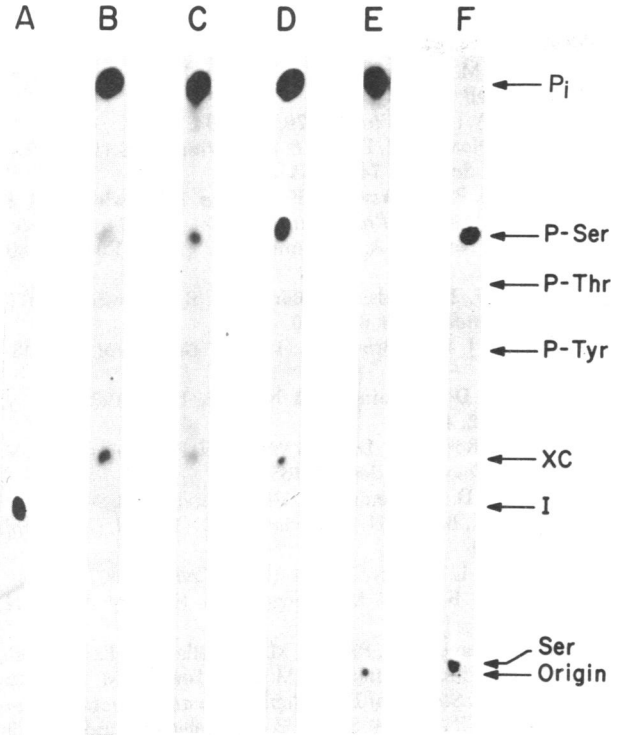

FIG. 3. High-voltage paper electrophoresis at pH 3.5 of the products of acid hydrolysis of ϕ29 p3-[^{32}P]DNA digested with micrococcal nuclease and Pronase. After purification by chromatography in DEAE-cellulose and a second treatment with the above enzymes, a product was isolated by preparative high-voltage paper electrophoresis. A sample of the product, spot I, was run as a control (A) or hydrolyzed in 5.8 M HCl at 110°C for 45 min (B), 75 min (C), and 90 min (D). Sample D was further digested with alkaline phosphatase (E). Nonradioactive standard *O*-phospho-L-serine was hydrolyzed as in D (F). Phosphoamino acids were run as internal standards, and their positions were located by ninhydrin staining. $^{32}P_i$ was run in parallel and xylene cyanol (XC) was used as the visual marker.

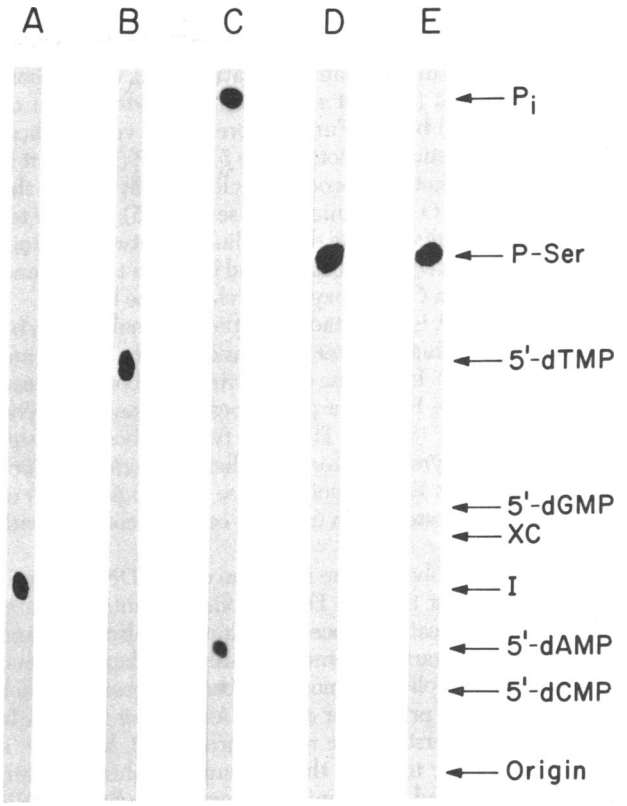

FIG. 4. High-voltage paper electrophoresis at pH 3.5 of the products of alkali hydrolysis of ϕ29 p3-[^{32}P]DNA digested with micrococcal nuclease and Pronase. After purification by chromatography in DEAE-cellulose and a second treatment with the above enzymes, a product was isolated by preparative high-voltage paper electrophoresis. A sample of the resulting product, spot I, was run as a control (A) or hydrolyzed for 3 hr at 37°C with 0.1 M NaOH before (B) or after (C) digestion with alkaline phosphatase. Nonradioactive *O*-phospho-L-serine was run as a control before (D) or after (E) treatment with 0.1 M NaOH as indicated above. 5'-Deoxyribonucleotides were run as internal standards and their positions were determined by UV absorbance. Phosphoserine was identified by ninhydrin staining. $^{32}P_i$ was run in parallel. Xylene cyanol (XC) was used as a visual marker.

yielded only $^{32}P_i$ (Fig. 3, lane E), as expected from a phosphoamino acid or a phosphopeptide. Because other phosphoamino acid candidates such as phosphothreonine and phosphotyrosine are well resolved from phosphoserine in this electrophoretic system (Fig. 3), the above results suggest that serine is the amino acid involved in the bond between protein p3 and ϕ29 DNA.

Alkali Hydrolysis of the Nucleotidyl-Peptide. Protein p3 can be removed from the DNA compound by alkali treatment (2), leaving a phosphate group at the 5' end of the DNA (unpublished data). This result was confirmed when the nucleotidyl-peptide resulting from the digestion of p3-DNA with micrococcal nuclease and Pronase (spot I, Fig. 4, lane A) was incubated in 0.1 M NaOH at 37°C for 3 hr, and the hydrolysis products were analyzed by high-voltage paper electrophoresis. A single radioactive spot was obtained (Fig. 4, lane B) that probably represents deoxyadenosine 3',5'-bisphosphate (the electrophoretic mobility of 5'-dADP under these conditions is slightly slower than that of 5'-dTMP). Digestion of spot I with phosphatase prior to the alkali treatment yielded P_i and a spot with the same electrophoretic mobility as 5'-dAMP (Fig. 4, lane C). This result indicates that the protein is directly linked to the DNA and that this bond is sensitive to alkali. *O*-Phosphoserine, however, was completely resistant to the above conditions of alkali hydrolysis (Fig. 4, lanes D and E). A possible explanation for these apparently contradictory results is discussed below.

DISCUSSION

The results presented above show the nature of the linkage between protein p3 and ϕ29 DNA. A serine residue is linked directly through a phosphate bond to a deoxyadenylyl residue shown to be the 5'-nucleotide at both ends of ϕ29 DNA. There are three possibilities for a covalent linkage of serine to a phosphate group. The most likely one would be a phosphoester, but a phosphoanhydride in the case of a COOH-terminal serine or a phosphoamide if the serine were NH_2-terminal are also possible.

Protein p3-DNA digested with micrococcal nuclease and pronase shows a resistance to acid similar to that of *O*-phosphoserine (Fig. 3, lanes D and F) and phosphoproteins, in which the phosphate group is covalently linked to serine (29, 30). Mild alkali treatment (37°C) hydrolyzes the digested protein p3-DNA (Fig. 4, lanes B and C), probably through a β-elimination mechanism (31), as expected in an *O*-phosphoserine linkage in which the amino group is blocked by a peptide bond (32). In contrast, under the same alkali conditions, *O*-phosphoserine, bearing a free amino group, is completely resistant as shown in Fig. 4, lane E (see also ref. 33). In any case, it is difficult to determine unequivocally the nature of the nucleotidyl-amino acid linkage solely from the analysis of the acid and alkali hy-

drolysis products, due to the effects of adjacent functional groups (34) or $O \rightarrow N$ linkage migration (32). Phage ϕ29 p3-DNA is resistant to hydroxylamine treatment (2), as are phosphodiester linkages (35) but not phosphoanhydride (31) or phosphoamide (32) bonds. Furthermore, snake venom phosphodiesterase digestion of spot II gave 5′-dAMP (Fig. 2); it is this enzyme, and not micrococcal nuclease, that digests the model compound *O*,5′-thymidylyl-L-serine (35). Taken together, all these results indicate that the linkage between protein p3 and ϕ29 DNA is a phosphoester bond between L-serine and 5′-dAMP, namely a *O*,5′-deoxyadenylyl-L-serine bond.

Adenovirus DNA is also attached to the terminal protein by a phosphoester linkage between L-serine and the 5′ end nucleotide, dCMP (36). In the case of poliovirus, the bond between the protein and the RNA is a phosphoester linkage between tyrosine and 5′UMP (37, 38). The same type of phosphodiester linkage involving tyrosine is formed when *Escherichia coli* or *Micrococcus luteus* DNA topoisomerase I or the A subunit of *M. luteus* DNA gyrase form a transient covalent complex with the DNA (39).

Protein p3 is involved in the initiation of ϕ29 DNA replication (8). Analysis of the ϕ29 DNA replicative intermediates indicates that replication proceeds by a mechanism of strand displacement and starts nonsimultaneously at either DNA end (21). A ϕ29 DNA replication model has been proposed in which protein p3 acts as a primer for ϕ29 DNA synthesis (21, 22). In order to fully understand the role of protein p3, however, it would be necessary to know the mechanism behind the formation of the p3 and the terminal 5′-nucleotide linkage. We suggest that protein p3 reacts with 5′-dATP, resulting in the adenylylation of a serine residue that provides the 3′-OH group needed for DNA elongation. For catalysis of the above reaction, several possibilities exist: catalysis by (*i*) protein p3 itself (*ii*) a ϕ29-induced DNA polymerase of unusual properties, or (*iii*) another phage-induced or bacterial protein.

We are grateful to C. Escarmís for providing fragments *Hha* I O and *Hin*dIII L labeled with ^{32}P. This investigation has been aided by Research Grant 1 RO1 GM27242-01 from the National Institutes of Health and by grants from Comisión Asesora para el Desarrollo de la Investigación Científica y Técnica and Comisión Administradora del Descuento Complementario (INP).

1. Sogo, J. M., Inciarte, M. R., Corral, J., Viñuela, E. & Salas, M. (1979) *J. Mol. Biol.* **127**, 411–436.
2. Salas, M., Mellado, R. P., Viñuela, E. & Sogo, J. M. (1978) *J. Mol. Biol.* **119**, 269–291.
3. Harding, N., Ito, J. & David, G. S. (1978) *Virology* **84**, 279–292.
4. Yehle, C. O. (1978) *J. Virol.* **27**, 776–783.
5. Ito, J. (1978) *J. Virol.* **28**, 895–904.
6. Carrascosa, J. L., Viñuela, E. & Salas, M. (1973) *Virology* **56**, 291–299.
7. Hawley, L. A., Reilly, B. E., Hagen, E. W. & Anderson, D. L. (1973) *J. Virol.* **12**, 1149–1159.
8. Mellado, R. P., Peñalva, M. A., Inciarte, M. R. & Salas, M. (1980) *Virology*, **104**, 84–96.
9. Rekosh, D. M. K., Rusell, W. C., Bellet, A. J. D. & Robinson, A. J. (1971) *Cell* **11**, 283–295.
10. Carusi, E. A. (1977) *Virology* **76**, 380–394.
11. Lee, F. L., Nomoto, A., Detjen, B. M. & Wimmer, E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 59–63.
12. Flanegan, J. B., Pettersson, R. F., Ambros, V., Hewlett, J. M. & Baltimore, D. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 961–965.
13. Golini, F., Nomoto, A. & Wimmer, E. (1978) *Virology* **89**, 112–118.
14. Sangar, D. V., Rowlands, D. J., Harris, T. J. R. & Brown, F. (1977) *Nature (London)* **268**, 648–650.
15. Burroughs, J. N. & Brown, F. (1978) *J. Gen. Virol.* **41**, 443–446.
16. Daubert, S. D., Bruening, G. & Najarian, R. C. (1978) *Eur. J. Biochem.* **92**, 45–51.
17. Stanley, J., Rottier, P., Davis, J. W., Zabel, P. & Kammer, A. V. (1978) *Nucleic Acids Res.* **5**, 4505–4522.
18. Daubert, S. D. & Bruening, G. (1979) *Virology* **98**, 246–250.
19. Mayo, M. A., Barker, H. & Harrison, B. D. (1979) *J. Gen. Virol.* **43**, 735–740.
20. Lechner, R. L. & Kelly, T. J., Jr. (1977) *Cell* **12**, 1007–1020.
21. Inciarte, M. R., Salas, M. & Sogo, J. M. (1980) *J. Virol.* **34**, 187–199.
22. Salas, M., García, J. A., Peñalva, M. A., Mellado, R. P., Hermoso, J. M., Escarmís, C., Inciarte, M. R. & Sogo, J. M. (1980) in *Mechanistic Studies of DNA Replication and Genetic Recombination*, ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Alberts, B. & Fox, C. F. (Academic, New York), Vol. 19, in press.
23. Inciarte, M. R., Lázaro, J. M., Salas, M. & Viñuela, E. (1976) *Virology* **74**, 314–323.
24. Sargent, M. G. (1973) *J. Bacteriol.* **116**, 397–409.
25. Méndez, E., Ramírez, G., Salas, M. & Viñuela, E. (1971) *Virology* **45**, 567–576.
26. Volkaert, G. & Fiers, W. (1977) *Anal. Biochem.* **83**, 222–227.
27. Randerath, K. & Randerath, E. (1965) *Anal. Biochem.* **13**, 575–579.
28. Reddi, K. K. (1967) *Methods Enzymol.* **12**, 257–262.
29. Krystal, G., Winn, P., Millward, S. & Sakuma, S. (1975) *Virology* **64**, 505–512.
30. Kabat, D. (1971) *Biochemistry* **10**, 197–203.
31. Taborsky, G. (1974) in *Advances in Protein Chemistry*, eds. Anfinsen, C. B., Edsall, T. J. & Richards, F. M. (Academic, New York), Vol. 28, pp. 1–210.
32. Shabarova, Z. A. (1970) in *Progress in Nucleic Acids Research and Molecular Biology*, eds. Davidson, J. N. & Cohn, W. E. (Academic, New York), Vol. 10, pp. 145–182.
33. Plimmer, R. H. A. (1941) *Biochem. J.* **35**, 461–469.
34. Kemp, B. E. (1980) *FEBS Lett.* **110**, 308–312.
35. Berkowitz, W. F. & Bendich, A. (1965) *Biochemistry* **4**, 1979–1983.
36. Desiderio, S. V. & Kelly, T. J., Jr. (1980) *J. Mol. Biol.*, in press.
37. Ambros, V. & Baltimore, D. (1978) *J. Biol. Chem.* **253**, 5263–5266.
38. Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4868–4872.
39. Wang, J. C. (1980) *J. Supramol. Struct.* Suppl. **4**, 305.