Volume 10 Number 19 1982 Nucleic Acids Research

# High level synthesis in *Escherichia coli* of the *Bacillus subtilis* phage $\phi$29 proteins p3 and p4 under the control of phage lambda $P_L$ promoter

Rafael P.Mellado and Margarita Salas

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

Received 12 August 1982; Accepted 17 September 1982

ABSTRACT

The Hind III G fragment from the Bacillus subtilis phage Ø29 DNA, inserted downstream from the bacteriophage λ promoter $P_L$ carried by a pBR322 derivative plasmid (pPLc28), directed the synthesis in E. coli of two proteins of apparent molecular weight 27 500 and 12 500. With the use of the recombinants obtained with the DNA from mutants sus3(91) and sus4(56), the two proteins were identified as a modified p3 (p3'), the protein covalently linked to the 5' ends of Ø29 DNA, and p4, responsible for the Ø29 late transcription, respectively. Under the best conditions used, proteins p4 and p3' were produced in E. coli from the cloned DNA fragments in an amount corresponding to approximately 30% and 6% of total de novo protein synthesis, respectively.

## INTRODUCTION

The Bacillus subtilis phage Ø29 contains a linear, double-stranded DNA of molecular weight 11.8 x $10^6$ daltons (1) with a protein covalently bound to the two 5' ends (2-5). This protein, p3, with a molecular weight of about 27 000 daltons (6,7) is the product of gene 3 (2), required for the initiation of the viral DNA replication (8). A new mechanism for the initiation of replication at the ends of Ø29 DNA has been proposed, whereby the terminal protein p3 acts as a primer by reaction with dATP, the 5' terminal nucleotide at both DNA ends, and formation of a covalent complex p3-dAMP that would provide the 3'OH group needed for polymerization by the DNA polymerase (8-11). The formation in vitro of a covalent complex between protein p3 and dAMP, supports the above mechanism (12).

It has been recently assumed that protein p3 is coded by the region of Ø29 DNA comprised between nucleotides 3686 and 2886 from the left end of the viral genome (13). This paper and the


0305-1048/82/1019-5773$ 2.00/0

accompanying one (14) demonstrate that this assumption is correct and they also show that between nucleotides 3952 and 3575 there is an overlapping reading frame which codes for the product of gene 4, p4, which controls the ϕ29 late transcription (1).

The fragment Hind III G of ϕ29 DNA which extends from nucleotide 2900 to 4049 (13,14) has been placed under the control of the phage λ thermoinducible leftward promoter $P_L$, harboured in the plasmid pPLc28 (15), so that the proteins coded by the cloned DNA could be overproduced in Escherichia coli. Comparison of the proteins synthesized upon induction of recombinant plasmids containing the wild-type DNA with those synthesized by recombinant plasmids carrying the equivalent DNA sequences from the ϕ29 nonsense mutants sus3(91) and sus4(56) showed that fragment Hind III G contains the complete coding sequence for p4 and that of p3 except the last five amino acids.

## MATERIALS AND METHODS

### Bacterial strains and plasmids

The E. coli strain HB101 (16) was obtained from N.E. Murray. E. coli K-12ΔH1Δtrp (17) has been described before (15). Plasmids pBR322 (18) and pPLc28 (15) were used as cloning vectors. The strain K-12ΔH1Δtrp and the plasmid pPLc28 were obtained from M. Zabeau.

### Enzymes

T4 DNA ligase and restriction enzymes (used according to the supplier) were purchased from New England Biolabs Inc. Fungal proteinase K was from Merck.

### DNA preparations

The ϕ29 lysis negative mutant sus14(1242) (19) was used as a source of wild-type phage DNA which was prepared after treatment of the phage with proteinase K as described by Inciarte et al. (20). DNA was also isolated from mutants sus3(91) and sus4(56) (21). Plasmid DNA was prepared according to Clewell (22). The fragment Hind III G was obtained from ϕ29 DNA by digestion with Hind III followed by separation in a 3.5% polyacrylamide gel (23). Recombinant DNA molecules were made by incubation of the DNA fragments with T4 DNA ligase under appropriate conditions (24) and recovered by transformation (25) as modified by Remaut et al.

(15). Bacterial colonies containing the desired recombinant plasmid were selected on the basis of their sensitivity to tetracycline when cloning in pBR322 or by appropriate restriction of plasmid DNA minipreparations (26) followed by analysis in 1.5% agarose gels (27) when cloning in pPLc28. They were further characterized in both cases by restriction analysis.

Peptide synthesis in transformed cells

The E. coli strain K-12ΔH1Δtrp harbouring the recombinant plasmids obtained using pPLc28 as a vector was grown at 30 ºC in L-broth without antibiotics to a density of 2 x $10^8$ cells/ml. The cells were collected and resuspended in their original volume of minimal medium (28). Incubation was continued at 30 ºC for 1 h before inducing one half of the culture by shifting at 42 ºC. At the indicated times 50 µl samples were labelled with 10 µCi of $^{35}$S-methionine (1200 Ci/mmol; The Radiochemical Centre, Amersham) for 10 min. Cells were disrupted by boiling for 5 min in a buffer containing 2% SDS, 5% β-mercaptoethanol and 3 M urea and analysed directly on a 20% polyacrylamide gel containing SDS (29). The gels were impregnated for fluorography (30), dried and autoradiographed with Fuji RX film exposed at -70 ºC. The relative production of the proteins was determined by densitometry of the autoradiographs obtained by exposure under linear conditions and expressed as percentage of the total labelled protein present on the gel track.

All the experiments involving recombinants were carried out following the NIH Guidelines.

## RESULTS AND DISCUSSION

### Isolation and characterization of recombinants

The fragment Hind III G of Ø29 DNA spans 1150 bp between nucleotides 2900 and 4049 from the left end of the viral genome (13) and is located in an early region transcribed from right to left (1). It begins 363 bp upstream the first triplet of the sequence assumed to code for protein p3 and ends 15 bp before the last triplet of the same sequence (13,14). Partially overlapping with this one, there is another open reading frame which starts 97 pb after the Hind III site at position 4049 spanning 378 bp with a coding capacity for a polypeptide with 125 amino acids.

which could be the product of the viral gen 4, involved in the control of late transcription (1), located in the genetic map next to gene 3, to the right (31).

The sequence believed to code for p3 contained in the Hind III G fragment lacks therefore the information for the last 5 amino acids at the carboxyl end of the molecule. To cover this deficiency, the Hind III G fragment from the Ø29 wild type DNA was cloned in a first step in the Hind III target of pBR322 (18). Thus, the recombinant carrying the insert with the viral sequence oriented clockwise from the EcoRI site of pBR322 (18) will provide enough information to add nine new amino acids to the prematurely ended protein p3 before finding a nonsense triplet in the pBR322 sequences, so that a polypeptide of 270 amino acids (protein p3') with a molecular weight similar to that of p3 could be synthesized. The recombinant plasmid produced in this way, pRMw5, was recovered from a population of transformed cells and the correct orientation of the insert confirmed by digestion with the endonuclease Hinf I. The recombinant plasmids pRMs36 and pRMs46 were obtained when the Hind III G fragment from the DNA of the Ø29 nonsense mutants sus3(91) and sus4(56), respectively, was cloned in pBR322. The construction of such recombinants is depicted in Figure 1.

To assess whether the region of Ø29 DNA carried in the Hind III G fragment codes for proteins p3 and p4, the Ø29 DNA sequences contained in the recombinants described above were placed under the control of the strong $\lambda P_L$ promoter so that they could be overexpressed in E. coli. Since the production of high levels of the proteins could be detrimental to the host cell, a thermoinducible system as that of plasmid pPLc28 as a vector and E. coli K-12ΔH1Δtrp as a host (15), was chosen.

The plasmid pPLc28 is a pBR322 derivative which carries the leftward promoter of phage $\lambda(P_L)$ inserted in such a way that $P_L$ transcription takes place in a clockwise direction (15). Downstream from the $P_L$ promoter the plasmid carries a linker fragment with single targets for the endonucleases Eco RI, Bam HI and Hind III in this order, so that any DNA fragment inserted in any of these targets will be efficiently expressed in E.coli as long as it contains its own ribosome binding sequences and

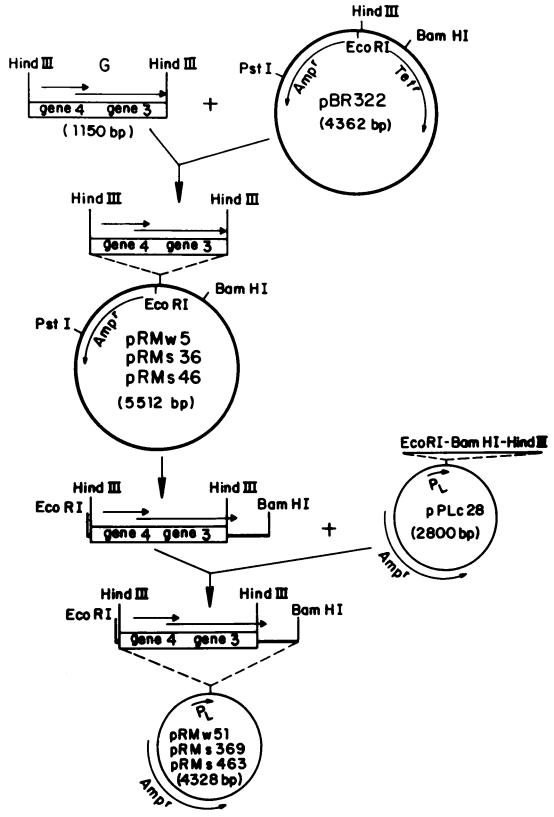


Fig. 1. Recombinant plasmids containing the Ø29 DNA fragment Hind III G. The arrows above enclosed genes 3 and 4 indicate their extension and direction of transcription in Ø29 DNA with respect to the $P_L$ promoter in pRMw51, pRMs369 and pRMs463. The direction of transcription of the ampicillin and tetracycline resistance genes is also indicated.

ATG initiator. The promoter activity of $P_L$ is repressed at low temperature by the thermosensitive repressor product of the cI857 mutation carried by the lysogen host, the normal $P_L$ activity being fully restored by heat induction (15).

The Ø29 sequences cloned in the plasmids pRMw5, pRMs36 and pRMs46 were transferred to pPLc28 by a triple digestion with the endonucleases Eco RI, Bam HI and Pst I and the resulting mixture ligated to pPLc28 linearized with Eco RI and Bam HI together. In this way the only Eco RI-Bam HI piece of DNA from the previous-

ly constructed recombinant plasmids includes the cloned viral sequences. This Eco RI-Bam HI fragment can be inserted in the Eco RI-Bam HI linearized pPLc28 molecule in the only possible orientation which places the sequences coding for the 270 and 125 amino acids proteins under the direct control of the $P_L$ promoter. The new recombinant plasmids, pRMw51, pRMs36 and pRMs463, obtained in this way are shown in Figure 1.

Inducible synthesis of the proteins coded in the cloned DNA

Bacteria harbouring the vector pPLc28 or the recombinant plasmids pRMw51, pRMs369 or pRMs463 were grown as described in Materials and Methods. After 1 h of incubation at 30 ºC one half of the cultures was shifted to 42 ºC and at various times after induction samples from the 42 ºC cultures were labelled with $^{35}$S-methionine for 10 minutes. As a control, samples from the cultures kept at 30 ºC were equally labelled at zero time and at the latest time of induction. Expression of the newly constructed recombinants would be expected to produce the polypeptides of 270 amino acids (protein p3') and 125 amino acids plus a putative third one of 152 amino acids corresponding to the fusion of the pBR322 sequences contained in the cloned Eco RI-Bam HI fragment with those of the plasmid pPLc28 (15,18). Shortly after induction the rate of synthesis of three polypeptides with apparent molecular weights of 27 500 (p3'), 15 000 (p15K), and 12 500 (p4) daltons increased markedly in the case of plasmid pRMw51 (Fig. 2 A and B, slots g-i). The three polypeptides were missing when the vector plasmid did not carry any insert (Fig. 2A and B, slots b-d). The protein of 12 500 daltons (p4) was not synthesized when the recombinant carried the sus4(56) sequences (plasmid pRMs463), being the other two polypeptides synthesized at the same rate (Fig. 2B, slots l-n). When the sus3(91) sequences were present in the recombinant plasmid (pRMs369) the synthesis of the 12 500 daltons protein took place at the same extent, but the 27 500 daltons polypeptide (p3') was lacking and, in addition the 15 000 daltons protein (p15K) was also not made (Fig. 2A, slots l-n), probably due to a polar effect of the sus 3 mutation. Neither in the case of plasmid pRMs369 nor pRMs463 the predicted sus fragments of proteins p3 and p4 (14) were de-

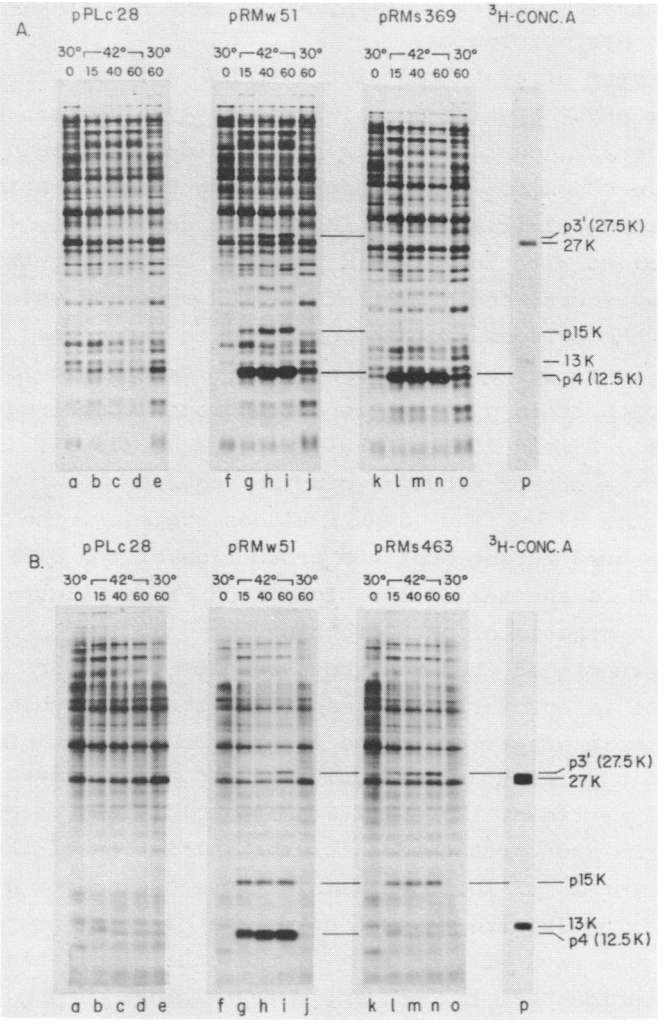


Fig. 2. Induced synthesis of proteins p3' and p4 directed by the recombinant plasmids. Bacterial cultures harbouring the vector and the recombinant plasmids were grown, labelled and subjected to SDS-gel electrophoresis as described in Materials and Methods. Panels A and B, slots a-e, cells carrying pPLc28; panels A and B, slots f-j, cells carrying pRMw51; panel A, slots k-o, cells carrying pRMs369; panel B, slots k-o, cells carrying pRMs463; (p) $^3$H-concanavalin A peptides of known molecular weights (32) as markers (Radiochemical Centre, Amersham). The starting time of each labelling period (in min) and the temperature of incubation are indicated above the slots. The molecular weight of the markers and the induced proteins, in K daltons, is given at the right.

tected in shorter electrophoretic runs, probably due to degradation by the host proteases.

The absence of synthesis of the 12 500 daltons protein with the plasmid pRMs436 which contains the sus4(56) mutant DNA sequences, clearly identifies this polypeptide as protein p4,which is therefore coded for by the open reading frame existing between nucleotides 3952 and 3574 in the Ø29 genome (13,14). The result obtained with the plasmid pRMs369, containing the sus3 (91) DNA sequences, confirms that the DNA sequence which codes for the 27 500 daltons protein contained in the cloned Hind III G fragment is a part of the Ø29 sequence responsible for the synthesis of protein p3, that is the sequence comprised between nucleotides 3686 and 2886 of the viral genome (13,14). Since there is not a proper ribosome binding sequence between those coding for the 27 500 and 15 000 daltons proteins, the polar effect of the sus3 mutant over the production of the p15K protein could be due to the use of the protein p3 Shine-Dalgarno sequence for the synthesis of both proteins.

It can be concluded from the above results that genes 3 and 4 are arranged in an overlapping manner in the Ø29 genome, being another example of saving in the use of the genetic information, since in this particular case only 1067 bp are utilized to code for two polypeptides, 125 and 266 amino acids long, respectively, being the sequences needed for the initiation of the translation of protein p3 included within those coding for protein p4 (13,14). The failure to synthesize one of these proteins does not affect the rate of synthesis of the other indicating that both polypeptides are independently translated.

Level of synthesis and stability of the induced proteins

To determine the level of synthesis of the proteins coded for by the plasmid pRMw51 upon induction, the cultures were incubated at 42 ºC in defined medium and aliquots were sampled out and labelled with $^{35}$S-methionine at several times after the shift-up. As can be seen in Figure 3, even at the latest time of induction (240 minutes) the synthesis of the induced polypeptides was still taking place. Densitometry of the labelled proteins at different times of induction permited to estimate the highest level of synthesis of the 12 500 daltons protein (p4) as about

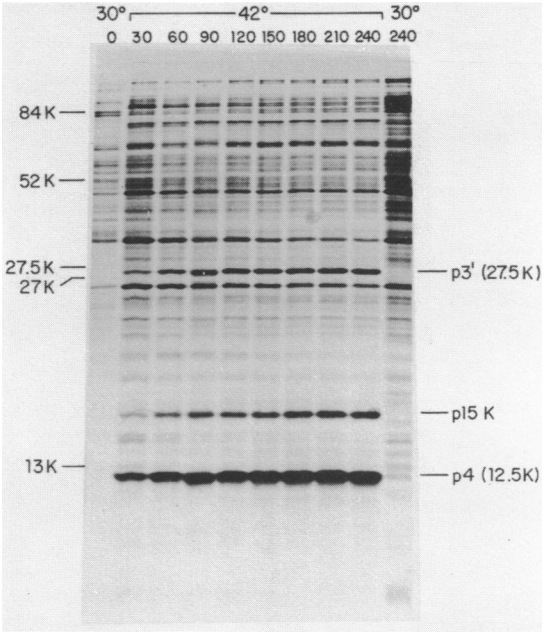


Fig. 3. Kinetic of synthesis of the induced polypeptides directed by the wild-type recombinant plasmid. Cultures of cells containing the plasmid pRMw51 were shifted at 42 ºC and labelled for 10 min with $^{35}$S-methionine at the indicated times as outlined in the legend to Fig. 2. The numbers at the left of the figure indicate the position of markers of known molecular weight, in K daltons.

30% of the de novo synthesized protein under the conditions used. This level was reached after 90-120 minutes of induction and kept without modification at successive induction times. Protein p3' behaved in a similar way reaching the maximum of synthesis (about 6% of the de novo synthesized protein) at 90 minutes of induction without changes at later induction times.

To determine the stability of the induced polypeptides a pulse-chase experiment was performed. Cultures of cells harbouring the plasmid pRMw51 were shifted to 42 ºC, induced for 1 hour, incubated 10 additional minutes in the presence of $^{35}$S-methionine and the labelled polypeptides chased by addition of a 1000-fold excess of cold methionine. Whereas the proteins p3' and p4 were stable even after a chase of 210 minutes (Figure 4),

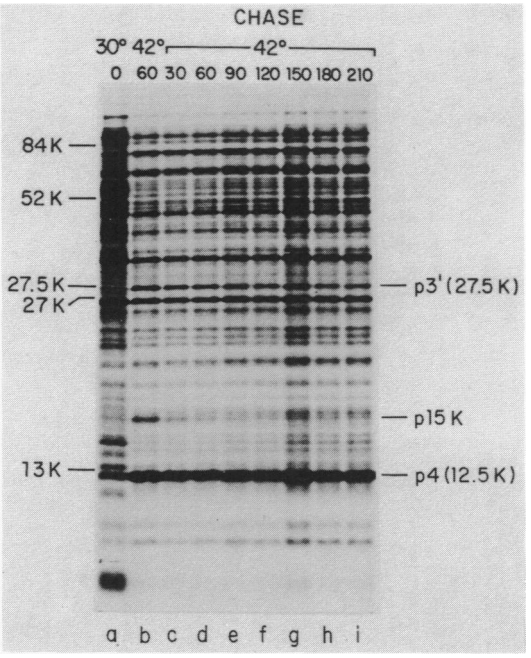


Fig. 4. Stability of the induced polypeptides directed by the wild-type recombinant plasmid. E. coli cultures carrying the plasmid pRMw51 were shifted at 42 ºC and labelled with $^{35}$S-methionine in a 10 min pulse after an induction period of 60 min. (b). The labelled proteins were chased by addition of a 1000 fold excess of cold methionine and samples removed at the times indicated above the slots (c-i). A culture incubated at 30 ºC was labelled at 0 time as a control (a). The numbers at the left of the figure indicate the position of markers of known molecular weight, in K daltons.

the p15K protein was not longer detectable after the first 30 minutes period of chase.

The stability of the induced polypeptides together with the possibility of reaching high levels of synthesis upon prolongated induction periods, makes the recombinant plasmid pRMw51 a good vehicle for the purification of protein p4 to study the mechanism for the control of Ø29 late transcription and for that of protein p3', a p3 protein with the carboxyl end modified. The functionality of p3' is presently being tested in an in vitro system for the initiation of replication (12). On the other hand, the complete sequence coding for protein p3 has been cloned using

a different approach which led to the synthesis in E. coli of a functional protein p3 (33).

ACKNOWLEDGEMENTS

We wish to thank Drs E. Remaut and W. Fiers for their permission to use the plasmid pPLc28. This investigation has been aided by research grant 1 R01 GM27242-03 from the National Institutes of Health and by grants from the Comisión Asesora para el Desarrollo de la Investigación Científica y Técnica and Fondo de Investigaciones Sanitarias. R.P.M. is recipient of a senior fellowship from the "Ministerio de Educación y Ciencia".

REFERENCES

1 Sogo, J.M., Inciarte, M.R., Corral, J., Viñuela, E. and Salas, M. (1979) J. Mol. Biol. 127, 411-436.
2 Salas, M., Mellado, R.P., Viñuela, E. and Sogo, J.M. (1978) J. Mol. Biol. 119, 269-291.
3 Harding, N., Ito, J. and David, G.S. (1978) Virology 84, 279-292.
4 Yehle, D.C. (1978) J. Virol. 27, 776-783.
5 Ito, J. (1978) J. Virol. 28, 895-904.
6 Carrascosa, J.L., Viñuela, E. and Salas, M. (1973) Virology 56, 291-299.
7 Hawley, L.A., Reilly, B.E., Hagen, E.W. and Anderson, D.L. (1973) J. Virol. 12, 1149-1159.
8 Mellado, R.P., Peñalva, M.A., Inciarte, M.R. and Salas, M. (1980) Virology 104, 84-96.
9 Inciarte, M.R., Salas, M. and Sogo, J.M. (1980) J. Virol. 34, 187-199.
10 Harding, N.E. and Ito, J. (1980) Virology 104, 323-338.
11 Sogo, J.M., García, J.A., Peñalva, M.A. and Salas, M. (1982) Virology 116, 1-18.
12 Peñalva, M.A. and Salas, M. (1982) Proc. Natl. Acad. Sci.USA, in press.
13 Yoshikawa, H. and Ito, J. (1982) Gene 17, 323-335.
14 Escarmís, C. and Salas, M., accompanying paper.
15 Remaut, E., Stanssens, P. and Fiers, W. (1981) Gene 15, 81-93.
16 Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
17 Bernard, H.U., Remaut, E., Hershfield, M.V., Das, H.K., Helinski, D.R., Yanofsky, C. and Franklin, N. (1979) Gene 5, 59-76.
18 Bolivar, F., Rodríguez, R.L., Greene, P.Y., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, Y.H. and Falkow, S. (1977) Gene 2, 95-113.
19 Jiménez, F., Camacho, A., de la Torre, J., Viñuela, E. and Salas, M. (1977) Eur. J. Biochem. 73, 57-72.
20 Inciarte, M.R., Lázaro, J.M., Salas, M. and Viñuela, E.(1976) Virology 74, 314-323.

21 Moreno, F., Camacho, A., Viñuela, E. and Salas, M. (1974) Virology 62, 1-16.
22 Clewell, D.D. (1972) J. Bacteriol. 110, 667-676.
23 Maxam, A. and Gilbert, W. (1980) Methods in Enzymol. 65,499-560.
24 Murray, N.E., Bruce, S.A. and Murray, K. (1979) J. Mol. Biol. 132, 493-505.
25 Lederberg, E.M. and Cohen, S.M. (1974) J. Bacteriol. 119, 1072-1074.
26 Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.
27 Sharp, P.A., Sudgen, B. and Sambrook, J. (1973) Biochemistry 12, 3055-3063.
28 Derom, C., Gheysen, D. and Fiers, W. (1982) Gene 17, 45-54.
29 Laemmli, U.K. (1970) Nature 227, 680-685.
30 Bonner, W.M. and Lasky, R.A. (1974) Eur. J. Biochem. 46, 83-88.
31 Mellado, R.P., Moreno, F., Salas, M., Reilly, B.E. and Anderson, D.L. (1976) J. Virol. 19, 495-500.
32 Wang, J.L., Cunningham, B.A. and Edelman, G.M.(1971) Proc. Natl. Acad. Sci. USA 68, 1130-1134.
33 García, J.A., Pastrana, R., Prieto, I. and Salas, M., submitted for publication.