JOURNAL OF BACTERIOLOGY, June 1989, p. 3579–3582
0021-9193/89/063579-04$02.00/0

Vol. 171, No. 6

# Overproduced Bacteriophage T4 Gene 33 Protein Binds RNA Polymerase

KELLY P. WILLIAMS,[1] REINHARD MÜLLER,[2] WOLFGANG RÜGER,[2] AND E. PETER GEIDUSCHEK[1]*

*Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093,[1] and Arbeitsgruppe Molekulare Genetik, Lehrstuhl Biologie der Mikroorganismen, Ruhr-Universität Bochum, D-4630 Bochum 1, Federal Republic of Germany[2]*

Received 15 December 1988/Accepted 4 March 1989

**Bacteriophage T4 gene 33 protein (gp33), which is required for viral late transcription, has been overproduced. The purified gp33 binds to RNA polymerase core from uninfected or T4-infected *Escherichia coli*, but the major *E. coli* transcription initiation factor, $\sigma^{70}$, competed effectively for this binding.**

The proteins encoded by bacteriophage T4 genes 55, 45, and 33 are specifically required for late transcription during viral development. Each interacts with *Escherichia coli* RNA polymerase (for reviews, see references 3 and 16). The gene 55 protein (gp55) is a member of the σ family of proteins (5, 7) necessary and sufficient for T4 late promoter recognition by RNA polymerase (2, 8), yet the major *E. coli* $\sigma^{70}$ is dominant over gp55 in competition for binding RNA polymerase in vitro (29). The mechanism by which gp55 might overcome this dominance during infection is not yet fully understood; a virus-encoded protein with an apparent molecular mass of 10 kilodaltons (kDa) might be involved (22, 23), but other possibilities have been suggested (16). gp45 is a T4 DNA polymerase accessory protein and is part of the T4 replisome (for a review, see reference 15). The biochemistry of gp45 function in late transcription is not understood but probably involves coordination of RNA polymerase with the replisome on templates of particular structure. The role of gp33 in late transcription is still obscure, but the gene has been cloned (6). We report here the effective overproduction of gp33 and describe some RNA polymerase-binding properties (8, 17, 21) of the radioactively pure protein.

An expression system based on the λ $p_L$ promoter under the control of the heat-sensitive λ *c*I857 repressor (18) had been used with success in the overproduction of T4 gp55 (4) but gave disappointing results for gp33 (6). Four possible sources of difficulty were apparent. First, gp33 might itself inhibit *E. coli* RNA polymerase. We therefore turned to a T4 RNA polymerase-based expression system in which T7 RNA polymerase synthesis is under control of λ *c*I857 repressor on plasmid pGP1-2 (26) and the gene of interest is under the control of a T7 RNA polymerase promoter, $p_{T7}$, on the compatible vector plasmid pGEM-1 (Promega Biotec). This system has the further advantage that, since T7 RNA polymerase is insensitive to rifampin and $p_{T7}$ is the principal (or sole) promoter in the cell that it can activate, the overexpressed protein can be radiolabeled with little competition from other protein synthesis and with high specificity. A second source of concern was a gene 33-internal inverted repeat sequence suspected of potential transcriptional termination (6). However, preliminary experiments (R. Henke and K. P. Williams, unpublished) indicated efficient transcription in vitro through this segment by T7 RNA polymerase. Furthermore, with the insertion of gene 33 (from pPLSB33 [6]) between $p_{T7}$ and the *bla* gene of pGEM-1, forming plasmid p33T, synthesis of β-lactamase was only slightly depressed (Fig. 1C, cf. lanes b and c). Nevertheless, the native gene 33 in p33T yielded relatively little gp33 in this system. The third and fourth sources of potential difficulty were a high frequency of rare *E. coli* codons in gene 33 and a poor site for translational initiation. It was hoped that reduction of competing translation would mitigate the former difficulty. The native translational initiation site was improved (24; L. Gold, personal communication) by positioning a strong Shine-Dalgarno sequence seven A or U upstream of the initiation codon, with minimal secondary structure in the vicinity (Fig. 1B); this was achieved by introducing (10) an *Xba*I cleavage site into p33T at the gene 33 initiation codon, permitting a convenient and precise replacement of the upstream sequence with a synthetic DNA fragment. gp33 synthesis with the resulting plasmid, pKW6, was improved relative to p33T (Fig. 1C). The intrinsically stronger translational initiation in pKW6 may compensate for ribosome drift (for a review, see reference 27) from a downstream stop codon (Fig. 1B) that might transcriptionally couple gene 33 to an upstream open reading frame (orfB [6]) during infection.

An inverse relationship between synthesis of gp33 and β-lactamase was observed among these plasmids, suggesting that the two cotranscribed translation units might compete during overexpression. β-Lactamase synthesis was further reduced by inserting the *t*φ terminator (from plasmid pET-3 [19]), active with T7 RNA polymerase on plasmids (25) between genes 33 and *bla* of pKW6. gp33 synthesis was indeed further improved in the resulting plasmid, pKW7 (Fig. 1C). gp33 is reasonably stable, with a half-life of ca. 60 min (data not shown) in the Lon and OmpT protease-deficient strain that we used (BL21 [25]).

BL21(pGP1-2, pKW7) was not totally stable at 30°C, perhaps due to selective pressure against the substantially "leaky" synthesis of gp33 that occurred in these cells at 30°C; we have documented one instance in which gene 33 spontaneously mutated to encode a severely truncated protein. Although we noted that this instability problem could, in principle, be solved by better control of leaky transcription of the T7 RNA polymerase gene or by coexpressing T7 lysozyme to inhibit the activity of leakage levels of T7 RNA polymerase (13), we chose to confront the problem by confirming the sequence of gene 33 at the time of overexpression.

gp33 becomes prominent among the proteins of BL21 (pGP1-2, pKW7) after induction (Fig. 2, lane c) and is

* Corresponding author.

soluble. We purified gp33 from a mixture of radiolabeled and unlabeled cells (Fig. 2; the procedure is detailed below). gp33 in the final fraction was 97% pure with respect to radioactivity and ca. 75% pure with respect to protein. The trace of RNA polymerase contaminating this preparation constituted less than 1% of gp33 on a molar basis and could subsequently be removed, together with most of the remaining contaminating proteins, by chromatography on Sephadex G-75. Although the calculated molecular mass of gp33 is 12.8 kDa, it has almost the same electrophoretic mobility as the ω subunit of *E. coli* RNA polymerase (10.1 kDa), reflecting its strongly acidic character.

The binding of $^{35}$S-labeled gp33 to certain forms of RNA polymerase was analyzed by gel exclusion chromatography. Proteins were incubated for 40 min at 25°C in 100 μl of 250 mM potassium acetate–10 mM magnesium acetate–10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), pH 7.8–5% glycerol–0.1% Brij 58–0.1 mM dithiothreitol and then subjected to chromatography on 3-ml Sephacryl S-200 columns at 2°C, taking ca. 100-μl fractions. The radioactivity profiles of such columns showed reasonable resolution of RNA polymerase-bound gp33 eluting in the void volume from free gp33, which was included. RNA polymerase was purified from *E. coli* DG156 infected (18 min, 30°C) with gene 33-deficient T4 (double-nonsense mutant [amN134-amC18]) and was used to quantitate binding parameters for gp33; its concentration, as well as that of the gp33, was determined by amino acid composition analysis (see below). T4-modified RNA polymerase contains one equivalent of T4 *rpbA* protein, and its α subunits are ADP ribosylated; our preparation was virtually free of $\sigma^{70}$, gp55, and gp33. The dissociation constant for gp33 bound to 33$^-$ T4 RNA polymerase core was 25 nM, and the stoichiometry

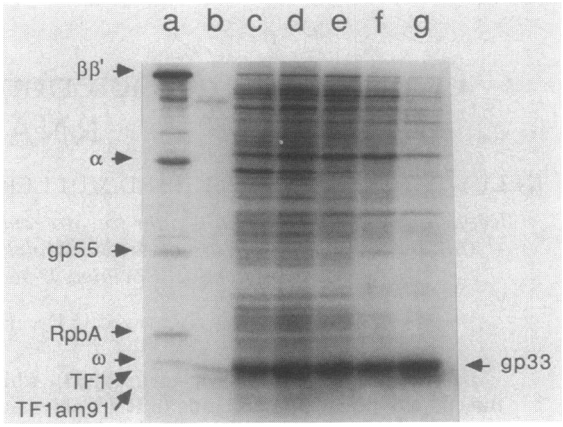


FIG. 2. Purification of gp33. Electrophoresis was done on a 14% sodium dodecyl sulfate-polyacrylamide gel (11) with Coomassie blue staining. Details of the procedure are provided in the text. Lanes contain crude extract of overproducing cells (lane c) and the same extract after lysis, sonication, and centrifugation (fraction 1) (lane d); gp33-containing fraction after streptomycin sulfate and ammonium sulfate precipitation (fraction 2) (lane e); flowthrough from heparin-Sepharose (fraction 3) (lane f); and pooled fractions eluting between 250 and 550 mM NaCl from DEAE-Sephadex (fraction 4) (lane g). Marker proteins were from T4-modified RNA polymerase (lane a) (9) and a mixture of wild-type (10.7 kDa) and am91 amber fragment (9.7 kDa) of TF1 protein of bacteriophage SPO1 (lane b) (prepared by M. Sayre) (20).

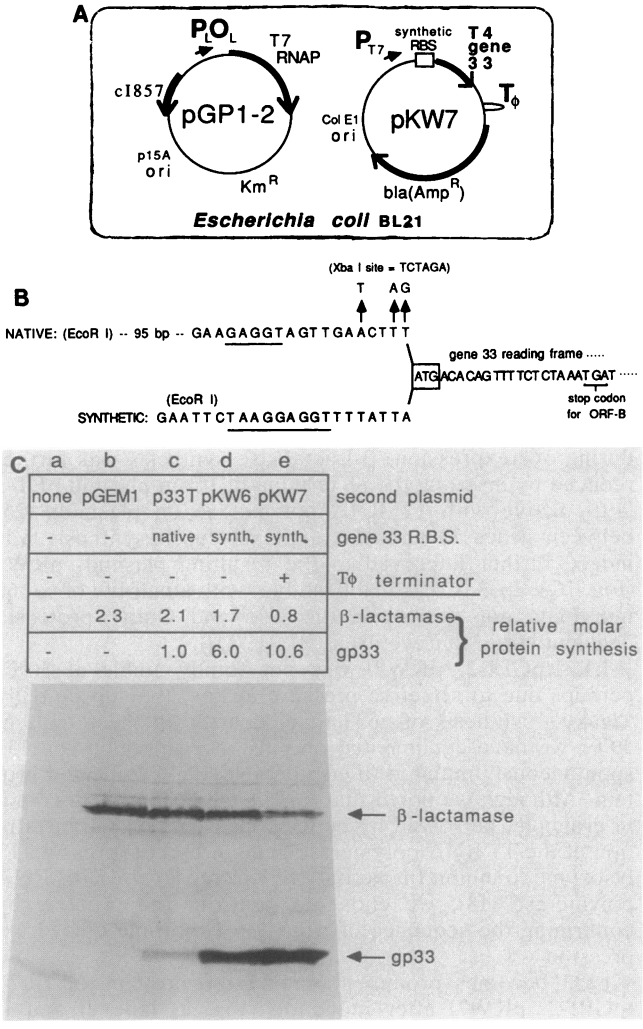


FIG. 1. Overexpression system for T4 gene 33. (A) The T4 gene 33 overexpression system consisted of *E. coli* BL21 cotransformed with pGP1-2 and pKW7. BL21 (25), a Met$^+$ derivative of B834, is deficient in the Lon and OmpT proteases. pGP1-2 (26) contains the bacteriophage T7 RNA polymerase (RNAP) gene under the control of the phage λ $p_L$ $o_L$ promoter and operator and also carries the gene encoding the heat-sensitive λ *c*I857 repressor. pKW7 contains T4 gene 33 with a synthetic ribosome-binding site (RBS) downstream of a T7 RNA polymerase promoter, $p_{T7}$, and upstream of *t*φ, a transcriptional terminator recognized by T7 RNA polymerase (25). (B) Nucleotide sequences encoding the native ribosome-binding site of gene 33 and the synthetic site present in pKW6 and pKW7 are shown; the longest contiguous matches to the Shine-Dalgarno sequence are underlined. The base substitutions that created an *Xba*I site (see text) are indicated. (C) Improvement of synthesis and specific radiolabeling of gp33. BL21(pGP1-2) (lane a) or BL21(pGP1-2) additionally transformed with the indicated compatible plasmid (lanes b to e) was grown, heat induced for T7 RNA polymerase production, treated with rifampin (400 μg/ml), labeled with [$^{35}$S]methionine (3.7 μCi in 1 ml of culture at an $A_{650}$ of 0.1), and processed for electrophoresis essentially as described elsewhere (26). Note that plasmids p33T, pKW6, and pKW7 differ from the vector pGEM-1 only within its multiple cloning site. Data on relative molar protein production are expressed in arbitrary units based on densitometric scans of the pictured autoradiogram of the gel and corrected for methionine content of the proteins (10 residues in β-lactamase and 3 in gp33).

was approximately one at saturation (Fig. 3A). We tested the binding capacity of two other forms of RNA polymerase from uninfected *E. coli* for gp33 (Fig. 3B). RNA polymerase core ($\sigma^{70}$-free) from uninfected cells had a capacity for gp33 similar to that of the 33⁻ T4-modified core, but the $\sigma^{70}$-saturated holoenzyme was unable to bind gp33 strongly, a pattern similar to that observed by Ratner (17). The apparent competition between gp33 and $\sigma^{70}$ for binding RNA polymerase core was further studied by first allowing gp33 to bind 33⁻ T4-modified RNA polymerase core and then adding purified $\sigma^{70}$ (5); previously bound gp33 was displaced from the core by $\sigma^{70}$ (Fig. 3C). gp33 and gp55 thus share certain qualitative features of interaction with RNA polymerase (29); they bind RNA polymerase in the absence of $\sigma^{70}$ but not in its presence, and these relationships obtain whether or not the RNA polymerase core is T4 modified (with *rpbA* protein and ADP-ribosylated α subunits.) These results suggest that gp33 alone would not facilitate the replacement of $\sigma^{70}$ by gp55 on RNA polymerase core. The availability of gp33 in substantial quantities allows further analysis of its mode of action in regulating viral gene activity.

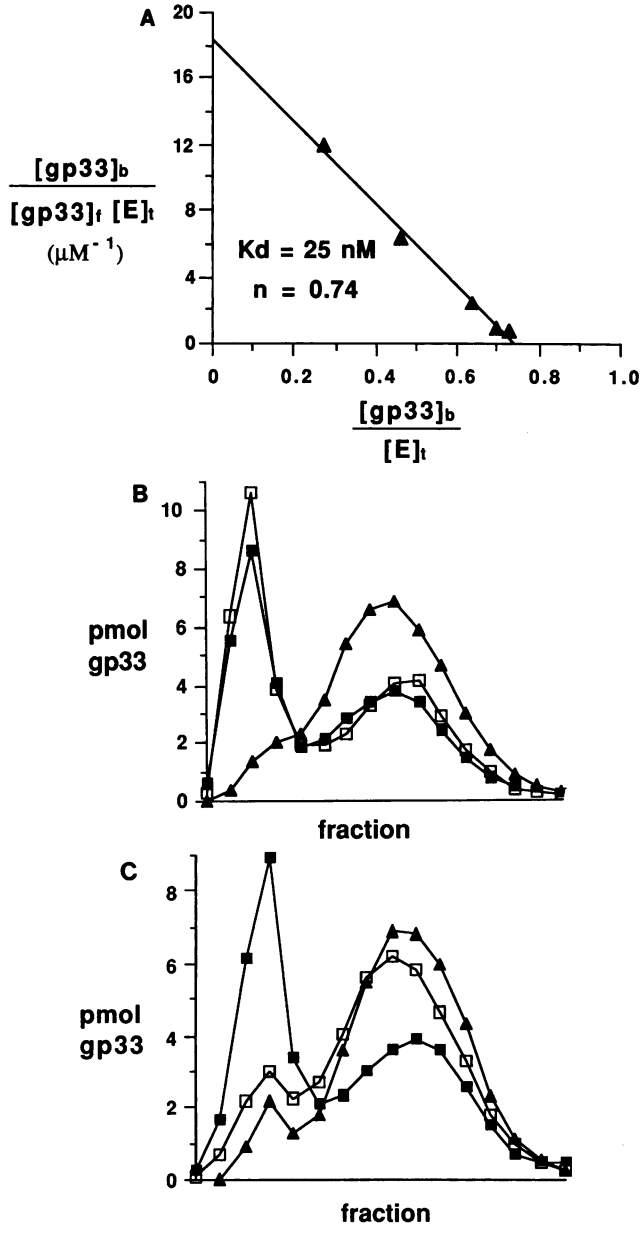


FIG. 3. RNA polymerase-binding properties of gp33. (A) Scatchard analysis. Radiolabeled gp33 (fraction 4, Fig. 2) was mixed with 27.5 pmol of 33⁻ T4-modified RNA polymerase core at various molar ratios, incubated, and subjected to gel exclusion chromatography as detailed in the text. Cross-contributions of the peaks of bound and free gp33 in the radioactivity profile were evaluated graphically by drawing out symmetrical peaks. The concentrations of free gp33, RNA polymerase-bound gp33, and total RNA polymerase are symbolized by $[gp33]_f$, $[gp33]_b$, and $[E]_t$, respectively; linear regression yielded the dissociation constant of the gp33-RNA polymerase complex (Kd) and the number of binding sites for gp33 on RNA polymerase (n). (B) RNA polymerase binding to gp33. gp33 (43 pmol) was mixed with 34 pmol of 33⁻ T4-modified RNA polymerase core (■), $\sigma^{70}$-free core from uninfected *E. coli* (□), or $\sigma^{70}$-saturated holoenzyme supplemented with 60 pmol of $\sigma^{70}$ (▲), incubated, and subjected to chromatography as described in the text. The leftmost peak represents gp33 bound to RNA polymerase and eluting with it in the void volume. (The result was similar when no extra $\sigma^{70}$ was added to the holoenzyme.) (C) Displacement of RNA polymerase bound to gp33 by $\sigma^{70}$. gp33 (43 pmol) was mixed with 34 pmol of 33⁻ T4-modified RNA polymerase core and incubated for 20 min at 25°C and then immediately subjected to chromatography (■) or mixed with 60 (□) or 200 pmol (▲) pmol of $\sigma^{70}$ and further incubated for 20 min at 25°C before chromatography.

**Purification of gp33.** gp33 was produced in BL21(pKW7, pGP1-2) by growing the cells in Luria-Bertani broth (12) with antibiotics to an $A_{650}$ of 1.0, shifting them to 42°C for 15 min, adding rifampin to 0.4 mg/ml, and shaking them for 10 min more at 42°C and then for 60 min at 30°C; the correct and unchanged gene 33 sequence was verified in plasmid DNA purified from a sample of these induced cells. A small batch of cells grown in parallel (5 ml; $A_{650}$, 0.5) was centrifuged, suspended in 10 ml of M9 medium supplemented with 20 μg of thiamine per ml and 100 μg each of 18 amino acids (cysteine and methionine excluded) per ml and shaken for 60 min at 30°C. gp33 synthesis was induced as described above, but 5 min after being shifted back to 30°C, 180 μCi of [$^{35}$S]methionine was added and shaking continued for 60 min. All purification steps (Fig. 2) were performed at 2 to 4°C. Both batches of cells were washed in lysis buffer (40 mM Tris hydrochloride [pH 8], 5% glycerol, 10 mM $MgCl_2$, 1 mM EDTA, 200 mM NaCl, 0.1 mM dithiothreitol). Labeled cells (7 mg) were mixed with 46 mg of unlabeled cells in 600 μl of lysis buffer supplemented with 25 μg of phenylmethylsulfonyl fluoride per ml. The cells were sonicated and centrifuged for 10 min at 15,000 × *g*, and then the pellet was suspended in 600 μl of lysis buffer, sonicated, and centrifuged as before. The supernatants were pooled (fraction 1) and centrifuged for 15 min at 95,000 rpm in the A-100/18 rotor of the Beckman air-driven centrifuge. The supernatant (S-100) was mixed with 0.25 volume of 10% streptomycin sulfate in lysis buffer, incubated for 15 min, and then centrifuged for 10 min at 15,000 × *g*. The supernatant was mixed with powdered ammonium sulfate to 65% saturation (400 mg of ammonium sulfate per ml of S-100), incubated for 15 min, and then centrifuged for 10 min at 15,000 × *g*. The pellet was suspended in B75 buffer (10 mM Tris hydrochloride [pH 8], 5% glycerol, 0.1 mM EDTA, 75 mM NaCl, 0.1 mM dithiothreitol) and reprecipitated with 65% saturated ammonium sulfate. The pellet was suspended in B75 buffer, and ammonium sulfate was removed by repeated concentration (Centricon-10; Amicon Corp.) and dilution. The final dilution, to 600 μl (fraction 2), was loaded onto a 1-ml heparin-Sepharose column (prepared by D. Ades), incubated

for 30 min, and then washed through. The flowthrough (fraction 3) was loaded onto a 1-ml DEAE-Sephadex A25 (Pharmacia, Inc.) column and eluted with a 4-ml linear gradient of B75-B750 buffer (B750 buffer is B75 with 750 mM NaCl instead of 75 mM NaCl.) Fractions eluting between 250 and 550 mM NaCl were pooled and concentrated (Centricon-10; Amicon) (fraction 4); the gp33 peak was at ca. 300 mM NaCl. Densitometric analysis of lane g showed that gp33 was ca. 76% pure. The concentration of gp33 was determined by subjecting duplicate samples to acid hydrolysis and by quantitation of the amino acids D, E, F, I, L, K, N, P, Q, and V in conjunction with the known sequence of gp33. The 24% contaminating proteins were assumed to have the average amino acid composition of all *E. coli* proteins (14).

We thank L. Gold, G. A. Kassavetis, and D. Herendeen for useful discussions and advice.

K.P.W. has been a trainee of a Public Health Service Cell and Molecular Biology Training grant from the National Institutes of Health. The research on this topic has been supported by grants from the National Institute of General Medical Sciences (E.P.G.) and the Deutsche Forschungsgemeinschaft (W.R.).

## LITERATURE CITED

1. **Burgess, R. R., and J. J. Jendrisak.** 1975. A procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving Polymin P precipitation and DNA-cellulose chromatography. Biochemistry **14**:4634–4638.
2. **Elliot, T., and E. P. Geiduschek.** 1984. Defining a bacteriophage T4 late promoter: absence of a "-35" region. Cell **36**:211–219.
3. **Geiduschek, E. P., and G. A. Kassavetis.** 1988. Changes in RNA polymerase, p. 93–115. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Publishing Corp., New York.
4. **Gram, H., and W. Rüger.** 1985. Genes 55, αgt, 47 and 46 of bacteriophage T4: the genomic organization as deduced by sequence analysis. EMBO J. **4**:257–264.
5. **Gribskov, M., and R. R. Burgess.** 1986. Sigma factors from *E. coli*, *B. subtilis*, phage SP01, and phage T4 are homologous proteins. Nucleic Acids Res. **14**:6745–6763.
6. **Hahn, S., U. Kruse, and W. Rüger.** 1986. The region of phage T4 genes 34, 33 and 59: primary structures and organization on the genome. Nucleic Acids Res. **14**:9311–9327.
7. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57**: 839–872.
8. **Horvitz, H. R.** 1973. Polypeptide bound to the host RNA polymerase is specified by T4 control gene 33. Nature (London) New Biol. **244**:137–140.
9. **Kassavetis, G. A., and E. P. Geiduschek.** 1984. Defining a bacteriophage T4 late promoter: bacteriophage T4 gene *55* protein suffices for directing late promoter recognition. Proc. Natl. Acad. Sci. USA **81**:5101–5105.
10. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**: 488–492.
11. **Laemmli, U.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
12. **Miller, J. H.** 1972. Experiments in molecular genetics, appendix I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Moffatt, B. A., and F. W. Studier.** 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell **49**:221–227.
14. **Neidhart, F. C.** 1987. Chemical composition of *Escherichia coli*, p. 3–6. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
15. **Nossal, N. G., and B. M. Alberts.** 1983. Mechanism of DNA replication catalyzed by purified T4 replication proteins, p. 71–81. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.
16. **Rabussay, D.** 1983. Regulation of gene expression, p. 167–173. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.
17. **Ratner, D.** 1974. The interaction of bacterial and phage proteins with immobilized *Escherichia coli* RNA polymerase. J. Mol. Biol. **88**:373–383.
18. **Remaut, E., P. Stanssens, and W. Fiers.** 1981. Plasmid vectors for high-efficiency expression controlled by the $P_L$ promoter of coliphage lambda. Gene **15**:81–93.
19. **Rosenberg, A. H., B. N. Lade, D.-S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56**:125–135.
20. **Sayre, M. H., and E. P. Geiduschek.** 1988. TF1, the bacteriophage SPO1-encoded type II DNA-binding protein, is essential for viral multiplication. J. Virol. **62**:3455–3462.
21. **Stevens, A.** 1972. New small polypeptides associated with DNA-dependent RNA polymerase of *Escherichia coli* after infection with bacteriophage T4. Proc. Natl. Acad. Sci. USA **69**: 603–607.
22. **Stevens, A.** 1976. A salt-promoted inhibitor of RNA polymerase isolated from T4-infected *E. coli*, p. 617–627. *In* R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Stevens, A.** 1977. Inhibition of DNA-enzyme binding by an RNA polymerase inhibitor from T4 phage-infected *Escherichia coli*. Biochim. Biophys. Acta **475**:193–196.
24. **Stormo, G. D.** 1986. Translation initiation, p. 195–224. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Stoneham, Mass.
25. **Studier, W. F., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
26. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**: 1074–1078.
27. **van Duin, J.** 1988. Single-stranded RNA bacteriophages, p. 117–167. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Publishing Corp., New York.
28. **Williams, K. P., G. A. Kassavetis, F. S. Esch, and E. P. Geiduschek.** 1987. Identification of the gene encoding an RNA polymerase-binding protein of bacteriophage T4. J. Virol. **61**: 597–599.
29. **Williams, K. P., G. A. Kassavetis, and E. P. Geiduschek.** 1987. Interactions of the bacteriophage T4 gene 55 product with *Escherichia coli* RNA polymerase. J. Biol. Chem. **262**:12365–12371.