JOURNAL OF VIROLOGY, Oct. 1978, p. 408-410
0022-538X/78/0028-0408$02.00/0

Vol. 28, No. 1
Printed in U.S.A.

Identification of Lysis Protein E of Bacteriophage φX174

THOMAS J. POLLOCK,† ETHEL S. TESSMAN, AND IRWIN TESSMAN*
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Received for publication 8 May 1978

The product of gene *E*, the lysis gene of φX174, has been identified as a distinct band in a sodium dodecyl sulfate-gel electropherogram. The position of the band is consistent with the molecular weight of 10,589 calculated from the nucleotide sequence of the gene. The band is eliminated by a nonsense mutation in gene *E*. It is estimated that roughly 100 to 300 molecules of E protein are made in an infected cell; this appears to be less than one-tenth the amount of protein made by gene *D*, in which gene *E* is wholly contained.

The process by which the single-stranded small DNA bacteriophages φX174 and S13 lyse their bacterial host is not understood, but it is known to require the product of phage gene *E* (6). Nonsense mutations in gene *E* appear to affect only lysis, thereby allowing large numbers of infective phage particles to accumulate during an extended latent period (6). Similar mutations have been isolated for S13 (12). The E protein lies within gene *D*, in a different reading frame (2).

We report here the identification of the product of the φX174 lysis gene. The E protein was identified by the elimination of a low-molecular-weight phage-specified band from gel patterns of radioactive proteins synthesized during an *amE3* infection (Fig. 1). The presence of E protein could only be detected after infection of Hcr^- host cells that were heavily UV irradiated to eliminate host protein synthesis. Even after irradiation of the host, a small amount of host protein migrated at the position of the E band in extracts of uninfected cells (Fig. 1, lane 1). However, in an *amE* infection (Fig. 1, lane 2), not only was the E protein band eliminated, but also the residual host protein band, presumably by the same mechanism that shuts off synthesis of most host proteins midway in the latent period after infection with φX174, S13, or G4.

Barrell et al. (2) found by restriction-fragment rescue experiments that the DNA sequence for the *E* gene overlaps the *D* region of the φX174 genome. The base changes caused by *amE* mutants were located on the φX174 sequence map by comparing the DNA sequences of *amE* mutants and wild-type phage. The location of two *amE* nonsense codons (one of which was *amE3*) showed the correct reading frame for the E protein and allowed Barrell et al. to identify the natural termination codon as being the first available in-phase nonsense codon after the *amE* mutations. The DNA sequence indicated two possible initiation codons. Barrell et al. considered one of these more likely than the other because it is preceded by a "Shine-Dalgarno sequence," implicated in ribosome binding (11). The molecular weight calculated from the gene *E* sequence is 10,589. (The other possible initiation codon would lead to an E protein that is 10 amino acids longer, with a molecular weight of 11,642.) Since the DNA sequence of gene *E* indicates that the E protein should contain 23% leucine, we labeled proteins with radioactive leucine.

The position of the E protein in the gel corresponded to a molecular weight of 10,700. The gel was calibrated (legend to Fig. 1) from the positions (lane 4) of proteins A, F, H, A^*_1, G_1, D, C, and K, whose molecular weights were all inferred from the DNA sequence. The identification of protein K of phage G4 and the location of gene *K* in the G4 and φX174 DNA sequences is described by Shaw et al. (10). The corresponding *K* gene product for φX174 was identified by Pollock et al. (8).

Burgess and Denhardt (4) previously reported an identification of the E protein. However, no single band was eliminated from their tube gel pattern. Instead there was about a 50% decrease in a broad band composed of several phage proteins smaller than the D protein. The mutant used was *am6*, now known to contain at least three nonsense mutations: an amber in *A*, an amber in *E*, and an opal in *K* (8, 9, 13). Therefore the reduction in peak height observed by Burgess and Denhardt was probably due largely to elimination of the K protein, which is more prominent than E (Fig. 1).

When corrected for the known leucine composition of proteins E and H (21 and 19, respec-

† Present address: Laboratory of Neurochemistry, National Institutes of Health, Bethesda, MD 20014.

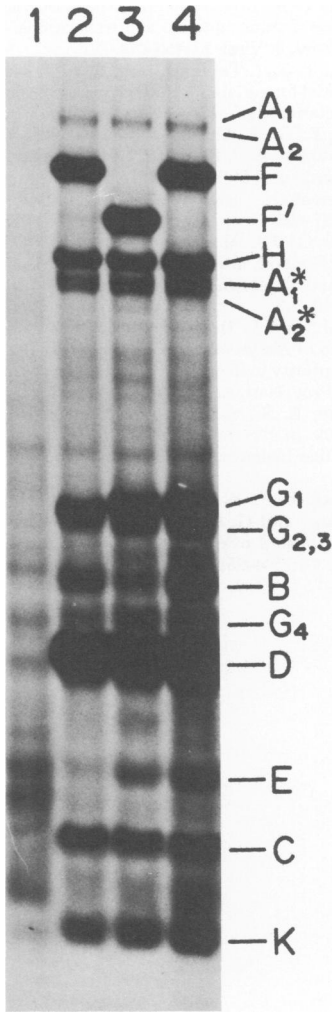

FIG. 1. *Fluorogram of sodium dodecyl sulfate-polyacrylamide gel of extracts of UV-irradiated φX-infected or uninfected cells. Escherichia coli C AP1 (Hcr⁻) was grown to 2 × 10⁸ cells per ml and irradiated at a UV dose of about 150 J/m². $MgSO_4$ was added to 10^{-2} M, phage were added to a multiplicity of 10 PFU/cell, and the cultures were aerated at 37°C. At 10 min after infection, [³H]leucine (60 Ci/mmol) was added to 50 μCi/ml. The cultures were chilled at 40 min after infection, processed as described previously (7), then subjected to electrophoresis on gels of 18.5% (wt/vol) acrylamide and 0.17% (wt/vol) N,N'-methylene-bisacrylamide. Lane 1, uninfected cells; lane 2, amE3; lane 3, amF57; lane 4, φX174 wild type. Identification of C and K proteins is described by Pollock et al. (8). The fluorogram was scanned with a Joyce Loebl Mark IIIC microdensitometer to locate the position of each band and also to determine the amount of E protein relative to H and D. The molecular weight of the E protein was calculated from the data in lane 4. For the known proteins the log of the molecular weight was plotted versus the position of the band in the gel, and a second-degree polynomial was fitted to the points by a least-squares analysis. The best-fitting curve turned out to be linear, the coefficient of the second-order term being negligible. The position of the E band on the curve corresponded to 10,700 daltons. The molecular weights of the proteins used to calibrate the gel were as follows: A, 58,650; F, 48,430; H, 34,380; A^*_1, 31,950; G_1, 19,030; D, 16,920; C, 10,070; K, 6,380. They were all obtained from the nucleotide sequence. The B protein was not used because its electrophoretic mobility is markedly anomalous (9).*

tively, inferred from the DNA sequence), densitometry of lanes 3 and 4 (Fig. 1) shows that the number of E molecules is 0.20 times that of H in UV-irradiated cells. The relative amount of E protein in normal cells could not be determined because the E band in gels was obscured by host proteins when the extracts were made from unirradiated cells. It is known that among the more prominent proteins the amount of protein D relative to F, G, and H can be increased two- to fourfold by irradiation of the cells (unpublished data). Therefore, within such limits it is now possible to estimate the number of E molecules in a normal infection. With 12 H molecules per virus particle (3) and assuming a burst size of about 100, there should be approximately 240 (0.20 × 100 × 12) molecules of E per cell. This is a very rough estimate inasmuch as we do not know if all H protein is assembled into virus particles and if all E protein is recovered.

The amount of E is in striking contrast to the amount of gene *D* product synthesized. Farber (5) estimated that about 32 monomers of D were present per infective particle produced in *amE3*-infected nonsuppressing cells. (As shown in Fig. 1, the *amE3* mutation does not greatly alter the amount of D protein synthesized.) We can then estimate that the expression of gene *E* (nested within gene *D*) is about 13-fold lower than for gene *D* itself. This can also be seen in Fig. 1. Although the great disparity in densities of D and E precludes an accurate comparison, densitometry shows that the number of E molecules is about 28-fold lower than that of D (corrected for the E/D leucine ratio of 1.9). Allowing even for a preferential expression of gene *D* in irradiated cells, there still would be a large excess of D over E in normal cells; it appears that there are 100 to 300 molecules of E made per cell. The nature of the regulatory event controlling expression of these two overlapping genes, whether it be at the level of transcription or translation, remains to be discovered, as does the mechanism of action of protein E in cell lysis.

This work was supported by Public Health Service grants from the National Institute of Allergy and Infectious Diseases to E.S.T. (AI-11853) and from the National Cancer Institute to I.T. (CA-22239). T.J.P. was a biophysics trainee supported by the National Institute of General Medical Sciences (Public Health Service grant GM-779).

LITERATURE CITED

1. **Anderson, E. H.** 1946. Growth requirements of virus-resistant mutants of *Escherichia coli* strain "B." Proc. Natl. Acad. Sci. U.S.A. **32**:120–128.
2. **Barrell, D. G., G. M. Air, and C. A. Hutchison III.** 1976. Overlapping gene in bacteriophage φX174. Nature (London) **264**:34–41.
3. **Burgess, A. B.** 1969. Studies on the proteins of φX174. II. The protein composition of the φX coat. Proc. Natl. Acad. Sci. U.S.A. **64**:613–617.
4. **Burgess, A. B., and D. T. Denhardt.** 1969. Studies on φX174 proteins. I. Phage-specific proteins synthesized after infection of *Escherichia coli*. J. Mol. Biol. **44**:377–386.
5. **Farber, M. B.** 1976. Purification and properties of bacteriophage φX174 gene D product. J. Virol. **17**:1027–1037.
6. **Hutchison, C. A., III, and R. L. Sinsheimer.** 1966. The process of infection with bacteriophage φX174. X. Mutations in a φX lysis gene. J. Mol. Biol. **18**:429–447.
7. **Pollock, T. J.** 1977. Gene *D* of bacteriophage φX174: absence of transcriptional and translational regulatory properties. J. Virol. **21**:469–474.
8. **Pollock, T. J., I. Tessman, and E. S. Tessman.** 1978. Potential for variability through multiple gene products of bacteriophage φX174. Nature (London) **274**:34–37.
9. **Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Slocomb, and M. Smith.** 1977. Nucleotide sequence of bacteriophage φX174 DNA. Nature (London) **265**:687–695.
10. **Shaw, D. C., J. C. Walker, F. Northrop, B. G. Barrell, G. N. Godson, and J. C. Fiddes.** 1978. Gene *K*, a new overlapping gene in bacteriophage G4. Nature (London) **272**:510–515.
11. **Shine, J., and L. Dalgarno.** 1975. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. **71**:1342–1346.
12. **Tessman, E. S.** 1967. Gene function in phage S13, p. 193–209. *In* J. S. Colter and W. Paranchych (ed.), The molecular biology of viruses. Academic Press Inc., New York.
13. **Weisbeek, P. J., W. E. Borrias, S. A. Langeveld, P. D. Baas, and G. A. van Arkel.** 1977. Bacteriophage φX174: gene *A* overlaps gene *B*. Proc. Natl. Acad. Sci. U.S.A. **74**:2504–2508.