

## Gene *K* of Bacteriophage $\phi$ X174 Codes for a Nonessential Protein

ETHEL S. TESSMAN,\* IRWIN TESSMAN, AND THOMAS J. POLLOCK†

*Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907*

Gene *K* of  $\phi$ X174, which overlaps genes *A*, *B*, and *C*, was found to be nonessential, although possibly beneficial for the growth of the phage. Viable mutants of gene *K* made less than 4% of the normal amount of *K* protein as judged by quantitative fluorography of sodium dodecyl sulfate-polyacrylamide gels; compared with the wild-type  $\phi$ X, *K* mutants had an identical latent period but a two- to threefold reduction in burst size.

In the small icosahedral DNA phage G4, related to  $\phi$ X174 and S13, a protein designated *K* has recently been identified by Shaw et al. (10). By comparing the N-terminal amino acid sequence of the *K* protein with the G4 DNA sequence, they located the *K* gene on the G4 genome. *K* overlaps genes *A*, *B*, and *C* (Fig. 1). Shaw et al. were unable to detect a *K* protein in extracts of  $\phi$ X-infected cells, but they pointed out that such a protein may exist because the region of the  $\phi$ X genome that starts at the end of *B* has considerable homology with all of the G4 *K* gene. In studies of  $\phi$ X proteins, we have reported that a protein having characteristics expected for one starting within the *B* termination codons and reading into the *C* region, but in a different translational frame, is observed in sodium dodecyl sulfate-polyacrylamide gels of extracts of infected cells (7). The presumptive  $\phi$ X *K* protein has a molecular weight of 8,000 calculated from its position in gels, compared with the molecular weight of 6,384 predicted by the  $\phi$ X sequence; its leucine to lysine incorporation ratio is  $3.2 \pm 0.5$ , compared with a ratio of 2.80 predicted by the sequence; it is eliminated by an ochre nonsense mutation in gene *C*, which would effect a serine  $\rightarrow$  leucine substitution in the gene *K* product; it is eliminated by an opal codon in the reading frame of the presumed gene *K*. We therefore concluded that  $\phi$ X does indeed also have a *K* protein. (An updated version of the  $\phi$ X sequence is used here [9].)

The functions of the nine other genes of  $\phi$ X are understood to varying degrees (reviewed in reference 12). Except for gene *E*, they are all essential for in vivo production of mature viruses; nonsense mutations in those genes are lethal. Gene *E*, a lysis gene, is a special exception; mature viruses are produced intracellularly, but the growth cycle is blocked because the phage

are not released from the cell (3). For gene *K*, nothing is known about its function.

We have evidence that gene *K* is not essential for growth of the virus because there are mutants of  $\phi$ X that are viable even when they make no detectable *K* protein. This was shown by electropherograms of extracts from cells infected with various mutants of  $\phi$ X (Fig. 2).

Lanes 1 and 7 of Fig. 2 show  $\phi$ X wild-type phage proteins labeled with [ $^3$ H]leucine. The *C* and *K* bands were previously identified (5), but a further identification of *K* was made simply by labeling with [ $^3$ H]tryptophan (lane 8). From the DNA sequence (2, 9) it is known that the *K* protein lacks tryptophan and is therefore identified by its absence in lane 8.

Lane 2, which contained an extract of  $\phi$ X*amD81*, shows the absence of both the *D* and *K* proteins. Another gene *D* mutant,  $\phi$ X*amD10*, makes the *K* protein (unpublished data), so we assume that in addition to the amber mutation in *amD81* there is also a cryptic mutation that eliminates *K*. This was confirmed by isolating two independent *am*<sup>+</sup> revertants of *amD81* (R1 and R2, lanes 3 and 4) that still lack *K*.

Another  $\phi$ X mutant that makes no *K* is *amA62*. This phage was derived from  $\phi$ X*am6* and is known to be a double amber mutant in genes *A* and *E* (13). In *amA62* the *E* mutation was eliminated, but the *A* mutation remains and produces a slight shortening of the *A* protein. Analysis of the nucleotide sequence (8) reveals that *amA62* has an amber mutation corresponding to the sixth amino acid from the 3' end of *A* and in addition two missense mutations in *A*. One of these, at nucleotide 57, creates an arginine  $\rightarrow$  opal nonsense mutation in the frame of the overlapping gene *K*. We isolated two different revertants of *amA62* that grow in a nonsuppressing host, a temperature-sensitive pseudo-revertant (R4) and one with wild-type sensitivity (R11). Both revertants still lack *K* (lanes 5 and 6).

† Present address: Laboratory of Neurochemistry, National Institutes of Health, Bethesda, MD 20205.

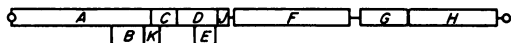


FIG. 1. The circular map of  $\phi X174$  in linear form. In  $\phi X$ , K uses an ATG start codon formed by the double TGA termination codons of B and overlaps 86 nucleotides of A and 89 of C.

The absence of K in all the revertants was not lethal; all revertants that do not make detectable K protein produce plaques on  $Sup^-$  (nonsuppressing) hosts, although the plaques are smaller than those of  $\phi X^+$ . To determine the maximum amount of K that could escape detection, the electropherograms were analyzed with a Joyce-Loebl microdensitometer (Fig. 3). The relative amounts and positions of the K and C proteins in the gels were determined from Fig. 3A. The K peak is apparently missing in Fig. 3B to F; the problem is to estimate how much K could be concealed in the shoulder of C. One could attempt to attribute the broadening of the C band at its base to a small K band. Each C band is indeed much wider at its base than a gaussian distribution would allow, but it is impossible for a K band to account for the widening because Fig. 3A shows that the peak has a relative standard deviation of 8%, which is much too narrow to explain the broadening of the C bands. There must be other background material that constitutes the broad and smooth shoulder.

Our approach was to assume that if there were a peak of K protein that amounted to as much as 40% of the magnitude of the shoulder at the expected K position, then it would be observed above the background. This is illustrated in Fig. 3F. In that way we calculated the maximum possible amounts of K in Fig. 3B to F, relative to the normal amount of K in wild-type extracts calculated by using the C peak as a reference, and found an average of  $3.9 \pm 0.3\%$ . We therefore concluded that there is at most 4% of the normal amount of K in these  $K^-$  mutants.

One-step growth experiments were done in a  $Sup^-$  host at  $37.0^\circ C$  to see whether  $\phi X$  growth was noticeably affected by the absence of K. Both  $\phi X^+$  and  $\phi X_{amA62-R11}$  (Fig. 4) had latent periods of  $13.0 \pm 0.5$  min, but  $\phi X^+$  had a burst size of 210 compared with only 88 for the mutant. And in another experiment, under the same growth conditions,  $\phi X^+$  and  $\phi X_{amD81-R1}$  both had a latent period of  $13.0 \pm 1.0$  min but a burst size of 190 for  $\phi X^+$  and 100 for the mutant. Thus, there seems to be roughly a twofold reduction in burst size, and that is consistent with the smaller mutant plaques.

The absence of the K protein was observed when the infected cells were grown in a minimal medium (Fig. 2), but the ability of the mutant phages to produce progeny was demonstrated in

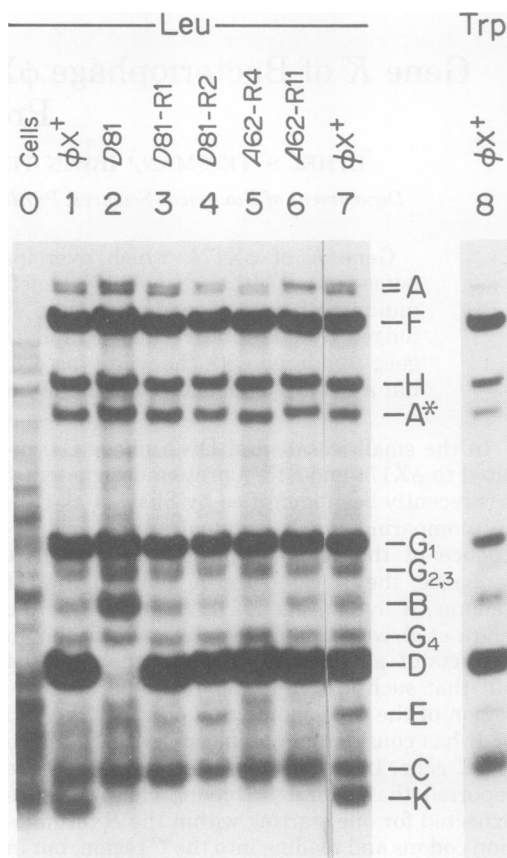


FIG. 2. Electropherogram showing absence of K protein in extracts of  $\phi X174$  mutants. *Escherichia coli* strain AP1, an *Hcr^-* derivative of strain C, was grown at  $37^\circ C$  to  $4 \times 10^8$  cells per ml in HFS-T medium consisting of 0.05 M Tris-hydrochloride (pH 7.7 at  $25^\circ C$ ), 0.02 M  $NH_4Cl$ , 0.01 M NaCl,  $10^{-3}$  M  $MgSO_4$ ,  $10^{-4}$  M  $CaCl_2$ ,  $5 \times 10^{-4}$  M potassium phosphate,  $10^{-5}$  M  $FeCl_3$ , and 0.2% glucose. The cells were irradiated with UV light ( $700$  to  $800 J/m^2$ ) to help reduce host protein synthesis and then infected with 10 PFU/cell. The cells were labeled with 20  $\mu Ci$  of either [ $^3H$ ]leucine (Amersham) or in one case [ $^3H$ ]tryptophan (Amersham) from 15 to 60 min after infection. Electrophoresis and fluorography have been described (4). The only significant differences were the use of a greater electric field (140 V/11 cm) and in the composition of the gel, which was, in the order of solution, 13.4% (wt/vol) acrylamide, 0.066% (wt/vol)  $N,N'$ -methylenebisacrylamide, 0.090% (vol/vol)  $N,N,N',N'$ -tetramethylethylenediamine, 0.090% (vol/vol) 2-mercaptoethanol, 26% (vol/vol) denaturing buffer (0.4 M Tris-acetate [pH 9.1], 2 M urea, 0.4% sodium dodecyl sulfate) and 0.24% (wt/vol) ammonium persulfate.

a complete medium (Fig. 4). For a more rigorous comparison we repeated the one-step growth experiments at  $37^\circ C$  in the HFS-T minimal medium used to make the extracts. The burst sizes

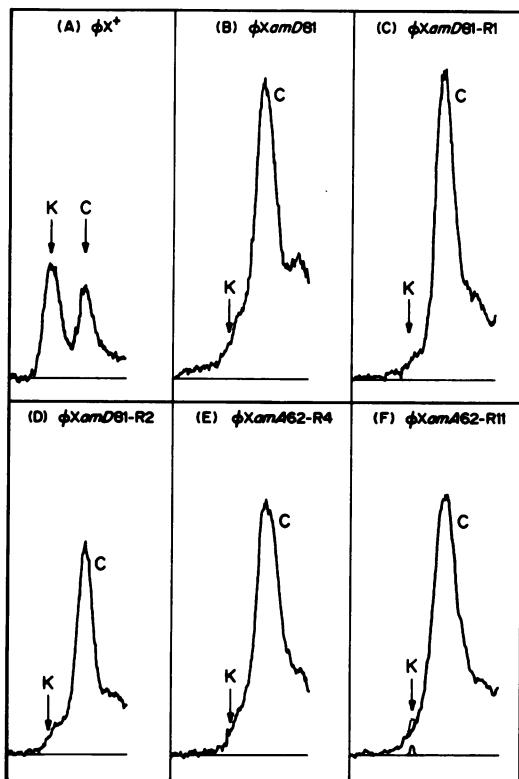


FIG. 3. Scan of original fluorogram films corresponding to Fig. 2 with a Joyce-Loebl Mark III C microdensitometer having a gray wedge with a maximum optical density of 1.60. (F) shows a small hypothetical K peak with the same relative standard deviation as the actual K peak in (A); the effect this peak would have on the observed shoulder of the C peak is depicted with a dotted curve.

at 60 min were as follows:  $\phi X^+$ , 52; A62-R11, 15; and D81-R1, 18. The latent periods were all  $15 \pm 1$  min. The results reinforce our conclusion that the mutants grow despite the lack of K protein; the reduction in burst size was threefold in this case.

Our experiments did not prove that gene K is responsible for the reduced burst size. The revertants of D81 and A62 could be pseudorevertants, and so their D or A genes need not be allelic to those in the  $\phi X^+$  standard. The K mutations could also affect the overlapping A or C genes. (In *ama62* and therefore also in the R4 and R11 revertants of *ama62* the gene K opal mutation is also a serine  $\rightarrow$  leucine mutation in the A protein [11]). At this stage, then, we prefer to emphasize that the absence of the K protein makes at most a two- to threefold reduction in the burst size and no noticeable effect on the latent period.

We estimated the relative amount of K made in a  $\phi X$ -infected cell. We compared C, K, and H by scanning the electropherogram corresponding to Fig. 1 with a Joyce-Loebl microdensitometer and determining the areas under the peaks (Table 1). From the areas and the leucine content of the proteins deduced from the DNA sequence (2, 9), the relative number of molecules for each protein was calculated. Our results indicated that in G4, K is made in about three

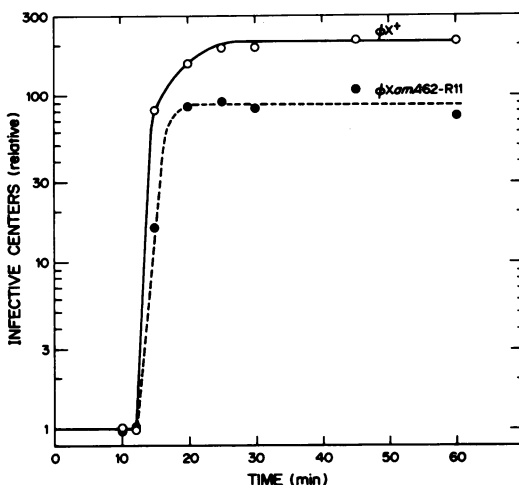


FIG. 4. One-step growth curves at  $37.0 \pm 0.1^\circ\text{C}$ . *E. coli* C (strain AP1) was grown to  $2 \times 10^8/\text{ml}$  in L broth (10 g of tryptone [Difco], 5 g of yeast extract [Difco], 5 g of NaCl, 1 liter of water) and  $\text{CaCl}_2$  was added to  $10^{-2}$  M. At zero time, phage were added ( $0.05$  PFU/ml) to 1 ml of cells in an 18-mm-outside-diameter tube and adsorbed for 5 min without additional aeration. The cultures were then diluted  $10^4$ - to  $10^6$ -fold in L broth and assayed for infective centers.

TABLE 1. Quantitative estimates of proteins C, K, and H<sup>a</sup>

Phage	Gene	Peak area	No. of leucines	Relative molar amt of proteins
$\phi X^+$	C	83	12	1.0
	K	109	14	1.1
	H	131	19	1.0
G4 <sup>+</sup>	C	100	13	3.5
	K	79	15	2.4
	H	42	19	1.0
G4 <sup>+</sup>	C	210	13	6.2
	K	149	15	3.8
	H	50	19	1.0

<sup>a</sup> Proteins J and K of G4 are made in approximately equimolar amounts (10). For comparison, we note that in  $\phi X$ , J, F, G, and H occur in the approximate ratios of 5:5:5:1 (1, 5).

times the molar amount of H; in  $\phi X$  we found that K and H are roughly equimolar.

The absolute amount of K can be crudely estimated in the same way we estimated that  $\phi X$  makes about 300 molecules of the E protein in an infected cell (6). Assuming 12 H molecules per virus and a burst size of 200, there would be 2,400 H molecules per infected cell and therefore 2,500 to 7,500 molecules of K depending on whether it is  $\phi X$  or G4. The large uncertainties in such a calculation have been discussed elsewhere (6).

It appears that either gene *K* is nonessential for growth of  $\phi X$  or else less than 4% of the normal K product, the limit of sensitivity of our experiments, could still provide most of the needed gene function, an unlikely though not impossible condition. But although *K* could be nonessential, it may be beneficial for phage growth. Its possible contribution to a two- to threefold increase in the phage burst size might give the gene a substantial selective advantage. Its function remains unknown.

Does *K* have a function at all? Since the reduced burst size has not been conclusively attributed to the absence of *K*, there is as yet no proof that *K* has any function. But it is difficult to argue that *K* is accidentally made simply because the nucleotide sequence is there. Not only do both G4 and  $\phi X$  make *K*, but extracts of S13-infected cells show a corresponding electrophoretic band. Furthermore, if the protein were of no use, it would seem, judging from the nucleotide sequence known for  $\phi X$  and G4, that the virus would have eliminated the ATG start codon, which apparently can be done without creating a missense mutation in gene *A* or preventing translational termination at the end of gene *B*. It appears likely, therefore, that *K* has a function.

We appreciate the assistance of Ruth Ann Gyure.

This work was supported by Public Health Service grant AI-11853 from the National Institute of Allergy and Infectious Diseases to E.S.T. and Public Health Service grant CA-22239 from the National Cancer Institute to I.T. T.J.P. was a bio-

physics trainee supported by the National Institute of General Medical Sciences (GM-779).

#### LITERATURE CITED

1. Burgess, A. B. 1969. Studies on the proteins of  $\phi X174$ . II. The protein composition of the  $\phi X$  coat. *Proc. Natl. Acad. Sci. U.S.A.* **64**:613-617.
2. Godson, G. N., B. G. Barrell, R. Staden, and J. C. Fiddes. 1978. Nucleotide sequence of bacteriophage G4 DNA. *Nature (London)* **276**:236-247.
3. Hutchison, C. A., III, and R. L. Sinsheimer. 1966. The process of infection with bacteriophage  $\phi X174$ . X. Mutations in a  $\phi X$  lysis gene. *J. Mol. Biol.* **18**:429-447.
4. Pollock, T. J. 1977. Gene *D* of bacteriophage  $\phi X174$ : absence of transcriptional and translational regulatory properties. *J. Virol.* **21**:468-474.
5. Pollock, T. J., and I. Tessman. 1976. Units of functional stability in bacteriophage messenger RNA. *J. Mol. Biol.* **108**:651-664.
6. Pollock, T. J., E. S. Tessman, and I. Tessman. 1978. Identification of lysis protein E of bacteriophage  $\phi X174$ . *J. Virol.* **28**:408-410.
7. Pollock, T. J., I. Tessman, and E. S. Tessman. 1978. Potential for variability through multiple gene products of bacteriophage  $\phi X174$ . *Nature (London)* **274**:34-37.
8. Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Y. Slocombe, and M. Smith. 1977. Nucleotide sequence of bacteriophage  $\phi X174$  DNA. *Nature (London)* **265**:687-695.
9. Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchison III, P. M. Slocombe, M. Smith, J. Drouin, T. Friedmann, and A. J. H. Smith. 1978. The nucleotide sequence of the DNA of  $\phi X174$ cs70 and the amino acid sequence of the proteins for which it codes, p. 659-669. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), *The single-stranded DNA phages*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
10. Shaw, D. C., J. E. Walker, F. D. 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. Smith, M., N. L. Brown, G. M. Air, B. G. Barrell, A. R. Coulson, C. A. Hutchison III, and F. Sanger. 1977. DNA sequence at the C termini of the overlapping genes *A* and *B* in bacteriophage  $\phi X174$ . *Nature (London)* **265**:702-705.
12. Tessman, E. S., and I. Tessman. 1978. The genes of the isometric phages and their functions, p. 9-29. *In* D. T. Denhardt, D. Dressler, and D. S. Ray (ed.), *The single-stranded DNA phages*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
13. Weisbeek, P. J., W. E. Borrias, S. A. Langeveld, P. D. Baas, and G. A. van Arkel. 1977. Bacteriophage  $\phi X174$ : gene *A* overlaps gene *B*. *Proc. Natl. Acad. Sci. U.S.A.* **74**:2504-2508.