# Presence of DNA, Encoding Parts of Bacteriophage Tail Fiber Genes, in the Chromosome of *Escherichia coli* K-12

ISOLDE RIEDE,\* MARIE-LUISE ESCHBACH, AND ULF HENNING

Max-Planck-Institut für Biologie, D-7400 Tübingen, Federal Republic of Germany

Received 19 February 1985/Accepted 31 May 1985

The classical T-even bacteriophages recognize host cells with their long tail fibers. Gene products 35, 36, and 37 constitute the distal moiety of these fibers. The free ends of the tail fibers, which are formed by the CO<sub>3</sub>H terminus of gene product 37, possess the host range determinants. It was found that 4 out of 10 different strains of Escherichia coli K-12 contained regions of chromosomal DNA which hybridized with a probe consisting of genes 35, 36, and 37 of the T-even phage K3. From one strain this homologous DNA, which was associated with an EcoRI fragment of about 5 kilobases, was cloned into plasmid pUC8. Two independently recovered hybrid plasmids had undergone a peculiar rearrangement which resulted in the loss of about 3 kilobases of cloned DNA and a duplication of both the vector and the remaining chromosomal DNA. The mechanisms causing this duplication-deletion may be related to that of transposases. The cloned DNA was capable of recombination with phage T4 gene 36 and a phage T2 gene 37 amber mutant. DNA sequencing revealed the existence of regions of identity between the cloned DNA and genes 36 and 37 of phage T2. In addition, after growth of a derivative of phage K3 on a strain harboring T2 DNA, it was found that this phage contained the same parts of the T2 tail fiber genes which had been recovered from the bacterial chromosome. There appears to be little doubt that the phage had picked up this DNA from the host. The possibility is considered that a repertoire of parts of genes 36 and 37 of various T-even-type phages is present in their hosts, allowing the former to change their host ranges.

The classical bacteriophages T2 and T4 and probably all T-even-type phages, recognize their host cell receptors with the free ends of their long tail fibers (1, 28, 35). The structural components of the T4 tail fibers are encoded by four linked genes, g34 to g37 (36). The proximal half fiber consists of a dimer of gene product (gp) 34. One copy of gp35 links this half to a dimer of gp36 which represents the proximal part of the distal half fiber. A dimer of gp37 completes the distal half fiber. The CO<sub>2</sub>H termini of the latter dimer are located at or near the free ends of the tail fibers (1). Several regions, including those responsible for the host range, of genes 37 of the very closely related phages T2, T4, and T6 are heterologous as shown by electron-microscopic-heteroduplex (8) and recombination analyses (1, 2). The three phages use different cellular receptors (7, 11, 35).

We have studied a set of 10 independently isolated T-eventype phages which use the Escherichia coli outer membrane protein OmpA as a receptor (21, 27). Several of these phages use different sites on this receptor protein (16, 27). DNA hybridization experiments have shown that the genes 37 of these phages are highly polymorphic: segments of T4 g37were found in some of the OmpA-specific phages but not in others. Segments of g37 of the OmpA-specific phage K3 cross-hybridized with DNA of most other such phages. However, these segments differed not only from one phage to another in size and location but also rather pronouncedly in the degree of homology (21). DNA sequence analyses performed since then have demonstrated that absence of cross-hybridization between areas of g37 of the OmpAspecific phages or of the classical T-even phages reflects major differences in the respective base and amino acid sequences of these phages (unpublished data).

#### MATERIALS AND METHODS

**Bacterial strains and phages.** Chromosomal DNA was isolated from the following *E. coli* K-12 strains: P400 (30), P400 *ompA2001* (16), KL 16-99 (9), DH1 (12), HB101 (3), JM103 (13),  $F^-Z^-M15$  (25), M3509 (32), UH100 (4), and UH201-3 (4). P400 *ompA2001* is the *ompA* mutant on which the host range mutant of phage K3 was isolated (see below). The T2 and T4 amber mutants used (gifts from E. D. Goldberg, E. Kellenberger, P. Tedesco, and W. B. Wood) are shown in Fig. 1 and referenced in the legend to Fig. 1. Cells were grown in L-broth (15) at 37°C.

Cloning of tail fiber genes and DNA sequencing. Chromosomal DNA was purified (17), and restriction enzyme digests were separated electrophoretically on 0.8% agarose gels. Transfer to nitrocellulose (31) and hybridization in the presence of 50% formamide under the conditions described by Wahl et al. (34) were performed as described previously (21). Initially, we almost always encountered difficulties in controlling the background radioactivity. We later performed prehybridizations and hybridizations in the presence of heparin (500  $\mu$ g/ml) instead of Denhardt solution and carrier DNA, as described by Singh and Jones (29); the superiority of the heparin method is shown in Fig. 2. The radioactive

The situation concerning receptor specificities and gene polymorphism is reminiscent of that of the immunoglobulins, and we had previously asked whether one phage might contain not only a complete g37 but also additional fragments of this gene, thus perhaps allowing the phage to change receptor specificities (27). So far there is no evidence for this. However, the data presented in this communication indicate that a repertoire of genes 36 and 37 or at least fragments thereof may be present in the chromosome of the bacterial host.

<sup>\*</sup> Corresponding author.



FIG. 1. Alignment of the T2 DNA in pTUT2C-18 with tail fiber genes. Top, Tail fiber genes of phage T4 and relevant mutant alleles (18, 19). The mutation sites 123 and 125 are in T2 (23), and their locations are not precisely known. Also, genes 37 differ in size in T4 and T2 (2). Middle, Part of the known restriction map of cloned T2 genes (21); bottom, cloned T2 DNA in pTUT2C-18. Horizontal arrows indicate the parts of the DNA which have been sequenced.

probe was prepared by nick translation (22) of the electrophoretically isolated HindII fragments of plasmid pTU K3-81 (21). These fragments carry genes 35, 36, and 37 from OmpA-specific phage K3 (21). Upon identification of the homologous chromosomal EcoRI fragment (about 5 kilobases [kb]), the digest was size fractionated by centrifugation into a sucrose gradient as previously described (21), and the appropriate fraction was ligated into plasmid pUC8 (33). Transformation into strain F<sup>-</sup>Z<sup>-</sup>M15 and colony hybridization (6) with the probe mentioned above yielded strains carrying plasmids pTUT2C-6 and pTUT2C-18. Marker rescue tests were done by placing 5 µl of serial dilutions of T2 or T4 amber mutants onto a lawn of cells in a soft-agar overlay. DNA sequences, using phages M13mp8 and M13mp9 (14), were determined by the dideoxy chain termination method (26). All DNA manipulations were performed essentially as described by Maniatis et al. (10). The enzymes were obtained from Boehringer Mannheim Biochemicals.

# RESULTS

The search for tail fiber genes in bacterial chromosomal DNA was provoked by the following accidental finding. From one of the OmpA-specific phages, K3, we had isolated host range mutants which were capable of infecting an ompA missense mutant (P400 ompA2001 resistant to the parent phage. We attempted to clone g37 from such mutants into plasmid pUC8 by using a probe consisting of a previously cloned DNA fragment which contained genes 36 and 37 from phage K3 (pTUK3-81; 21). Positive clones were found, and in one, as verified by DNA sequencing, at least part of g37 of the host range mutant was present. In another one, however, restriction enzyme analysis and DNA sequencing revealed that g36 and part of g37 were derived not from phage K3 but from phage T2. This result was reproducible. A second clone with the same T2 sequences was obtained from a K3 host range mutant isolated independently on the same host. We therefore began to search for the origin of this DNA. This T2 material could not be detected in any of the many clones of tail fiber genes previously isolated from phage K3 (21), an

indication that K3 does not normally carry T2 DNA. We therefore turned to the bacterial chromosomal DNA.

Cloning of parts of T2 tail fiber genes from the bacterial chromosome. Chromosomal DNA from E. coli strain P400 ompA2001 (the host used to isolate the host range mutants) was restricted with EcoRI. The probe described above was then hybridized to this DNA under stringent conditions. A positive signal was indeed found and was localized to a restriction fragment of about 5 kb (see Fig. 2). After size fractionation of the DNA fragments, the appropriate fraction was ligated into plasmid pUC8. Transformation into strain M15 yielded colonies which were capable of hybridizing with the same probe. We expected to find hybrid plasmids containing an insert of about 5 kb. Restriction analysis revealed, however, that in two independent isolates (pTUT2C-6 and pTUT2C-18) the same duplication-deletion events had occurred causing the loss of about 3 kb of the cloned DNA (Fig. 3). Apparently the hybrid plasmids generated in vitro were unstable in vivo and led to the rearrangement shown in Fig. 3. (Interestingly, the same rearranged plasmids were recovered from the K3 host range mutants containing the T2 DNA).

The tail fiber genes from phage K3, present in the probe used to obtain the two clones from chromosomal DNA, possess areas which are highly homologous to those of phage T2 and T4 (21). Strain  $F^{-}Z^{-}M15$ , harboring the hybrid plasmids, did not carry an amber suppressor. Therefore, marker rescue tests with amber mutants of these two phages were possible. The mutant alleles used, together with their respective locations, are shown in Fig. 1. Recombination between mutant phage and plasmid to produce wild-type progeny was unambiguous with both T4 amE1 (g36) and T2 am123 (g37). All other tests were negative. Thus, DNA



FIG. 2. Autoradiogram of DNA hybridizations. The restriction enzyme digests were separated electrophoretically in 0.8% agarose, transferred to nitrocellulose, and hybridized with the radioactive K3 DNA probe described in the text. Lanes: 1–5, hybridizations performed in Denhardt solution with carrier DNA; Lanes 6–8, hybridizations performed in the presence of heparin (see text). Lanes: 1, strain P400 *Eco*RI digest; 2, strain P400 *ompA2001 Eco*RI digest; 3, strain P400, having lost the T2 DNA (see text), *Eco*RI digest; 4, strain P400 *ompA2001 Eco*RI-*Hind*III digest; 5, same strain *Hind*III digest (which appears to be a partial one); 6, same strain *PsI* digest; 7, same strain *PvuI* digest; 8, strain P400 *Eco*RI digest. In lane 3, 2.5  $\mu$ g of DNA was applied to the gel; in all other cases, 0.5  $\mu$ g was used. Numbers at the left margin, size of marker fragments in kb.

which encoded at least parts of genes 36 and 37 and which was homologous in phages T2, T4, and K3 had been recovered.

It was found by DNA sequencing (see below) that in pTUT2C-18, the EcoRI site was hybrid; i.e., it was derived partly from such a restriction site in g37 of phage T2 and partly from this site in the plasmid. Therefore, EcoRI must have cleaved within the part of g37 present in the bacterial chromosome. We propagate our phages on strain P400, which does not reduce the extent of phage DNA modification, and this DNA cannot be cleaved at all with EcoRI. Therefore, the cloned DNA could not have originated from contaminating phage. To stress this point, we have cleaved the chromosomal DNA also with *Hind*III, *PvuI*, and *PstI*. All enzymes generated fragments hybridizing with the probe (Fig. 2), and none of them is able to digest DNA stemming from phage T2 DNA.

The T2 DNA in the bacterial chromosome is not entirely stable. The *Eco*RI fragment described above was found in chromosomal DNA of strain P400 *ompA2001*. This mutant was isolated in 1975, and the cells used for the present experiments came from a culture kept frozen since that time. The same *Eco*RI fragment which could hybridize with the tail fiber gene probe was found in the DNA from parental strain P400 stemming from a culture also frozen since 1975. It was absent, however, in the DNA from strain P400 which had been subcultured for an unknown number of generations (Fig. 2). Obviously, the chromosomal tail fiber gene insert may be lost, probably in toto.

DNA sequence analysis. The 1.5-kb SmaI-EcoRI fragment from pTUT2C-18 was transferred to phages M13mp8 and M13mp9 and sequenced for about 200 base pairs from both ends. The same operation was performed with the 0.45-kb Smal-HindIII fragment (Fig. 1). The sequences starting at the SmaI site were identical (in both directions) to the sequence known for this part of g36 of phage T2, and the sequence starting at the HindIII site was identical to the sequence known for this region of  $g_{35}$  of this phage (20). Sequence determination of g37 of phage T2 has recently been completed (K. Drexler, manuscript in preparation), and the sequence starting at the EcoRI site was also identical to the corresponding sequence (g37) of this phage. It should be pointed out that in pTUT2C-18 the *Hin*dIII site, as well as the EcoRI site, is hybrid; i.e., it is derived partly from the HindIII site in the T2 DNA and partly from that in the cloning area in pUC8. This site arose by the deletionduplication event described in the previous section (cf. Fig. 3). Therefore, more DNA (upstream from the HindIII site shown in Fig. 1) encoding phage tail fiber protein must have been originally present but subsequently lost from the cloned DNA.

Only a small portion of the *SmaI-Eco*RI fragment has thus far been sequenced. Therefore, we do not know whether all of it is identical with the T2 gene. Despite this uncertainty and for convenience, we have used the term T2 DNA for the cloned material.

### DISCUSSION

The presence of at least parts of phage tail fiber genes in the bacterial chromosome is not unique to strain P400. Using the probe from phage K3, we analyzed chromosomal DNA from nine additional *E. coli* strains (see above) digested with *Eco*RI. In three of these (DH1, HB101, and  $F^{-}Z^{-}M15$ ), sequences homologous to the HK3 DNA were clearly present. Also, T-even phage tail fiber genes apparently



FIG. 3. Expected and recovered structures of plasmid pTUT2C-18. A chromosomal EcoRI fragment of about 5 kb was ligated into pUC8 (thin line, vector DNA), and a plasmid (A) was expected. The cloning area of pUC8 is indicated by several restriction sites between the EcoRI and HindIII sites. B, The plasmid actually found. It lacked about 3-kb of cloned material (indicated by  $\Delta$  in panel A), including most of the cloning region of pUC8, and it represented a duplication of vector and remaining insert DNA.

vagabonding are not without precedence. Computer analysis of the DNA sequence of a  $\lambda$  phage has shown that this phage contains a stretch of nucleotides possessing two rather unexpected open reading frames (5). One which codes for 318 amino acid residues exhibits striking homology with the CO<sub>2</sub>H-terminal end of gp37 of phage T4, and it is this area which should contain the host range property. The other codes for 195 residues and shares homology with the NH<sub>2</sub>terminal end of gp38 of phage T4 (g38, located immediately downstream from g37, encodes a protein which is required for efficient dimerization of gp37 without being present in the tail fiber).

The chromosomal-DNA insert appears to be unstable. It can be lost from the chromosome, and, in addition, it was found on the plasmid only in the partially deleted state shown in Fig. 3. Because of the latter situation and because EcoRI cleaved within g37, the flanking sequences could not be recovered. We therefore do not know whether the chromosomal insert may have the character of a transposon (see below) and how much of g37 is present in the chromosome. The region encoding the host range area of gp37 should be situated further downstream from the EcoRI site in question. Thus, unfortunately we do not know whether the chromosomal tail fiber gene contains the host range property of phage T2.

The simplest explanation for the formation of pTUT2C-18 would be duplication of the original plasmid followed by a recombination event between the *Hin*dIII sites in the cloned DNA and those in the vector, thus eliminating DNA as indicated in Fig. 3. It is also possible that recombination occurred at both *Hin*dIII sites followed by deletion of the *Hin*dIII-*Hin*dIII segment. Other mechanisms are conceivable; whatever they may be, we have never observed such a rearrangement in rather numerous cloning experiments involving other bacterial or phage DNA. It therefore is tempting to assume that plasmids pTUT2C-18 or pTUT2C-6 arose by a process similar to that of transposase action and utilized the *Hin*dIII recognition sequence.

The presence of the T2 DNA in the host range mutants of phage K3 probably has nothing to do with the altered host range of these mutants, which remains specific for the OmpA protein. Also, this protein is not involved in any way in the process of T2 infection. We have performed immunoelectron microscopy with the host range mutant by using antiserum against phage K3. The pattern of immunoglobulin G (IgG) bound to the tail fibers was that characteristic of K3 and not that of T2 (27). Thus, it is almost certain that the host range mutation and the presence of T2 DNA in the phage are not causally related but are coincidental. The parent K3 phage for these experiments had been propagated, from single plaque isolates, on strain P400 which was later found to possess the T2 genes, and the host range mutants were obtained on strain P400 *ompA2001* also harboring the T2 material. Isolation of the host range mutants and cloning of the K3 genes have been performed by different researchers, and the cloner used as host strain P400 not possessing the T2 genes. We believe that there is little doubt that the host range mutants picked up these genes from the host, although these genes do not appear to be of any use to the phages.

In summary, at least parts of the genes encoding distal tail fiber polypeptides of T-even phages could be present in bacterial chromosomal and in phage  $\lambda$  DNA. Also, such gene fragments from phage T2 were found in the DNA of host range mutants of the OmpA-specific phage K3. The possibility exists that a phage can change its host range by picking up a corresponding part of the relevant gene from its host. Furthermore, T-even-type phages, although closely related, often more or less strongly exclude each other upon mixed infection, thus permitting only very limited recombination or no recombination to occur (24). This genetic isolation mechanism could be overcome when phage DNA was available on the bacterial chromosome or on plasmids. Finally, it is possible that the presence of phage tail fiber genes in bacterial chromosomes may at least contribute to the astonishing heterogeneity which we have found to prevail in genes 36 and 37 of various T-even-type phages (20, 21). For such chromosomal genes there is no selection for the production of functional fibers. Multiple mutations can accumulate, and at least parts of the altered genes could be compatible with the tail fiber structure, whereas a single mutation may not.

# ACKNOWLEDGMENTS

We thank H. Schwarz for performing the immunoelectron microscopy and Sheila MacIntyre for critically reading the manuscript.

The financial support of the Fonds der Chemischen Industrie and gifts of ampicillin by Bayer AG are gratefully appreciated.

#### LITERATURE CITED

- 1. Beckendorf, S. K. 1973. Structure of the distal half of the bacteriophage T4 tail fiber. J. Mol. Biol. 73:37-53.
- Beckendorf, S. K., J. S. Kim, and I. Lielausis. 1973. Structure of bacteriophage T4 genes 37 and 38. J. Mol. Biol. 73:17–35.
- 3. Boyer, H. W., and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41:459–472.
- Bremer, E., S. T. Cole, I. Hindennach, U. Henning, E. Beck, C. Kurz, and H. Schaller. 1982. Export of a protein into the outer membrane of *Escherichia coli* K-12. Eur. J. Biochem. 122:223-231.
- George, D. G., L.-S. L. Yeh, and W. C. Barker. 1983. Unexpected relationships between bacteriophage λ hypothetical proteins and bacteriophage T4 tail-fiber proteins. Biochem. Biophys. Res. Commun. 115:1061-1068.
- Hanahan, D., and M. Meselson. 1980. Plasmid screening at high colony density. Gene 10:63–67.
- 7. Hantke, K. 1978. Major outer membrane proteins of *E. coli* K12 serve as receptors for the phages T2 (protein 1a) and 434 (protein 1b). Mol. Gen. Genet. 164:131-135.
- 8. Kim, J. S., and N. Davidson. 1974. Electron microscope heteroduplex study of sequence relations of T2, T4, and T6

bacteriophage DNAs. Virology 57:93-111.

- Low, B. 1968. Formation of merodiploids in matings with a class of rec<sup>-</sup> recipient strains of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA 60:160-167.
- 10. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 11. Manning, P. A., and P. Reeves. 1978. Outer membrane proteins of *Escherichia coli* K-12: isolation of a common receptor protein for bacteriophage T6 and colicin K. Mol. Gen. Genet. 158:279-286.
- 12. Meselson, M., and R. Yuan. 1968. DNA restriction enzyme from E. coli. Nature (London) 217:1110-1114.
- 13. Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun sequencing. Nucleic Acids Res. 9:309-321.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene 19:269-276.
- Miller, J. H. (ed.). 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Morona, R., M. Klose, and U. Henning. 1984. Escherichia coli K-12 outer membrane protein (OmpA) as a bacteriophage receptor: analysis of mutant genes expressing altered proteins. J. Bacteriol. 159:570-578.
- 17. Nakamura, K., R. M. Pirtle, and M. Inouye. 1979. Homology of the gene coding for outer membrane lipoprotein within various gram-negative bacteria. J. Bacteriol. 137:595–604.
- Oliver, D. B., and R. A. Crowther. 1981. DNA sequence of the tail fibre genes 36 and 37 of bacteriophage T4. J. Mol. Biol. 153:545-568.
- Revel, H. R. 1981. Molecular cloning of the T4 genome: organization and expression of the tail fiber gene cluster 34-38. Mol. Gen. Genet. 182:445-455.
- Riede, I., K. Drexler, and M.-L. Eschbach. 1985. The nucleotide sequences of the tail fiber gene 36 of bacteriophage T2 and of genes 36 of the T-even type *Escherichia coli* phages K3 and Ox2. Nucleic Acids Res. 13:605-616.
- Riede, I., M.-L. Eschbach, and U. Henning. 1984. DNA sequence heterogeneity in the genes of T-even type *Escherichia coli* phages encoding the receptor recognizing protein of the long tail fibers. Mol. Gen. Genet. 195:144–152.
- Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113:237-251.
- Russel, R. L. 1974. Comparative genetics of the T-even bacteriophages. Genetics 78:967-988.
- Russel, R. L. 1974. Partial exclusion between T-even bacteriophages; an incipient genetic isolation mechanism. Genetics 78:989-1014.
- Rüther, U., M. Koenen, K. Otto, and B. Müller-Hill. 1981. pUR222, a vector for cloning and rapid chemical sequencing of DNA. Nucleic Acids Res. 9:4087–4098.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74:5463-5467.
- Schwarz, H., I. Riede, I. Sonntag, and U. Henning. 1983. Degrees of relatedness of T-even type *E. coli* phages using different or the same receptors and topology of serologically cross-reacting sites. EMBO J. 2:375–380.
- 28. Simon, L. D., and T. D. Anderson. 1967. The infection of *Escherichia coli* by T2 and T4 bacteriophages as seen in the electron microscope. Virology 32:279–297.
- Singh, L., and K. W. Jones. 1984. The use of heparin as a simple cost-effective means of controlling background in nucleic acid hybridization procedures. Nucleic Acids Res. 2:5627-5638.
- Skurray, R. A., R. E. W. Hancock, and P. Reeves. 1974. Conmutants: class of mutants in *Escherichia coli* K-12 lacking a major cell wall protein and defective in conjugation and adsorption of a bacteriophage. J. Bacteriol. 119:726–735.
- Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

- 32. Timmis, K. N., I. Andrés, and M. Achtman. 1978. Fertility repression of F-like conjugative plasmids: physical mapping of the R6-5 finO and finB cistrons and identification of the finO protein. Proc. Natl. Acad. Sci. U.S.A. 75:5836-5840. 33. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13
- mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259-260.
- 34. Wahl, G. M., M. Stern, and G. R. Starck. 1979. Efficient

transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization by using dextransulfate. Proc. Natl. Acad. Sci. U.S.A. 76:3683-3687.

- 35. Wilson, J. H., R. B. Luftig, and W. B. Wood. 1970. Interaction of bacteriophage T4 tail fiber components with a lipopolysaccharide fraction from *Escherichia coli*. J. Mol. Biol. **51**:423-434. 36. Wood, W. B., and H. R. Revel. 1976. The genome of bacterio-
- phage T4. Bacteriol. Rev. 40:847-868.