Journal of Supramolecular Structure 2:329–336 (1974)

PROTEIN FUSION DURING THE ASSEMBLY OF PHAGE LAMBDA HEADS

Roger W. Hendrix* and Sherwood R. Casjens†

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

INTRODUCTION

The assembly of many large bacteriophages, including λ and other tailed phages, is qualitatively more complex than the assembly of structurally simple viruses. If a simple virus, such as tobacco mosaic virus, is chemically disassembled, the resulting parts can spontaneously reassemble under appropriate conditions to yield infectious virions (1). While assembly of the λ capsid undoubtedly includes this type of simple self assembly, higher order processes are involved as well. Correct λ assembly requires, in addition, direct participation of a host cell function (2, 3), the presence during assembly of phage-coded proteins which are absent from the mature virion (4), and chemical processing of structural proteins of the virion (4, 5). Presumably these processes allow a higher level of structural complexity and better control over the steps of assembly than would otherwise be possible. In this article we describe a new type of protein processing reaction, protein fusion, which occurs during λ head assembly, prior to the time that DNA enters the head.

*Current address: Department of Biophysics and Microbiology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

†Current address: Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

MATERIALS AND METHODS

A previous report of this work contains a description of the methods and materials used (16).

RESULTS

Proteins of λ Heads

The phage genes involved in λ head assembly are well characterized genetically, and the proteins encoded by all but two of them have been identified (5). Figure 1 is a physical map of the region of the λ genome containing the genes involved in head assembly. The bars below the map represent the calculated sizes of the genes, based on the known sizes of the corresponding proteins. Clearly, nearly all of the coding capacity of the head region of the genome must be taken up in coding for the proteins that have been identified as products of the various head genes. We were consequently puzzled by the fact that heads contain three additional species of phage-specific protein. Figure 2 is an SDS-polyacrylamide gel pattern of highly purified λ heads. Three of the bands, pB, pE, and pD, correspond to the proteins coded by the B, E, and D genes. In addition there are three bands, h3, X1, and X2, with molecular weights that are different from those of any of the identified head proteins.

One attractive explanation for the presence of h3, X1, and X2 in heads was that they arose by chemical processing of some of the primary protein products of the head genes. To test this possibility we made tryptic fingerprints of the phage head proteins, derived either from the virion or from phage-infected lysates. This technique entails isolating individual ^{35}S-labeled proteins as bands on an SDS-polyacrylamide gel, cutting out the bands and eluting the protein, digesting with trypsin, and displaying the radioactive peptides by electrophoresis and chromatography on thin layer plates.

Figure 3 shows such tryptic fingerprints of proteins pE, pC, X1, and X2. pE is the major structural component of the heads and is present in about 420 copies per head, pC is absent from mature heads, and X1 and X2 are minor head components present in about 6 copies each. The fingerprints of X1 and X2 are nearly identical, differing only in 3 spots. Consequently, we conclude that the X1 and X2 proteins are related to each other by proteolytic cleavage.The minor differences between the fingerprints presumably represent the results of the cleavage, which reduces the molecular weight of X1 from 31,000 to 29,000.

X1 and X2 Are Fusion Products

A further comparison of the 4 fingerprints in Fig. 3 reveals an unusual relationship between them: 5 of the spots in the X1 fingerprint have mobilities that are identical to spots in the pC fingerprint and 5 other spots of X1 have mobilities that are identical to spots in the pE fingerprint. Thus X1 and X2 may contain amino acid sequences derived from both pE and pC. There are some spots in both pC and pE which are absent from X1, so it appears that X1 contains a subset of the amino acid sequences of pC and a subset of those of pE. Figure 4 is a schematic representation of the fingerprints in Fig. 3. The numbering is meant to indicate the correspondences between the spots. Since spot 2X has about the same mobility as spots from both pE and pC, it is not clear whether it

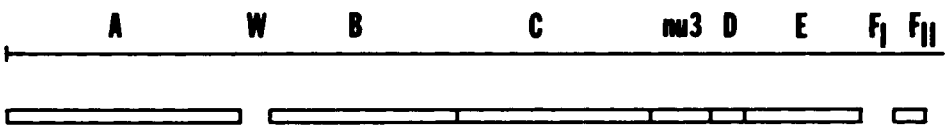

Fig. 1. Physical map of the head region of the λ chromosome.

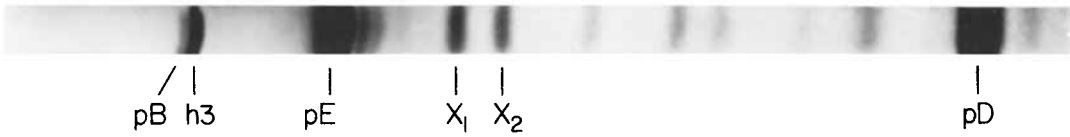

Fig. 2. Autoradiogram of sodium dodecyl sulfate polyacrylamide gel of ^{35}S-labeled λ heads.

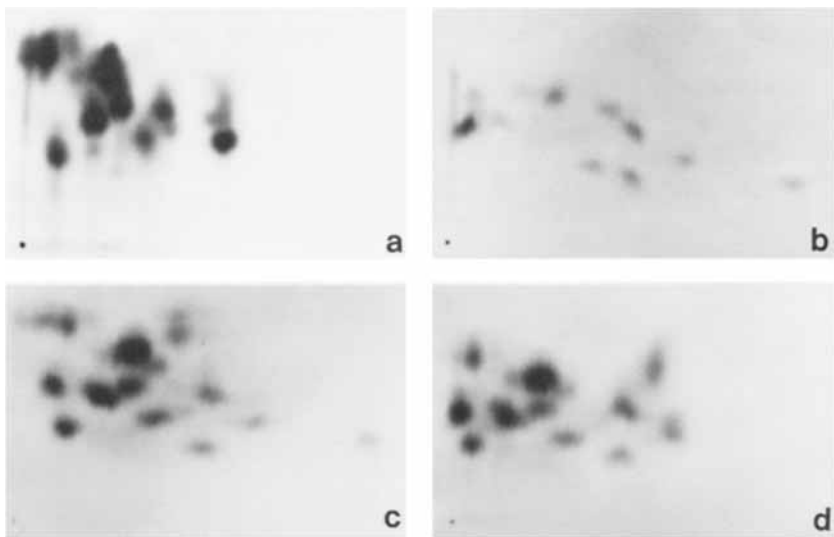

Fig. 3. Autoradiograms of tryptic fingerprints of ^{35}S-labeled λ proteins. a) pE; b) pC; c) X1; d) X2.

corresponds to one or both of them. Spot 18, the only spot in the X1 fingerprint that is not shown in either the pE or the pC fingerprints, may in fact correspond to a pE spot. Although no spot in position 18 is seen in the pE fingerprint in Fig. 3a, such a spot is frequently seen in other pE fingerprints, e.g., the pE fingerprints in Fig. 5a and b both have faint spots in that position.

To prove that the spots in the X1 and X2 fingerprints really represent amino acid sequences derived from pE and pC, we sought evidence that a mutational change in the amino acid sequence of pE would produce a corresponding change in the amino acid sequences of X1 and X2. If a λ mutant carrying an amber mutation in the E gene is grown in a host carrying an amber suppressor, then a mutant pE protein will be synthesized with an amino acid substitution at the position corresponding to the amber codon. The finger-

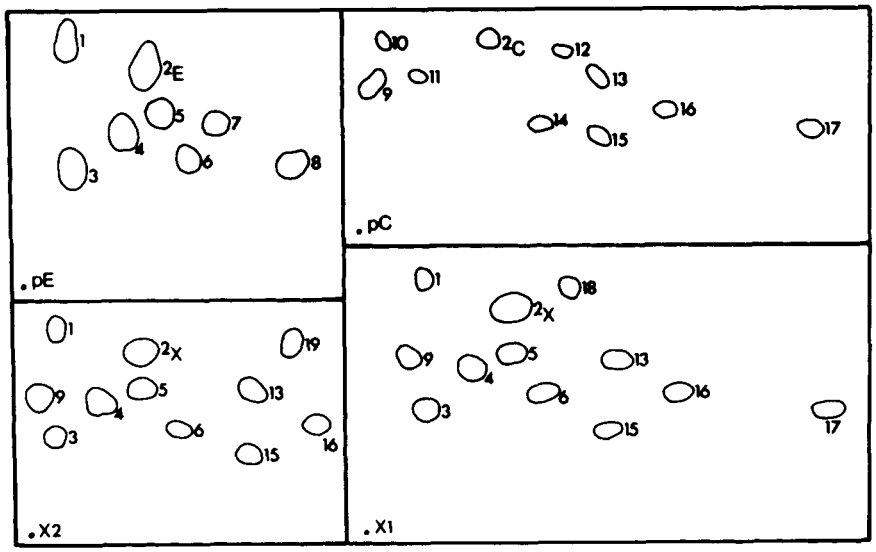

Fig. 4. Schematic drawing of the fingerprints in Fig. 3. Corresponding numbers indicate spots of identical mobilities.

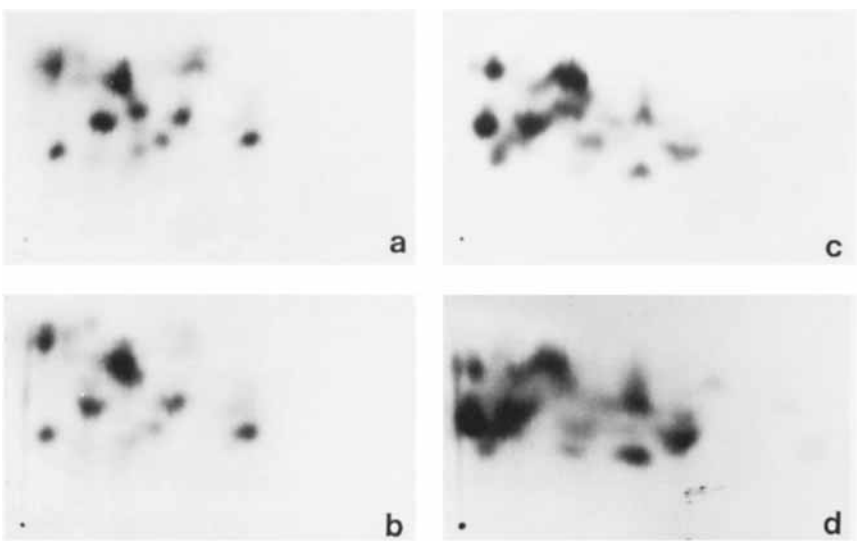

Fig. 5. Autoradiograms of tryptic fingerprints of ^{35}S-labeled λ proteins. a) pE from a wild-type infection; b) pE from an infection of sup$^+$F cells by λ Eam4; c) X2 from a wild-type infection; d) X2 from an infection of sup$^+$F cells by λ Eam4.

print of such a protein is shown in Fig. 5. Figure 5a is a tryptic fingerprint of wild-type pE, and Fig. 5b is a fingerprint of an altered pE produced after infection by λ Eam4 of a SupF host, which inserts tyrosine at the amber codon (6). The two fingerprints are identical with the exception of a change in the position of spot 6 (arrow) which appears to have run slightly faster in the chromatographic direction in the case of the mutant pE. Thus spot 6 must be the peptide encoded by the part of the E gene containing the amber mutation. In a similar experiment using an amber suppressor that inserts serine, spot 6 did not change mobility, suggesting that the wild-type amino acid at this position is either serine or another amino acid with similar electrophoretic and chromatographic properties.

Since spot 6 is one of the pE spots that is present in X1 and X2, the X1 and X2 fingerprints should show the same change as the pE fingerprint when derived from the amber infection of the tyrosine-inserting cell. Figure 5c shows a fingerprint of wild-type X2 and Fig. 5d shows a fingerprint of X2 purified from the same lysate as the mutant pE in Fig. 5b. Again the two fingerprints are identical with the exception of spot 6. Consequently, X1 and X2 must in fact contain sequences from pE and not simply possess fortuitously similar fingerprints. Figure 6 is a schematic representation of the postulated relationships between pE, pC, X1, and X2.

Origin of Other Proteins in Heads

In addition to X1 and X2, SDS gels of λ heads show several other bands which cannot be primary gene products. Experiments to be described elsewhere (4) show that these also result from chemical changes to primary products of head genes. Tryptic fingerprints show that h3, which behaves like a processed protein in pulse-chase experiments (5), is a proteolytic cleavage product of pB (4). Recently, we have found that the remaining minor bands that are visible in Fig. 2 are in fact not present in the phage head. Rather, they are degradation products of pE which are produced by a small amount of nonenzymatic cleavage of pE during boiling in an SDS-containing buffer prior to electrophoresis (Hendrix, unpublished). It is now possible, therefore, to account for all of the bands on an SDS gel of λ heads and consequently to completely describe the protein composition of λ heads.

DISCUSSION

We describe in this paper a new type of protein processing reaction which occurs during the assembly of phage λ heads. Parts of two separate phage-coded proteins, pC and pE, become linked to each other during the process of assembly. There are two alternative fusion products (X1 and X2) produced in approximately equal amounts, and these are related to each other by proteolytic cleavage. Both X1 and X2 are smaller than either pC or pE, and they are presumably proteins with an entirely different set of properties than either of the precursor proteins.

What is the structure of the fusion products? Analysis of the relative intensities of the tryptic spots of X1 and X2 shows that the sequence contributed by pC and by pE are present in equimolar amounts (16). This fact is most simply interpreted as meaning that each fusion product (X1 or X2) consists of one copy of a fragment of pC linked to one copy of a fragment of pE. The nature of the linkage is still unclear, but it is very

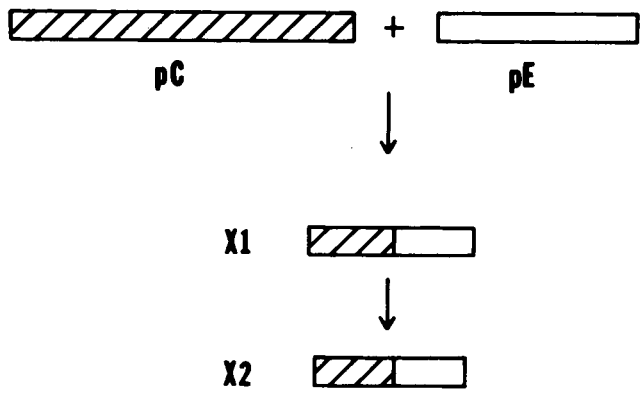

Fig. 6. Schematic representation of the fusion reaction. The lengths of the bars are proportional to the apparent molecular weights of the proteins.

likely a covalent bond, since it is resistant to prolonged boiling in a buffer containing 2% sodium dodecylsulfate and 10% 2-mercaptoethanol, or a buffer containing 6 M guanidine HCl and 10% 2-mercaptoethanol. Further, it is unlikely to be a disulfide bond, because of its resistance to 2-mercaptoethanol. The bond might well be a peptide bond, joining the pC and pE sequences into a continuous polypeptide chain. Alternatively, it could be a crosslinking bond between amino acid side chains of the two components. Experiments are currently underway to test these possibilities.

The fusion of pC and pE proposed here is consistent with several other observations. Unprocessed pC is absent from mature heads but can be found in infected lysates (5). In lysates pC disappears in a pulse-chase experiment and its disappearance requires the presence of a functional pE (5). The protein composition of the head-related particle petit λ is also consistent with the fusion scheme. Petit λ is a hollow protein shell which has been shown to be a precursor to heads (7, 8). Petit λ particles derived from a wild-type infection or from most mutant infections contain X1 and X2 and no pC. However, the petit λ produced after infection by λ Cam$^-$ is missing X1 and X2, as is expected if pC is a precursor to the fusion products. X1 and X2 are also absent from the petit λ produced when λ infects a host carrying the groE mutation. Instead, these groE$^-$ petit λ particles contain unprocessed pC (4). Apparently the fusion reaction is blocked in the groE host.

When does the fusion reaction occur? Since the groE$^-$ petit λ contains pC and so must have a site where it binds, it seems reasonable to assume that nascent wild-type petit λ also has a binding site for unprocessed pC. Thus the fusion of pC and pE would occur after both proteins are assembled into a nascent petit λ structure. Recently, Kaiser et al. (7) have shown that the fusion reaction must occur prior to the packaging of the DNA. Consequently our picture of the place of the fusion reaction in the assembly process is the following: pE is assembled into a structure (petit λ) containing about 420 pE monomers. Subsequently, about 12 copies of pC join the structure and each pC monomer reacts with a pE monomer to produce either X1 or X2. In the final series of steps, the DNA enters the head (7, 8), the head expands to normal size, a second major capsid protein, pD, is added (9), the DNA is cut to genome size (10), and the heads are prepared for tail attachment (11).

It is of interest to know what sort of mechanism is involved in the production of X1 and X2, which have MWs of 31,000 and 29,000, respectively, from the fusion of pC, (MW 64,000) and pE (MW 38,000). Possibly pC and pE are first cleaved and two of the resulting fragments are then fused to produce X1 or X2. Alternatively, pC and pE might first be fused to produce a large intermediate, which would be subsequently cleaved. Or pC and pE might "recombine" by transpeptidation to yield X1 or X2 in a single step. Recent evidence (4) suggests that the second alternative, fusion followed by cleavage, may be the correct one. The SDS gel patterns of groE$^-$ petit λ show, in addition to a strong band of unprocessed pC, a faint band at about 100,000 daltons. This is the molecular weight that would be expected for a protein composed of intact pC joined to intact pE. If tryptic fingerprinting experiments bear out this structure for the 100,000 dalton protein, then it will seem probable that X1 and X2 are produced by cleavage of that protein. The reason for the production of both X1 and X2 rather than a single fusion product is obscure.

Various functional roles have been suggested for the protein processing steps that occur during phage assembly, but in no case is the function known with certainty. Presumably processing reactions tend to make the assembly irreversible, and this may be an important part of their biological function. Processing reactions might also exert control over the order of assembly, e.g., by creating a binding site in situ for the next protein to be assembled. Processing reactions may also be involved in DNA packaging. In T4 assembly some protein cleavages occur concomitantly with DNA condensation (12). However, since the pC-pE fusion occurs prior to DNA packaging, it is presumably not directly involved in this function.

A striking feature of the pC-pE fusion is that it occurs about 12 times per head structure. Since the icosahedral head has 12 "corners" (centers of five-fold symmetry), we speculate that the reaction might occur once on each corner. Possibly the fusion alters the geometry or properties of the corners. We have found that petit λ in which the fusion is blocked has a different sedimentation rate than petit λ containing the fusion products (4). This change could be caused, for example, by a stiffening of the corners, although we cannot yet rule out the possibility that the change in sedimentation rate might instead be due to structural changes unrelated to the fusion.

The fusion of pC with pE appears to be a new type of protein processing reaction. Although there are several examples of covalent bond formation between proteins, these generally involve either bonds between identical or nearly identical subunits (e.g., collagen and fibrin, as reported in this issue) or interchain disulfide bonds (e.g., antibodies). In E. coli the only documented fusion-type reaction involves the fusion of fragments of β-galactosidase to give a wild-type enzyme (13). The reaction described here for λ involves the products of two distinct genes with entirely different properties which fuse and ultimately produce proteins that are smaller than either of them. We have read with interest a recent report (14) in which tryptic fingerprints are used to show sequence relationships among the four nonhistone proteins of polyoma virions. One possible interpretation of these results is that polyoma protein P3 is a fusion product. Another recent report (15) which can be interpreted in terms of protein fusion describes an increase in apparent molecular weight of one subunit of the E. coli RNA polymerase upon infection by phage T3. Thus it seems possible that protein fusion reactions may be found to be important in a number of biological systems.

ACKNOWLEDGMENTS

We thank Drs. Ron Davis, Dale Kaiser, Paul Schedl, and Mike Syvanen for useful discussion. This work was supported by NIH grant #A104509 to Dale Kaiser and was carried out during the tenures of a Damon Runyan Postdoctoral Fellowship to R. Hendrix and an NIH predoctoral training grant to S. Casjens.

REFERENCES

1. Fraenkel-Conrat, H., and Williams, R. C., Proc. Nat. Acad. Sci. U.S. 41:690 (1955).
2. Georgopoulos, C., Hendrix, R., Casjens, S., and Kaiser, D., J. Mol. Biol. 76:45 (1973).
3. Sternberg, N., J. Mol. Biol. 76:25 (1973).
4. Hendrix, R., and Casjens, S. (in preparation).
5. Murialdo, H., and Siminovitch, L., Virology 48:785 (1972).
6. Smith, J., Abelson, J., Clark, B., Goodman, H., and Brenner, S., Cold Spring Harbor Symp. Quant. Biol. 31:479 (1966).
7. Kaiser, A. D., Syvanen, J. M., and Masuda, T., J. Supramol. Struct. 2:318 (1974).
8. Hohn, B., Wurtz, M., Klein, B., Lustig, A., and Hohn, T., J. Supramol. Struct. 2:302 (1974).
9. Casjens, S., Hohn, T., and Kaiser, A. D., Virology 42:496 (1970).
10. Mackinlay, A., and Kaiser, A. D., J. Mol. Biol. 39:679 (1969).
11. Casjens, S., Hohn, T., and Kaiser, D., J. Mol. Biol. 64:551 (1972).
12. Laemmli, U. K., and Favre, M., J. Mol. Biol. 80:575 (1973).
13. Apte, B., and Zipser, D., Proc. Nat. Acad. Sci. U.S. 70:2969 (1973).
14. Friedmann, T., Proc. Nat. Acad. Sci. U.S. 71:257 (1974).
15. Dharmgrongarfama, B., Mahadik, S. P., and Srinivasan, P. R., Proc. Nat. Acad. Sci. U.S. 70:2845 (1973).
16. Hendrix, R. W., and Casjens, S. R., Proc. Nat. Acad. Sci. U.S. 71:1451 (1974).