Bacteriophage f1 Gene II and X Proteins

ISOLATION AND CHARACTERIZATION OF THE PRODUCTS OF TWO OVERLAPPING GENES*

T. S. Benedict Yen§ and Robert E. Webster‡

From the Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710

We have isolated and characterized the 2 major proteins of a dense complex which accumulate in Escherichia coli cells infected with bacteriophage f1 under conditions where the phage gene V protein is inactive (Webster, R. E., and Rementer, M. (1980) J. Mol. Biol. 139, 393-405). The amino acid composition and NH₂- and COOH-terminal sequences of the larger polypeptide (estimated molecular weight of 46,000) correspond to those predicted from the DNA sequence for the f1 gene II protein. The other polypeptide (estimated molecular weight of 14,000) has the amino acid composition and COOH-terminal sequence predicted for the f1 X protein, which previously had been found only as a product of an in vitro transcription-translation reaction. The X protein contains N-formylmethionine, cross-reacts with antibodies against gene II protein, and is present in wild type f1-infected bacteria. Thus, X protein is the product of f1 gene X (10), which is contained entirely in, and translated in phase with, gene II.

The genome of the male-specific filamentous coliphage (f1, fd, or M13) is a circular single-stranded DNA molecule of known sequence† (1, 2). Studies with conditional lethal mutants together with the knowledge of the DNA sequence have defined 9 genes (3). The products of 2 of these genes have been shown to be necessary for the replication of the phage DNA. Upon infection, host enzymes convert the incoming single-stranded DNA to a supercoiled circular double-stranded form, termed replicative form I DNA. The product of the phage gene II is then necessary for the synthesis of daughter double-stranded molecules. One of its major functions is the cleavage of the parental strand of the replicative form I DNA at a specific site (4). When sufficient amounts of the product of gene V have accumulated, replication becomes asymmetric and progeny single-stranded DNA accumulates (5, 6). This switchover is thought to result from the inability of gene V protein to bind the newly synthesized single-stranded DNA and form an elongated gene V protein-phony DNA complex (7-10). The synthesis of single-stranded DNA requires active gene II protein as well as the gene V protein (11).

The products of phage genes III, VI, VII, VIII, and IX are found in the mature virion (3, 12, 13), which is formed by extrusion of the DNA through the cell envelope without lysis or death of the host (14). Genetic studies have indicated that the products of 2 additional genes (I and IV) are necessary for phage production, but no function has been assigned to these proteins, nor have they been isolated (3).

When double-stranded DNA of f1 or M13 is used to direct protein synthesis in vitro in a coupled transcription-translation system, the polypeptides made can be identified as the products of genes I, II, III, IV, V, and VIII (15, 16). Another protein, with an estimated molecular weight of 12,000, is also synthesized and has been named the X protein (16, 17). Use of various restriction fragments to direct the in vitro synthesis system has shown that this protein is coded by the 3' portion of gene II DNA (17-19). Since DNA of R21, an amber mutant in gene II, fails to program the synthesis of X protein (17), it must be translated in phase with gene II protein. Hence, X protein should be identical with the carboxyl quarter of gene II protein. In fact, examination of the DNA sequence reveals a possible ribosome binding site around codons 296-298 of gene II, right before a potential initiating methionine at codon 300 (Fig. 1) (20). This should give rise to an 111-residue polypeptide of molecular weight 12,681, in good agreement with the estimated value for X protein. Several groups have, therefore, proposed the existence of a 10th gene that codes for X protein and is contained entirely within gene II (3).

Until recently, however, no evidence could be found for the existence of X protein in vivo, even by radiolabeling in UV-irradiated cells (21) or in minicells (22). In 1980, Webster and Rementer (23) reported the isolation of a dense complex from f1-infected bacteria under conditions where gene V protein is inactive. Analysis of the dense complex by SDS-polyacrylamide gel electrophoresis showed that it comprised 5 major polypeptides. Three of these migrated at the positions expected for gene II protein, gene IV protein, and the X protein. The identification of the gene II protein was supported by one-dimensional chymotryptic peptide mapping.

In this paper, we describe the isolation and characterization of the gene II and X proteins from the dense complex. Amino acid compositions and end group analyses confirm the sequence of these proteins predicted from the DNA sequence and show the absence of major post-translational processing. We also show that the X protein is synthesized in wild type f1-infected bacteria and, therefore, is the product of a 10th gene, which overlaps gene II.

---

MATERIALS AND METHODS

Materials—[131]-I-protein A of Staphylococcus aureus (100 mCi/µg), L-[4,5-3H]lysine (60 Ci/mmol), L-[4,5-14C]lysine (>300 mCi/mmol), and L-[35S]cysteine (600 Ci/mmol) were obtained from New England Nuclear, while L-[methyl-3H]methionine (15 Ci/mmol) was from ICN. Pierce supplied the trifluoroacetic acid, phenylisothiocyanate, and dansyl chloride; Calbiochem supplied pronase and BSA; and Sigma

---

*This work was supported by Public Health Service Grant GM 3067. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†Recipient of Medical Scientist Training Grant GM 0717-06.

‡Recipient of an American Cancer Society Faculty Research Award. To whom correspondence should be addressed.

1 D. F. Hill and G. B. Petersen, personal communication.

---

The abbreviations used are: dansyl, 5-dimethylaminonaphthalene-1-sulfonyl; BSA, bovine serum albumin; IgG, immunoglobulin G; SDS, sodium dodecyl sulfate.
supplied carboxypeptidases A and B, leucine aminopeptidase, rabbit serum albumin, N-acetyl-L-methionine, N-formyl-L-methionine, and L-threonyl-L-asparagine. Guanidinium chloride was from Heico. Polyamide plates and nitrocellulose sheets (45-μm pore) were manufactured by Schleicher and Schuell. Fotodyne made theFixion 50× 8 chromatography plates. Lin et al. (13) have given the sources of all other materials used.

Barbara, Bacteriophage, and Media—The Escherichia coli strains used were K38 (sup') (24) and DS410/pMC7/F104 (minA, minB), a mimic-producing strain obtained from D. Stueger, Duke University. Bacteriophage fl strains used were wild type, R111 (temperature-sensitive mutation in gene V), and R16 (amber mutation in gene V) (24). Tryptone-yeast medium, containing 10 g of tryptone, 1 g of yeast extract, and 8 g of NaCl/liter (pH 7.0), was used for growing bacteria and phage. For radioabsorbing, the MTPA salts medium of Viñuela et al. (25) was used, but supplemented with 0.01 M glucose, 10 μg/ml of thiamine, and 1 mM concentration of each amino acid except the one used for labeling.

Isolation of Dense Complex—K38 cells were grown in tryptone-yeast medium at 37 °C to a density of 2×10^9 cells/ml, infected with R16 at a multiplicity of infection of 20, and shaken at 42 °C. After 45 min, the cells were cooled to 4 °C and harvested by centrifugation at 14,000 × g for 8 min. The washed crude membrane fraction from spheroplasts was isolated by the method of Osborne et al. (26). The material from this fraction was resuspended in 25% (w/v) sucrose and 5 mM EDTA (pH 8.0) by extrusion through a 23-gauge needle and subjected to sucrose gradient centrifugation as described by Osborn et al. (26) but on a modified step gradient containing equal volumes of 55%, 52%, 50%, 45%, 40%, and 35% (w/v) sucrose in 5 mM EDTA. The band of interest was visually localized, removed with a syringe, and concentrated by centrifuging at 250,000 × g for 2 h after dilution with 3 volumes of H_2O. The pellets of dense complex were stored at −20 °C.

The same procedure was used for obtaining radiolabeled dense complex, but the cells were grown in MTPA medium minus the amino acid to be used for labeling, and 0.05–0.6 mM of the radioactive amino acid was added to 200 ml of culture at 25 min after infection. The sucrose gradients were collected from the bottom, and the radioactive peaks were located by scintillation counting.

Purification of Proteins from the Dense Complex—Dense complex pellets were washed with 0.2 M NaCl and 0.01 M Tris-HCl (pH 8.0), dissolved in 7.35 M guanidinium chloride, 0.05 M β-mercaptoethanol, and 0.02 M Tris-HCl (pH 8.2), and applied to a Sephacyr S-200 column (1.2 cm diameter × 80 cm long) eluted with the latter solution. The materials in the peak fractions were pooled and dialyzed against 3 liters of H_2O for 24 h; essentially all of the protein precipitated. The precipitates were collected by centrifugation at 20,000 × g for 20 min, dissolved in 7.5% (w/v) SDS, 1% (v/v) β-mercaptoethanol, 5% (v/v) glycerol, and 0.125 M Tris-HCl (pH 6.8), and the solution was heated to 95 °C for 10 min immediately before electrophoresis. Samples containing the putative gene II protein were electrophoresed on 3-mm thick SDS slab gels containing 10.5% acrylamide and 0.18% bisacrylamide according to the procedure of Laemmli (27). Samples containing the putative X protein were similarly electrophoresed on gels with 14% acrylamide and 0.4% bisacrylamide. The gels also contained 8 μM urea because the proteins being electrophoresed were intended for end group analyses.

After electrophoresis, each gel was sliced horizontally at 3-mm intervals. Each piece was diced into small fragments and incubated in 10 volumes of H_2O at 50 °C for 20 h. One-hundred μl of each eluate were lyophilized, redissolved in 25 μl of protein sample buffer, and then electrophoresed on SDS-urea polyacrylamide gels. These gels were stained by a silver method (28) modified as follows. After fixing in 7% (v/v) acetic acid and 50% (v/v) methanol for 30 min, the gels were equilibrated into H_2O by washing 4 times for 15 min each in H_2O, soaked in 8% (w/v) glutaraldehyde for 1 h, and re-equilibrated with H_2O. After soaking in 0.9% (w/v) AgNO_3, 0.02 M NaOH, and 16.5% (v/v) ethanol for 30 min, the gels were again re-equilibrated in H_2O and then soaked in 0.006% (w/v) citric acid and 0.0185% (v/v) formaldehyde. Upon development of color, the gels went through a final wash in H_2O. Under these conditions, 20 μg of protein is detectable. The eluates determined by this method to containing the proteins of interest were pooled and lyophilized. The dried protein/SDS/buffer mixtures were each redissolved in 1 ml of H_2O and desalted on a Sephadex G-50 column (1 cm diameter × 25 cm long). The column was eluted with 0.01 M N-ethylmorpholinium acetate and 0.1% (w/v) SDS (pH 8.0), and the protein peak was collected and lyophilized.

Amino Acid and End Group Analyses—For amino acid analysis, the lyophilized proteins were hydrolyzed in 6 N HCl for 6 h or 72 h at 110 °C. The hydrolysates were dried and applied on a Beckman 118C analyzer. Since polyacrylamide gels contain a background level of amino acids (29), eluate from an equal size slice of gel adjacent to that containing the polypeptide of interest was desalted, hydrolyzed, and analyzed in an identical manner, and the values so obtained were subtracted from the sample values.

Manual NH_2-terminal sequence analysis with the dansyl-Edman procedure was performed exactly as described by Weiner et al. (30). Carboxypeptidases A and B were used to determine COOH-terminal sequences as follows (31). Typically, 10–500 μg of the desalted and lyophilized protein were dissolved in 0.1 ml of 0.01 M N-ethylmorpholinium acetate and 2% (w/v) SDS (pH 8.0), heated to 90 °C for 10 min, and diluted to 0.5 ml with 0.01 M N-ethylmorpholinium acetate (pH 8.0). Carboxypeptidase A at a weight ratio of substrate to enzyme of 50 was added to the solution and the mixture was incubated at 40 °C for 10 min. Carboxypeptidase B at the same substrate to enzyme ratio was added, and the incubation was continued for 1 h. At 0, 90, and 150 min, 0.15-ml aliquots were removed and chilled to 4 °C following addition of 0.2 ml of glacial acetic acid, and the resulting protein precipitates were removed by centrifugation. The supernatant from each sample was dried down under vacuum, and the released amino acids were analyzed on a Beckman 118C analyzer or on polyamide plates after dansylation (30). A parallel sample with the substrate omitted was used as blank to subtract out residues released from the carboxypeptidases.

Immunological Studies—Antibodies to gene II and X proteins were induced in rabbits by the following method, which requires very little protein. An SDS-urea polyacrylamide gel on which the dense complex had been electrophoresed was stained with Coomassie blue, and the bands of interest (5–20 μg of protein) were cut out. Each slice was minced in a Potter-Elvehjem homogenizer with 1 ml of 0.01 M NaPO_4, and 0.15 M NaCl (pH 7.4), and the protein contained therein was cross-linked to rabbit serum albumin by the addition of 0.5 mg of rabbit serum albumin and 100 μl of 8% glutaraldehyde. After mixing, 20 μg of cyanogen bromide-activated Sepharose 4B (32) were added and the resulting mixture was again homogenized. One-half of the homogenate was mixed with 0.5 ml of Freund’s adjuvant containing 1 mg/ml of heat-killed Mycobacterium butyricum and then injected subcutaneously into a rabbit at 4 sites. 0.1 ml of the other half of the homogenate, stored at −4 °C, was injected intravenously into the

FIG. 1. Top, linear representation of the gene II region. Methionine codons are represented by circles (C). The chevron points to the site of the amber mutation in R21 (3). The dotted line represents the presumed start position for the X protein. IG, the intergenic region; II, gene II; V, gene V. Bottom, Partial sequence of gene II. The predicted amino acid sequences at the NH_2 and COOH termini of gene II and X proteins are given above the nucleotide sequence. The presumed initiating methionines are overlined. The Shine and Delgarno sequence for X protein is underlined (20).
rabbit 6 weeks later and at 2-week intervals thereafter. Bleedings were done at 5-7 days after boosting. Serum and IgG were prepared by standard techniques (33).

Nitrocellulose blotting (34, 35) was used for detection of unlabeled antigens. Proteins separated by SDS-urea polyacrylamide gel electrophoresis were electrophoretically transferred to nitrocellulose sheets at 5 V for 4-6 h. The sheets were washed with rinse buffer (0.01 M Tris-HCl, 0.01 M EDTA, 0.15 M NaCl, and 0.1% (v/v) Triton X-100, pH 8.0) for 15 min and then with rinse buffer containing 2% (w/v) BSA for 2 h. The sheets were probed with 25-100 µl of serum or IgG in 10 ml of rinse buffer containing BSA for 15 h. After washing with rinse buffer for 1 h, the sheets were reacted with [125I]-protein A (0.5 µCi in 15 ml of rinse buffer containing BSA) for 6 h. A final washing for 2 h was done in rinse buffer with the NaCl concentration raised to 1.1 M. The sheets, after drying under vacuum, were placed on film for autoradiography.

RESULTS

Purification and Identification of the Gene II and X Proteins—Webster and Remender (23) showed that a dense complex accumulates in gene V amber mutant-infected bacteria. This complex contains large amounts of gene II protein and a protein of the molecular weight expected for the X protein. In order to obtain large amounts of this dense complex for isolation of these 2 proteins, we fractionated membranes from R16-infected bacteria by sucrose step gradient centrifugation as described under "Experimental Procedures." A dense complex (HH), not present in the membrane fraction from uninfected bacteria (Fig. 2A), was resolved from the R16-infected cell membranes (Fig. 2B). Analysis of this dense complex by SDS-urea polyacrylamide gel electrophoresis (Fig. 4A) confirmed the presence of large amounts of proteins (p46 and p14) migrating at the positions expected for gene II and X proteins.

The dense complex was collected and washed with 0.2 M NaCl and 0.01 M Tris-HCl (pH 8.9) to remove any lysozyme still present. Attempts to solubilize the proteins in the complex with solutions containing 1 M LiCl, 0.01 N HCl, 0.01 N NaOH, 1% Triton X-100, 0.1% sodium deoxycholate, or 0.1% Sarkosyl NL were unsuccessful. Only SDS or 7.35 M guanidinium chloride would solubilize them, so these solvents were used in further purification of the putative gene II and X proteins.

After dissolution in guanidinium chloride, the complex was subjected to gel filtration on Sephacryl S-200 in guanidinium chloride (Fig. 3). Analysis of the eluate revealed 2 major peaks, 1 of which (peak L) contained the putative gene II protein (p46) and the other (peak S) contained the putative X protein (p14), as judged by SDS-urea polyacrylamide gel electrophoresis (Fig. 4B). The proteins in each peak were quantitatively precipitated by extensive dialysis against water and collected by centrifugation. The L peak proteins were fractionated by SDS-polyacrylamide gel electrophoresis using 10.5% acrylamide gels, while the S peak proteins were fractionated using 14% acrylamide gels as described under "Experimental Procedures." The gels were sliced, the protein was eluted from each slice, and an aliquot of each eluate was electrophoresed on a SDS-urea polyacrylamide gel. We further analyzed only those eluates which contained a single band migrating at the positions expected for the putative gene II protein (p46) or X protein (p14) (Fig. 4C).

The amino acid compositions were determined for 3 independently isolated samples of p46 and p14, and the averaged results are shown in Table I. Within experimental error, the compositions of p46 and p14 agreed with those predicted from the DNA sequence for gene II and X proteins, respectively. This strongly suggested that p46 and p14 were the gene II and X proteins, P14, thus, should have the same amino acid sequence as the COOH-terminal 27% of p46 (Fig. 1). It was, therefore, probable that p46 and p14 would cross-react immunologically. To test this, we injected rabbits with p46 or p14 which had been cross-linked to rabbit serum albumin and Sepharose 4B as described under "Experimental Procedures." The sera were tested for the presence of specific antibodies by determining whether IgG from each serum could mediate binding of [125I]-protein A to p46 and p14 which had been
Bacteriophage f1 Gene II and X Proteins

**Fig. 4.** SDS-urea polyacrylamide gels of p46 and p14 at various stages of purity. A, the dense complex was electrophoresed on a SDS-urea gel with a gradient of 15-22% acrylamide by the procedure of Laemmli (27) and the proteins were stained with Coomassie blue. B, the materials in the L and S peaks from Fig. 3 were electrophoresed on a 14% acrylamide SDS-urea gel, and the proteins were stained with Coomassie blue. C, p46 eluted from preparative gels as described under “Experimental Procedures” was electrophoresed on a 10.5% acrylamide SDS-urea gel; eluted p14 was electrophoresed on a 14% acrylamide gel. Only 1 band was detectable by Coomassie blue staining on each gel. The chevrons indicate the positions of molecular weight standards. From top to bottom they are: BSA, deoxyribonuclease I, and lysozyme, with molecular mass of 66, 31, and 14 kilodaltons, respectively.

**Table I**

| Amino acid compositions of purified p46 and p14 | p46<sup>1</sup> | Gene II protein<sup>1</sup> | p14<sup>1</sup> | X protein<sup>1</sup> |
|-----------------------------------------------|----------------|---------------------------|----------------|-------------------|
| Aminos acid                                  | mol residue/mol protein | mol residue/mol protein | mol residue/mol protein | mol residue/mol protein |
| Asx<sup>4</sup>                               | 47.1            | 47                         | 14.0           | 14                |
| Threonine                                    | 20.5            | 22                         | 5.0            | 5                 |
| Serine                                       | 44.6            | 46                         | 10.5           | 12                |
| Glx<sup>1</sup>                               | 34.8            | 31                         | 7.3            | 6                 |
| Proline                                      | 16.7            | 16                         | 5.5            | 6                 |
| Glycine                                      | 16.9            | 17                         | 4.7            | 4                 |
| Alanine                                      | 27.4            | 28                         | 7.5            | 7                 |
| Valine                                       | 24.8            | 26                         | 10.4           | 10                |
| Methionine                                   | 8.8             | 7                          | 3.3            | 3                 |
| Isoleucine                                   | 21.2<sup>1</sup> | 22                         | 6.1<sup>1</sup> | 6                 |
| Leucine                                      | 48.5            | 47                         | 8.1            | 8                 |
| Tyrosine                                     | 13.4<sup>1</sup> | 15                         | 5.2<sup>1</sup> | 6                 |
| Phenylalanine                                | 25.1            | 25                         | 5.5            | 7                 |
| Histidine                                    | 8.6             | 9                          | 1.1            | 1                 |
| Lysine                                       | 20.7            | 21                         | 5.2            | 5                 |
| Arginine                                     | 19.7            | 20                         | 7.0            | 7                 |

<sup>1</sup> Determined as described under “Experimental Procedures” calculated for a total of 399 residues. This excludes the 11 cysteine and tryptophan residues contained in the 410 residues of gene II protein.

<sup>2</sup> Predicted from the f1 DNA sequence.<sup>3</sup>

<sup>3</sup> Determined as described under “Experimental Procedures” and calculated for a total of 107 residues. This excludes the 4 cysteine and tryptophan residues contained in the 111 residues of X protein.

<sup>4</sup> Asx, the sum of aspartic acid and asparagine.

<sup>5</sup> Glx, the sum of glutamic acid and glutamine.

<sup>6</sup> Extrapolated to zero hour of hydrolysis.

<sup>7</sup> Taken only from 72 h of hydrolysis.

Transferred to nitrocellulose paper (34, 35). Fig. 5 shows that either p46 or p14 antibody reacted with both p46 and p14. These data confirmed our identification of p46 and p14 as gene II and X proteins, respectively.

**End Group Analyses of Gene II and X Proteins**—We determined the NH<sub>2</sub> and COOH termini of the proteins in order to ascertain whether they were the same as those predicted from the DNA sequence. Such an analysis also would reveal whether any post-translational cleavage had occurred. Dansyl-Edman analysis of approximately 1 nmol of gene II protein gave the NH<sub>2</sub>-terminal sequence Met-Ile-Asx<sup>4</sup>. Only 1 α-amino-dansylated amino acid, of approximately the expected intensity, was detected during each cycle of the Edman degradation, confirming the lack of gross contamination. Thus, synthesis of gene II protein started at the methionine residue predicted from the DNA sequence (see Fig. 1), and only the formyl group was removed in the mature protein.

Carboxypeptidase A digestion of approximately 3 nmol of gene II protein for 90 min released 2 amino acids detected by amino acid analyzer, 3.2 nmol of alanine, and 2.4 nmol of isoleucine. A subsequent hour of carboxypeptidase B treatment released in addition 3.1 nmol of lysine. The COOH-terminal sequence can, thus, be deduced as Lys-(Ala,Ile), again agreeing with the gene II DNA sequence, which gives the sequence Lys-Ile-Ala before the stop codon (see Fig. 1). The sequence Lys-(Ala,Ile) occurs at only 1 other place in gene II, at codons 33-35.

COOH-terminal analysis was similarly performed on approximately 1 nmol of X protein, but the hydrolyzed amino acids were dansylated and then separated on polyamide plates. Carboxypeptidase A treatment released approximately equal amounts of alanine and isoleucine, and further carboxypeptidase B digestion released about the same amount of lysine.

During the preparation of this manuscript, Meyer et al. (36) published their studies on the NH<sub>2</sub> terminus of gene II protein from the dense complex using radiolabeling methods. Their results were identical with ours, except they also found trace amounts of the protein starting at the next methionine (residue 4 in Fig. 1). The dansyl-Edman method we employed may not have been sensitive enough to detect this minor species.
lysine, as judged visually by the intensity of the fluorescent dansyl derivatives. X protein, therefore, had the same COOH terminus as gene I1 protein.

No NH2-terminal dansyl derivatives of several samples of X protein could be detected, even after 3 cycles of Edman degradation. Leucine aminopeptidase also did not release any amino acids. This was a strong indication of a blocked NH2 terminus. One possibility was the retention of the initiating amino acids. This was a strong indication of a blocked NH2 terminus. As an alternative, we used pronase to proteolyze [35S]methionine-labeled X protein to amino acids (39) without any treatment to remove the NH2-terminal blocking group. Thin layer cation exchange chromatography (40) was then used to fractionate the digest. The majority of the radioactivity was in a peak co-migrating with authentic methionine (B in Fig. 6B), as was the case for the gene II protein proteolyzed in the same manner (Fig. 6A). Unlike the gene II protein digest, the X protein digest contained an additional radioactive peak co-migrating with authentic N-formylmethionine (D in Fig. 6B). This material also co-migrated with N-formylmethionine on thin layer cellulose chromatograms (data not shown). When the putative [35S]N-formylmethionine was isolated, treated with 3 N HCl at 37 °C for 6 h, and then rechromatographed, only [35S]methionine was detected (Fig. 6C). Under these conditions, authentic N-formylmethionine was totally hydrolyzed to methionine, whereas less than 10% of N-acetylmethionine was hydrolyzed (data not shown). We, therefore, deduced N-formylmethionine to be the NH2 terminus of X protein.

These data showed that the X protein was a fragment of gene II protein initiated at an internal AUG codon of gene II RNA and ending at the gene II stop codon. The amino acid composition and molecular weight estimated on SDS-urea polyacrylamide gels agreed best with those calculated for the polypeptide initiating at codon 300 in the gene II sequence (see Fig. 1). However, since both of these procedures are subject to fairly large errors, we used an independent method to locate the exact start of X protein in the gene II sequence.

An accurate determination of the ratios of methionine to lysine and of methionine to cysteine would be adequate to assign the start codon because almost all possible starts for the X protein would yield different ratios of these residues (see Table II). We, therefore, double labeled R16-infected cells with [3H]methionine and either [35S]cysteine or [14C]lysine, isolated the dense complex as described above, and electrophoresed the polypeptides on a 16.8% acrylamide and 0.45% bisacrylamide SDS-urea gel. Each lane was cut into 2.5-mm pieces, and the amount of H and 35S in each slice was determined with a liquid scintillation spectrometer as previously described (13). Fig. 7 shows that gene II and X proteins were the major radioactive species. The ratio of methionine to cysteine residues of the X protein was then calculated with the formula, $\frac{R_X = \frac{R_{II} - C_{II}}{C_{II}}}{C_{II}}$, where $R_X$ is the ratio of methionine to cysteine residues in gene II protein, which is precisely known from the DNA sequence of gene II, $C_{II}$ is the experimentally determined ratio of 3H to 35S in the X protein peak of the gel, and $C_{II}$ is the ratio of 3H to 35S in the gene II protein peak. An analogous equation, but with different constants substituted, was used to determine the ratio of methionine to lysine residues in the X protein. As Table II shows, the numbers so derived were

![Fig. 6. Chromatograms of pronase digests of p46 and p14. A. [35S]methionine-labeled p46 was eluted from SDS-polyacrylamide gels as described under "Experimental Procedures." The eluate was made 0.1 mg/ml in N-formylmethionine and 0.5 mg/ml in pronase and incubated at 37 °C for 70 h. An aliquot, along with standards, was spotted on a Fixion thin layer ion exchange plate and developed with 0.4 m sodium citrate (pH 3.3) at 37 °C. The lane containing labeled material was cut into 1-cm slices, which were individually incubated at 37 °C for 10 min. The radioactivity in each slice was then determined in a scintillation counter (13). Standards were visualized with ninhydrin or I2 vapor. The material migrating near the origin probably represents oligopeptides. B. [35S]methionine-labeled p14 was digested and chromatographed as in A. C, the radioactive material in the p14 digest co-migrating with N-formylmethionine on a Fixion plate was eluted with 2 ml of 0.01 M N-ethylmorpholinum acetate (pH 8.5). Two ml of 6 N HCl was added to the eluate, and the mixture was incubated at 37 °C for 6 h and then dried down. The residue was dissolved in water and rerun on a Fixion plate. O, origin; A, methionylasparagine; B, methionine; C, N-acetylmethionine; D, N-formylmethionine; F, front.

![Fig. 7. Radioactivity profile of double labeled dense complex electrophoresed on polyacrylamide gel. [3H]Methionine- and [35S]cysteine-labeled dense complex was electrophoresed on a 10.8% acrylamide SDS-urea gel, which was sliced into 2.5-mm pieces. The amount of H and 35S in each slice was determined in a scintillation spectrometer as described under "Experimental Procedures." The positions of stained gene II and X proteins are indicated.

![TABLE II]

| Amino acids measured | Predicted ratio of proteins starting at the following methionine residue in gene II protein | Observed ratio for X protein* |
|----------------------|-----------------------------------------------|-----------------------------|
|                      | 1     | 4     | 218   | 392   | 375   |                     |                      |                     |
| Methionine/cysteine  | 1.17  | 1.00  | 1.00  | 2.00  | 1.50  | 1.00  | 1.00  | 1.43 (±0.09)        |
| Methionine/lysine    | 0.33  | 0.29  | 0.25  | 0.36  | 0.60  | 1.00  | 1.00  | 0.54 (±0.05)        |

*a The locations of the methionine residues in the 410 residues of gene II protein (see Fig. 1) are based on the DNA sequence. * Each ratio was determined from 4 experiments as described in the text. The numbers given are the mean ± the range.
Bacteriophage f1 Gene II and X Proteins

quite reproducible and fit only the polypeptide starting at codon 300. This was the same initiating codon deduced for in vitro synthesized X protein (3).

Synthesis of X Protein in the Normal Infection Cycle—Since the dense complex was formed in cells infected with a mutant phage at high temperature, it may be argued that X protein resulted from aberrant transcription or translation. We, therefore, looked for X protein synthesis under more usual conditions of infection. When we radiolabeled minicells derived from bacteria infected with R111 (temperature-sensitive mutation in gene V) at the permissive temperature of 34 °C and subjected the total cell proteins to SDS-urea polyacrylamide gel electrophoresis, a band co-migrating with X protein from the dense complex was detectable by fluorography (Fig. 8A).

We also looked for the X protein in wild type f1-infected K38 cells by reacting nitrocellulose blots of total cellular protein with anti-gene II protein IgG and 125I-protein A as described under “Experimental Procedures.” Both gene II and X proteins were detectable at 37 °C as early as 5 min after infection (Fig. 8B, lane 2). The increased intensity of the gene II and X protein bands at higher temperatures (Fig. 8B, lanes 5 and 8) confirmed the temperature-induced overproduction of these proteins, which was suggested by the presence of the dense complex in wild type f1-infected cells at 42 °C (23).

Whereas there was more gene II protein per cell at 16 min than at 5 min after infection at 37 °C, the amount of X protein per cell was the same at both times (Fig. 8B, lanes 2 and 3). During the same time period at 42 °C, the amount of both gene II and X proteins per cell increased (Fig. 8B, lanes 8 and 9).

DISCUSSION

The X protein previously has been detected only as the product of an in vitro transcription-translation reaction directed by double-stranded DNA of f1 or M13 (16, 17). Correlation of the results of these experiments with the knowledge of the phage DNA sequence led to the proposal that the X protein might be the product of a gene located entirely within gene II (3). In this paper, we present evidence that the X protein is synthesized in vivo and consists of the COOH-terminal 111 amino acids of the gene II protein. The X protein, when isolated from the dense complex found in infected cells with nonfunctional gene V protein (23), has the expected amino acid composition along with the COOH-terminal sequence predicted for the protein from the DNA sequence. Antibodies against either gene II or X protein cross-react with the other protein and can be used to detect the presence of X protein in wild type infected bacteria. Finally, the X protein isolated from the dense complex contains N-formylmethionine, indicating that it is a primary translation product and does not arise from proteolysis of the gene II protein. Thus, the X protein appears to be the product of gene X (ten), an overlapping gene of f1, and we propose the X protein be renamed the gene X (ten) protein.

Several overlapping genes occur in the genome of the isometric single-stranded DNA bacteriophage, such as φX174 (3). One of these genes, A*, codes for a polypeptide that results from an in-phase translational start within gene A (43, 44); hence, the A* protein is identical with the COOH half of the A protein. The A protein is required for specific cleavage and replication of φX174 replicative form I DNA (45, 46). This close similarity to the f1 gene II and gene X system may lend support to Denhardt's (47) proposal that the filamentous and isometric phages are related. On the other hand, using the Queen and Korn computer program (48), we could detect no significant nucleotide or amino acid sequence homology between φX174 gene A and f1 gene II (data not shown).

The f1 protein A* is not fully characterized but has been reported to be a single chain DNase (49, 50) and also to be possibly present in the mature virion (43). The function, if any, of gene X protein is also unknown. We could not detect it in the mature virion immunologically, nor could we demonstrate nuclease activity in the dense complex (data not shown). The fact that gene II and gene X proteins are found together in the dense complex may be evidence that both proteins act in concert during DNA replication. However, the active gene II protein purified by Meyer and Geider (51) is not associated with gene X protein. Thus, there does not seem to be strong interactions between the 2 proteins in solution. The gene X protein reaches its maximal concentration in wild type f1-infected cells by 5 min after infection, so it may be needed only for double-stranded DNA synthesis, which is largely switched off by 20 min after infection (3). We are now purifying soluble, nonadenatured gene X protein as a first step in resolving this question.

Both gene II and gene X proteins are present in infected cells in much greater quantities at 42 °C than at 37 °C. The increase is larger than expected from thermal acceleration of

Fig. 8. X protein in R111-infected minicells (A) and f1-infected K38 cells (B). A, male minicell-producing E. coli (DS410/pMC7/F104) were left uninfected or were infected with R111 at 34 °C, and after 2 h, the minicells were purified with 3 diffusion centrifugations and 1 sucrose step gradient (41). The minicells (<0.01% normal cell contamination) were labeled with 0.5 mCi of [35S] cysteine for 10 min at 34 °C, harvested, and electrophoresed on a 16.8% acrylamide and 0.45% bisacrylamide SDS-urea gel, which was prepared for fluorography with salicylic acid (42). Lane 1 (uninfected cells) shows labeled proteins presumably coded by the 2 resident plasmids. Lane 2 (R111-infected cells) shows, in addition, bands corresponding to f1 gene V, gene X, and possibly gene II proteins. Gene VIII protein has no cysteine residue and, hence, was not labeled. Radioactive phage, dense complex, and purified gene V protein were run in parallel lanes, and the migration positions of the products of genes II, III, V, VIII, IX, and X thus determined are indicated by Roman numerals. B, K38 cells were grown at 37 °C to a density of 2 × 108 cells/ml, split into 3 20-ml cultures which were further incubated at 37 °C (lanes 1–3), 39 °C (lanes 4–6), or 42 °C (lanes 7–9), and infected with wild type f1 after 10 min to allow equilibration. At 3 min before (lanes 1, 4, and 7), 5 min after (lanes 2, 5, and 8), and 16 min after infection (lanes 3, 6, and 9), 5-ml aliquots were removed from each culture and poured over 2 ml of a frozen solution containing 0.1 m NaClN, 0.01 m o-isocyanacetic acid, 0.1 m NaCl, 1 mM EDTA, and 10 mM Tris-HCl (pH 7.6), and the poisoned bacteria were harvested by centrifugation at 17,000 × g for 4 min at 4 °C and then dissolved in protein sample buffer. The proteins from equal numbers of cells (approximately 2 × 107 cells/lane) were electrophoresed on a 16.8% acrylamide and 0.39% bisacrylamide SDS-urea gel and transferred to a nitrocellulose sheet, which was then reacted with anti-gene II protein IgG and 125I-protein A as described under “Experimental Procedures.” The positions of gene II and X proteins from the dense complex are marked with Roman numerals.
biosynthetic reactions, but may be explained by the proposals that the gene V protein represses the synthesis of gene II protein and that the wild type gene V protein is thermolabile (17). Gene X expression may be similarly turned off by gene V protein. Therefore, the dense complex may not reflect the accumulation of a normal replicative apparatus, but may be merely an aberrant structure resulting from overproduction of these proteins in the absence of active gene V protein. Prouty et al. (52) have reported finding similar dense proteinaceous material in _E. coli_ under abnormal growth conditions. Especially pertinent is that these proteins were found to be denatured and soluble only in SDS, as is the case for the dense complex (23 and this paper).

If the dense complex is an abnormal structure, then one may ask whether the gene II or gene X protein in the complex is processed normally. It can be argued that the processing of this protein is accessible to the deformylating enzyme. Last, may play a role in phage protein localization or regulation.

**Acknowledgments**—We thank W. Konigsberg for advice about amino acid analysis, J. Reynolds for use of the Beckman 118C, L. Hamilton for technical assistance, and M. Outlaw for assistance in the preparation of this manuscript.

**REFERENCES**

1. Beck, E., Sommer, R., Auerwald, E. A., Kurz, C., Zink, R., Osterberg, G., Schaller, H., Sugimoto, K., Sugisaki, H., Oka-moto, T., and Takanami, M. (1978) _Nucleic Acids Res._ 5, 4495-4503
2. van Wezenbeek, P. M. G. F., Hulsebos, T. J. M., and Schoemakers, J. G. G. (1981) _Cell_ 15, 1-11
3. Denhardt, D. T., Dressler, D., and Ray, D. (eds) (1978) _Cold Spring Harbor Laboratory, Cold Spring Harbor, New York_
4. Meyer, T. F., Geider, K., Kurz, C., and Schaller, H. (1979) _Nature_ 285, 365-367
5. Salstrom, J. S., and Pratt, D. (1971) _J. Mol. Biol._ 61, 489-501
6. Mazur, B. J., and Model, P. (1973) _J. Mol. Biol._ 78, 285-300
7. Alberts, B., Frey, L., and Delius, H. (1972) _J. Mol. Biol._ 68, 139-152
8. Oey, J. L., and Knippers, R. (1972) _J. Mol. Biol._ 68, 125-138
9. Webster, R. E., and Cashman, W. F. (1971) _Biochim. Biophys. Acta_ 254, 247-253
10. Prouty, W. F., and Korn, L. J. (1980) _Bacteriophages, Denhardt, D. T., Dressler, D., and Ray, D. S., ed._ 507-530, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
11. van den Hondel, C. A., Konings, R. N. H., and Schoenmakers, J. G. G. (1975) _Virology_ 57, 467-477
12. Schaller, H., Beck, E., and Takanami, M. (1978) in _The Single-Stranded DNA Phages_ (Denhardt, D. T., Dressler, D., and Ray, D. S., eds) pp. 139-163, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
13. Henry, T. J., and Pratt, D. (1968) _Proc. Natl. Acad. Sci. U. S. A._ 55, 800-807
14. Smits, M. A., Simons, G., Konings, R. N. H., and Schoenmakers, J. G. G. (1978) _Biochim. Biophys. Acta_ 521, 27-34
15. Webster, R. E., and Rementer, M. (1980) _J. Mol. Biol._ 139, 393-406
16. Lyons, L. B., and Zinder, N. D. (1972) _Virology_ 49, 45-60
17. Viñuela, E., Algranati, I. D., and Ochoa, S. (1967) _Eur. J. Biochem._ 1, 3-11
18. Osborn, M. J., Gander, J. E., Parisi, E., and Carson, J. (1972) _J. Biol. Chem._ 247, 3962-3972
19. Laemmli, U. K. (1970) _Nature_ 227, 680-685
20. Meril, C. R., Switzer, R. C., and van Keuren, M. L. (1979) _Proc. Natl. Acad. Sci. U. S. A._ 76, 4355-4359
21. Brown, W. E., and Howard, G. C. (1980) _Anal. Biochem._ 101, 294-298
22. Weinier, A. M., Platt, T., and Weber, K. (1972) _J. Biol. Chem._ 247, 3242-3251
23. Ambler, R. P. (1972) _Methods Enzymol._ 25, 143-154
24. March, S. C., Parikh, I., and Cquelacecas, P. (1974) _Anal. Biochem._ 60, 149-152
25. Goding, J. W. (1978) _J. Immunol. Methods_ 20, 241-253
26. Towbin, H., Staehelin, T., and Gordon, J. (1979) _Proc. Natl. Acad. Sci. U. S. A._ 76, 4350-4354
27. Vaessen, R. T. M. J., Krieke, J., and Groot, G. S. P. (1981) _FEBS Lett._ 124, 193-196
28. Meyer, T. F., Beyreuther, K., and Geider, K. (1980) _Mol. Gen. Genet._ 180, 489-494
29. Pilet, N. A., Brewer, H. B., and Anderson, W. F. (1974) _J. Biol. Chem._ 249, 5227-5235
30. Takeda, M., and Webster, R. E. (1968) _Proc. Natl. Acad. Sci. U. S. A._ 60, 1487-1494
31. Nomoto, M., Narahashi, Y., and Murakami, M. (1960) _J. Biochem._ 48, 593-602
32. Dénényi, T., Hazai, L., Ferenczi, S., and Bati, J. (1971) _Acta Biochim. Biophys. Biol._ 6, 385-388
33. Meagher, R. B., Tait, R. C., Bethlac, M., and Boyer, H. W. (1977) _Cell_ 10, 521-536
34. Chamberlain, J. P. (1979) _Annu. Rev. Biochem._ 48, 132-135
35. Linsey, E., and Hayashi, M. (1974) _Nature_ 249, 345-348
36. Godson, G. N., Fiddes, J. C., Barrell, B. G., and Sanger, F. (1978) in _The Single-Stranded DNA Phages_ (Denhardt, D. T., Dressler, D., and Ray, D. S., eds) pp. 51-86, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
37. Francke, D., and Ray, D. S. (1972) _J. Mol. Biol._ 61, 565-586
38. van den Hondel, C. A., Konings, R. N. H., and Schoenmakers, J. G. G. (1978) in _The Single-Stranded DNA Phages_ (Denhardt, D. T., Dressler, D., and Ray, D. S., eds) pp. 570-584
39. Horuiuchi, K., Vovis, G. F., and Model, P. (1978) in _The Single-Stranded DNA Phages_ (Denhardt, D. T., Dressler, D., and Ray, D. S., eds) pp. 113-137, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York