Structure of the lc and nmpC Outer Membrane Porin Protein Genes of Lambdoid Bacteriophage*

Andrew J. Blasband, William R. Marcotte, Jr., and Carl A. Schnaitman

From the Department of Microbiology, University of Virginia School of Medicine, Charlottesville, Virginia 22908

The lc gene of the lambdoid bacteriophage PA-2 and the nmpC gene located on a defective lambdoid prophage in the 12-min region of the Escherichia coli K12 chromosome have been sequenced. The porin proteins encoded by these two genes were almost identical, with only 4 of the 365 residues of the precursor forms of the proteins being different. The Lc and NmpC proteins were strongly homologous to the OmpC, OmpF, and PhoE proteins, with greater than 56% of the residues identical in each case. Sequencing of the region flanking the lc gene allowed precise positioning of this gene with respect to the rightward cos site of the phage and to sequences which are homologous between PA-2 and λ. In wild-type strains of E. coli K12, the nmpC gene is not expressed and contains an IS5 insertion near the 3' end of the coding region. This insertion deletes 18 residues from the COOH terminus of NmpC protein and adds 8 residues from an open reading frame extending into IS5 sequence. Expression of this form of the gene in an expression vector plasmid demonstrated that this altered protein is still capable of being transferred to the outer membrane. Plasmid expression experiments using lc-nmpC hybrid genes show that it is the presence of the IS5 insertion which prevents expression of the porin in wild-type E. coli K12. In the nmpC mutant which expresses the protein, there has been a precise excision of the IS5 which regenerates a COOH terminus of NmpC protein which is identical to that of the lc protein. Blot hybridization detected no mRNA transcripts from the wild-type nmpC gene; although transcripts were readily detected from the lc gene in PA-2 lysogens and from the nmpC mutant which has excised the IS5. This indicates that IS5 affects the production or stability of transcripts from the adjacent nmpC gene.

Porin proteins are major proteins found in the outer membranes of Gram-negative bacteria, mitochondria, and chloroplasts, where they form channels for the nonspecific permeation of small solutes. In Escherichia coli where this class of proteins has been studied most, the proteins have monomeric molecular weights of about 37,000 (1) and are present as trimers that appear to form three solute channels (2). E. coli K12 normally expresses two porins, the OmpC and OmpF proteins. The relative amounts of these two proteins are determined by conditions of growth, with expression controlled primarily by medium osmolarity, temperature, and carbon source (3, 4). Under conditions of phosphate starvation, E. coli K12 expresses a third porin, the PhoE protein, which appears better suited to allow the permeation of phosphorylated compounds (5).

A porin gene can also be found in the genomes of certain lambdoid bacteriophage, and this gene is expressed in the lysogenic state. We isolated a new lambdoid phage termed PA-2 from a porcine strain of E. coli and found that strains of E. coli K12 lysogenic for this phage produced a new porin which was distinct from the OmpC, OmpF, and PhoE proteins (6, 7). When this protein is expressed, the expression of OmpC and OmpF proteins is reduced substantially (6, 8). The phage porin was initially called "protein 2." It is now termed Lc protein as we have adopted the mnemonic lc for the locus on the phage genome encoding the protein.

Phage PA-2 has a different immunity, host range, and site of chromosomal attachment than λ (6, 9). When heteroduplexes between λ and PA-2 were examined, the two phage were found to be about 70% homologous. Regions of nonhomology were found in the regions corresponding to the immunity region, the int-att region, and the J-gene region which would account for the differences mentioned above (35). An additional region of nonhomology was found near the right cos site in the region of λ which includes gene Q and the S and R lysis genes. Genetic mapping indicated that the lc locus of PA-2 lay in this region (10), and the construction of lc-ompC hybrid genes which expressed hybrid porin proteins showed that this locus included the structural gene for the porin protein. These studies also showed that the gene was transcribed in the opposite direction from the late phage genes such as S and R.

Phage PA-2 uses the OmpC protein as its receptor (6), and the ability of this phage to cause expression in the lysogenic state of a new porin which strongly inhibits expression of OmpC protein is probably of value to the phage. When such a lysogen is induced, there is little OmpC protein present on the cell surface to neutralize progeny phage. Other lambdoid phage that use OmpC protein as receptor might also be expected to carry a porin gene similar to the lc gene of PA-2. This is indeed the case. Chang and co-workers (10) have isolated three new lambdoid phage which have DNA homologous to the lc region of PA-2 DNA, but are otherwise not identical to PA-2. These phage also use OmpC protein as receptor and produce a protein similar to the Lc protein.

Moreover, there is a gene termed nmpC on the E. coli K12 chromosome which encodes a porin nearly identical to the Lc protein, and it has recently been shown that this gene lies within a defective prophage located at 12 min on the E. coli K12 genetic map (10). This defective prophage includes DNA

* This research was supported by Grant GM18006 from the National Institutes of Health and Grant PCM-82-159112 from the National Science Foundation. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EMBL Data Bank with the accession number J02590.

† T. Gregg and C. Schnaitman, unpublished data.
which is homologous to the \( lc \) region of PA-2, as well as DNA which is unique to this prophage and DNA which is homologous to \( \lambda \). It is this defective prophage which is the source of the DNA substitutions in the \( \lambda \) mutant strains (11).

The nmpC gene is not expressed in wild-type strains of \( E. \ coli \) K12. The nmpC locus was originally identified in a single mutant strain (CS384) which expressed NmpC protein in the outer membrane (12). Studies of DNA from the nmpC region from both CS384 and its wild-type parent have shed some light on why the gene is not expressed. Both heteroduplex examination of DNA from \( \lambda \) phages derived from CS384 and its parent and Southern blot analysis of chromosomal DNA from CS384 and its parent have shown that, in the wild-type parental strain, there is an IS5 insertion near the 3' end of the nmpC coding region (10). This IS5 insertion is absent in DNA from strain CS384, suggesting that the IS5 insertion may be the reason the gene is not expressed.

In the present report, we describe the cloning and sequencing of the \( lc \) region of the phage PA-2 genome and the nmpC region of the \( E. \ coli \) K12 genome. We also show a comparison of the deduced amino acid sequence of the \( LC \) and NmpC proteins to the sequences of the OmpC, OmpF, and PhoE proteins.

### EXPERIMENTAL PROCEDURES

### RESULTS AND DISCUSSION

Sequencing of the \( lc \) and nmpC Regions—Fig. 1 is a summary restriction map of the right end of the PA-2 genome including the \( lc \) locus and the nmpC locus from the 12-min region of the genome including the IS5 insertion 3' to nmpC. Fig. 2 shows the sequencing strategy used for \( lc \), and areas which were sequenced from both strands are shown by the arrows. The sequencing strategy used for nmpC was similar except that areas where there was good agreement with the corresponding \( lc \) sequence were sequenced from only one strand. The complete sequences are given in the "Appendix."

As shown in Fig. 1, the 2800-bp base segment of the PA-2 genome which was sequenced included three complete open reading frames plus a portion of the NH2-terminal end of the \( Rz \) gene. The \( lc \) coding sequence is slightly longer than 1 kilobase and is read from right to left, as shown in Fig. 1. The \( lc \) open reading frame encodes the 365-residue precursor form of \( LC \) protein. Open reading frame Orf-1 and Orf-2 are read from left to right, the direction of transcription of the late phage genes \( S \), \( R \), and \( Rz \) of \( B \).

Fig. 3 shows a comparison of the overlapping junction between the lysis genes \( R \) and \( Rz \) of \( \lambda \) (13) to the similar junction in phage PA-2 between Orf-2, which is not homologous to \( \lambda \) sequence, and the PA-2 \( Rz \) gene, which is very strongly homologous to that of \( \lambda \). The homology between the \( \lambda \) and PA-2 \( Rz \) genes begins at the second base of the initiation codon since the \( Rz \) gene initiates with AUG, whereas the PA-2 \( Rz \) initiates with the less commonly used GUG. This is analogous to a situation seen in another pair of related phage. The \( A \) cistron of MS2 initiates with GUG, whereas the homologous R17 \( A \) cistron initiates with AUG (14). The fact that third-base substitutions are seen when the \( \lambda \) and PA-2

---

**Fig. 1. Structure of the \( lc \) region of phage PA-2 and the nmpC region of \( E. \ coli \) K12.** The top half shows the right end of the phage PA-2 chromosome. The heavy lines indicate regions homologous to \( \lambda \) and the location in \( \lambda \) of genes \( O, P \), and \( Rz \). The location of \( lc \) is shown, and the directions of transcription of these and other genes noted below are indicated by arrows. The expanded portion of the PA-2 genome shows the region which was sequenced and indicates, in addition to \( lc \), the two significant open reading frames and the PA-2 sequence homologous to the \( Rz \) gene of \( \lambda \). The bottom half shows a section of the 12-min region of the \( E. \ coli \) K12 chromosome surrounding the nmpC locus. The nmpC coding region and the IS5 insertion are indicated by dark and open bars, respectively. The expanded region shows the portion of the nmpC region which was sequenced and indicates, in addition to nmpC, the open reading frame present in this region. \( Rz \), EcoRV; \( A \), AccI; \( E \), EcoRI; \( S \), SacI; \( Bn \), BanII; \( Bg \), BglII; \( Dr \), Drai; \( Hp \), Hpal; \( S \), Sall; \( B \), BamHI; \( H \), HindIII; \( b \), bases; \( kb \), kilobases.

**Fig. 2. Sequencing strategy.** The portion of the \( lc \) region of the right end of the PA-2 genome which was sequenced is shown. The \( lc \) coding region is indicated by heavy line. The orientation of this fragment shown is the opposite of that shown in Fig. 1 and is the same as that of the sequence presented in the "Appendix." The numbers indicate distance in kilobases from cosR. Arrows indicate the direction of sequencing, and dots indicate the start of sequencing. Abbreviations of restriction sites are as for Fig. 1 except for the following: \( S \), Sau3A; \( M \), MspI; \( T \), TaqI.

\( Rz \) sequences are compared (two are seen in Fig. 3) indicates that this portion of \( Rz \) is, in fact, from PA-2 and that the junction is real and not an artifact created by recombination or gene conversion during the construction of the hybrid phage AIC-1 (10). The region of homology which begins at this junction extends to cosR.

The sizes of Orf-2 and \( LR \) proteins are nearly identical (165 and 157 residues, respectively), which suggests that Orf-2 may be a functional analog of \( LR \) protein; however, a comparison of the amino acid sequence of these proteins both visually and with the aid of a computer showed no significant homology. Thus, the function of Orf-2 remains unknown. We also cannot ascribe a function to Orf-1, which begins just prior to the 3'
end of the \( \ell \) coding sequence.

In agreement with our earlier observations (10), an IS5 element is inserted near the 3' end of the coding region of nmpC in DNA from wild-type E. coli. Cloning and sequencing of a chromosomal fragment from the wild-type strain CS180 has allowed us to define the location of the IS5 more precisely. Fig. 4 shows that the junction between IS5 and nmpC lies with the codon for threonine at residue 323. The open reading frame extending across the junction into IS5 adds 8 residues not found in \( \ell \) and deletes 18 residues found at the COOH-terminal end of \( \ell \) protein, resulting in a mature protein of 332 residues as compared to 342 for \( \ell \) protein.

The IS5 insertion near the 3' end of the nmpC coding region eliminates the NH2-terminal portion of Orf-1 as it is found in \( \ell \). However, an open reading frame extending out of IS5 is in-frame with Orf-1, leading to an IS5-nmpC hybrid Orf-1. The hybrid Orf-1 of nmpC is slightly shorter; 57 residues as compared to 68 residues for the Orf-1 of \( \ell \). Of these 57 residues, 34 at the COOH terminus of the \( \ell \) protein are from nmpC and 23 are from IS5.

The \( \ell \) gene has one feature which is different from the nmpC, ompF, and phoE genes in that the coding region is not followed by a typical rho-independent termination signal (1, 15, 16). Three large stem-loop structures can be formed 3' to the coding region, and these might function as pause sites for rho-dependent termination.

Sequence 5' to the coding regions is strongly conserved between \( \ell \) and nmpC. In sequence extending 467 base pairs 5' from the initiation codons, we found only 19 single-base substitutions of unrelated amino acids and these were in regions which did not appear to contain any major structural features or consensus sequences. The expressions of both \( \ell \) and NmpC protein are subject to catabolite repression (6, 12). Sequences located 320 and 110 base pairs 5' to the initiation codon of \( \ell \) are the only sequences in reasonable agreement with the consensus CRP-binding sequence found near other catabolite-repressible genes and operons (17).

The BanII–EcoRI fragment from the 5' end of the \( \ell \) coding region (see Fig. 1) was end-labeled at the EcoRI site and used to determine the location of the IS5 more precisely. This region of homology extends to the right end of the \( \ell \) gene. The lower portion shows the corresponding region from PA-2 and shows the COOH terminus of Orf-2. Asterisks indicate termination codons. The underlined sequence shows where PA-2 and \( \ell \) are homologous. This region of homology extends to cosR.

**FIG. 3.** Junction between the \( \ell \) region of PA-2 and sequence homologous to the right end of the \( \ell \) genome. The top portion of the figure shows \( \ell \) sequence in the region encoding the COOH terminus of the \( \ell \) gene product and the NH2-terminal region of the B gene product. The lower portion shows the corresponding region from PA-2 and shows the COOH terminus of Orf-2. Asterisks indicate termination codons. The underlined sequence shows where PA-2 and \( \ell \) are homologous. This region of homology extends to cosR.

**Orf-2**

**FIG. 4.** Comparison of the 3' ends of the \( \ell \) and nmpC genes. The numbers refer to residues of \( \ell \) protein, as shown in Fig. 5. The end of the IS5 insertion into nmpC is indicated by the dashed line, and the underlined sequence indicates the 4-base consensus IS5 target site (CTAA) and the 16-base inverted duplication as described by Engler and van Bree (30). Asterisks indicate termination codons.

**Homology of the \( \ell \) and NmpC Proteins to Other Porins—**

Fig. 5 shows a comparison of the precursor form of the \( \ell \) protein to the precursor forms of the OmpF, OmpC, and PhoE proteins of E. coli K12. The comparison was done by computer using the FASTP program as described by Lipman and Pearson (18). A comparison against the 3309 protein sequences in the National Biomedical Research Foundation protein library yielded no proteins other than the above which shared significant homology. The \( \ell \) protein was compared to the recently sequenced yeast mitochondrion outer membrane porin (19). Although a few short overlaps were found, these were not statistically significant enough to indicate a relationship between the two proteins (data not shown).

The results in Fig. 5 confirm and extend the observations

**Fig. 5.** Evolutionary comparison of the E. coli K12 porin proteins. The top line shows the \( \ell \) porin, and the next three lines show OmpF, OmpC, and PhoE, respectively. Capital letters show sequence identity, dashes indicate deletions, and plus signs indicate substitution of a related amino acid (positive relatedness odds or a score zero or greater on the mutation data matrix of Barker and Dayhoff (20)). Zeros indicate substitutions of unrelated amino acids (negative relatedness odd). Lower-case letters indicate insertions of unrelated amino acids. Homologous regions (3 of the 4 residues identical at that position) are underlined. The numbers refer to \( \ell \) residue numbers.
of Mizuno et al. (1) of strong homology between the amino acid sequences of the E. coli porins. The 342-amino acid Lc protein and its 23-amino acid leader peptide showed 58.3% identity to the OmpF precursor, 58.9% identity to the OmpC precursor, and 56.5% identity to the PhoE precursor. A striking feature of the homology between these proteins is its patchy nature, interspersing regions which are very strongly conserved with regions which show little sequence conservation. There is a strongly conserved sequence at the NH₂ terminus (residues 1-18) and at the COOH terminus (residues 331-342) as well as regions within the protein, for example, residues 38-60 and 286-314, where sequence is strongly conserved. There are also short regions, for example, the regions around residues 120 and 165, where there appears to be little conservation of sequence. It should be noted that the lc coding region exhibits a very strongly biased codon usage, essentially identical to that described by Mizuno et al. (1), indicative of a very strongly translated protein.

As anticipated from previous chemical comparison of the proteins (7) and DNA hybridization results (10), there was little difference either at the protein or DNA level between the lc and nmpC genes. Within the coding regions, the DNA sequence was more than 95% identical, with the majority of the changes being third-base substitutions. The leader peptide sequences of both proteins are identical. Within the 324 residues of NmpC protein prior to the site of the IS5 insertion, only 4 residues differed from those of Lc protein. Three of these, substitution of glutamine for glycine at residue 126, alanine for threonine at residue 128, and valine for phenylalanine at residue 131 of the mature protein, fall within a region where sequence is poorly conserved between the other porins. The fourth substitution, lysine for asparagine at residue 302, may account for the slight difference in tryptic-chymotryptic peptides noted during the previous chemical comparison of the Lc and NmpC proteins.

As noted previously (10,12), there is a single nmpC mutant, strain CS384, which has lost the IS5 insertion and which expresses the NmpC protein in the outer membrane. We have cloned and sequenced (data not shown), a fragment from CS384 which includes the 3' end of the nmpC coding sequence. The sequence of this fragment indicates that excision of the IS5 in CS384 was precise and that the deduced amino acid sequence from the site of the IS5 insertion to the COOH terminus of the expressed NmpC protein is identical to that of Lc protein.

The homology between the various porins at the amino acid level is reflected in extensive homology between the porin genes at the DNA level. Even third-base changes are kept to a minimum by the strong codon bias of these genes. This raises an interesting question. In a cell which contains at least four homologous porin genes, what acts to prevent homologous recombination and subsequent homogenotization between these genes? The observation reported here that minor differences are observed in the sequences of the lc and nmpC genes even though phage PA-2 has been propagated on E. coli K12 for many years suggests that recombination and homogenotization between these genes has not occurred.

Expression of lc and nmpC Cloned into Multicopy Plasmid Vectors—In order to initiate the study of the regulation of the lc gene and to obtain more information about the way in which the IS5 insertion prevents expression of NmpC protein in wild-type E. coli K12, we examined the outer membrane proteins produced by strains carrying the various plasmid constructions which are summarized in Fig. 6.

Fig. 6. Plasmid constructions used to examine Lc and NmpC protein expression. The solid lines indicate vector sequence, whereas the open bars indicate cloned inserts from the nmpC locus and the stippled bars indicate cloned inserts from the lc locus. Plasmid pLC6 contains the entire lc coding region, plus sequence 5' to the coding region extending to the BglII site (see Fig. 1). The BglII site was converted to a HindIII site during construction. Plasmid pLC4 is similar, except that the region 5' to the coding region is longer, extending to the HpaI site. The HpaI site was converted to a HindIII site during construction. Plasmid pBM11.0 contains a nmpC-lc hybrid insert in which the 5'-untranslated region and the coding region to the Act1 site are from nmpC and the 3' portion of the coding region and 3'-untranslated region are from lc. pBM13.0 is a similar construction, except that the longer 5' region from the HpaI site to the BglII site is from nmpC and the entire coding region is from lc. Plasmid pBM6.2 contains a lc-nmpC hybrid insert in which the lc portion consists of a long region 5' to the coding sequence and almost the entire coding sequence, whereas the nmpC portion consists of the 3' end of the coding sequence plus most of IS5. Plasmid pBM8.1 has an insert of the BglII fragment from nmpC containing the coding region and most of IS5. Plasmid pBM5.0 consists of the same fragment inserted into the tac expression vector. The tac promoter (ptac) and the rrrB sequence which provides two terminators are indicated. Restriction site abbreviations are as for Fig. 1.
When outer membrane proteins from a strain carrying plasmid pLC6, which includes the EcoRV-BglII fragment of \( lc \), were examined, the amount of Lc protein produced was about the same as the amount produced by a single copy of the intact gene present in a PA-2 lysogen (Fig. 7, lanes B and D). The amount of Lc protein produced by strains carrying this construction is not strongly influenced by temperature or catabolite repression, which are both known to affect expression of Lc protein in PA-2 lysogens (6, 8). Strains carrying plasmid pLC4, which includes the EcoRV-HpaI fragment of \( lc \), produced considerably more Lc protein. When grown at 37 °C (Fig. 7, lane C), the amount of Lc protein produced is more than twice that produced by a PA-2 lysogen. There is a reduction in the amount of OmpA protein produced, similar to that reported in strains expressing higher copy numbers of the ompC gene (3). This suggests that sequence between the BglII site and the HpaI site is necessary for full expression of \( lc \). When plasmids which carried an additional 350 base pairs 5′ to the HpaI site were examined, there was no additional production of Lc protein over that seen with plasmid pLC4. These results are consistent with a regulatory role for one or both of the CRP-binding sites located upstream from the \( lc \) transcriptional start site. The BglII site lies within the CRP-binding site closest to the 5′ end of \( lc \) mRNA; and thus, both CRP-binding sites are absent in pLC6 and present in pLC4. None of these plasmids resulted in gross overproduction of Lc protein. This is in contrast to what was observed with strains carrying plasmids containing the 2.6-kilobase HindIII fragment including the ompC gene (36). When the ompC gene is introduced into \( E. coli \) K12 on a multicopy plasmid, there is a vast overproduction of OmpC protein which leads to a complete suppression of the other major proteins including other porins and the OmpA protein. The observation that Lc protein is not similarly overproduced indicates that its expression may be self-regulated. This is in agreement with the observations of Fralick and Diedrich (8), who examined Lc protein expression as a function of growth temperature in a strain carrying two copies of the \( lc \) region of PA-2. They found that, at low temperature, at which expression of \( lc \) is not maximal, the diploid strain produced twice as much Lc protein as a haploid strain; whereas at 40 °C, which results in maximal expression, the amount of protein produced by haploid and diploid strains was the same. It is quite reasonable that the \( lc \) gene should be self-regulated since, during the establishment of lysogeny, a cell may contain many copies of the phage and overexpression of Lc protein may be deleterious. Because of the high copy number, Lc protein expression from plasmid pLC4 is much greater than that from a PA-2 lysogen at low temperature or under conditions of catabolite repression, so it was not possible to determine whether the cloned fragment contains all of the information necessary for regulated expression of Lc protein.

The expression of NmpC protein in wild-type strains which contain the ISS insertion is completely null. No protein resembling the NmpC protein is found in the outer membrane, even in strains deficient in the other porins. More significantly, \( nmpC \) phage which carry the \( nmpC \) locus do not yield porin* plaques on a porin-deficient indicator strain (10). This is a very sensitive test which will detect porin activity at levels below that at which the protein can be detected in stained gels of outer membrane proteins.

Since the ISS insertion has deleted residues equivalent to residues 351–342 at the COOH terminus of Lc protein which seem to be strongly conserved among the other porins, one explanation for the null phenotype of \( nmpC \) might be that the truncated protein is missing sequence necessary for transport into the outer membrane. In order to test this, we constructed plasmid pBM5.0, which places the \( nmpC \) coding region downstream from a strong, inducible tac (trp-lac hybrid) promoter (Fig. 6). This plasmid was introduced into a derivative of strain JM101 which carries a strong lacI’ mutation to prevent expression in the absence of inducer. Fig. 8 shows a radiograph of protein from cells which were induced briefly with isopropyl-\( \beta \)-D-thiogalactoside and then given a short pulse label with \( ^{35}S \)methionine. After a short chase, the cells were broken with a French press and separated by centrifugation into an outer membrane fraction and a supernatant which contained the soluble cytoplasmic and periplasmic proteins and most of the cytoplasmic membrane. In the absence of inducer, neither the outer membrane fraction nor the supernatant contained NmpC protein. After induction, a large amount of newly labeled protein was found in the outer membrane fraction. When prepared for electrophoresis by a method which involves boiling briefly (5 min) in a solution containing sodium dodecyl sulfate and 2-mercaptoethanol, about half of the labeled protein migrated slightly faster than Lc or NmpC protein, as expected since the protein is slightly truncated. The remainder of the protein migrated more slowly, at a position on the gel characteristic of undenatured porin trimer (7). Prolonged boiling in sodium dodecyl sulfate/2-mercaptoethanol solution shifted some of the label from the trimer form to the monomeric form, but it was not possible to convert all of the labeled protein into the monomeric form.

These results indicate that the truncated NmpC protein made from the \( nmpC \) gene with the ISS insertion contains the information necessary for export to the outer membrane. However, the protein in the outer membrane is at least partly in an altered form so that trimers become cross-linked or are otherwise modified so that they cannot readily be denatured to the monomeric form. We suggest that the abnormal property of the truncated NmpC protein is due to the loss of a sequence at the COOH terminus which is strongly conserved among the porins. An alternative which cannot be ruled out is that the 8 residues from ISS and its junction which are at the COOH terminus prevent the truncated protein from assembling properly in the outer membrane. It should be noted (see Fig. 4) that the addition of these residues alters the charge of the COOH terminus with respect to that of the other porins.

Since the truncated form of NmpC protein can be expressed and exported to the outer membrane when the gene is pro-
volved with a strong promoter, we are left with two other hypotheses to explain the null phenotype. First, since the truncated NmpC protein resulting from the IS5 insertion is abnormal and may be deleterious to the cell, a secondary hypothesis to explain the null phenotype. First, since the provided with a strong promoter, we are left with two other hypotheses seems unlikely since there is so little difference between the sequence of the promoters of the inserted genes. For example, a strong promoter in IS5 may result in transcription into NmpC which interferes with transcription from the nmpC promoter or with the translation of NmpC mRNA. The first of these hypotheses seems unlikely since there is little difference between the sequence of the 5' regions of lc and nmpC. To test these hypotheses further, we constructed the series of plasmids shown in Fig. 6 which carry hybrid inserts consisting of portions of both lc and nmpC. There was no expression of NmpC protein from plasmid pBM8.1 (Fig. 7, lane E), which is entirely nmpC and which carries all of the coding region plus most of the IS5 insertion; nor was there expression from plasmid pBM8.2, which carries the promoter region and the 5' end of the coding region from lc fused to the 3' end of nmpC including the IS5 sequence (Fig. 7, lane F). However, expression was observed from plasmids pBM11.0 and pBM13.0, which carry the promoter regions of nmpC fused to the coding region and 3' end of lc (Fig. 7, lanes G and H). Expression from these plasmids was comparable to that from plCM6 and plCM4, which include comparable inserts which are entirely from Ic.

These results strongly suggest that the null phenotype of nmpC in wild-type strains is solely a consequence of the IS5 insertion, and this insertion prevents expression of the gene rather than preventing translation of the protein product. In order to study this further, lc and nmpC transcripts were examined by blot hybridization as shown in Fig. 9. Two strand-specific probes were made by synthesizing 32P-labeled DNA from M13 templates carrying nmpC restriction fragments. The labeled restriction fragments were cut out and gel-purified. Fig. 9A shows a blot probed with the internal Acc1-EcoRI fragment of nmpC in which the labeled strand was complementary to nmpC mRNA. The blot was then stripped of probe and re-probed with the restriction fragment from the EcoRI site in IS5 to the EcoRI site in nmpC. The labeled strand was the opposite strand to the first probe and thus would detect transcripts originating at a promoter within IS5 and extending into nmpC in the opposite direction from nmpC mRNA. The results are shown in Fig. 9B.

The first probe hybridized to two closely spaced bands (Fig. 9A, lane B) of RNA from strain CS384. This strain is the nmpC mutant which has lost the IS5 insertion and expresses NmpC protein. Identical bands (lane D) were detected in RNA from strain CS1385, which is a PA-2 lysogen of a strain deleted for the nmpC locus and expresses Lc protein. No transcripts were detected in RNA from strain CS180 (lane A), which is the parent of CS384 and contains the IS5 insertion, or in strain CS457 (lane C), which is a derivative of CS384.
deleted for the nmpC locus. No transcripts were detected with the opposite strand probe (Fig. 9B), although the probe did hybridize to the molecular weight standard (lane M). Thus, there is no indication of an anti-sense transcript originating in IS5 as the explanation of the null phenotype of CS180 and other wild-type strains which have the IS5 insertion into nmpC. The fact that no transcript was detected in RNA from CS180 indicates that the presence of the IS5 insertion affects either the production or the stability of nmpC mRNA and the null phenotype is not the result of rapid protein turnover.

It is likely that the two transcripts from both lc and nmpC have identical 5' ends but different 3' ends. Only a single band was detected when the 5' end of lc mRNA was mapped by primer extension. As noted previously, there are three stem-loop structures 3' to lc which might serve as pause sites, and it is possible that two of these represent rho-dependent termination signals. If the transcripts terminated at the first base after the first and last stem-loop structures, the transcripts would be 1198 and 1244 bases, respectively. The two bands seen in the blots of lc and nmpC mRNA shown in Fig. 9 have sizes in the range of 1200-1400 bases.

Acknowledgments—We thank M. M. Smith and W. R. Pearson for computer analysis of DNA and protein sequences, respectively.

REFERENCES
1. Mizuno, T., Chou, M.-Y., and Inouye, M. (1983) J. Biol. Chem. 258, 6932-6940
2. Enger, A., Massalki, A., Schindler, H., Dorset, D. L., and Rosenbusch, J. P. (1985) Nature 317, 643-645
3. Schnaitman, C. A., and McDonald, G. A. (1984) J. Bacteriol. 159, 555-565
4. Lundrigan, M. D., and Earhart, C. F. (1984) J. Bacteriol. 157, 262-268
5. Korteland, J., De Graaff, P., and Lugtenberg, B. (1984) Biochim. Biophys. Acta 778, 311-316
6. Schnaitman, C., Smith, D., and Forn de Salsas, M. (1975) J. Virol. 15, 1121-1130
7. Lee, D. R., Schnaitman, C. A., and Pugsley, A. P. (1979) J. Bacteriol. 138, 861-870
8. Fralick, J., and Diedrich, D. (1982) Mol. Gen. Genet. 188, 139-142
9. Pugsley, A. P., Littman-Louth, D., and Schnaitman, C. A. (1979) J. Virol. 29, 808-811
10. Highton, P. J., Chang, Y., Marcotte, W. R., Jr., and Schnaitman, C. A. (1985) J. Bacteriol. 162, 256-262
11. Kaiser, K. (1980) Mol. Gen. Genet. 179, 547-554
12. Pugsley, A. P., and Schnaitman, C. A. (1978) J. Bacteriol. 135, 1118-1129
13. Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A. (1983) Lambda II, Appendix I, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
14. Gold, L., Frihnow, D., Schneider, T., Shinedling, S., Singer, B. S., and Stormo, G. (1981) Annu. Rev. Microbiol. 35, 365-403
15. Inokuchi, K., Mutoh, N., Matsuyma, S., and Mizushima, S. (1982) Nucl. Acids Res. 10, 6957-6968
16. Overbeeke, N., Bergmans, H., van Mansfeld, F., and Lugtenberg, B. (1983) J. Mol. Biol. 163, 513-552
17. de Crombrugghe, E., Busby, S., and Bue, H. (1984) Science 224, 831-838
18. Lipman, D. J., and Pearson, W. R. (1985) Science 227, 1425-1441
19. Mihara, K., and Sato, R. (1985) EMBO J. 4, 769-774
20. Barker, W. C., and Dayhoff, M. O. (1972) in Atlas of Protein Sequence and Structure (Dayhoff, M. O., ed) Vol. 5, p. 103, National Biomedical Research Foundation, Washington, D.C.
21. Davis, R. W., Botstein, D., and Roth, J. R. (1980) A Manual for Genetic Engineering: Advanced Bacterial Genetics, p. 206, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sei. U. S. A. 74, 5463-5467
23. Biggin, M. D., Gibson, J. J., and Hong, G. F. (1983) Proc. Natl. Acad. Sei. U. S. A. 80, 3963-3965
24. Miller, J. H. (1972) Experiments in Molecular Genetics, p. 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Silhavy, T. J., Benaan, M. L., and Enquist, L. W. (1984) Experiments with Gene Fusions, pp. 177-181, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Maniatis, T., Fritsch, E. F., and Sambrino, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 90-92, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Dretzen, G., Bellard, M., Sassone-Corsi, P., and Chambon, P. (1981) Anal. Biochem. 112, 296-298
28. Myers, R. Mihara, K., and Sato, R. (1985) Nucleic Acids Res. 13, 513-517
29. Dager, M., and Ehrlich, S. D. (1979) Gene (Amst.) 6, 23-28
30. Engler, J. A., and van Bree, M. P. (1981) Gene (Amst.) 14, 155-163
31. Staden, R. (1980) Nucleic Acids Res. 8, 3673-3694
32. Aiba, H., Azthy, S., and de Crombrugghe, B. (1981) J. Biol. Chem. 256, 11905-11910
33. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-506
34. Domdey, H., Apostol, B., Newman, A., Brody, E., and Abelson, J. (1984) Cell 39, 611-621
35. Myers, R. J. (1981) Ph.D. thesis, University of Edinburgh
36. McDonald, G. (1984) Ph.D. thesis, University of Virginia
FIG. 11. Nucleotide sequence of the nmpC region of E. coli K12. Numbering: base 1 corresponds to the BglII site 5' to the NmpC protein initiation codon. The numbering is counterclockwise with respect to the E. coli K12 genetic map. The sequence includes a portion of the IS5 insertion, beginning at base 1161 and extending rightward to the end of the sequence. Open reading frames: the precursor form of NmpC protein begins at base 125 and terminates within the IS5 sequence at base 1190. Orf-1 extends leftward from bases 1134 to 1064.

**Continued on next page.**
**Phage Porin Genes**

**Supplement**

**STORAGE** of the J, and MACG cystic membrane pore proteins genes of LAB strains BACTERIOPHAGE

**By Andrew J. Blazina, William E. Mericote, Jr.,**

*and Earl A. Schneeman*

**EXPERIMENTAL PROCEDURES**

**Fragmentation and Collection Conditions.** Except as specified, cultures were grown in LB medium. Bacterial strains were E. coli, and E. coli CSH4. maca367 was a derivative of CSH4. Detailed genotypes of these strains are given in reference 12. Strain MACG was used as the host for PI spotting phage, and also as the host for the expression vector plasmid pAM2 and its derivative plasmid pAM2. The observation was used as a source of DNA or sequencing was lacI(1)-10(1) which is a derivative of pBR322 which carries pAM2 DNA from d2 through the BglII region.

**Isolation of Outer Membrane.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**DNA Sequence Determination.** DNA sequence analysis was performed by the enzymatic method of Maxam and Gilbert (26). DNA fragments were digested with BglII and RsaI and the 5' and 3'-end sequences were obtained by the dideoxy method of Sanger et al. (27). Sequence determination of the genome was done in both directions, as shown in Figure 2. Sequence of the MACG locus was done in both directions, except that the restriction sites identified in the J region and MACG regions were determined to define the desired restriction fragments into MACG. A similar strategy was used for the degenerate region of the MACG, which was used for MACG. Sequence computer analysis was as described by Staden (28).

**Isolation of DNA and Treatment used in Phage Assays.** E. coli chromosomal DNA was prepared as described by Sambrook et al. (25), and plasmid DNA was prepared by the cleaved lysis method or alkaline lysis method described by Maniatis et al. (26). DNA restriction fragments obtained after electrophoresis on 0.7% agarose gels were cloned from the gel onto M13-cuticular base essentially as described by Dotto et al. (21). Plasmids were constructed as follows. The lacI(1)-10(1) was cloned by using the appropriate restriction enzyme at the appropriate genome region for MACG, 

**Isolation of DNA from the Cystic Locus.** Plasmid pMY1 DNA and chromosomal DNA from E. coli strain CHS4 were digested in completion with BglII.

**Phage DNA Preparation.** The plasmid DNA was purified by electrophoresis and ligased as described above. After transformation and growth overnight, ampicillin resistant colonies were lifted onto Col/Polyscreen Filter (New England Nuclear) and treated as recommended by the manufacturer. As a probe, the 2.3-kb BglII fragment of the lacI region was labeled by nick translation. Plasmids were isolated from colonies which matched positive autoradiographic signals and sequenced for inserts. Plasmid pAM2 contained the 2.3-kb BglII fragment of the lacI(1)-10(1) region (Fig. 1) and plasmid pAM2 contained the 2.3-kb BglII fragment containing most of the amac1 coding region and most of the 155 region.

**Phage Transformation.** The J plasmids were constructed beginning with the 5.6-kb HindIII fragment from the right end of lacI. This was digested to completion with EcoRI and then ligated with the 2.1-kb HindIII fragment from pBR322. Because of the degeneracy of the lacIJ sites, this yields a 2.6-kb BglII fragment with an insert of 4.7 kb which contains an intact lacI locus. This plasmid, designated plasmid 1, was digested with HindIII and BglII, and the 0.55-kb fragment was gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Plasmid DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.

**Isolation of outer membranes.** Isolation of outer membrane and analysis of the proteins on Polyacryamide-GELS were as described by Schneeman and McDonald (13).

**Phage Transformation.** Phage DNA obtained from pAM2 was digested with HindIII and BglII and gel purified. The HindIII site was filled in by incubation with Klenow fragment and 125 dATP complementary fragments for 30 min at 37°C. This was in the presence of HindIII ligase (obtained from Pharmacia-P-L Biochemicals, Milwaukee, Wl). The resulting DNA was transformed into strain CM77 and ampicillin resistant colonies were selected. This construction resulted in retention of the HindIII used to digest the vector chromosome of the CHS4 site within the lacI region.