JOURNAL OF BACTERIOLOGY, Sept. 1991, p. 5523–5531
0021-9193/91/175523-09$02.00/0

Vol. 173, No. 17

A Gene for a New Lipoprotein in the *dapA-purC* Interval of the *Escherichia coli* Chromosome

JEAN BOUVIER,[1] ANTHONY P. PUGSLEY,[2] AND PATRICK STRAGIER[1]†*

Institut de Microbiologie, Bâtiment 409, Université Paris-Sud, 91405 Orsay Cedex,[1] and Unité de Génétique Moléculaire, Institut Pasteur, 75724 Paris Cedex 15,[2] France

Received 23 April 1991/Accepted 30 June 1991

Cloning and sequence analysis of the region located downstream of the *dapA* gene of *Escherichia coli* has revealed the presence of an open reading frame that is cotranscribed with *dapA*. This gene codes for a 344-amino-acid polypeptide with a potential signal sequence characteristic of a lipoprotein. When this gene, called *nlpB*, is expressed from a multicopy plasmid in bacteria grown in the presence of [^{3}H]palmitate, a labeled 37-kDa protein is produced. A slightly larger precursor molecule is detected when minicells expressing *nlpB* are treated with globomycin, a specific inhibitor of lipoprotein signal peptidase. Therefore, the *nlpB* gene encodes a new lipoprotein, designated NlpB. This lipoprotein is detected in outer membrane vesicles prepared from osmotically lysed spheroplasts and appears to be nonessential, since a strain in which the *nlpB* gene is disrupted by insertion of a chloramphenicol resistance gene is still able to grow and shows no discernible NlpB phenotype. The putative transcription termination signals of the *dapA-nlpB* operon overlap the promoter of the adjacent *purC* gene.

Escherichia coli contains several lipoproteins which are synthesized as precursors with signal peptides (for a review, see reference 46). These precursors are modified by the addition of a glyceryl group to the cysteine residue present at the signal peptide cleavage site. This is followed by O acetylation of the glycerylcysteine. This lipid modification is required for processing by the lipoprotein signal peptidase, the product of the *lspA* gene, which is selectively inhibited by the antibiotic globomycin (11). Once processed, the lipoproteins remain membrane associated through their covalently attached lipid moiety.

In addition to the major (Braun's) lipoprotein encoded by the *lpp* gene (20) and the peptidoglycan-associated lipoprotein encoded by the *pal* gene (4), several other membrane-associated lipoproteins have been characterized (12). One of these (NLP4) is likely to be identical to NlpA (lipoprotein 28), the 28-kDa product of the *nlpA* gene (50). Disruption of the *nlpA* gene, which has not been mapped on the *E. coli* chromosome, has shown that NlpA is not essential for growth (47).

J.-C. Patte and his coworkers have been involved in cloning and analyzing the structure and the expression of the genes of the diaminopimelate-lysine pathway in *E. coli* (2, 3, 9, 34–36, 39). In the course of those studies, the beginning of a putative gene was found immediately downstream from *dapA*, and it was proposed that this gene could be the second gene of the *dapA* operon (36). In this article, we report the cloning and the complete sequence of this gene, which is indeed cotranscribed with *dapA*. Because the predicted amino-terminal sequence of its product shows strong similarities with the signal peptide of known prelipoproteins, experiments which demonstrate that this gene, henceforth called *nlpB*, actually encodes a new lipoprotein were performed.

* Corresponding author.
† Present address: Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France.

MATERIALS AND METHODS

Bacterial strains and media. The presence of the *dapA* or *purC* gene on recombinant plasmids was checked with the *E. coli* K-12 strains RDA8 (a *dapA*::Mucts derivative of our reference strain, RM4102 [*araD139* Δ*lacU169 rpsL thiA*] [36]) and H624 (F$^-$ *thr-1 leuB6 metB1 purC60 rpsL126 tsx-71* [24]), provided by C. Richaud and J. Parker, respectively. Recombination with linear DNA was carried out with strain JC7623 (AB1133 *tsx-33 recB21 recC22 sbcB15* [44]), obtained from H. de Reuse. The insertion of DNA fragments upstream of the chromosomal *malPQ* operon was performed with strain pop2239 (C600 Δ*malA510* [32]), provided by O. Raibaud. Strains JB50 (a highly transformable recombination-deficient derivative of C600 provided by J. Brevet), JM83, JM101, and JM109 (49) were used for identifying recombinant plasmids and producing DNA for sequencing. They were grown in LB medium (19) in the presence of the appropriate antibiotics (50 μg of ampicillin per ml, 10 μg of tetracycline per ml, and 25 μg of chloramphenicol per ml).

Recombinant DNA procedures. The recombinant DNA procedures were previously described (2). In order to facilitate cloning of various DNA fragments in the *Eco*RI site of the promoter-probe plasmid pOM41, a pUC18 (49) derivative in which the multiple-cloning site is bordered by two *Eco*RI sites was constructed. This plasmid, pSB118, was constructed in three steps. First, the 31-bp *Eco*RI-*Hin*dIII fragment from pBR322 was cloned in pUC19 (49). The resulting plasmid was cut with *Eco*RI and *Cla*I (which cuts 6 bp upstream from the *Hin*dIII site), treated with the Klenow fragment of DNA polymerase, and ligated. This step maintains the *Eco*RI site (but destroys the *Cla*I site) and reconstitutes the reading frame of the *lacZ* α fragment. Finally, the 222-bp *Hin*dIII-*Nde*I fragment from the latter plasmid was exchanged with the equivalent fragment from pUC18. The resulting plasmid, pSB118, is almost identical to pUC18 but carries an extra *Eco*RI site located 3 bp downstream of *Hin*dIII. A similar construction, in which the roles of pUC18 and pUC19 were reversed, yielded pSB119, in which the multiple-cloning site is bordered by two *Hin*dIII sites.

Construction of *malPQ* fusions. The 1,357-bp *Nsi*I-*Pst*I and 1,202-bp *Nhe*I-*Pst*I DNA fragments carrying the proximal part of the *nlpB* gene and its upstream region, with or without the *dapA* promoter, were cloned in pSB118 (in the *Pst*I site and between the *Xba*I and *Pst*I sites, respectively). The resulting *Eco*RI fragments were then cloned into pOM41 (42) and recombined onto the chromosome of strain pop2239 upstream of the *malPQ* operon by selection for growth on M63 agar plates (19) supplemented with maltose as the sole carbon source. The Mal^+ clones were reisolated twice on LB agar (19) in the absence of antibiotics, which led to loss of the plasmid. Amylomaltase activity in sonicated extracts was measured as described previously (32) and expressed as nanomoles of glucose produced per minute per milligram of protein.

Disruption of the chromosomal *nlpB* gene. The *dapA-nlpB* region was cloned as a 2,312-bp *Nsi*I-*Eco*RI fragment from plasmid pLC17-30 (21) between the *Pst*I and *Eco*RI sites of pBR322. The desired plasmid was selected as providing tetracycline resistance and diaminopimelate prototrophy to strain RDA8 and named pDA34. The 1,870-bp *Pst*I fragment from plasmid pSKS114 (38) carrying the Tn*9* chloramphenicol resistance *cat* marker was subcloned into the *Pst*I site of the versatile cloning vector pJRD184 (10). The *cat* gene was thus bracketed between two *Mlu*I sites, one originating from pJRD184 and the other one already present downstream of *cat* in the Tn*9* sequence. The *cat* marker was then cloned as an *Mlu*I fragment into the *Bss*HII site located at codon 95 of the *nlpB* gene carried by plasmid pDA34. Since this plasmid contains two *Bss*HII sites, one in *nlpB* and the other in *dapA*, pDA34 was partially digested with *Bss*HII and the desired recombinant plasmid was selected as providing chloramphenicol resistance and diaminopimelate prototrophy to strain RDA8. One plasmid in which the *cat* marker was inserted in the *nlpB* gene and transcribed in the same orientation was kept and called pDA37. This plasmid contains long stretches of adjacent DNA upstream (1,267 bp) and downstream (1,045 bp) of the *cat* marker which allow homologous recombination into the chromosome of the exonuclease-deficient strain JC7623 (44). Plasmid pDA37 was linearized with *Sal*I (which cuts outside the cloned *E. coli* fragment) and used to transform competent JC7623 cells. Transformants were selected for chloramphenicol resistance and screened for tetracycline sensitivity. The absence of free plasmids was checked by making plasmid minipreparations from 12 clones and transforming strain JB50 for chloramphenicol resistance. Three clones were found to be devoid of plasmid and appeared to be chromosomal recombinants. A P1 lysate was grown on one of these clones and used to transduce strain H624 for purine prototrophy. Ninety-one percent of the transductants were chloramphenicol resistant, confirming the genetic linkage between the *purC* gene and the *cat* marker.

Sequencing procedures. The nucleotide sequence was determined by the chemical procedure of Maxam and Gilbert (17). Restriction fragments were ^{32}P labeled either at their 5′ end with T4 polynucleotide kinase or at their 3′ end with the Klenow fragment of DNA polymerase. Uniquely end-labeled DNA fragments were purified after a secondary restriction cleavage or after DNA strand separation. The complete sequence of the 1,361-bp *Sph*I-*Eco*RI fragment shown in Fig. 2 was obtained for both strands, and all restriction sites used for labeling were cross-checked.

Protein analysis. Lipoproteins were labeled with [^{3}H]palmitate (200 μCi/ml) in cells growing in M63 minimal medium (19) containing 0.4% Casamino Acids and 0.4% glucose. Isopropyl-β-D-galactoside (IPTG; 0.2 mM) was added to induce expression of the *lacZ-nlpB* operon fusion in pDA25 (a pUC19 derivative carrying the 1,361-bp *Sph*I-*Eco*RI fragment containing the *nlpB* gene). Incubation was continued for 2 h before the cells were harvested and either resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer (12.5% glycerol, 5% SDS, 1% 2-mercaptoethanol, 0.1 M Tris Cl [pH 8.0]) and heated to 100°C for analysis or subjected to subcellular fractionation. For subcellular fractionation, the cells were converted into spheroplasts, which were then lysed by centrifugation and resuspension in 10 mM Tris Cl (pH 7.4) containing 1 mM $MgCl_2$, 10 μg of DNase I per ml, and 10 μg of RNase per ml followed by sonication for 5 s (27, 30). Cell debris was removed by low-speed centrifugation, and the membranes were pelleted by centrifugation at 90,000 × *g* for 2 h at 4°C. The membranes were resuspended in 10 mM Tris Cl (pH 7.4) containing 1 mM EDTA, applied to the top of a 35 to 55% sucrose step gradient (10 mM Tris Cl, pH 7.4), and centrifuged to equilibrium in a Beckman VTi65 rotor at 140,000 × *g* for 16 h. Samples collected from the bottom of the gradient were treated with 15% trichloroacetic acid to precipitate the proteins, which were then dissolved in SDS-PAGE sample buffer.

Samples were examined by SDS-PAGE on Tris-glycine-buffered gels (31) containing 10% acrylamide and 0.26% bisacrylamide or 9% acrylamide, 0.24% bisacrylamide, and 8 M urea. After electrophoresis, the gels were stained with Coomassie blue, destained, treated with Amplify (Amersham), dried, and exposed to Kodak X-Omat film at −80°C. The radioactivity in the gels was quantified by cutting out the radioactive bands, dissolving them in Protosol (Du Pont), and scintillation counting.

The sucrose concentration in samples from gradients was measured with a refractometer. The positions of the outer membrane, intermediate-density, and cytoplasmic membrane vesicles in the gradient were distinguished by their distinctive protein profiles after SDS-PAGE. For internal controls, we used membranes prepared under identical conditions from derivatives of strain PAP105 (16) carrying pJG311 (encoding the outer membrane Lpp-BlaM hybrid protein), pKY201 (encoding a cytoplasmic membrane/intermediate-density vesicle NlpA-BlaM hybrid [48]), pRW83 (a pBR325 derivative encoding the BlaZ lipoprotein β-lactamase from *Bacillus licheniformis* [22]), pCHAP656 (which encodes the enzyme pullulanase from *Klebsiella oxytoca* [16]), or pCHAP875, which is similar to pCHAP656 except that the codon for Asp at position +2 is changed to an Asn codon (27). IPTG (0.2 mM) was used to induce expression of the *nlpA-blaM* genes in pKY201 and the *pulA* gene in pCHAP656 and pCHAP875. Membranes prepared as described above from unlabeled or [^{3}H]palmitate-labeled cells were analyzed in parallel with those containing the radioactive NlpB protein, and fractions collected from the gradients were either assayed for β-lactamase content with nitrocefin (23) or pullulanase content (16) or analyzed by SDS-PAGE. Previous studies (27) indicated that the BlaM hybrids are stable and that the enzyme activities accurately reflect their position in the gradients.

The effect of globomycin on the processing of pre-NlpB in minicells carrying pDA25 was determined. The minicells were prepared from strain AR1062 (*thr leu ara azi fhuA lacY tsx min gal xyl thi rpsL hsdR* pDA25) as described elsewhere (29) and labeled in M63 glucose medium at 37°C with 20 μCi of [^{35}S]methionine per ml with or without prior treatment for 5 min with 0.1 mg of globomycin per ml. After 5 min, the

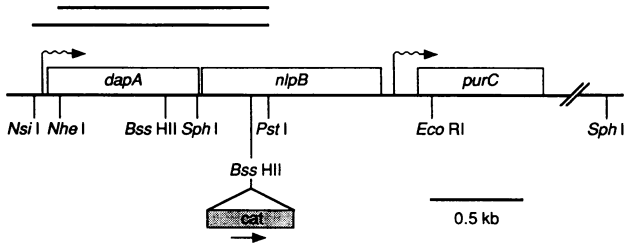

FIG. 1. Genetic organization and simplified physical map of the *dapA-purC* region. Only the restriction sites relevant to the experiments described in the text are shown. The size of the *purC* reading frame is derived from reference 41. The wavy arrows define transcription initiation sites for *dapA* (36) and *purC* (41). The thick bars show the two fragments that were used for construction of *nlpB-malPQ* fusions. The site of disruption of the *nlpB* gene with a chloramphenicol resistance marker is indicated by the shaded *cat* box (not to scale).

minicells were harvested, resuspended in sample buffer, and analyzed by SDS-PAGE on 10% acrylamide gels followed by autoradiography as described above.

Nucleotide sequence accession number. The nucleotide sequence presented in Fig. 2 has been deposited in the EMBL-GenBank data library under accession number X57402.

RESULTS

Cloning of the *nlpB* gene. The *dapA purC dapE* region of the *E. coli* chromosome has been characterized by Parker (24) and Richaud et al. (36) and is located at 2,600 kb on the *E. coli* restriction map (15, 18). As shown in Fig. 1, the *purC* gene includes an *Eco*RI site located about 1.4 kb downstream of *dapA* and is transcribed in the same counterclockwise orientation as *dapA* (24, 36). Sequencing of the *dapA* gene has revealed the presence of a putative open reading frame starting 16 bp after the *dapA* stop codon (36). The largest insert in plasmids carrying the *dapA* region did not extend further than a *Pst*I site located 400 bp downstream of *dapA* (36). In order to further characterize the region around *dapA*, we cloned a fragment of DNA overlapping the end of the *dapA* gene and containing the whole *purC* gene. This was done by subcloning a 3.4-kb *Sph*I fragment from plasmid pLC17-30, which contains this region of the *E. coli* chromosome (21), into pBR322 and selecting purine prototrophs of strain H624 (see Fig. 1). Deletion by *Eco*RI led to a plasmid containing a 1.36-kb fragment which no longer complemented the *purC* mutation of strain H624. This insert, which covers the *dapA-purC* interval, was sequenced on both strands as described in Materials and Methods.

Nucleotide sequence of the *nlpB* locus. The sequence shown in Fig. 2 confirms the existence of a large open reading frame located immediately downstream of *dapA*. This reading frame could code for a 344-amino-acid polypeptide (M_r, 36,840) with a potential 24-residue-long signal sequence followed by a cysteine residue which could become the amino terminus of the mature protein (M_r, 34,369). The presence of this cysteine residue and the sequence surrounding it (Leu Leu Ala Ala Cys Ser Ser) suggest that the product of this reading frame could be a lipoprotein, since such features are characteristic of the cleavage site of lipoprotein signal sequences (26). Experiments described below demonstrated that this is actually the case. Thus, we call the gene *nlpB* (new lipoprotein B) and designate its product NlpB. The sequence of NlpB does not contain any long hydrophobic stretches that could act as cytoplasmic membrane anchor sequences. No significant similarity with other sequences present in the GenBank-EMBL data bases could be found. The *nlpB* reading frame is immediately followed by an inverted repeat associated to a run of T's which could be a rho-independent transcription termination signal and which overlaps the promoter of the adjacent *purC* gene (41).

Expression of the *nlpB* gene. The very short intergenic region between *dapA* and *nlpB* suggests that both genes are part of the same operon. To further analyze the relationship between transcription of *dapA* and *nlpB*, we followed the strategy devised by Raibaud and coworkers (32, 42), which allows promoter activity in a cloned DNA fragment to be assessed by integration into the chromosome upstream of the *malPQ* operon and assay of amylomaltase, the *malQ* product. As detailed in Materials and Methods, two DNA fragments were used, one starting at the *Nsi*I site located 68 bp upstream of the *dapA* transcription start (36) and the other one starting at the *Nhe*I site located at codons 22 and 23 in *dapA*; both fragments end at the *Pst*I site located at codons 124 and 125 in *nlpB* (Fig. 1). Two strains were constructed, containing either the large fragment *Nsi*I-*Pst*I or the small fragment *Nhe*I-*Pst*I cloned upstream of the chromosomal *malPQ* operon. These two strains were grown in M63-glucose liquid medium, and amylomaltase assays were performed. The level of expression of the *nlpB-malPQ* fusion is four times higher when the *dapA* promoter is present in the cloned fragment than when it is absent (amylomaltase specific activity of 185 versus 41 nmol of glucose produced per min per mg of protein, respectively), confirming that *dapA* and *nlpB* are cotranscribed. However, there is still significant activity in the absence of the *dapA* promoter, indicating the presence of another promoter for *nlpB* in the *dapA* coding sequence (the background level of amylomaltase specific activity is about 1 nmol of glucose produced per min per mg of protein in the absence of any insert cloned upstream of *malPQ*). Our attempts to map this secondary promoter precisely were unsuccessful. Neither promoter responds to the addition of lysine to the growth medium (amylomaltase specific activities of 168 and 42 nmol of glucose produced per min per mg of protein for the strains containing the large or the small fragment cloned upstream of *malPQ*, respectively).

Inactivation of the *nlpB* gene. In an attempt to identify the function of the *nlpB* product, we disrupted the chromosomal *nlpB* coding sequence with a chloramphenicol resistance marker that could easily be selected for. As described in Materials and Methods, a plasmid in which the *cat* marker was inserted at codon 95 of the *nlpB* gene with large regions of adjacent DNA on both sides was constructed. The linearized plasmid was used to transform strain JC7623, in which the *cat* marker could recombine into the chromosome (44). The *nlpB*::*cat* mutation was then moved into different genetic backgrounds by P1 transduction. The *nlpB* mutation appears to be cryptic in all conditions tested: it does not create any auxotrophy, it does not modify the uptake of [^{3}H]lysine or [^{3}H]diaminopimelate, it does not change the growth characteristics in a large range of osmolarities, and it does not alter the membrane-derived oligosaccharide content of the cells (data not shown).

The *nlpB* product is a lipoprotein. As discussed above, the sequence of the extreme N terminus of the *nlpB* gene product resembles those of signal peptides of a class of exported proteins, the lipoproteins, in which the N-terminal Cys residue immediately after the signal peptide cleavage

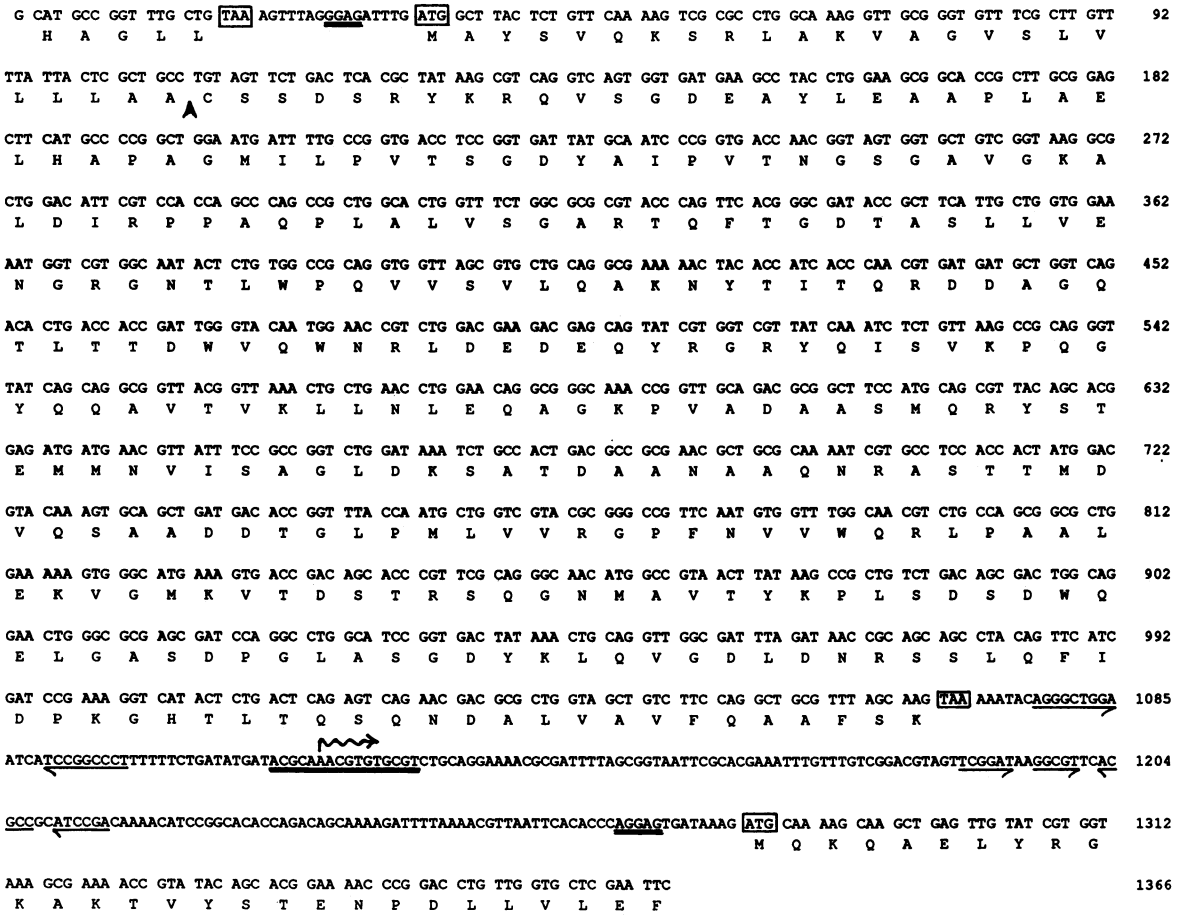

FIG. 2. Nucleotide sequence of the *dapA-purC* interval. The sequence of the *Sph*I-*Eco*RI fragment is shown with the predicted sequence of the encoded polypeptides. A putative signal peptide cleavage site is indicated by the arrowhead. Translation initiation and termination codons are boxed, and ribosome binding sites are underlined twice. Palindromic sequences in mRNA (potential *nlpB* transcription termination signal and REP sequence [8]) are indicated by convergent half-arrows. The *purC* transcription start is shown by a wavy arrow. A presumed PurR binding site (37) is underlined.

site is fatty acylated. To determine whether this was the case for NlpB, cells carrying the chromosomal *nlpB* gene, the *nlpB*::*cat* mutation, or the *nlpB* gene cloned in pDA25 were labeled with [^{3}H]palmitate. As shown in Fig. 3, amplification of *nlpB* gene expression resulted in the synthesis of large amounts of a ca. 37-kDa protein that was strongly labeled by the palmitate. Optimal separation of NlpB from other lipoproteins of similar sizes was obtained on urea-containing gels, which clearly show that the band corresponding to NlpB is a minor protein in the wild-type strain and is absent from the mutant. Thus, we conclude that NlpB is a lipoprotein. One of the comigrating lipoproteins may be the 36-kDa product of the *rplA* gene (40).

As an alternative approach to determine whether NlpB is a typical bacterial lipoprotein, we labeled minicells carrying pDA25 in the presence or absence of globomycin, a specific inhibitor of lipoprotein signal peptidase (11). As shown in Fig. 4, about 50% of the [^{35}S]methionine label incorporated into NlpB in the presence of globomycin was present in a slightly larger protein when examined by SDS-PAGE. This protein presumably corresponds to a precursor form of NlpB.

It is interesting to note that high-level expression of *nlpB* seems to reduce the synthesis or the level of [^{3}H]palmitate incorporation of some but not all lipoproteins (Fig. 3). This does not appear to be due simply to overproduction of a lipoprotein, since production of the BlaZ lipoprotein of *B. licheniformis* at comparable levels did not have the same effect (data not shown).

The NlpB protein is located in outer membrane vesicles. Lipoproteins can be found in several different locations in gram-negative bacteria (Table 1). According to the studies of Yamaguchi et al. (48) and our own studies (27), lipoproteins with an Asp residue located immediately after the Cys residue are located on the periplasmic face of the cytoplasmic membrane (27, 28, 30) and appear in vesicles whose density is intermediate between those of cytoplasmic and outer membrane-derived vesicles when cells are lysed by conversion to spheroplasts followed by osmotic shock (27, 30). Replacement of this Asp by Ser, Glu, or Asn causes these proteins to be sorted to the periplasmic face of the outer membrane (27, 48), in agreement with the fact that lipoproteins with these or other amino acids except Asp in position +2 are located in the outer membrane (Table 1). The NlpB protein has a Ser residue at positions +2 and +3 but an Asp at position +4. Since it is not known whether an Asp residue must be specifically located at position +2 in order to

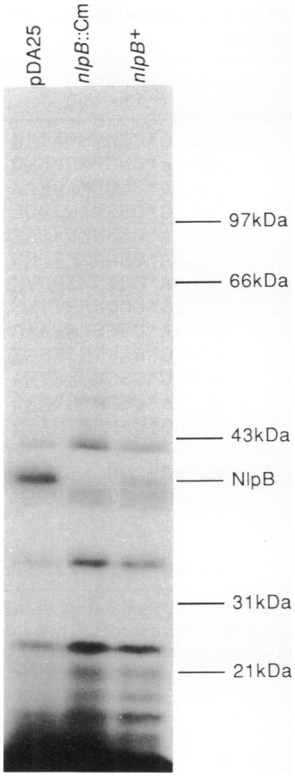

FIG. 3. [^{3}H]palmitate labeling of NlpB in wild-type and mutant *E. coli* and in the mutant strain carrying pDA25 after growth in the presence of IPTG. Proteins were resolved on a 9% acrylamide gel containing 8 M urea. Note that several lipoproteins appear to be less abundant in the strain that overproduces NlpB. Cm, chloramphenicol resistance gene.

function as the lipoprotein sorting signal, it was of interest to determine the subcellular location of NlpB.

In order to identify the NlpB protein, we used IPTG-induced, [^{3}H]palmitate-labeled cells carrying pDA25. Membranes were prepared by an osmotic lysis technique that was shown elsewhere to minimize lipoprotein redistribution (27, 30). Membranes prepared from strains producing large amounts of other, well-characterized lipoproteins (see Materials and Methods) were mixed with those containing NlpB and centrifuged to equilibrium in sucrose gradients from which fractions were collected and examined by SDS-PAGE and autoradiography. As shown in Fig. 5A, the radioactive band of NlpB is present in the dense fractions corresponding to the outer membrane in which the major outer membrane porins and OmpA protein are enriched (Fig. 5B). Its fractionation behavior is identical to that of two other lipoproteins that are presumably also located in the outer membrane (the 28- and 22-kDa lipoproteins in Fig. 5A; the 28-kDa lipoprotein is probably PAL [4, 12], and the 22-kDa lipoprotein is probably NLP7 [12]) and to that of pullulanase with substitution of the Asp at position +2 by Asn (PulA:D2N in Fig. 5A). The NlpA-BlaM hybrid protein and the 50-kDa lipoprotein were located in less dense (intermediate-density) vesicles (Fig. 5A).

Quantification of this and other similar autoradiographs confirmed this fractionation behavior and showed that Lpp-BlaM and BlaZ were located in the dense outer membrane fractions and that pullulanase (PulA) was mainly located in

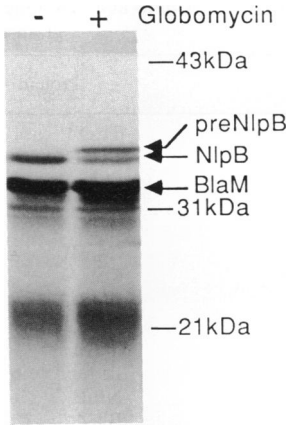

FIG. 4. Effect of globomycin on [^{35}S]methionine-labeled NlpB produced in minicells harboring pDA25 (see Materials and Methods for conditions). The positions of precursor (preNlpB) and mature NlpB and molecular size markers are indicated.

the intermediate-density fractions (Fig. 6). A similar result was obtained when the activities of the lipoprotein enzymes (in fractionated, unlabeled membranes) were compared with that of the [^{3}H]palmitate-labeled NlpB in another gradient similar to that in Fig. 5A (not shown; see references 27 and 30 for examples).

These data are essentially the same as those reported previously with PulA, Lpp-BlaM, and NlpA-BlaM (27, 30). Lipoproteins in the intermediate-density fractions are consistently poorly resolved from those in the outer membrane. It can also be seen in Fig. 5A and 6 that pullulanase, its derivative PulA:D2N, NlpA-BlaM, and Lpp-BlaM are all also found in fractions near the top of the gradient. This phenomenon was reported previously and was attributed to the release of these proteins from the membrane to form protein micelles that are less dense than membrane vesicles (27, 28, 30). Small amounts of other lipoproteins were also detected in these light fractions (Fig. 5A and 6). These complications notwithstanding, the data clearly show that NlpB is highly enriched in outer membrane vesicles and is therefore probably located in the outer membrane.

DISCUSSION

The region located downstream of *dapA* has been cloned. It contains a new gene, *nlpB*, immediately followed by the *purC* gene. The *nlpB* gene should encode a 344-amino-acid polypeptide. According to the sequence of this open reading frame, the product of the *nlpB* gene should be a precursor that is processed by the globomycin-sensitive lipoprotein signal peptidase and modified by the incorporation of fatty acids. The results in Fig. 3 and 4 show this to be the case and furthermore show that a lipoprotein of the appropriate size is missing in a strain with a mutation in *nlpB* and overproduced in the same strain harboring a plasmid with the *nlpB* reading frame under the control of the *lac* promoter (see below for further discussion). While this article was in preparation, the sequences of the *purC* locus and of the *dapA-purC* intergenic region were published (41). There are several differences in the sequence reported in that article which introduce successive frameshifts in the *nlpB* reading frame. In the sequence in reference 41, there is an extra C at position 184 (position 1315 in reference 41), which eliminates an *Alu*I site

TABLE 1. Amino acid sequences around the signal peptide cleavage sites of lipoproteins whose locations in gram-negative bacteria have been rigorously determined[a]

Location	Protein	Species or strain (relevant characteristic)	Sequence (−5 to +10)	Reference
Outer membrane	TraT	*Escherichia coli* K-12 (F^+)	LALSG*CGAMSTAIKK	13
	17-kDa protein	*Rickettsia rickettsii*	SMLQA*CNGPGGMNKQ	1
	H.8	*Neisseria gonorrhoeae*	LALAA*CGGEKAAEAA	45
	Lpp	*E. coli* K-12	TLLAG*CSSNAKIDQL	20
	PAL	*E. coli* K-12	MAIAA*CSSNKNASNE	4
	OsmB	*E. coli* K-12	MSLSA*CSNWSKRDRN	14
	Cnl	*E. coli* K-12[pCHAP4(ColN)]	MTLSA*CQVNHIRDVK	25
	PulS	*E. coli* K-12 (*pulS*$^+$)	VLLSG*CQQNRPTTLS	7
	PAL	*Haemophilus influenzae*	AALAA*CSSSNNEAAG	6
	PCN	*H. influenzae*	FSVTG*CANTEIFSGE	6
	Lipoprotein 1	*Pseudomonas aeruginosa*	VLATC*CSSHSKETEA	5
	NlpB	*E. coli* K-12	LLVVA*CSSDSRYKRQ	This study
	BlaZ	*E. coli* K-12 (*blaZ*$^+$)	VALAG*CANNQTNASQ	22
Cytoplasmic membrane or intermediate vesicles (periplasmic face)	Pullulanase	*E. coli* K-12 (*pulA*$^+$)	VLLSG*CDNGSSSSSS	16
	NlpA	*E. coli* K-12	ILLAG*CDQSSSDAKH	50
Photosynthetic membrane	Cytochrome *c*	*Rhodopseudomonas viridis*	SLVAG*CFEPPPATTT	43

[a] Included in this list are all lipoproteins whose locations were determined by subcellular fractionation, including buoyant-density centrifugation of membrane vesicles, or by accessibility (in whole cells) to externally added antibodies. Lipoproteins whose location was determined solely by detergent treatment of membrane preparations are not included. The *pulA* and *pulS* genes are from *K. oxytoca* UNF5023, and the *blaZ* gene is from *B. licheniformis*. Charged amino acids are underlined, and the asterisks indicate the lipoprotein signal peptidase cleavage site.

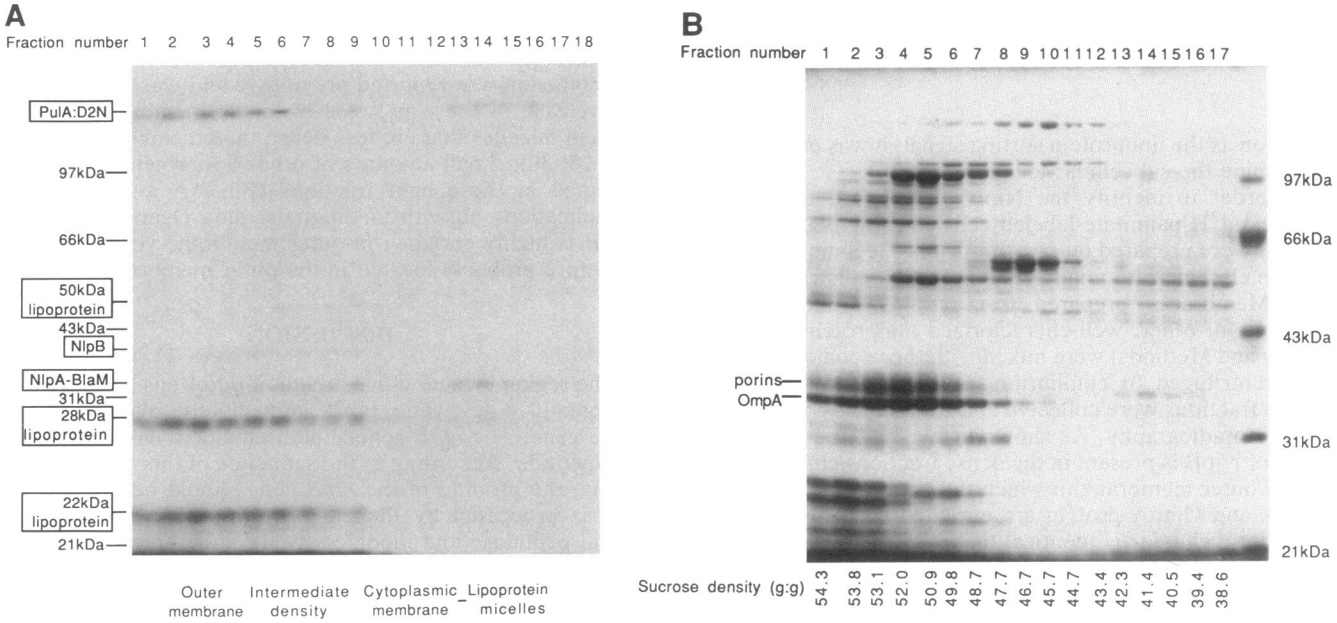

FIG. 5. Separation of membranes prepared from cells carrying pDA25 and from other strains carrying plasmids coding for the high-level production of other, previously characterized lipoproteins by isopycnic sucrose gradient centrifugation. (A) Autoradiograph displaying the [^{3}H]palmitate-labeled proteins in the different fractions collected from the gradients after separation by SDS-PAGE. The membranes were prepared from IPTG-induced cells carrying pDA25 (NlpB), pCHAP875 (pullulanase with an Asp-to-Asn substitution at position +2 [PulA:D2N]), and pKY201 (NlpA-BlaM). The positions of these and other abundant lipoproteins are noted, as are the positions of molecular size markers. (B) Coomassie brilliant blue-stained gel used to separate proteins in fractions from a gradient that was loaded with membranes from unlabeled cells and run in parallel with that shown in panel A. The positions of molecular size markers, the major outer membrane porins, and OmpA protein and the densities of sucrose (percent sucrose [wt/wt]) in the fractions from the gradient shown in panel B are indicated, as are the approximate boundaries of the outer membrane, intermediate-density, and cytoplasmic membrane vesicles in the gradient shown in panel A.

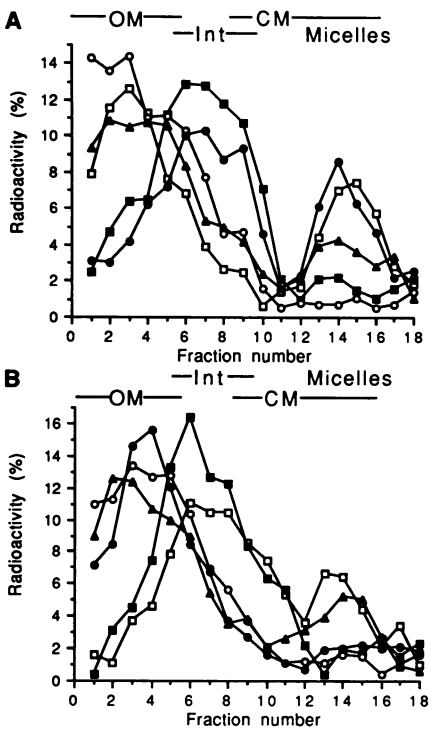

FIG. 6. Fractionation of NlpB and marker lipoproteins by isopycnic sucrose gradient centrifugation. Membrane fractions were mixed and applied to sucrose gradients, which were then centrifuged. Fractions collected from the gradients were electrophoresed on an SDS-polyacrylamide gel, which was then autoradiographed, and the radioactivity in the bands was quantified as described in Materials and Methods. Results are plotted as percent recovery of radioactivity in each fraction. The quantification of radioactivity in two gels that were run in parallel is shown; similar profiles were obtained in several successive experiments. The approximate boundaries of the outer membrane (OM), intermediate-density (Int), and cytoplasmic membrane (CM) vesicles were determined from the protein profile in the Coomassie blue-stained gel (Fig. 5B). The approximate position of protein micelles is also indicated. Symbols: (A) □, PulA:D2N (pullulanase derivative with Asp-to-Asn substitution at position +2 [27]); ■, 50-kDa lipoprotein; ○, NlpB; ●, NlpA-BlaM; ▲, 28-kDa lipoprotein; (B) □, PulA (pullulanase); ■ and ○, same as in panel A; ●, BlaZ; ▲, Lpp-BlaM.

that is definitely present. An A residue, a G, and a C are missing at positions 226, 276, and 435 (positions 1358, 1407, and 1565 in reference 41), respectively. The GC sequence at positions 690 and 691 (1819 and 1820 in reference 41) is switched to CG in the sequence in reference 41. We have carefully checked all of these positions, and we have no doubt that the sequence shown in Fig. 2 is correct. Since a null mutation in *nlpB* is cryptic, it is possible that some of these discrepancies are due to undetected mutations in the DNA cloned by Tiedeman et al. (41).

The sequence shown in Fig. 2 indicates a quite compact genetic organization. The *nlpB* coding sequence starts only 16 bp after the *dapA* termination codon, and both genes are cotranscribed, although the *nlpB* gene appears to be also partially expressed from a secondary promoter located in the *dapA* coding sequence. Transcription at these two promoters is not regulated by the concentration of lysine in the growth medium. This was already known for the promoter of *dapA*, which codes for the first enzyme of the lysine-diaminopimelate-specific biosynthetic pathway, with regulation at that step being exerted only at the level of the activity of dihydrodipicolinate synthetase, the *dapA* product (36). A potential rho-independent transcription termination signal is present 6 bp downstream of the *nlpB* translation stop codon. This transcription terminator overlaps the *purC* promoter in such a way that presumably only 15 bp in that region are not transcribed (between the run of T's at positions 1098 to 1103 and the *purC* transcription start at position 1119). This density of genetic information is reinforced by the presence of a putative binding site for the *purR* product (37) around the *purC* transcription initiation site. Conversely, the *purC* transcript contains an unusually long 164-nucleotide noncoding leader sequence whose role remains to be investigated.

The function of NlpB is unclear. It is not involved in the uptake of lysine or diaminopimelate, and it is not required for lysine-mediated retroinhibition of the dihydropicolinate synthetase (33, 36). Therefore, the genetic arrangement of *nlpB* and *dapA* in the same operon could be fortuitous and could have no physiological consequences, since the *dapA* promoter is constitutively expressed (36). NlpB, like NlpA (47), is not essential for growth.

As predicted from its amino-terminal sequence, the *nlpB* product is a lipoprotein. It can be labeled with radioactive palmitate (Fig. 3), and it is processed from a larger precursor by the lipoprotein signal peptidase, as demonstrated by the effect of globomycin (Fig. 4). Subcellular fractionation data show this protein to be located in the outer membrane (Fig. 5A and 6). The *nlpB* product is thus probably not the same as NLP3, a previously-described lipoprotein of similar size which was found associated with the cytoplasmic membrane (12). Only two of the many lipoproteins of gram-negative bacteria that have been characterized so far, NlpA from *E. coli* (48, 50) and the *K. oxytoca* enzyme pullulanase when it is produced in the absence of its cognate secretion factors (references 27 and 30 and Table 1), are known not to be located in the outer membrane. Each of these proteins have an Asp residue at position +2 of its mature sequence (Table 1). This residue appears to be critical for sorting to those parts of the envelope that form intermediate-density vesicles; its replacement by Ser, Asn, or Glu causes these proteins to be sorted to the outer membrane (27, 48). None of the other characterized lipoproteins have an Asp residue at position +2, but NlpB has an Asp residue at position +4. Our data therefore suggest that Asp must be located at position +2 (or possibly +3) to constitute (part of) the lipoprotein sorting signal. It would be of interest to determine the corresponding sequence of the 50-kDa lipoprotein that also appears to be located in intermediate-density vesicles (Fig. 5A).

It is worth noting that cytochrome *c* of *Rhodopseudomonas viridis* is located in the cytoplasmic (photosynthetic) membrane and yet does not have an Asp residue within the first 10 residues of its mature sequence (Table 1). It may be that lipoprotein sorting signals are different in this bacterium or that there are distinct signals for different regions of the cell envelope. According to our data, *E. coli* does not have lipoproteins that fractionate with cytoplasmic membrane vesicles (Fig. 5A and B and references 27 and 30). This conflicts with data presented earlier which show that there are several lipoproteins in cytoplasmic membrane-derived vesicles (12) and with data showing that NlpA is located in cytoplasmic membrane vesicles (47, 48). This discrepancy may be due to the poor resolution of the intermediate-density fractions from the overlapping outer and cytoplasmic membrane fractions and to the tendency of some lipopro-

teins to be released from the membrane to form protein micelles when cells are lysed.

ACKNOWLEDGMENTS

We are grateful to Catherine Richaud and to Olivier Raibaud for helpful discussions and expert advice, to Joseph Gennity for pKY201 and pJG311, and to J. Oliver Lampen for pRW83.

This work was supported by grants from the Centre National de la Recherche Scientifique (UA136 and UA1149), the Institut Pasteur, and the Institut National de la Santé et de la Recherche Médicale (841025).

REFERENCES

1. **Anderson, B. E., B. R. Baumstark, and W. J. Bellini.** 1988. Expression of the gene encoding the 17-kilodalton antigen from *Rickettsia rickettsii*: transcription and posttranslational modification. J. Bacteriol. **170**:4493–4500.
2. **Bouvier, J., C. Richaud, F. Richaud, J. C. Patte, and P. Stragier.** 1984. Nucleotide sequence and expression of the *Escherichia coli dapB* gene. J. Biol. Chem. **259**:14829–14834.
3. **Cassan, M., C. Parsot, G. N. Cohen, and J. C. Patte.** 1986. Nucleotide sequence of *lysC* gene encoding the lysine-sensitive aspartokinase III of *Escherichia coli* K12. J. Biol. Chem. **261**:1052–1057.
4. **Chen, R., and U. Henning.** 1987. Nucleotide sequence of the gene for the peptidoglycan-associated lipoprotein of *Escherichia coli* K12. Eur. J. Biochem. **163**:73–77.
5. **Cornelis, P., A. Bouia, A. Belarbi, A. Guyonvarch, B. Kammerer, V. Hannaert, and J. C. Hubert.** 1989. Cloning and analysis of the gene for the major outer membrane lipoprotein from *Pseudomonas aeruginosa*. Mol. Microbiol. **3**:421–428.
6. **Deich, R. A., B. J. Metcalf, C. W. Finn, J. E. Farley, and B. A. Green.** 1988. Cloning of genes encoding a 15,000-dalton peptidoglycan-associated membrane lipoprotein and an antigenically related 15,000-dalton protein from *Haemophilus influenzae*. J. Bacteriol. **170**:489–498.
7. **D'Enfert, C., and A. P. Pugsley.** 1989. *Klebsiella pneumoniae pulS* gene encodes an outer membrane lipoprotein required for pullulanase secretion. J. Bacteriol. **171**:3673–3679.
8. **Gilson, E., J. M. Clément, D. Perrin, and M. Hofnung.** 1987. Palindromic units: a case of highly repetitive DNA sequences in bacteria. Trends Genet. **3**:226–230.
9. **Haziza, C., P. Stragier, and J. C. Patte.** 1982. Nucleotide sequence of the *asd* gene of *Escherichia coli*: absence of a typical attenuation signal. EMBO J. **1**:379–384.
10. **Heusterspreute, M., V. Ha Thi, S. Emery, S. Tournis-Gamble, N. Kennedy, and J. Davison.** 1985. Vectors with restriction banks. IV. pJRD184, a 3793-bp plasmid vector having 43 unique cloning sites. Gene **39**:299–304.
11. **Hussain, M., S. Ichihara, and S. Mizushima.** 1980. Accumulation of glyceride-containing precursor of the outer membrane lipoprotein in the cytoplasmic membrane of *Escherichia coli* treated with globomycin. J. Biol. Chem. **255**:3707–3712.
12. **Ichihara, S., M. Hussain, and S. Mizushima.** 1981. Characterization of new membrane lipoproteins and their precursors of *Escherichia coli*. J. Biol. Chem. **256**:3125–3129.
13. **Jalajakumari, M. B., A. Guidolin, H. J. Buhk, P. A. Manning, L. M. Ham, A. L. M. Hodgson, K. C. Cheah, and R. A. Skurray.** 1987. Surface exclusion genes *traS* and *traT* of the F sex factor of *Escherichia coli* K-12. Determination of the nucleotide sequence and promoter and terminator activities. J. Mol. Biol. **198**:1–11.
14. **Jung, J. U., C. Gutierrez, and M. R. Villarejo.** 1989. Sequence of an osmotically inducible lipoprotein gene. J. Bacteriol. **171**: 511–520.
15. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
16. **Kornacker, M. G., and A. P. Pugsley.** 1990. Molecular characterization of *pulA* and its product, pullulanase, a secreted enzyme of *Klebsiella pneumoniae* UNF5023. Mol. Microbiol. **4**:73–85.
17. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
18. **Médigue, C., J. P. Bouché, A. Hénaut, and A. Danchin.** 1990. Mapping of sequenced genes (700kbp) in the restriction map of the *Escherichia coli* chromosome. Mol. Microbiol. **4**:169–187.
19. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Nakamura, K., and M. Inouye.** 1979. DNA sequence of the gene for the outer membrane lipoprotein of *Escherichia coli*: an extremely AT-rich promoter. Cell **18**:1107–1117.
21. **Neidhardt, F. C., V. Vaughn, T. A. Phillips, and P. L. Bloch.** 1983. Gene-protein index of *Escherichia coli* K-12. Microbiol. Rev. **47**:231–284.
22. **Neugebauer, K., R. Sprengel, and H. Schaller.** 1981. Penicillinase from *Bacillus licheniformis*: nucleotide sequence of the gene and implications for the biosynthesis of a secretory protein in a Gram-positive bacterium. Nucleic Acids Res. **9**:2577–2588.
23. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of β-lactamases by using a chromogenic cephalosporin substrate. Antimicrob. Agents Chemother. **1**:283–288.
24. **Parker, J.** 1984. Identification of the *purC* gene product of *Escherichia coli*. J. Bacteriol. **157**:712–717.
25. **Pugsley, A. P.** 1988. The immunity and lysis genes of ColN plasmid pCHAP4. Mol. Gen. Genet. **211**:335–341.
26. **Pugsley, A. P., C. Chapon, and M. Schwartz.** 1986. Extracellular pullulanase of *Klebsiella pneumoniae* is a lipoprotein. J. Bacteriol. **166**:1083–1088.
27. **Pugsley, A. P., and M. G. Kornacker.** Secretion of the cell surface lipoprotein pullulanase in *Escherichia coli*. Cooperation or competition between the specific secretion pathway and the lipoprotein sorting pathway. J. Biol. Chem., in press.
28. **Pugsley, A. P., M. G. Kornacker, and A. Ryter.** 1990. Analysis of the subcellular location of pullulanase produced by *Escherichia coli* carrying the *pulA* gene from *Klebsiella oxytoca* strain UNF5023. Mol. Microbiol. **4**:59–72.
29. **Pugsley, A. P., and B. Oudega.** 1987. Methods for studying colicins and their plasmids, p. 105–161. *In* K. Hardy (ed.), Plasmids, a practical approach. IRL Press, Oxford.
30. **Pugsley, A. P., I. Poquet, and M. G. Kornacker.** 1991. Two distinct steps in pullulanase secretion by *Escherichia coli* K12. Mol. Microbiol. **5**:865–873.
31. **Pugsley, A. P., and C. A. Schnaitman.** 1979. Factors affecting the electrophoretic mobility of the major outer membrane proteins of *Escherichia coli* in polyacrylamide gels. Biochim. Biophys. Acta **581**:163–178.
32. **Raibaud, O., M. Mock, and M. Schwartz.** 1984. A technique for integrating any DNA fragment into the chromosome of *Escherichia coli*. Gene **29**:231–241.
33. **Richaud, C., J. Bouvier, and P. Stragier.** Unpublished data.
34. **Richaud, C., W. Higgins, D. Mengin-Lecreulx, and P. Stragier.** 1987. Molecular cloning, characterization, and chromosomal localization of *dapF*, the *Escherichia coli* gene for diaminopimelate epimerase. J. Bacteriol. **169**:1454–1459.
35. **Richaud, C., F. Richaud, C. Martin, C. Haziza, and J. C. Patte.** 1984. Regulation of expression and nucleotide sequence of the *Escherichia coli dapD* gene. J. Biol. Chem. **259**:14824–14829.
36. **Richaud, F., C. Richaud, P. Ratet, and J.-C. Patte.** 1986. Chromosomal location and nucleotide sequence of the *Escherichia coli dapA* gene. J. Bacteriol. **166**:297–300.
37. **Rolfes, R. F., and H. Zalkin.** 1988. *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis: cloning, nucleotide sequence and interaction with the *purF* operator. J. Biol. Chem. **263**:19653–19661.
38. **Shapira, S. K., J. Chou, F. Richaud, and M. J. Casadaban.** 1983. New versatile vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzymatically active carboxy-terminal portion of β-galactosidase. Gene **26**:71–82.
39. **Stragier, P., F. Richaud, F. Borne, and J. C. Patte.** 1983. Regulation of diaminopimelate decarboxylase synthesis in *Esch-*

erichia coli. I. Identification of a *lysR* gene encoding an activator of the *lysA* gene. J. Mol. Biol. **168**:307–320.

40. **Takase, I., F. Ishino, M. Wachi, H. Kamata, M. Doi, S. Asoh, H. Matsuzawa, T. Ohta, and M. Matsuhashi.** 1987. Genes encoding two lipoproteins in the *leuS-dacA* region of the *Escherichia coli* chromosome. J. Bacteriol. **169**:5692–5699.
41. **Tiedeman, A. A., D. J. DeMarini, J. Parker, and J. M. Smith.** 1990. DNA sequence of the *purC* gene encoding 5'-phosphoribosyl-5-aminoimidazole-4-*N*-succinocarboxamide synthetase and organization of the *dapA-purC* region of *Escherichia coli* K-12. J. Bacteriol. **172**:6035–6041.
42. **Vidal-Ingigliardi, D., and O. Raibaud.** 1985. A convenient technique to compare the efficiency of promoters in *Escherichia coli*. Nucleic Acids Res. **16**:5919–5926.
43. **Weyer, K. A., W. Schäfer, F. Lottspeich, and H. Michel.** 1987. The cytochrome C subunit of the photosynthetic reaction center from *Rhodopseudomonas viridis* is a lipoprotein. Biochemistry **27**:2909–2914.
44. **Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker.** 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. J. Bacteriol. **161**:1219–1221.
45. **Woods, J. P., S. M. Spinola, S. M. Strobel, and J. G. Cannon.** 1989. Conserved lipoprotein H.8 of pathogenic *Neisseria* consists entirely of peptapeptide repeats. Mol. Microbiol. **3**:43–50.
46. **Wu, H. C.** 1987. Posttranslational modification and processing of membrane proteins in bacteria, p. 37–71. *In* M. Inouye (ed.), Bacterial outer membranes as model systems. Academic Press, Inc., New York.
47. **Yamaguchi, K., and M. Inouye.** 1988. Lipoprotein 28, an inner membrane protein of *Escherichia coli* encoded by *nlpA*, is not essential for growth. J. Bacteriol. **170**:3747–3749.
48. **Yamaguchi, K., F. Yu, and M. Inouye.** 1988. A single amino acid determinant of the membrane localization of lipoproteins in E. coli. Cell **53**:423–432.
49. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
50. **Yu, F., S. Inouye, and M. Inouye.** 1986. Lipoprotein-28, a cytoplasmic membrane lipoprotein from *Escherichia coli*. Cloning, DNA sequence, and expression of its gene. J. Biol. Chem. **261**:2284–2288.