

Escherichia coli prlC Encodes an Endopeptidase and Is Homologous to the *Salmonella typhimurium opdA* Gene

CHRISTOPHER A. CONLIN,¹ NANCY J. TRUN,^{2†} THOMAS J. SILHAVY,²
AND CHARLES G. MILLER^{1*}

Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801,¹ and
Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544²

Received 13 April 1992/Accepted 14 July 1992

Mutations at the *Escherichia coli prlC* locus suppress the export defect of certain *lamB* signal sequence mutations. The *Salmonella typhimurium opdA* gene encodes an endoprotease that can participate in the catabolism of certain peptides and is required for normal development of phage P22. Plasmids carrying either the wild-type (pTC100 *prlC*⁺) or suppressor alleles of *prlC* complemented all phenotypes associated with an *S. typhimurium opdA* mutation. A plasmid carrying an amber mutation in *prlC* [*prlC31*(Am)] was unable to complement except in an amber suppressor background. Tn1000 insertions which eliminated the ability of pTC100 (*prlC*⁺) to complement *opdA* mapped to the region of the plasmid shown by deletion analysis and subcloning to contain *prlC*. The nucleotide sequence of a 2.7-kb fragment including this region was determined, revealing an open reading frame encoding a 77-kDa protein. The sequences of this open reading frame and its putative promoter region were very similar (84% nucleotide sequence identity and 95% amino acid identity) to those of *S. typhimurium opdA*, showing that these genes are homologs. The nucleotide sequence of the *prlC1* suppressor allele was determined and predicts an in-frame duplication of seven amino acids, providing further confirmation that the *prlC* suppressor phenotype results from changes in the endopeptidase *OpdA*.

One strategy used to identify components of the protein secretion apparatus of cells has been to isolate mutations which allow export of proteins with defective signal sequences. One class of suppressor mutations isolated in this way affects the *prlC* locus (7). The first *prlC* suppressor was selected as a mutation which allowed a strain carrying a 36-bp deletion (*lamBs60*) in the *lamB* signal sequence to grow on maltodextrins. LamB is an outer membrane protein that serves as the phage lambda receptor and is required for uptake of maltodextrins. Mutations at *prlC* are rare among *lamBs60* revertants; only one allele (*prlC1*) was obtained in the original selection, which produced mainly *prlA* mutations (7). More recently, Trun and Silhavy (19) have used localized mutagenesis to isolate additional *prlC* alleles. One of the strongest of the *prlC* suppressor mutations, *prlC8*, obtained in this way has been characterized in detail. Although strains carrying *prlC8* efficiently remove the mutant *lamBs78* signal sequence, very little of the LamB protein is exported to the outer membrane (20). Trun and Silhavy (19) have cloned wild-type, suppressor, and amber alleles of *prlC* and shown that *prlC* is dispensable for growth. The *prlC* locus has been mapped to approximately 71 map units on the *Escherichia coli* chromosome (19).

The *opdA* gene was first identified by Vimr et al. (22) as the locus of mutations which prevented a peptidase-deficient *Salmonella typhimurium* strain from using *N*-acetyl-L-Ala₄ (AcAla₄) as a sole nitrogen source. The *opdA* gene has been mapped to 76 min on the *S. typhimurium* chromosome by cotransduction with transposon insertions near *asd*. It has been cloned and sequenced and shown to encode a 680-amino-acid protein (4). Oligopeptidase A, the product of the *opdA* gene, is an endoprotease which has been purified from

both *E. coli* and *S. typhimurium* (12). Oligopeptidase A is the major soluble enzyme in *E. coli* which is able to hydrolyze free lipoprotein signal peptide in vitro (12).

Several facts suggested that *prlC* and *opdA* might be related. *prlC* and *opdA* map in the same region of the chromosome. The molecular masses of their gene products are identical (68 kDa), as determined by denaturing polyacrylamide gel electrophoresis. In addition, the results of Novak and Dev (11) indicate that oligopeptidase A can recognize a signal sequence, although it can cleave this sequence only after the release of signal sequence from the precursor protein. These observations led us to investigate the relationship between the two loci. The results presented in this article show that *E. coli prlC* and *S. typhimurium opdA* are indeed homologous genes.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. All *Salmonella* strains used were derivatives of *S. typhimurium* LT2. *E. coli* DH5 (9) was used to cultivate plasmids for sequencing, and *E. coli* NT113 [F⁻ *araD139* Δ(*argF-lac*)U139 *rpsL150 relA1 ffbB5301 ptsF25 deoC1 lamBs60 recA56 srlC300::Tn10*] (19) was used to test for *lamBs60* suppression. Bacteria were routinely grown in LB medium (14) at 37°C. Plasmids and other markers were transferred between *S. typhimurium* strains by using the generalized transducing phage P22 (HT12/4 *int103*) (13). Plasmids pTA108, pTC100 (*prlC*⁺), pTC101, pTC102, pTC105, pTC107, pTC108, pTC111, pTC112, and pTC118 were described by Trun and Silhavy (19). Plasmids pTC103, pTC106, pTC109, pTC113, pTC114, and pTC115 were obtained by exchange of chromosomal *prlC* suppressor alleles with *prlC*⁺ in pTC100 (18, 19). Plasmid pTC231 is a pTC101 derivative carrying the *prlC* gene with both the *prlC1* allele and the *prlC31* amber mutation. This amber mutation elim-

* Corresponding author.

† Present address: Laboratory of Molecular Biology, National Cancer Institute, Bethesda, MD 20892.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or description
<i>S. typhimurium</i>	
TN1201 <i>dcp-1 zcf-845::Tn10 zhg-848::Tn5 opdA1</i>
TN1292 <i>dcp-1 zcf-845::Tn10 zhg-848::Tn5</i>
TN1727 <i>leuBCD485 pepN90 pepA16 pepB11 supQ302(AproAB pepD) pepP1 pepQ1 pepT1 dcp-1 zcf-845::Tn10 zhg-848::Tn5 opdA1</i>
TN3031TN1727/pJG13
TN3032TN1727/pJG70
TN3033TN1727/pTC100
TN3034TN1727/pTC108
TN3035TN1727/pTC231
TN3067 <i>leuA414(Am) hisC527(Am) supJ60 zhg-848::Tn5 opdA1</i>
TN3072TN3067/pTC231
TN3334 <i>leuBCD opdA10::MudJ dcp-102::MudJ(lacZ::Tn10dCam)</i>
TN3392 <i>purC7 recA1 srl-202::Tn10/F'42 finP30/pTC100</i>
TN3393 <i>purC7 recA1 srl-202::Tn10/F'42 finP30</i>
TN3948TN3334/pTC100
TN3949TN3334/pTC102
TN3950TN3334/pTC108
TN3951TN3334/pTC112
TN3952TN3334/pTC118
TN3953TN3334/pTC101
TN3954TN3334/pTC103
TN3955TN3334/pTC105
TN3956TN3334/pTC106
TN3957TN3334/pTC107
TN3958TN3334/pTC109
TN3959TN3334/pTC111
TN3960TN3334/pTC113
TN3961TN3334/pTC114
TN3962TN3334/pTC115
Plasmids	
pTC100 <i>prlC</i> ⁺ in pSC101-based vector pTA108
pTC101As pTC100 but with <i>prlC1</i>
pTC102As pTC100 but with <i>prlC2</i>
pTC103As pTC100 but with <i>prlC3</i>
pTC105As pTC100 but with <i>prlC5</i>
pTC106As pTC100 but with <i>prlC6</i>
pTC107As pTC100 but with <i>prlC7</i>
pTC108As pTC100 but with <i>prlC8</i>
pTC109As pTC100 but with <i>prlC9</i>
pTC111As pTC100 but with <i>prlC11</i>
pTC112As pTC100 but with <i>prlC12</i>
pTC113As pTC100 but with <i>prlC13</i>
pTC114As pTC100 but with <i>prlC14</i>
pTC115As pTC100 but with <i>prlC15</i>
pTC118As pTC100 but with <i>prlC18</i>
pTC231As pTC100 but with both <i>prlC1</i> and <i>prlC31(Am)</i>
pTC1692.7-kb <i>prlC1</i> -containing fragment from pTC101 in pTA108
pCM1912.9 kb <i>prlC1</i> -containing fragment from pTC169 in pBR322
pJG13 <i>opdA</i> ⁺ in pBR328
pJG70 <i>opdA</i> ⁺ in pBR328

inated the suppressor activity of the *prlC1* allele (18). Deletion analysis of pTC101 identified a 2.7-kb *HaeIII-PstI* fragment which allowed utilization of maltodextrins as a sole carbon source by NT113 (18). This fragment was subcloned into vector pTA108 to generate pTC169. A 2.9-kb *BamHI* fragment of pTC169 containing *prlC1*, generated by the cleavage of *BamHI* sites which flank the insert in the vector, was cloned into the *BamHI* site of pBR322 to create plasmid pCM191.

Characterization of phenotypes of *prlC* plasmids. The ability of plasmids to suppress the *lamB* signal sequence mutations by allowing NT113 to use maltodextrins as a sole carbon source was determined as previously described (19). Phage P22 forms pinpoint plaques on *opdA* strains (4, 21). Complementation of this small-plaque phenotype was tested by plating P22H5, a clear-plaque mutant, on plasmid-containing strains. An *opdA dcp* strain is unable to use AcAla₄ as a sole nitrogen source or to hydrolyze AcAla₄ (22) or benzyloxycarbonyl Ala-Ala-Leu *p*-nitroanilide (Z-AALpNA) (4). Utilization of AcAla₄ as a sole nitrogen source, AcAla₄ hydrolysis by cell extracts, and Z-AALpNA hydrolysis by cell extracts were determined as previously described (4, 22).

Tn1000 ($\gamma\delta$) insertions in pTC100. Tn1000 insertions in pTC100 were isolated by the method of Guyer (8). Briefly, plasmid pTC100 was introduced into strain TN3393 to make strain TN3392. Strain TN3393 contains F'42, which carries at least one uncharacterized Tn1000 insertion. Four-milliliter samples of exponentially growing cultures of TN3392 and TN1201 were combined and incubated at 37°C without shaking for 1 h. These cells were vortexed vigorously, pelleted by centrifugation, and resuspended in 0.5 ml of LB broth. Samples were plated on LB agar containing both ampicillin (50 μ g/ml) and kanamycin (50 μ g/ml) and incubated overnight at 37°C. Individual transconjugants were patched to LB agar containing ampicillin and kanamycin and replica plated to NN-plus-glucose medium (22) containing 1.5 mM AcAla₄ as the sole nitrogen source. Single colonies of strains which failed to use AcAla₄ were isolated. Plasmid DNA was isolated from these strains, and the locations of the Tn1000 insertions were determined by restriction mapping.

DNA sequencing and manipulations. DNA sequence was determined by using the dideoxynucleotide chain termination method (15), double-stranded plasmid templates (1), and Sequenase (United States Biochemical Corp.) according to the manufacturer's directions. Template DNA was isolated by using Qiagen columns according to the manufacturer's directions. Primers for the plasmid pBR322 sequence adjacent to the *BamHI* site were obtained from New England Biolabs. Primers for sequencing from the Tn1000 insertions were those described by Liu et al. (10). In addition, because of the high degree of homology between *prlC* and *opdA*, oligonucleotide primers which had been synthesized to sequence *opdA* were also used to sequence *prlC*. Restriction enzymes and T4 DNA ligase were purchased from Bethesda Research Laboratories (Life Technologies Inc.) and used according to the manufacturer's directions. Other manipulations were performed by standard techniques (14).

Maxicell analysis. Maxicell analysis was performed as previously described (16), except that cells were irradiated with 150 ergs of UV light per mm².

Nucleotide sequence accession number. The *prlC* sequence has been assigned GenBank accession number M93984.

RESULTS

Complementation of *opdA* by *prlC* plasmids. *Salmonella* strains carrying *opdA* mutations fail to support the normal growth of phage P22. P22 makes tiny plaques on an *opdA* strain (5, 21). This small-plaque phenotype is complemented by *opdA*⁺ plasmids (4). Plasmids carrying wild-type (pTC100) and 14 independent signal sequence mutant suppressor alleles of the *E. coli prlC* locus were introduced into TN1727 (*opdA1*) or TN3334 (*opdA10::MudJ*), which are *S.*

TABLE 2. AcAla₄ hydrolysis by soluble extracts of strains containing *prc* or *opdA* plasmids

Strain (relevant genotype)	Parent	Plasmid	Sp act ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)
TN3031	TN1727	pJG13 (<i>opdA</i> ⁺)	1.3
TN3033	TN1727	pTC100 (<i>prc</i> ⁺)	0.94
TN3034	TN1727	pTC108 (<i>prc</i> 8)	0.44
TN3035	TN1727	pTC231 [<i>prc</i> (Am)]	0.002
TN1292 (<i>opdA</i> ⁺)	None	None	0.02
TN1727 (<i>opdA</i>)	None	None	0.001

typhimurium strains containing *opdA* mutations in the chromosome, to form the strains listed in Table 1. The ability of P22 to plate on these strains was tested. Normal large plaques formed on lawns of TN1727 and TN3334 containing plasmids carrying wild-type or suppressor alleles of *E. coli prc*. However, tiny plaques identical to those formed on TN1727 (no plasmid) were formed on TN3035 [pTC231, *prc*(Am)]. Plasmid pTC231 was introduced into TN3067, which contains both the amber suppressor *supJ60* and *opdA1*, to form TN3072. P22 formed normal large plaques on TN3072. To confirm that the plasmid present in TN3072 still carried the amber mutation, it was transduced from TN3072 back into TN1727 (which contains no suppressor), where it once again failed to complement. These results show that the defect for phage P22 development conferred by an *opdA* mutation in *S. typhimurium* can be overcome by the presence of a plasmid carrying wild-type and suppressor alleles of *E. coli prc* but not by the same plasmid carrying a *prc* amber mutation. Thus, *prc*, or possibly a gene downstream from *prc* in an operon, can supply the function lost as a result of a mutation in *S. typhimurium opdA*.

S. typhimurium strains containing mutations in *opdA* and *dcp* are unable to use AcAla₄ as a sole nitrogen source (dipeptidylcarboxypeptidase, the product of *dcp*, also hydrolyzes AcAla₄) (22). In order to determine whether *E. coli prc* plasmids could restore AcAla₄ utilization to an *opdA* strain, plasmids containing wild-type *prc* (pTC100) and the 14 suppressor alleles were introduced into TN3334. All of these plasmids conferred the ability to use AcAla₄ as a sole nitrogen source. To establish that AcAla₄ utilization was the result of restoring peptidase activity, extracts of strains carrying pTC100 or pTC108 were assayed for activity toward two OpdA substrates, AcAla₄ and Z-AALpNA. These plasmids were introduced into TN1727, a strain carrying chromosomal mutations in *opdA* and *dcp*. The specific AcAla₄-hydrolyzing activity in extracts of these strains was determined (Table 2). Extracts of the strains carrying pTC100 (*prc*⁺) or pTC108 (*prc*8) had much more AcAla₄-hydrolyzing activity than the parent strain without any plasmid. Strain TN3035 containing pTC231 [*prc*1 *prc*31(Am)] had a specific activity equivalent to that of the parent, indicating that the amber mutation which prevents the suppression of the *lamBs60* signal sequence mutation also eliminated AcAla₄-hydrolyzing activity. Extracts of peptidase-deficient strains carrying *prc*⁺ (pTC100) or *opdA*⁺ (pJG70) plasmids (strains TN3033 and TN3032, respectively) were also assayed for the ability to hydrolyze Z-AALpNA, a chromogenic peptide substrate of OpdA (4). Whereas the specific activity of TN1727, the peptidase-deficient host strain without any plasmid, was <0.002 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, the activities of TN3032 (pJG70) and TN3033 (pTC100) were 0.05 and 0.01 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, respectively. Thus, the *prc* plasmids complemented not only the

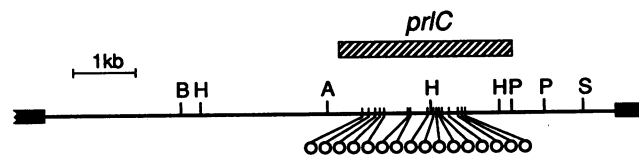


FIG. 1. Location of Tn1000 insertions in pTC100. Tn1000 insertions were generated in pTC100 as described in the text and were localized on the plasmid by restriction mapping and sequencing out from the Tn1000 ends. The insertions, indicated by open circles, eliminated the ability of pTC100 to complement the AcAla₄ utilization defect of the *opdA* strain TN1201. The thin line represents *E. coli* chromosomal DNA on pTC100; the thick lines at the ends indicate vector DNA. The hatched bar above the line indicates the 2.7-kb region to which *prc* was localized by deletion analysis of pTC101 (18). Restriction enzyme sites: A, *Ava*I; B, *Bgl*II; H, *Hpa*I; P, *Pst*I; S, *Sal*I.

P22 growth defect characteristic of *opdA* strains but also the *opdA* AcAla₄ and Z-AALpNA hydrolysis defects. Similarly, the *prc*(Am) mutation, which eliminated LamB signal sequence mutant suppressor activity, also eliminated the ability of the *prc* plasmid to complement both the *opdA* peptide-hydrolyzing and P22 growth defects.

Localization of *opdA*-complementing activity and *prc* suppressor activity on pTC100. To localize the region of pTC100 required for complementation of *opdA*, Tn1000 insertions in pTC100 were isolated, and the resulting plasmids were screened for their ability to complement the AcAla₄ utilization defect conferred by *opdA*. Of 191 insertions characterized, 30 failed to complement the *opdA* AcAla₄ utilization defect. Sixteen of these 30 insertions were localized by restriction enzyme analysis and found to lie in a 1.4-kb region of pTC100 (Fig. 1). This 1.4-kb region was within the 2.7-kb region which was shown by deletion analysis of pTC101 to contain *prc* (18). A 2.9-kb *Bam*HI fragment of pTC169 containing the 2.7-kb region shown to encode *prc*1 was subcloned into pBR322. This plasmid, pCM191, complemented the *opdA* small-plaque defect in TN1201 and suppressed the *lamBs60* signal sequence mutation by allowing growth of NT113 on maltodextrins. These results indicated that the 2.7-kb fragment contains all of the information necessary for both *opdA* complementation and suppressor activity.

Nucleotide sequence of *prc*. The deletion analysis and subcloning described above identified a 2.7-kb region which must contain the *prc* gene (Fig. 1). Of this 2.7-kb region, 2.6 kb was sequenced. The wild-type *prc* gene from pTC100 was sequenced with primers to the Tn1000 insertions which eliminated *opdA* complementation and primers originally synthesized for sequencing *S. typhimurium opdA*. Additional flanking sequence was obtained from pCM191 by using primers to the vector sequence. The complete sequence is shown in Fig. 2. This sequence contains an open reading frame (ORF) (positions 337 to 2377) encoding a 680-amino-acid protein. This ORF is preceded by a possible ribosome binding site. The region from nucleotide 243 to 2578 showed 84% identity to the *opdA* sequence. The predicted amino acid sequence of the 680-amino-acid protein showed 95% identity to the predicted amino acid sequence of OpdA (Fig. 2). Nucleotide sequence similarity extended 3' of the ORF to the end of the *opdA* sequence. Sequence similarity also extended 5' to the coding region until immediately after a putative σ^{32} promoter which is present in both the *opdA* and the *prc* sequences (Fig. 2) (4). (Experiments are in progress to determine whether *prc* and *opdA* tran-

<u>prlC</u>	CCCCTGCGTGGGTGTCGAGATAGAGAAACGGTTT	34
<u>opdA</u>	.G.TGC.GAACCC....CC.GC.CAGCGTCAGGC	
<u>prlC</u>	ATCTTTCTCTTTCAGCGACTCGATGATCAGGCTCTGAACGGTATGTTTAAGGACGTGGCGTGGTTCGCCAGCGTAAAGCTGTGGCGATAACTGAGCATG	134
<u>opdA</u>	G.TAA.GGAAGAG.CG.CG.TACC.CGAG.ATATCCGTT.TCGG.CGAG.CCGA.GATCAACA...CTGGCAGCGTCT.GATCT.TCCGCTGCTTA.GA.	
<u>prlC</u>	GGGTAAAGGTGTTCCGGTAAGTAAAAATCGGCCGTATTTCGGGGCGCACAAAAGCGTATCAGGACAGTATACCGAAAAGAGGCCCGCCGCGAAAGCGT	234
<u>opdA</u>	AT.GCGA.AAAAA.GC.CTGACGTTGTACC.GGC.TA.AAAACGTA.GCCGG.ATG..GTTA.AG.GGCGGCTT.TT.CA..GATAC..A..TCT.TAC	
<u>prlC</u>	AACGTTTTCATTGAAATTCACCTACACTTAACCCCATGCTACA-CACATTATGTAACGCCTGTGAGCGCTTCCTTAACTCTTTAACCGACTGCGC	333
<u>opdA</u>	GCT.TCCA.....C.CTT...ATC.....T...G.T.T...CGC.AG.G....AAAA.C.....C.AC.C...TC..CTA.....	
<u>OpdA</u>Ser...Ser.....Ala...Lys.....Ala.....	
<u>PrlC</u>	MetThrAsnProLeuLeuThrProPheGluLeuProProPheSerLysIleLeuProGluHisValValProAlaValThrLysAlaLeuAsnAspCys	
<u>prlC</u>	GAATGACGAATCCGTTACTGACTCCCTTTGAATTGCCTCGTTCCTAAAATTCTCCCGAACATGTCGTTCCAGCCGTGACTAAGGCATTGAACGACTG	433
<u>opdA</u>	AT....C....A....A..GT.T...TC.C....G..T....GC....AAA....G....G..G..T..G..C..C..A..G....GC...T..	
<u>OpdA</u>	...AlaAla.....Gly.....His.....Ser.....Ala.....	
<u>PrlC</u>	ArgGluAsnValGluArgValValAlaGluGlyAlaProTyrThrTrpGluAsnLeuCysGlnProLeuAlaGluValAspAspValLeuGlyArgIle	
<u>prlC</u>	CCCGAAAACGTTGAGGCGCTAGTAGCCGAAAGGGGACCGTACACCTGGGAAAATCTCTGCCAGCCGCTGGCGGAAGTGACGATGTTGGGGCGTATC	533
<u>opdA</u>	...G.CGGC..A..AG...T..G..GC.T..C..G.....T.G.....C.....T.....CC.....TC.....T	
<u>OpdA</u>Ile.....	
<u>PrlC</u>	PheSerProValSerHisLeuAsnSerValLysAsnSerProGluLeuArgGluAlaTyrGluGlnThrLeuProLeuLeuSerGluTyrSerThrTrpVal	
<u>prlC</u>	TTCTCCCGGTGAGCCACTGAACTCGGTGAAAAATAGCCGAACTCGGTGAAGCTACGAACAAACCTGCCGCTGTGCGAATACAGCACCTGGG	633
<u>opdA</u>G..AA.T.....T.A.....T.....G.....G.....G..G.....T.....G..G.....	
<u>OpdA</u>Asn.....	
<u>PrlC</u>	GlyGlnHisGluGlyLeuTyrLysAlaTyrArgAspLeuArgAspGlyAspHisTyrAlaThrLeuAsnThrAlaGlnLysLysAlaValAspAsnAla	
<u>prlC</u>	TAGGGAACATGAAGGCTGTATAAAGCGTATCGCGACCTGCGGATGGCGATTACGCCACGCTGAACACGGCGCAGAAAGCGGTTGATAACGC	733
<u>opdA</u>	.T..C..G.....T...C..C.....C.....C.....T.....T..C.....	
<u>OpdA</u>Ile.....	
<u>PrlC</u>	SerAsnAsnValLeuAspAlaThrMetGlyTrpThrLysLeuValThrAspGluAlaGluLeuAlaGlyMetProGluSerAlaLeuAlaAlaLysAla	
<u>prlC</u>	AGCAACAACGCTCGATGCGACAATGGGCTGGACAAACTCGTTACCGACGAAAGCGGAGCTGGCGGGATGCCAGAAAGCGGCTGGCTGCGCAACG	933
<u>opdA</u>T..T..G.....C..T.....G..G..A.C.....T.....C.....A....G.....T.....G..A..T...	
<u>OpdA</u>Gln.....LeuLeu.....	
<u>PrlC</u>	GlnAlaGluAlaLysGluLeuGluGlyTyrPheValThrLeuAspIleProSerTyrLeuProValMetThrTyrCysAspAsnGlnAlaLeuArgGlu	
<u>prlC</u>	CCCAGCCGAAAGCGAAGGCTGGAAGTTATTTCTGACGCTGGATATCCCAAGCTACTTGGCGGATGACCTACTGCGACAACCGGCTGCGTGA	1033
<u>opdA</u>G....C....A....A.....CC.GC....C.....G.....TC.....C.....GT.....C....	
<u>OpdA</u>Pro.....LeuArg.....	
<u>PrlC</u>	GluMetTyrArgAlaTyrSerThrArgAlaSerAspGlnGlyProAsnAlaGlyLysTrpAspAsnSerLysValMetGluGluIleLeuAlaAlaAla	
<u>prlC</u>	AGAGATGTATCGCGCTTACAGCACCGCGCTCCGATCAAGCCCAACCGCGTAAATGGGATAACAGCAAGGTGGAAGAGATCCTCGCCGCTGC	1132
<u>opdA</u>C...TTC...T.....G..G..G.....C.....CC.....A...T..T.....G	
<u>OpdA</u>HisGlu.....	
<u>PrlC</u>	HisGluLeuAlaGlnLeuLeuGlyPheGluAsnTyrAlaPheLysSerLeuAlaThrLysMetAlaGluAsnProGlnGlnValLeuAspPheLeuThr	
<u>prlC</u>	TCACGAACTGGCGCAACTGCTGGGCTTTGAAACTACGCTTTAAATCCCTTGCCACTAAAATGGCAGAAAACCGCAGCAGGTGCTGGATTCTTAACC	1232
<u>opdA</u>T..T.....C....T..T...CA.G...G..G....C.....G.....T.....C.....C.....	
<u>OpdA</u>Glu.....	
<u>PrlC</u>	AspLeuAlaLysArgAlaArgProGlnGlyGluLysGluLeuAlaGlnLeuArgAlaPheAlaLysAlaGluPheGlyValAspGluLeuGlnProTrpAsp	
<u>prlC</u>	GATCTGGCAAAACGCGCGCTCCACAAGGGGAAAAAGAGCTGGCGCAATTGCGTGCTTTGCCAAAGCCGAATTTGGCGTCTGATGAGTGCAGCCGTTGG	1332
<u>opdA</u>	...T...G.....C....G..G..G.....C..GC....C....C.....T..G...C...A.....	
<u>OpdA</u>IleAlaTyrTyrSerGluLysGlnLysGlnHisLeuTyrSerIleSerAspGluGlnLeuArgProTyrPheProGluAsnLysAlaValAsnGlyLeu	
<u>PrlC</u>	ATATCGCTTACTACAGCGAAAAACAAAACAGCACCTTACAGCATCAGTGACGAACAGCTGCGTCCGTACTCCCGAAAACAAAGCGGTTAACGGCCT	1432
<u>opdA</u>G.....T.....G.....G.....C.....C.....T..G...A.....C...G..T.....	
<u>OpdA</u>Thr.....Glu.....	
<u>PrlC</u>	PheGluValValLysArgIleTyrGlyIleThrAlaLysGluArgLysAspValAspValTrpHisProHisValArgPhePheGluLeuTyrAspGlu	
<u>prlC</u>	GTGGAAGTGGTAAAGCGTATTACGGCATCACCGCTAAAGAGCGTAAAGATGTTGATGTCTGGCATCCGATGTACGTTCTTCGAACTGTATGACGAA	1532
<u>opdA</u>G..A.....T..C.....CT..C.....C...G..A..G.....T.....	
<u>OpdA</u>Ser.....His.....	
<u>PrlC</u>	AsnAsnGluLeuArgGlyThrPheTyrLeuAspLeuTyrAlaArgGluAsnLysArgGlyGlyAlaTrpMetAspAspCysValGlyGlnMetArgLysAla	
<u>prlC</u>	AATAACGAACTGCGCGTACGTTCTACCTCGATCTGTATGCCCGTAAAAACAAGCGCGGGGGCGTGGATGGATGACTGCGTAGGCCAGATGCGTAAAG	1632
<u>opdA</u>G.....C.GC....T..C.....C..G..C...C...A.....C.....T..C.....	

FIG. 2. Comparison of *prlC* and *opdA* nucleotide and deduced amino acid sequences. The *prlC* and *opdA* nucleotide and predicted amino acid sequences are shown aligned with each other. Only bases or amino acids which differ between the two are indicated. The overlined *prlC* nucleotide sequence indicates the putative σ^{32} promoter. The underlined *prlC* nucleotide sequence indicates the region which is duplicated in *prlC1*.

<i>OpdA</i>Thr.....	
<i>PrlC</i>	AspGlySerLeuGlnLysProValAlaTyrLeuThrCysAsnPheAsnArgProValAsnGlyLysProAlaLeuPheThrHisAspGluValIleThr	
<i>prlC</i>	CTGATGGTCTTTGCCAAAACCGGTCCGGTATTTGACTTGAACCTCAACCGCCGGTAAATGGTAAACCGCGCTGTCTACTCACGACGAAGTGATCAC	1732
<i>opdA</i>	.G..C..CA.A.....G.....T...C...C.....T.....T....G..C.....C..T....T..C..T.....	
<i>OpdA</i>	
<i>PrlC</i>	LeuPheHisGluPheGlyHisGlyLeuHisHisMetLeuThrArgIleGluThrAlaGlyValSerGlyIleSerGlyValProTrpAspAlaValGlu	
<i>prlC</i>	CCTCTTCCACGAGTTCGGTCACGGCTGCACCATATGCTGACCCGCATCGAAACCGCTGGTGTTCGGTATCAGCGGTGTGCCGTGGGATCGGTCGAA	1832
<i>opdA</i>	...G.....T....T.....T.....T...G....C..G..C.....C.....C.....	
<i>OpdA</i>Ala.....	
<i>PrlC</i>	LeuProSerGlnPheMetGluAsnTrpCysTrpGluProGluAlaLeuGlyPheIleSerGlyHisTyrGluThrGlyGluProLeuProLysGluLeuLeu	
<i>prlC</i>	CTGCCGAGTCAGTTTATGAAAACCTGGTCTGGGAGCCGGAGCCGCTGGGTTTATCTCTGGTCACTATGAAACCGGGAACCGCTGCCGAAAGAGTTGC	1932
<i>opdA</i>A.....	
<i>OpdA</i>Asn.....	
<i>PrlC</i>	AspLysMetLeuAlaAlaLysAsnTyrGlnAlaAlaLeuPheIleLeuArgGlnLeuGluPheGlyLeuPheAspPheArgLeuHisAlaGluPheArg	
<i>prlC</i>	TGGATAAAATGCTGCCGCGAAGAACCACAGCGCGGCTGTTTATTCTGCGTCAGTGGAGTTCGGCTGTTGATTCCGCCCTTATGCCGAGTCCG	2032
<i>opdA</i>A.....T.....T.....T.....T...G....G..A..TAA	
<i>OpdA</i>Gln.....Phe.....Gln.....Thr.....	
<i>PrlC</i>	ProAspGlnGlyAlaLysIleLeuGluThrLeuAlaGluIleLysLysLeuValAlaValValProSerProSerTrpGlyArgPheProHisAlaPhe	
<i>prlC</i>	CCCGATCAGGGGCAAAAATCCTCGAAACTCTGCCAGAAATCAAGAACTGGTTCGGTGGTCCATCTCCGTCTGGGCGCTTCCCGCACGCTTTC	2132
<i>opdA</i>	T..AC..G..A..A..G....T..T..G..G..CTTT....T..A...A...C.....G..A...A.A....T..C..T..A..T..G...	
<i>OpdA</i>Tyr.....	
<i>PrlC</i>	SerHisIlePheAlaGlyGlyTyrAlaAlaGlyTyrTyrSerTyrLeuTrpAlaAspValLeuAlaAlaAspAlaPheSerArgPheGluGluGlyIle	
<i>prlC</i>	AGCCATATTTCCCGGGTGTATGCCGAGTTACTACAGTACTCTGAGCGTGGCTGAGTACTGGCGGAGATGCTTCTCGCGCTTGAGGAAGAGGCA	2232
<i>opdA</i>C..T..T..C..C.....G.....C.....T.....C.....G..C....AT..C.....G..A....	
<i>OpdA</i>Thr.....Glu.....	
<i>PrlC</i>	PheAsnArgGluThrGlyGlnSerPheLeuAspAsnIleLeuSerArgGlyGlySerGluGluProMetAspLeuPheLysArgPheArgGlyArgGlu	
<i>prlC</i>	TTTTCAACCGTGAACCGGGCAGTCTTCTCCGCAACATTCTGAGCCGCGGGTTCGAAGAGCCGATGGATCTGTTCAACCGCTTCCGTTGGTCTGGA	2332
<i>opdA</i>G.....T.....T...C...CT..C..T.....A..C..T.....C.....	
<i>OpdA</i>	
<i>PrlC</i>	ProGlnLeuAspAlaMetLeuGluHisTyrGlyIleLysGly	
<i>prlC</i>	ACCCGAGCTGGATGCGATGCTGGAGCATTACGGCATTAGGGCTGATCATTAGTAAATCTGCTTAATTGATGAAACAGGCACCGGAGACGCTGCCT	2432
<i>opdA</i>A..A.....C.....G.....A.....-T..AC...C.....G.....G...G..G..AC....GC....	
<i>prlC</i>	ATCTGTTCTGGCGCCGCTGGGGCTGGAGCACGATGAAGACAACCTGATGGCGCTGGTGAACGCCGGAACATCTGGAATTGCGCAAGCGTGATGAG	2532
<i>opdA</i>T.....A.....T..C.....C.....A.G.....C.....GC.....CG..C..C..A	
<i>prlC</i>	CCAAAACCTGGCGCATCTTTGTTGATTTTGGCGGAGCGATGGCGCACCGCAAATTCGGCGGTGGTCCGGTGAGGCGGTGGCGAAAGCGGTGG	2632
<i>opdA</i>	...G..G..C.....T..T..C..C.....C.....C.....	
<i>prlC</i>	GCATTAAGCGGATTATTTGCCGGATGTTGGTGGATGCCAC	2672

FIG. 2—Continued.

scription is dependent on the heat shock sigma factor σ^{32} .) A search of the GenBank data base revealed no similarity between the sequence upstream of either *prlC* or *opdA* and any *E. coli* or *S. typhimurium* sequences. However, the upstream region of the *prlC* sequence from nucleotide 1 to 132 was 67% identical to an uncharacterized ORF, ORFJ, from *Haemophilus influenzae* which may be involved in DNA transformation competence (17).

Nucleotide sequence of the *prlC1* allele. Nucleotide sequencing of 300 bp of the *prlC1* allele from pCM191 identified a 21-bp duplication at position 1422. This 21-bp duplication results in the duplication of seven amino acids in the protein. The region of the protein affected by *prlC1* is conserved among the related metalloproteinases encoded by *prlC*, *opdA*, and *dcp* and rat metalloproteinase EP 24.15 (4). Maxicell analysis demonstrated that the PrlC protein made by strains carrying pTC101 (*prlC1*) is slightly larger than that made by strains carrying pTC100 (*prlC*⁺) (18), thus supporting this identification of the *prlC1* allele.

DISCUSSION

The major conclusion of this article is that *E. coli prlC* and *S. typhimurium opdA* are homologous genes. Every line of investigation described above supports this conclusion. It is also clear that the *E. coli prlC* gene codes for a protease. The sequence similarity and the observation that a subclone

carrying only one major ORF restores protease activity in *S. typhimurium* and suppressor activity in *E. coli* make this conclusion unambiguous.

Two lines of evidence suggest that *E. coli prlC* and *S. typhimurium opdA* are not located at precisely corresponding positions on the genetic maps of the two organisms. The map data for *prlC* indicate a location at 71 map units (19). *S. typhimurium opdA* has been mapped at 76 map units (4). The nucleotide sequences of the genes diverge strongly in the region 5' to the putative promoters. Southern hybridization data (not shown) indicate that a probe internal to the *opdA* coding region identifies only one fragment in each organism. A probe containing the *prlC* coding sequence and about 500 bases 5' to the coding region identified the same fragment as the *opdA* probe in *E. coli*. In *S. typhimurium*, this *E. coli* probe identified an additional fragment. These results suggest that the sequence in the dissimilar 5' region is present in *S. typhimurium* but at a different location on the chromosome. They also indicate that only one *prlC* or *opdA* gene is present in each organism.

It is possible that the previously observed toxicity of constructs carrying *prlC* (19) was caused by sequences present in the dissimilar 5' region. High-copy-number plasmids carrying *S. typhimurium opdA* and several kilobases on either side of the gene have been easily isolated, and these plasmids lead to high-level overproduction of OpdA. These plasmids could be introduced into *E. coli* with no apparent

consequences for growth (4). In contrast, the *prlC* locus could not be cloned in high-copy-number plasmids and appeared to cause reduced λ phage yields when present on a phage λ vector (19). The *prlC* region could be cloned in lower-copy-number vectors (19), and it is these plasmids that were used for most of the experiments described in this article. All of these earlier cloning experiments would have included a significant segment of DNA 5' to the coding region. However, a plasmid (pCM191) carrying very little of this 5' region in a high-copy-number vector conferred no growth defect on either *S. typhimurium* or *E. coli*.

The mechanism by which mutation in a protease leads to suppression of a localization defect is not obvious. A reasonable hypothesis is that the *prlC* suppressor mutations alter the specificity of OpdA, allowing it to cleave the mutant signal sequence from the precursor. This might in some way permit some of the cleaved protein to be correctly localized. Several characteristics of the *prlC* suppressor mutations are consistent with this hypothesis. (i) All of the *prlC* suppressor mutations are dominant to the wild-type allele, suggesting that they alter the activity of the gene product they affect (18, 19). (ii) *prlC* suppressors are quite allele specific. Only a subset of signal sequence mutations are suppressed, and all of these affect the hydrophobic region of the signal peptide (7, 20). (iii) Some *prlC* alleles allow very efficient processing of the mutant precursor, LamBs78, but the amount of protein correctly localized is apparently very small. None of the *prlC* suppressor mutations appears to allow efficient localization of the mutant *lamB* protein (18, 20). Indeed, the existence of correctly localized protein is inferred from the restoration of the wild-type phenotype, growth on maltodextrins, and sensitivity to phage lambda, not from any physical detection of correctly localized protein. The proposed role for OpdA in the processing of the bacteriophage P22 gene 7 protein is also consistent, although the sequence removed from this viral protein has no obvious amino acid sequence similarity to the LamB signal peptide region (5).

The major flaw in this hypothesis is that although it provides an explanation of how the mutant precursor might be processed, it does not explain how this processing leads to correct localization. One possibility is that the protein containing the mutant signal peptide is less efficiently localized than a protein with no signal peptide at all. This idea implies that the secretion apparatus is able to recognize the mature protein. At least one component of the apparatus, SecB, appears to be able to do this (2, 3). The observation that *prlC*-dependent localization absolutely requires *secB* but processing does not is consistent with this idea (20).

An alternative hypothesis for the mechanism of *prlC* suppressor action is based on the observation that OpdA appears to be a signal peptide peptidase (11, 12). Although the in vitro evidence indicates that the *E. coli* lipoprotein signal peptide is attacked by OpdA only after it is cleaved from the precursor (11), it seems possible that the peptidase or a mutant variant might be able to recognize and bind to the signal peptide while it is still attached to the precursor protein. If so, it might act to maintain the precursor in a conformation competent for secretion in a manner similar to that proposed for the chaperone, SecB (2). This is consistent with the observation that both *prlA* and *prlC* suppressor mutations allow localization without cleavage of LamBs60 (6). It should be noted that these two hypotheses are not mutually exclusive. It is possible that the pathway leading to precursor cleavage is different from that leading to localization. The suppressor protease might be responsible for most of the cleavage, but proper localization might be achieved by

a separate pathway in which binding of the mutant protease to the precursor is the important step.

Obviously, there are other possibilities for mechanisms by which the *prlC* suppressor mutations might act. Further work will be necessary to illuminate the problem and to learn what role the PrlC or OpdA protein plays in protein secretion and cellular metabolism.

ACKNOWLEDGMENTS

This work was supported by grant AI10333 from the National Institute for Allergy and Infectious Diseases, awarded to C.G.M., and by grant GM34821 from the National Institute for General Medical Sciences, awarded to T.J.S. C.A.C. was supported while at Case Western Reserve University by Public Health Service training grant GM08056.

We gratefully acknowledge the contributions to this work of Ann Kukral, Judith Miller, and Ann Flower.

REFERENCES

- Chen, E. J., and D. H. Seeburg. 1985. Supercoil sequencing: a fast simple method for sequencing plasmid DNA. *DNA* 4:165-170.
- Collier, D. N., V. A. Bankaitis, J. B. Weiss, and P. J. Bassford, Jr. 1988. The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein. *Cell* 53:273-283.
- Collier, D. N., S. M. Strobel, and P. J. Bassford, Jr. 1990. SecB-independent export of *Escherichia coli* ribose-binding protein (RBP): some comparisons with export of maltose-binding protein (MBP) and studies with RBP-MBP hybrid proteins. *J. Bacteriol.* 172:6875-6884.
- Conlin, C. A., and C. G. Miller. 1992. Cloning and nucleotide sequence of *opdA*, the gene for oligopeptidase A from *Salmonella typhimurium*. *J. Bacteriol.* 174:1631-1640.
- Conlin, C. A., E. R. Vimr, and C. G. Miller. 1992. Oligopeptidase A is required for normal phage P22 development. *J. Bacteriol.* 174:5869-5880.
- Emr, S. D., and P. Bassford, Jr. 1982. Localization and processing of outer membrane and periplasmic proteins in *Escherichia coli* strains harboring export-specific suppressor mutations. *J. Biol. Chem.* 257:5852-5860.
- Emr, S. D., S. Hanley-Way, and T. J. Silhavy. 1981. Suppressor mutations that restore export of a protein with a defective signal sequence. *Cell* 23:79-88.
- Guyon, M. S. 1983. Uses of the transposon $\gamma\delta$ in the analysis of cloned genes. *Methods Enzymol.* 101:362-369.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Liu, L., W. Whalen, A. Das, and C. M. Berg. 1987. Rapid sequencing of cloned DNA using a transposon for bidirectional priming: sequence of the *Escherichia coli* K-12 *avtA* gene. *Nucleic Acids Res.* 15:9461-9469.
- Novak, P., and I. K. Dev. 1988. Degradation of a signal peptide by protease IV and oligopeptidase A. *J. Bacteriol.* 170:5067-5075.
- Novak, P., P. H. Ray, and I. K. Dev. 1986. Localization and purification of two enzymes from *Escherichia coli* capable of hydrolyzing a signal peptide. *J. Biol. Chem.* 261:420-427.
- Roth, J. R. 1970. Genetic techniques in studies of bacterial metabolism. *Methods Enzymol.* 17A:3-35.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. *Experiments with gene fusions*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Tomb, J. F., H. el-Hajj, and H. O. Smith. 1991. Nucleotide sequence of a cluster of genes involved in the transformation of *Haemophilus influenzae* Rd. *Gene* 104:1-10.

18. **Trun, N. J.** 1988. Ph.D. thesis. Princeton University, Princeton, N.J.
19. **Trun, N. J., and T. J. Silhavy.** 1987. Characterization and *in vivo* cloning of *prlC*, a suppressor of signal sequence mutations in *Escherichia coli* K12. *Genetics* **116**:513-521.
20. **Trun, N. J., and T. J. Silhavy.** 1989. PrlC, a suppressor of signal sequence mutations in *Escherichia coli*, can direct the insertion of the signal sequence into the membrane. *J. Mol. Biol.* **203**:665-676.
21. **Vimr, E. R.** 1981. Ph.D. thesis. Case Western Reserve University, Cleveland, Ohio.
22. **Vimr, E. R., L. Green, and C. G. Miller.** 1983. Oligopeptidase-deficient mutants of *Salmonella typhimurium*. *J. Bacteriol.* **153**:1259-1265.