JOURNAL OF BACTERIOLOGY, Nov. 1990, p. 6217–6222
0021-9193/90/116217-06$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 11

Genes on the 90-Kilobase Plasmid of *Salmonella typhimurium* Confer Low-Affinity Cobalamin Transport: Relationship to Fimbria Biosynthesis Genes

CLEMENT R. RIOUX,† MARY JANE FRIEDRICH, AND ROBERT J. KADNER*

Department of Microbiology, School of Medicine, and Molecular Biology Institute, University of Virginia, Charlottesville, Virginia 22908

Received 24 April 1990/Accepted 13 August 1990

A cloned fragment of *Salmonella typhimurium* DNA complemented the defect in cobalamin uptake of *Escherichia coli* or *S. typhimurium btuB* mutants, which lack the outer membrane high-affinity transport protein. This DNA fragment did not carry *btuB* and was derived from the 90-kb plasmid resident in *S. typhimurium* strains. The cobalamin transport activity engendered by this plasmid had substantially lower affinity and activity than that conferred by *btuB*. Complementation behavior and maxicell analyses of transposon insertions showed that the cloned fragment encoded five polypeptides, at least two of which were required for complementation activity. The nucleotide sequence of the coding region for one of these polypeptides, an outer membrane protein of about 84,000 Da, was determined. The deduced polypeptide had properties typical of outer membrane proteins, with an N-terminal signal sequence and a predicted preponderance of β structure. This outer membrane protein had extensive amino acid sequence homology with PapC and FaeD, two *E. coli* outer membrane proteins involved in the export and assembly of pilus and fimbria subunits on the cell surface. This homology raises the likelihood that the observed cobalamin transport did not result from the production of an authentic transport system but that overexpression of one or more outer membrane proteins allowed leakage of cobalamins through the perturbed outer membrane. These results also suggest that the 90-kb plasmid carries genes encoding an adherence mechanism.

Uptake of vitamin B_{12} (CN-Cbl) and other cobalamins in *Escherichia coli* requires the *btuB*-encoded outer membrane transport protein and the *tonB*-encoded energy-coupling protein for active transport across the outer membrane and the *btuCD* products for passage across the cytoplasmic membrane (3, 4, 13). *Salmonella typhimurium* contains the analogous *btuB* and *tonB* genes for the high-affinity transport system. Evidence for the presence of an additional low-affinity cobalamin transport system was obtained by the cloning of a fragment of *S. typhimurium* DNA that complemented *btuB* mutants of either host (24). The presence of this cloned fragment in plasmid pCRR10 led to production of an M_r-84,000 outer membrane protein and elevated cobalamin binding and transport, although these activities were much lower than those provided by the cloned $btuB^+$ gene. Southern hybridization analysis revealed that the 6-kb insert in pCRR10 did not hybridize to DNA from *E. coli* K-12 or from *S. typhimurium* χ3344 that was cured of the 90-kb plasmid resident in most isolates of this species. These results indicated that the low-affinity cobalamin transport system was encoded by the 90-kb plasmid, thus accounting for the absence of this transport activity from *E. coli* strains.

The presence of the 90-kb plasmid in *S. typhimurium* has been associated with several virulence traits, including adherence to and invasion of HeLa cells, the ability to colonize spleen and liver, and resistance to normal human serum (5–7, 12, 20, 29). A restriction map of this plasmid has been described (17), and several of the virulence and replication functions have been localized by subcloning studies (20). However, the functions associated with substantial portions of this plasmid have not been defined.

This report describes the structure and coding properties of the DNA fragment cloned in plasmid pCRR10. The polypeptides encoded by this fragment were identified by maxicell analysis, and their role in cobalamin transport was demonstrated from the complementation behavior of transposon insertion mutants. The hypothesis had been proposed that the complementing genes encode an outer membrane transport protein that normally carries an unidentified ligand with high affinity, but carries CN-Cbl as a poor substrate (24). To examine this hypothesis, the nucleotide sequence of the gene for the outer membrane protein was determined to compare the deduced polypeptide with other outer membrane transport proteins.

MATERIALS AND METHODS

Bacterial strains and plasmids. The *E. coli* K-12 strains have been previously described (9, 24). Complementation testing and maxicell analysis were carried out in strain RK5016 [Δ(*argF-lac*)*U169 araD139 relA1 rpsL150 flbB5301 deoC1 thi gyrA219 non metE70 argH1 recA56 btuB461*] (9).

Plasmid pCRR10 carries a 6-kb partial *Sau*3A fragment of *S. typhimurium* DNA inserted in the *Bam*HI site of pBR322 (24). The insert in pCRR10 was excised by cleavage at the *Cla*I and *Sph*I sites in pBR322 flanking the insert and ligated into the same sites in pACYC184 (1), to yield plasmid pCRR11 (chloramphenicol resistance). For nucleotide sequence determination, bacteriophages M13mp18 and M13mp19 were propagated in strain JM101 (22).

Media and growth conditions. Complementation by transposon insertions in pCRR11 of the $BtuB^-$ phenotype of strain RK5016 was tested on minimal medium A supplemented with glucose (0.5%), arginine (100 μg/ml), and either

* Corresponding author.
† Present address: Biotechnology Research Institute, National Research Council Canada, Montreal, Quebec, Canada H4P 2R2.

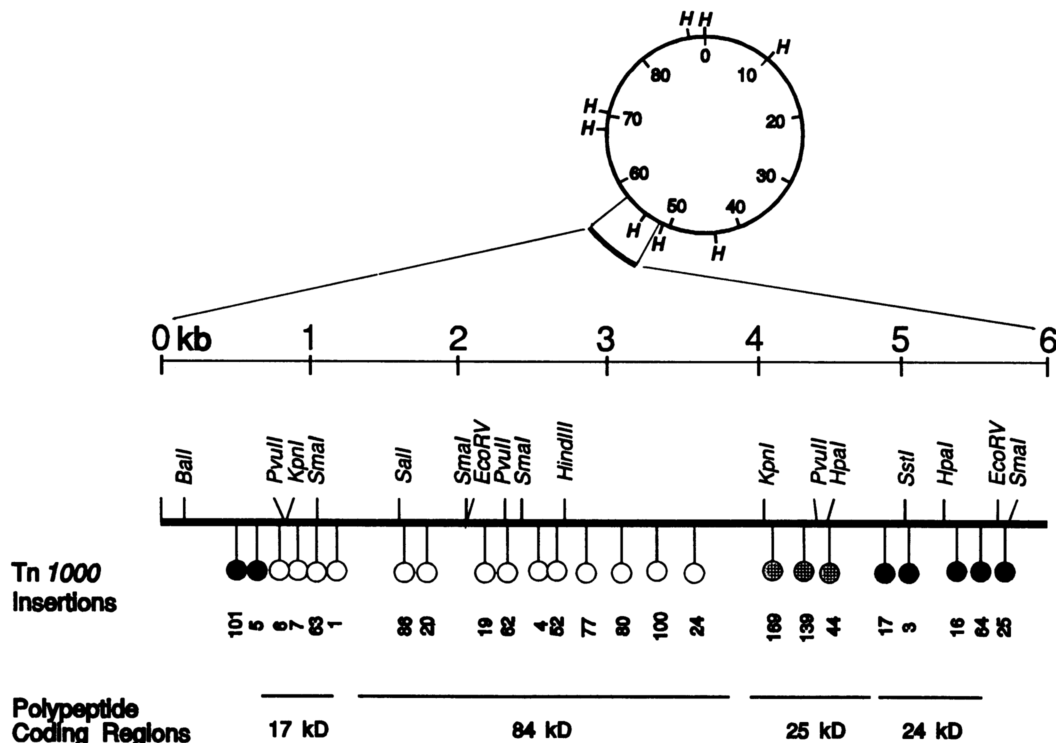

FIG. 1. Restriction map of the insert in plasmid pCRR10. At the top is represented the 90-kb plasmid, with the location of *Hin*dIII cleavage sites represented by H, as determined by Michiels et al. (17). The origin of the insert in pCRR10 on the 90-kb plasmid is indicated by the expanded arc, and its restriction map is shown. There were no sites for *Bam*HI, *Bgl*II, *Cla*I, *Eco*RI, *Sph*I, *Xba*I, or *Xho*I. Below the map are shown the approximate sites of transposon Tn*1000* insertion and their isolation numbers. The complementation behavior of the transposon insertions is indicated as full complementation (●), partial complementation (◍), or no complementation (○). The bottom lines indicate the approximate locations of polypeptide-coding regions, based on the effect of transposon insertions on production of insert-specified polypeptides (see Fig. 2).

CN-Cbl (5 nM) or methionine (100 μg/ml) (18, 24). Rich medium was L broth (18). The antibiotic ampicillin (50 μg/ml) or chloramphenicol (15 μg/ml) was used for selection and maintenance of plasmid-containing strains.

Genetic techniques. Insertions of transposon Tn*1000* into plasmid pCRR11 were isolated by conjugation with an F′ plasmid as previously described (3, 9).

Maxicell analysis of plasmid-coded polypeptides. The procedure of Sancar et al. (25) was used for labeling of plasmid-coded polypeptides with [^{35}S]methionine as previously described (3). Polypeptides were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis as described by Lugtenberg et al. (15). Gels were stained with Coomassie blue and then subjected to autoradiography.

Recombinant DNA techniques. Standard methods were used for plasmid isolation, restriction endonuclease analysis, and ligation (11, 16). Plasmid transformation was done by the method of Hanahan (8). The nucleotide sequence of the insert in pCRR10 was determined by subcloning of overlapping restriction fragments into the replicative forms of M13 vectors. Nested deletions were generated by treatment with exonuclease III and S1 nuclease, using the Erase-A-Base kit (Bio-Rad Laboratories, Inc., Richmond, Calif.). Single-stranded M13 derivatives were used as templates in dideoxy-chain termination reactions by the method of Sanger et al. (26). These reactions used the universal primer and phage T7 DNA polymerase (Pharmacia, Inc.). To reduce sequence ladder compression, some reactions were carried out with 2′-deoxyribo-deazaguanosine triphosphate or dITP in place of GTP.

Sequence data were compiled with the DBUTIL program of Staden (28) and were compared with translated sequences in the GenBank, NBRF, and Swiss data bases with the FASTA program of Lipman and Pearson (14). The MULTALIN program of Corpet (2) was used for sequence alignment.

Nucleotide sequence accession number. The sequence reported has been assigned GenBank accession number M37853.

RESULTS

Structure of the insert cloned in pCRR10. The restriction map of the 6-kb insert in plasmid pCRR10 was determined (Fig. 1). This restriction map and the sizes of the genomic fragments to which the insert hybridizes on Southern blots were compared with the restriction map for the 90-kb *S. typhimurium* plasmid described by Michiels et al. (17) and Norel et al. (20). On the basis of the sizes of the two *Hin*dIII fragments that hybridized to pCRR10 (24), the absence of *Bam*HI and *Bgl*II sites in the insert, and the location of the *Sal*I site, the origin of the insert in pCRR10 was unambiguously placed at coordinates 50.8 to 56.8 kb, between the *vir* and *repB* regions.

Complementation properties of transposon insertion mutants. To identify regions of the insert essential for cobalamin transport activity, transposon Tn*1000* insertions were isolated in the insert cloned in pACYC184 (pCRR11). The sites of the transposon insertions were estimated by restriction endonuclease mapping and were distributed fairly

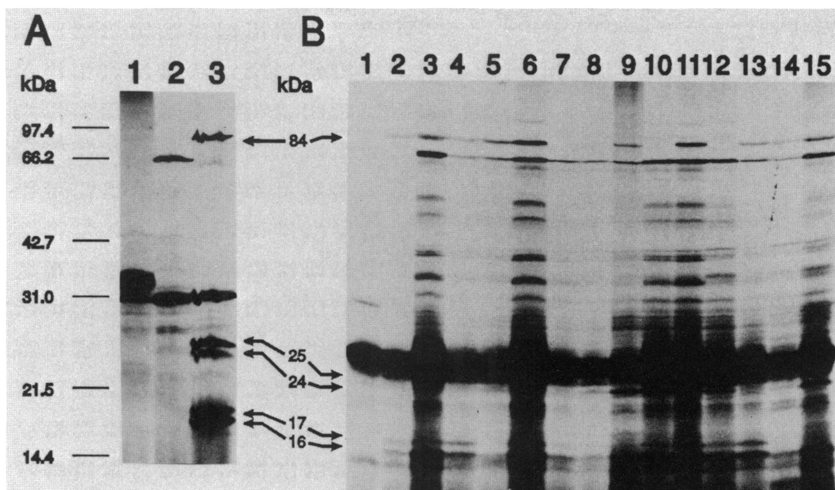

FIG. 2. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of polypeptides labeled in maxicells with [^{35}S]methionine. The host strain was RK5016. (A) Resident plasmids were pBR322, expressing β-lactamase and the tetracycline resistance determinant (lane 1); pKH3-8, expressing BtuB (lane 2); and pCRR10 (lane 3). (B) Resident plasmids were pACYC184, expressing the chloramphenicol and tetracycline resistance determinants (lane 1); pCRR11 (lane 2); and the transposon Tn*1000* insertions 1 (lane 3), 5 (lane 4), 6 (lane 5), 17 (lane 6), 20 (lane 7), 24 (lane 8), 44 (lane 9), 62 (lane 10), 64 (lane 11), 77 (lane 12), 101 (lane 13), 139 (lane 14), and 169 (lane 15). The mobility of molecular weight standards is indicated on the left side; the locations of insert-specified polypeptides are indicated by the arrows.

evenly throughout the insert (Fig. 1). Each mutant derivative was introduced by transformation into *E. coli btuB* strain RK5016, and the growth of chloramphenicol-resistant transformants on 5 nM CN-Cbl was tested. As summarized in Fig. 1, transposon insertions at either end of the insert (0 to 0.5 kb and 5 to 6 kb) did not impair the ability of the plasmids to complement the defect in CN-Cbl uptake. Insertions in the region from 0.5 to 3.5 kb eliminated complementation, and insertions between 4 and 5 kb resulted in reduced growth with CN-Cbl. Thus, a substantial portion of the cloned insert was required for transport activity.

Identification of polypeptide products. Five insert-specified polypeptides were identified by labeling of maxicells of strain RK5016 carrying pCRR10 with [^{35}S]methionine (Fig. 2A, lane 3). These products had apparent molecular weights of 84,000, 25,000, 24,000, 17,000, and 16,000 and were labeled roughly to the same extent as the vector-encoded β-lactamase. The effect of transposon insertions on production of these polypeptides was determined in derivatives of pCRR11 (Fig. 2B). With this plasmid, expression of five insert-specified polypeptides was much weaker than that of chloramphenicol acetyltransferase. Plasmids carrying Tn*1000* insertions programmed synthesis of several transposon-specific polypeptides, and in most of them synthesis of one or more insert-specific polypeptides was eliminated.

The Btu$^+$ insertions 101 and 5, at the left end of the insert, did not affect synthesis of any of the insert-specific polypeptides. The Btu$^-$ insertion 6 eliminated the 17-kDa and possibly the 16-kDa species, while the Btu$^-$ insertion 1 did not appear to affect any of the polypeptides. The Btu$^-$ insertions 20, 62, 77, and 24, which lie between the two *Kpn*I sites, blocked synthesis of the 84-kDa polypeptide. Insertions 169, 139, and 44, which all displayed weak complementation activity, appeared to eliminate the 25- and 17-kDa polypeptides. Finally, Btu$^+$ insertion 17 eliminated synthesis of the 24-kDa polypeptide. The extent of the coding regions for the insert-specified polypeptides, based on this analysis, is diagrammed in Fig. 1. These results localized the coding region for the 84-kDa outer membrane protein and showed that its synthesis is necessary but not sufficient for cobalamin transport. The 17-kDa and perhaps the 16-kDa polypeptides also appeared to be necessary for transport activity. It is not known whether these two species are related by posttranslational processing or are products of separate genes, since the coding capacity of the region is near the maximum needed for two separate polypeptides of their size. The reason that insertions between coordinates 4 and 5 kb also affected synthesis of the 17-kDa polypeptide is discussed below.

Nucleotide sequence of the gene for the outer membrane protein. To compare the structure of the 84-kDa outer membrane protein with those of other outer membrane transport proteins, the nucleotide sequence of the region between the *Kpn*I sites at coordinates 0.9 and 4.1 kb was determined. The sequence was determined for both strands, with an average of 6.1 gel readings for each sequence character. A single open reading frame of the proper size for the 84-kDa protein was found in the location expected from the transposon analysis (Fig. 3).

This open reading frame of 2,409 nucleotides is directed from left to right on the map of Fig. 1, and it encodes an 802-amino-acid polypeptide with a molecular weight of 86,387. This gene is preceded by a possible promoter region located between 39 and 68 bp upstream from the start of the coding region; this sequence, TTAACA-N_{17}-TGCAAT, matches the consensus promoter sequence at 9 of 12 positions. Transcription mapping will be necessary to show whether this sequence is active, but the fact that upstream transposon insertions displayed no polar effect on protein expression in maxicells suggests that this gene might have its own promoter. The initiating GUG codon lies 7 nucleotides downstream from a likely Shine-Dalgarno sequence, AGGGG (27). Codon usage within the coding sequence was extremely broad, as expected for a weakly expressed bacterial protein. All but two of the possible codons (CGA and AGA) were used at least once. The 3′ end of the coding region overlaps with another open reading frame in the sequence UAAGAUG. This distal reading frame encodes a polypeptide with deduced M_r of 24,600, which is probably the 25-kDa polypeptide observed in maxicells.

```
                                                                        50               S-D                                    100
CCTGTTTAACAGGAAATTGTATTTTCCGTGCAATTTCCTCCGGAGAGGGACTGCCTGAGTCAGGGGGGTTGTT GTG TCA TTT CAT CAC CGG GTA TTT AAA CTG TCG GCG CTG
                                                                          Met Ser Phe His His Arg Val Phe Lys Leu Ser Ala Leu

                                            150                                                     200
AGT CTC GCC TTA TTT TCT CAC CTA TCT TTT GCC|AGC ACT GAC TCA GAG CTG AAC CTG GAT TTC CTG CAG GGA ATG AGC GCC ATC CCC TCC GTA TTA
Ser Leu Ala Leu Phe Ser His Leu Ser Phe Ala|Ser Thr Asp Ser Glu Leu Asn Leu Asp Phe Leu Gln Gly Met Ser Ala Ile Pro Ser Val Leu

                                                    250                                                             300
AAA TCC GGC TCG GAT TTT CCG GCC GGA CAG TAT TAT GTC GAC GTC ATT GTT AAC CAG GAA AAC GTG GGT AAA GCC CGT TTG TCC ATT ACG CCA CAG
Lys Ser Gly Ser Asp Phe Pro Ala Gly Gln Tyr Tyr Val Asp Val Ile Val Asn Gln Glu Asn Val Gly Lys Ala Arg Leu Ser Ile Thr Pro Gln

                                                            350                                                             400
GAA GAA TCA GCA AAT GCC CTG TGC CTG TCG CCG GAG TGG CTG AAA GCT GCC GGG GTT CCT GTC CGC CTG GAG GGA TAT GCC TCC ACG CTG AAT GCT
Glu Glu Ser Ala Asn Ala Leu Cys Leu Ser Pro Glu Trp Leu Lys Ala Ala Gly Val Pro Val Arg Leu Glu Gly Tyr Ala Ser Thr Leu Asn Ala

                                                                450
GCC GGG CAG TGC TAT GTC CTC AGC CGC AAC CCC TAT ACC AGG GTG GAC TTC AGC TAT GGC TCC CAG AGC TTG GTG TTC AGT ATT CCC CAG TCG TTC
Ala Gly Gln Cys Tyr Val Leu Ser Arg Asn Pro Tyr Thr Arg Val Asp Phe Ser Tyr Gly Ser Gln Ser Leu Val Phe Ser Ile Pro Gln Ser Phe

 500                                                      550
CTG GTC GGT AAA ACG GAC CCC AGC CGC TGG GAC TAC GGC GTG CCG GCG GCA CGC CTG AAG TAC TCC GCC AAC GCC TCG CAG ACG TCC GGG CAA AGC
Leu Val Gly Lys Thr Asp Pro Ser Arg Trp Asp Tyr Gly Val Pro Ala Ala Arg Leu Lys Tyr Ser Ala Asn Ala Ser Gln Thr Ser Gly Gln Ser

     600                                                           650
ACC AGT GCC TAT GCA AAT GCC GAC CTG ATG GTC AAC CTC GGA CGC TGG GTG CTG GCC AGT AAC ATG AGC GCA TCC CGG TAC GCT GAC GGC TCC GGT
Thr Ser Ala Tyr Ala Asn Ala Asp Leu MET Val Asn Leu Gly Arg Trp Val Leu Ala Ser Asn Met Ser Ala Ser Arg Tyr Ala Asp Gly Ser Gly

          700                                                              750
GAG TTC ACC GCC CGG GAT ATC ACG CTG TCC ACC GCC ATC AGC CAA GTG CAG GGG GAC CTG CTG CTC GGT AAA TCC CAG ACC CGC AGC GCC CTG TTC
Glu Phe Thr Ala Arg Asp Ile Thr Leu Ser Thr Ala Ile Ser Gln Val Gln Gly Asp Leu Leu Leu Gly Lys Ser Gln Thr Arg Ser Ala Leu Phe

                 800                                                             850
TCT GAT TTC GGC TTT TAC GGG GCG GCA CTG CGC TCC AAC AGT AAC ATG CTG CCG TGG GAG GCC CGC GGG TAT GCC CCG CTT ATC ACC GGG GTG GCG
Ser Asp Phe Gly Phe Tyr Gly Ala Ala Leu Arg Ser Asn Ser Asn Met Leu Pro Trp Glu Ala Arg Gly Tyr Ala Pro Leu Ile Thr Gly Val Ala

                     900                                                                 950
AAC TCC ACC TCC CGC GTC ACC ATC AGC CAG AAC GGG TAC GCC GTG TAC TCA AAA GTG GTG CCG CCC GGT CCG TAC CAG CTG GAT GAT GTC CGC TCC
Asn Ser Thr Ser Arg Val Thr Ile Ser Gln Asn Gly Tyr Ala Val Tyr Ser Lys Val Val Pro Pro Gly Pro Tyr Gln Leu Asp Asp Val Arg Ser

                         1000                                                                1050
GTG GGG AAC GGT GAC CTG GTG GTG ACC GTG GAG GAT GCG TCC GGC CAC AAA ACC ACC ACC GTC TAT CCG GTC ACC ACC CTG CCG ACG CTG CTG CGT
Val Gly Asn Gly Asp Leu Val Val Thr Val Glu Asp Ala Ser Gly His Lys Thr Thr Thr Val Tyr Pro Val Thr Thr Leu Pro Thr Leu Leu Arg

                             1100                                                                1150
CCC GGG GAG GTT GAA TAC AAT GTG GCG GTC GGT CGC AAG TCC AGT AAC TAT AAG CTG AAA AAA CCG TTC GCT GAC GGT GAA AAC GGC ATG TTC TGG
Pro Gly Glu Val Glu Tyr Asn Val Ala Val Gly Arg Lys Ser Ser Asn Tyr Lys Leu Lys Lys Pro Phe Ala Asp Gly Glu Asn Gly Met Phe Trp

                                 1200                                                                1250
ATG GGA AGC GTG GGG TAC GGC TTT GAT TCC ACC ACG CTG AAT GCC GCT TCT ATC CTG CAC GGT AAA TAC CAG GCC GGC GGG GTG AGC GTG ACG CAG
Met Gly Ser Val Gly Tyr Gly Phe Asp Ser Thr Thr Leu Asn Ala Ala Ser Ile Leu His Gly Lys Tyr Gln Ala Gly Gly Val Ser Val Thr Gln

                                     1300                                                                1350
GCC CTG GGC GGC TTC GGC GCC GTG TCC GCC GGG ATG AAC CTG TCA CAG GCG AAA TAC GAC AAC GGA GAC AAC AAA CGC GGG CAC AGC GTC AGC GCC
Ala Leu Gly Gly Phe Gly Ala Val Ser Ala Gly Met Asn Leu Ser Gln Ala Lys Tyr Asp Asn Gly Asp Asn Lys Arg Gly His Ser Val Ser Ala

                                         1400                                                                1450
AAG TAT GCC AAA AGC TTC TCT GAC AGC TCG GAC CTG CAG TTG CTG GCC TAC CGT TAC CAG AGT AAG GGG TAT GTG GAG TTC GCA GAC TTC TAC AGT
Lys Tyr Ala Lys Ser Phe Ser Asp Ser Ser Asp Leu Gln Leu Leu Ala Tyr Arg Tyr Gln Ser Lys Gly Tyr Val Glu Phe Ala Asp Phe Tyr Ser

                                             1500                                                                1550
ACA GAC CGG TAT ACC CGC TAC AAC ACA AAA TCA CGC TAT GAG ATG CGC TTC TCG CAA CGC TTG GGG AAC AGT AAC CTG AAT CTG GCC GGC TGG CAG
Thr Asp Arg Tyr Thr Arg Tyr Asn Thr Lys Ser Arg Tyr Glu MET Arg Phe Ser Gln Arg Leu Gly Asn Ser Asn Leu Asn Leu Ala Gly Trp Gln

                                                 1600
GAG GAC TAC TGG TGG ATG AAA GGA AAG GCC ATC GGA GGC GAT GTT TCC CTC AGT ACC ACC ATT CTG GAC GGT GTG TCG GTC TTC CTG AAC GGC AGC
Glu Asp Tyr Trp Trp Met Lys Gly Lys Ala Ile Gly Gly Asp Val Ser Leu Ser Thr Thr Ile Leu Asp Gly Val Ser Val Phe Leu Asn Gly Ser

                                                     1700
TAC AGT AAA CGC CCG TAC CTG GAC AAA CCG GAC TAC AGC ACG TCG CTC TCC TTT AGC ATT CCG TTC ACC CTG GGT GGG ATT CGC CAT TAC AGC AGT
Tyr Ser Lys Arg Pro Tyr Leu Asp Lys Pro Asp Tyr Ser Thr Ser Leu Ser Phe Ser Ile Pro Phe Thr Leu Gly Gly Ile Arg His Tyr Ser Ser

 1750                                                     1800
ACC GGA CTG AGC TAC AAC AGC AGC GGC AGG ATG GGG ATG AAC AGC GGG GTG TCG GCA AGT CCG ACG GAC CGC CTG AGC TAC GGC CTG AAC ACC AAC
Thr Gly Leu Ser Tyr Asn Ser Ser Gly Arg Met Gly Met Asn Ser Gly Val Ser Ala Ser Pro Thr Asp Arg Leu Ser Tyr Gly Leu Asn Thr Asn

     1850                                                     1900
CTG AGT GAT AAG GGC GAC CGC AGC CTG AGC GGC AAC CTC TCG TAT GGC TTT GAT GCC ATC CAG ACC AAC ATG ATG CTG TCG CAG GGA CGT GAT AAC
Leu Ser Asp Lys Gly Asp Arg Ser Leu Ser Gly Asn Leu Ser Tyr Gly Phe Asp Ala Ile Gln Thr Asn Met Met Leu Ser Gln Gly Arg Asp Asn

         1950                                                     2000
ACC ACC GTG TCA GGC AGC GTG AGC GGC ACA ATT CTG GGC ACG GCA GAC AGC GGC CTG ATG ATG ACG AAG GAA ACC GGT AAC ACG CTG GGC GTG GCG
Thr Thr Val Ser Gly Ser Val Ser Gly Thr Ile Leu Gly Thr Ala Asp Ser Gly Leu Met Met Thr Lys Glu Thr Gly Asn Thr Leu Gly Val Ala

             2050                                                     2100
CGC ATT CCG GGG GTG AAG GGC GTG AGG ATT AAC GGC TCA GCC CCC ACC AAC AGC AAG GGA TAC ACC GTG GTG AAC CTG TCG GAT TAC TCC CTG AAC
Arg Ile Pro Gly Val Lys Gly Val Arg Ile Asn Gly Ser Ala Pro Thr Asn Ser Lys Gly Tyr Thr Val Val Asn Leu Ser Asp Tyr Ser Leu Asn

                 2150                                                     2200
CGG GTC AGC GTG GAC ATG GAA AAC GTG CCG GAT GAC CTG GAG CTG CAG ACC ACC TCC TTT AAC GTG GTG CCG ACG GAA AAA GCT GTC GTG TAC CGG
Arg Val Ser Val Asp Met Glu Asn Val Pro Asp Asp Leu Glu Leu Gln Thr Thr Ser Phe Asn Val Val Pro Thr Glu Lys Ala Val Val Tyr Arg

                     2250                                                     2300
GAG TTT GGC GCC GAG CAT GTG CTG CGC TAC ATC CTG CGG GTG AAG GAG CGT GAC GGA CGG ATA TTG AAC GGG GGC AGC GCG CAG ACG GAG CAG GGA
Glu Phe Gly Ala Glu His Val Leu Arg Tyr Ile Leu Arg Val Lys Glu Arg Asp Gly Arg Ile Leu Asn Gly Gly Ser Ala Gln Thr Glu Gln Gly

                         2350                                                     2400
CTG GAT GCC GGG TTC ATT GCC GGT AAC GGC GTC CTG CTG ATG AAT ATG CTG AGC GCG CCG TCA CGG GTC AGC GTC GAG CGG GGG GAC GGC AGT GTC
Leu Asp Ala Gly Phe Ile Ala Gly Asn Gly Val Leu Leu Met Asn Met Leu Ser Ala Pro Ser Arg Val Ser Val Glu Arg Gly Asp Gly Ser Val

                             2450                                                     2500
TGC CAT TTT TCA GTG AAA GGT ATT GTG CCT AAT ACC GGC AAA GTT CAG GAG GTT TAT TGT GAA TAAGATG ATG AAG TGG GGA CTG GTG TCC CTG CTG
Cys His Phe Ser Val Lys Gly Ile Val Pro Asn Thr Gly Lys Val Gln Glu Val Tyr Cys Glu TER
                                                                                     Met Met Lys Trp Gly Leu Val Ser Leu Leu

                                 2550                                                     2600
TCC CTG GCC GTC AGC GGG CAG GCC ATG GCA|GCC TTT GTG CTG AAC GGC ACG CGT TTT ATC TAT GAG GAA GGG AGA AAG AAC ACC TCA TTT GAG GTG
Ser Leu Ala Val Ser Gly Gln Ala Met Ala|Ala Phe Val Leu Asn Gly Thr Arg Phe Ile Tyr Glu Glu Gly Arg Lys Asn Thr Ser Phe Glu Val
```

FIG. 3. Nucleotide sequence of the coding region for the outer membrane polypeptide. The sequence shown comprises a 2,609-bp fragment that extends between nucleotides 1351 and 3960 on the map shown in Fig. 1. The deduced amino acid sequence corresponds to a protein of 86,387 Da. The termination codon at nucleotide 2480 is followed by a reading frame for a polypeptide of approximately 24,600 Da. A potential −10 and −35 promoter region for the outer membrane protein is underlined, and the putative ribosome-binding sequence (S-D) is indicated. Vertical bars indicate probable sites of leader peptide cleavage.

Properties and homologies of the outer membrane protein. The amino-terminal portion of the deduced polypeptide is typical of bacterial signal sequences. There are 4 basic residues among the first 9 amino acids, followed by 15 nonpolar residues. Cleavage is likely to occur after the sequence Ser-Phe-Ala, resulting in removal of the first 24 residues and leaving a mature 778-amino-acid polypeptide of 83.8 kDa, which is very close to the value determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The general features of the processed polypeptide are typical for an outer membrane protein. It has a fairly polar character in that charged residues comprise 18.6% of the total (estimated isoelectric point, 8.61) and nonpolar residues represent 37.7%. The hydropathy profile reveals the presence of few segments with sufficient length and hydrophobic character to be likely membrane-spanning regions (not shown). Various secondary-structure predictions indicate a preponderance of β structure. Only two regions, including the signal sequence, display α-helical propensity.

Sequence homology searches revealed strong similarity of this outer membrane protein only to two *E. coli* proteins, PapC (21) and FaeD (19). PapC and FaeD are outer membrane proteins of about 87 kDa that are involved in the export and assembly of pilin or fimbria subunits. The *pap* operon is responsible for synthesis of the P-type pilus found on most strains of *E. coli* associated with urinary tract infections (23). The *fae* operon encodes the K88 antigen responsible for adherence to intestinal epithelia of some enterotoxinogenic *E. coli* strains that cause porcine diarrhea. Relatedness extended over the entire length of all three proteins. When the sequences were aligned with gaps inserted to maximize homology by using the MULTALIN program of Corpet (2), the *S. typhimurium* protein shared 31.6% identity with FaeD and 24.5% identity with PapC (Fig. 4). With this alignment, PapC and FaeD shared 21.0% identity, suggesting that the *S. typhimurium* protein was more closely related to both *E. coli* proteins than they were to each other.

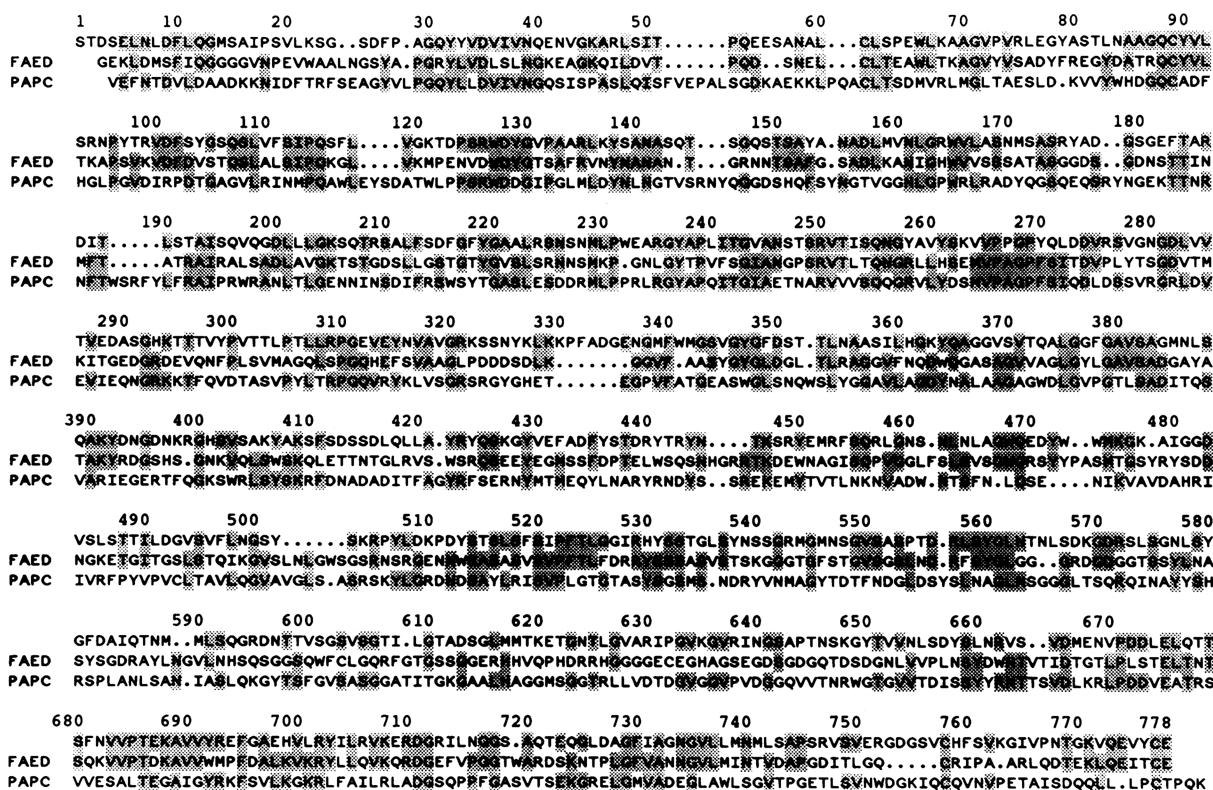

FIG. 4. Homologies between the amino acid sequences of the *S. typhimurium* outer membrane protein (top line) and the *E. coli* FaeD (19) and PapC (21) proteins. Sequences were aligned by the program of Corpet (2), with a gap penalty of 8. Identical residues are identified by shading.

DISCUSSION

S. typhimurium possesses the BtuB/TonB-dependent system for active transport of cobalamins across the outer membrane. The presence of a second, low-affinity uptake system was suggested from the cloning of a DNA fragment that complemented partially the BtuB$^-$ phenotype in *E. coli* or *S. typhimurium*. This DNA fragment encodes at least five polypeptides, and the transposon insertions that blocked synthesis of two of these polypeptides (84 and 17 kDa) also eliminated complementation activity. Loss of the 25-kDa polypeptide led to decreased complementation activity. It was possible that the 84-kDa polypeptide was an outer membrane transport protein that carried cobalamins with low affinity. Since pCRR10-dependent cobalamin transport did not require *tonB* function (24), the low-molecular-weight polypeptides might be required for energy coupling.

The relatedness of the outer membrane protein to PapC and FaeD suggests a different model for the transport activity. The PapC and FaeD proteins are thought to be necessary for export and assembly of the pilin and lectin subunits on the bacterial cell surface. The *pap* and *fae* operons comprise 10 to 12 genes encoding the pilin and lectin subunits and proteins involved in subunit export. If the region cloned in pCRR10 is part of a homologous operon with the gene for the outer membrane protein in the middle, this cloned region is likely not to include either end of the operon. Thus, the normal promoter is probably absent and gene expression is initiated from weak internal promoters. This possibility is consistent with the low level of protein labeling in maxicells with pCRR11 and the absence of transcription polarity resulting from transposon insertions.

The low level of cobalamin transport in response to pCRR10 could result from disruption of the outer membrane permeability barrier by insertion of elevated amounts of the 84-kDa protein. The fact that the 17-kDa polypeptide is also necessary for complementation suggests that (i) these two polypeptides must be expressed to disrupt the outer membrane or (ii) insertion of the smaller polypeptide disrupts the outer membrane and the larger polypeptide is needed for its insertion. The presence of pCRR10 in cells does not result in overt disruption of the outer membrane or markedly deranged barrier function, insofar as these strains grow on MacConkey medium, which is lethal for mutant cells whose outer membrane is grossly permeable to hydrophobic detergents (24).

The presence on the 90 kb plasmid of a gene so closely related to other fimbriae-assembly genes suggests that an intact *fim* operon resides on this plasmid. This possibility is consistent with observed correlation between the presence of the 90 kb plasmid and adhesive, invasive, and virulent phenotypes (5–7, 12, 20, 29). The other polypeptides encoded by pCRR10 are similar in size to products of the *pap* and *fae* operons. Transposon insertions in the gene for the M_r-25,000 protein 3′ to the gene for the outer membrane protein resulted in decreased production of at least two polypeptides of M_r 25,000 and 17,000. This M_r-25,000 polypeptide is homologous to the *papD* product, which is a molecular chaperone responsible for effective transport of the lectin subunits across the periplasmic space (10). Absence of the M_r-25,000 polypeptide might result in decreased production or stability of the pilin proteins by preventing their normal export. The structure of the region surrounding

the insert and the effect of this region on cellular adhesive properties are under study.

ACKNOWLEDGMENTS

We are indebted to Roy Curtiss III for providing strains.

This work was supported by Public Health Service grant GM19078 from the National Institute of General Medical Sciences and by a postdoctoral fellowship to C.R.R. from the Fonds de la Recherche en Sante du Quebec.

LITERATURE CITED

1. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:** 1141–1156.
2. **Corpet, F.** 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16:**10881–10890.
3. **DeVeaux, L. C., D. S. Clevenson, C. Bradbeer, and R. J. Kadner.** 1986. Identification of the BtuCED polypeptides and evidence for their role in vitamin B_{12} transport in *Escherichia coli*. J. Bacteriol. **167:**920–927.
4. **Friedrich, M. J., L. C. DeVeaux, and R. J. Kadner.** 1986. Nucleotide sequence of the *btuCED* genes involved in vitamin B_{12} transport in *Escherichia coli* and homology with components of periplasmic-binding-protein-dependent transport systems. J. Bacteriol. **167:**928–934.
5. **Gulig, P. A., and R. Curtiss III.** 1987. Plasmid-associated virulence of *Salmonella typhimurium*. Infect. Immun. **55:**2891–2901.
6. **Gulig, P. A., and R. Curtiss III.** 1988. Cloning and transposon insertion mutagenesis of virulence genes of the 100-kilobase plasmid of *Salmonella typhimurium*. Infect. Immun. **56:**3262–3271.
7. **Hackett, J., P. Wyk, P. Reeves, and V. Mathan.** 1987. Mediation of serum resistance on *Salmonella typhimurium* by an 11-kilodalton polypeptide encoded by the cryptic plasmid. J. Infect. Dis. **155:**540–549.
8. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
9. **Heller, K., B. J. Mann, and R. J. Kadner.** 1985. Cloning and expression of the gene for the vitamin B_{12} receptor in the outer membrane of *Escherichia coli*. J. Bacteriol. **161:**896–903.
10. **Holmgren, A., and C.-I. Bränden.** 1989. Crystal structure of chaperon protein PapD reveals an immunoglobulin fold. Nature (London) **342:**248–251.
11. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9:**2989–2998.
12. **Jones, G. W., D. K. Rabert, D. M. Svinarich, and H. J. Whitfield.** 1982. Association of adhesive, invasive, and virulent phenotypes of *Salmonella typhimurium* with autonomous 60-megadalton plasmids. Infect. Immun. **38:**476–486.
13. **Kadner, R. J., M. D. Lundrigan, and K. Heller.** 1987. Sequences and interactions of proteins participating in the transport of iron and vitamin B_{12} in *Escherichia coli*, p. 85–97. *In* G. Winklemann, D. van der Helm, and J. B. Neilands (ed.), Iron transport in microbes, plants and animals. VCH Publishers, Weinheim, Federal Republic of Germany.
14. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
15. **Lugtenberg, B., I. Meijers, P. Reeves, P. Van der Hoeck, and L. Van Alphen.** 1975. Electrophoretic resolution of the major outer membrane protein of *Escherichia coli* into four bands. FEBS Lett. **58:**254–258.
16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Michiels, T., M. Y. Popoff, S. Durviaux, C. Coynault, and G. Cornelis.** 1987. A new method for the physical and genetic mapping of large plasmids: application to the localization of the virulence determinants on the 90kb plasmid of *Salmonella typhimurium*. Microb. Pathogen. **3:**109–116.
18. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mooi, F. R., J. Classen, D. Bakker, H. Kuipers, and F. K. de Graaf.** 1986. Regulation and structure of an *Escherichia coli* gene coding for an outer membrane protein involved in export of K88ab fimbrial subunits. Nucleic Acids Res. **14:**2443–2457.
20. **Norel, F., C. Coynault, I. Miras, D. Hermant, and M. Y. Popoff.** 1989. Cloning and expression of plasmid DNA sequences involved in *Salmonella* serotype *typhimurium* virulence. Mol. Microbiol. **3:**733–743.
21. **Norgren, M., M. Båga, J. M. Tennent, and S. Normark.** 1987. Nucleotide sequence, regulation, and functional analysis of the *papC* gene required for cell surface localization of Pap pili of uropathogenic *Escherichia coli*. Mol. Microbiol. **1:**169–178.
22. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26:**101–106.
23. **Parry, S. H., and D. M. Rooke.** 1985. Adhesins and colonization factors in *Escherichia coli*, p. 79–155. *In* M. Sussman (ed.), The virulence of *Escherichia coli*. Academic Press, Inc., Orlando, Fla.
24. **Rioux, C. R., and R. J. Kadner.** 1989. Two outer membrane transport systems for vitamin B_{12} in *Salmonella typhimurium*. J. Bacteriol. **171:**2986–2993.
25. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for the identification of plasmid-coded proteins. J. Bacteriol. **137:**692–693.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome-binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
28. **Staden, R.** 1980. A new computer method for the storage and manipulation of DNA gel reading data. Nucleic Acids Res. **8:**3673–3694.
29. **Van den Bosch, J. L., D. K. Rabert, D. R. Kurlandsky, and G. W. Jones.** 1989. Sequence analysis of *rsk*, a portion of the 95-kilobase plasmid of *Salmonella typhimurium* associated with resistance to the bactericidal effect of serum. Infect. Immun. **57:**850–857.