JOURNAL OF BACTERIOLOGY, May 1992, p. 2754–2762
0021-9193/92/092754-09$02.00/0

Vol. 174, No. 9

Structure and Function of the *uhp* Genes for the Sugar Phosphate Transport System in *Escherichia coli* and *Salmonella typhimurium*

MICHAEL D. ISLAND, BEI-YANG WEI, AND ROBERT J. KADNER*

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

Received 17 October 1991/Accepted 18 February 1992

Expression of the *Escherichia coli* sugar phosphate transport system, encoded by the *uhpT* gene, is regulated by external glucose 6-phosphate through the action of three linked regulatory genes, *uhpABC*. The nucleotide sequence of the *uhp* region cloned from *Salmonella typhimurium* was determined. The deduced Uhp polypeptide sequences from the two organisms are highly related. Comparison with the corrected sequence from *E. coli* revealed that the four *uhp* genes are closely spaced, with minimal intergenic distances, and that *uhpC* is nearly identical in length to *uhpT*, both of which have substantial sequence relatedness along their entire lengths. To facilitate analysis of *uhp* gene function, we isolated insertions of a kanamycin resistance (Km) cassette throughout the *uhp* region. In-frame deletions that removed almost the entire coding region of individual or multiple *uhp* genes were generated by use of restriction sites at the ends of the Km cassette. The phenotypes of the Km insertions and the in-frame deletions confirmed that all three regulatory genes are required for Uhp function. Whereas the deletion of *uhpA* completely abolished the expression of a *uhpT-lacZ* reporter gene, the deletion of *uhpB* or *uhpC* resulted in a partially elevated basal level of expression that was not further inducible. These results indicated that UhpB and perhaps UhpC play both positive and negative roles in the control of *uhpT* transcription. Translational fusions of the *uhpBCT* genes to topological reporter gene *phoA* were generated by making use of restriction sites provided by the Km cassette or with transposon Tn*phoA*. The alkaline phosphatase activities of the resultant hybrid proteins were consistent with models predicting that UhpC and UhpT have identical transmembrane topologies, with 10 to 12 transmembrane segments, and that UhpB has 4 to 8 amino-terminal transmembrane segments that anchor the polar carboxyl-terminal half of the protein to the cytoplasmic side of the inner membrane.

Cells of *Escherichia coli* possess an inducible active transport system which is required for growth on glucose 6-phosphate (Glu6P), fructose 6-phosphate (Fru6P), and several other phosphorylated sugars and related compounds (4, 7, 21). This transport system, encoded by the *uhpT* gene, consists of a single polypeptide chain of 463 amino acids. It carries out an electroneutral exchange process that couples the accumulation of sugar phosphates to the downhill release of P_i (1, 26). The glycerol 3-phosphate transporter, GlpT, is similar to UhpT in both amino acid sequence and overall transport mechanism (1, 8).

The production of UhpT is induced specifically by extracellular Glu6P but not by intracellular Glu6P formed during the metabolism of glucose or other carbon sources (5, 33). The induction of *uhpT* transcription is regulated by the *uhpABC* genes, located immediately upstream of *uhpT* at 82.1 min on the *E. coli* genetic map. Previous studies have shown that all three *uhp* regulatory genes are required for *uhpT* expression and support a model for *uhp* regulation in which the membrane-localized UhpB and UhpC proteins respond to the presence of external Glu6P and convert the UhpA protein to a form able to activate *uhpT* transcription (30, 31).

Analysis of the deduced amino acid sequences of UhpA and UhpB showed that they possess regions with homology to the receiver and transmitter modules, respectively, of two-component regulatory systems (14, 31). Members of these families of regulatory proteins are widespread in bacteria and mediate responses to a variety of environmental signals by means of protein phosphorylation and phosphate transfer to the receiver protein, which functions in many cases as a transcriptional activator (28). A previous study concluded that UhpC was about 20 kDa in size and homologous to the middle third of UhpT (30). The homology between UhpC and UhpT suggested that UhpC could act as a receptor for Glu6P in the signaling process. Consistent with this hypothesis, the Uhp^- phenotype conferred by transposon mutations in *uhpC* could be overcome by additional mutations linked to the *uhp* region (13, 30). The expression of *uhpT* in these UhpC-independent variants was *trans*-dominant and constitutive.

Subsequently, several errors were found in the published sequence of *uhp*; the corrected sequence is reported here and was corroborated by the cloning and sequencing of the *uhp* region from *Salmonella typhimurium*. The *uhp* regions from *E. coli* and *S. typhimurium* are highly related, and the four *uhp* genes are closely spaced, with minimal intergenic intervals. UhpC is almost identical in length to UhpT, and these proteins share considerable sequence relatedness and have similar hydropathic distributions, suggesting very similar topological dispositions in the cytoplasmic membrane.

To define more precisely the role of the *uhp* regulatory genes, we isolated insertions of a kanamycin resistance (Km) cassette in numerous sites within the *uhp* locus. The locations and phenotypes of these insertion mutations support the new open reading frame assignments for *uhpC* and are consistent with those of the other three genes. In addition, in-frame deletions were generated by use of the unique restriction sites introduced by the Km insertions and allowed the determination of the null phenotypes resulting from the loss of a regulatory gene independently or in combination.

* Corresponding author.

TABLE 1. Bacterial strains, plasmids, and phages used in this study

Strain, plasmid, or phage	Relevant characteristics	Reference or source
Strains		
E. coli		
XL-1 Blue	*recA1 endA1 gyrA96 thi hsdR17* (r_K^- m_K^+) *supE44 relA1 lac* [F' *proAB lac*^{q} Δ*M15* Tn*10*]	Stratagene
MC4100	Δ(*argF-lac*)*U169 araD139 rpsL150 relA flbB5301 deoC1 ptsF25 rbsR*	24
JC7623	*recB21 recC22 sbcB15 leu arg his thr pro ara*	20
Y1090r$^-$	Δ(*argF-lac*)*U169* Δ*lon araD139 rpsL150 supF trpC22*::Tn*10* [pMC9] *hsdR* (r_K^- m_K^+)	Stratagene
CC118	*araD139* Δ(*ara-leu*)*7697* Δ*lacX74 phoAD20 galE galK thi rpsE rpoB argE*(Am) *recA1*	16
RK4353	Δ(*argF-lac*)*U169 araD139 thi gyrA219 relA rpsL150 non*	23
RK6894	RK4353 Δ(*ilv-uhpABCT'*)*2095 polA1 zig621*::Tn*10*	30
RK9332	RK4353 Δ*uhp*(A15–A189) *metE*	This work
RK9334	RK4353 Δ*uhp*(B60–B489)::Km	This work
RK9335	RK4353 Δ*uhp*(C41–C437)::Km	This work
RK9336	RK4353 Δ*uhp*(B60–C437)::Km	This work
RK9337	RK4353 Δ*uhp*(T1–T427)::Km	This work
RK9338	RK6894 λRZ5-P_T*lac*[Km]	This work
RK9339	MC4100 λRZ5-P_T*lac*	This work
S. typhimurium ST422	*recA metA22 metE551 trpC2 ilv-452 flaA66 rpsL120 xyl-404 hsdT6 galE496*	K. E. Sanderson
Plasmids		
pKB18	6.3-kb *S. typhimurium uhp* region in Bluescript pKS$^+$	This work
pKB23	2.6-kb *S. typhimurium uhp'CT* region in pKS$^+$	This work
pKB24	3.7-kb *S. typhimurium uhpABC'* region in pKS$^+$	This work
pMI29	6.0-kb *E. coli uhp* region in Bluescript pKS$^+$	This work
pRS415:P_T*RsaI*	*uhpT-lacZ* operon fusion plasmid	17
pSWFII	*phoA* fusion cassette (frame 1)	6
pKS-*phoA*	*phoA* fusion cassette (frame 2)	This work
pMLB1113	*lacZ* fusion vector (*lacI*^{q})	M. Berman
Phages		
λRZ5	Ap *lacZ* fusion vector	22
λRZ5-P_T*lac*	Ap *uhpT-lacZ* fusion phage	17
λRZ5-P_T*lac*[Km]	Km *uhpT-lacZ* fusion phage	This work

This approach was not complicated by the polar effects that transposon insertions often have on genes distal to their site of insertion. We show that whereas *uhpA* was absolutely required for the transcription of a *uhpT-lacZ* reporter gene when all genes were present in single copies, the absence of *uhpB* or *uhpC* resulted in an elevated basal level of expression. Induction by Glu6P, however, was blocked by the loss of any of these regulatory genes. These results indicate that UhpB and possibly UhpC play both positive and negative roles in the control of *uhpT* transcription.

The new restriction sites associated with the Km insertions also allowed the construction of fusions of the *uhp* genes to the topological reporter, PhoA, encoding the mature portion of periplasmic alkaline phosphatase (16). Analysis of the enzymatic activity expressed by fusions of different lengths provided extended support for the previous topological map of UhpT (15) and showed that the topology of UhpC is likely to be identical to that of UhpT. Finally, the properties of the UhpB-PhoA fusions were consistent with the presence of multiple transmembrane segments in the nonpolar amino-terminal half of UhpB and with the location of the carboxyl-terminal half in the cytoplasm.

MATERIALS AND METHODS

Plasmids, phages, and bacterial strains. The strains used in this work are described in Table 1. Plasmid constructions and manipulations were performed with *E. coli* XL-1 Blue (Stratagene, La Jolla, Calif.). Plasmid pMI29 was constructed by cloning the 6.0-kb *Eco*RI-*Bam*HI fragment from pRK10 (30), containing the entire *uhp* region, into Bluescript plasmid pKS$^+$ (Stratagene) from which several restriction sites in the polylinker region had been removed. Plasmid pRS415:P_T*RsaI* contains the *E. coli uhpT* promoter region (nucleotides 3472 to 3848) fused to *lacZ* in pRS415 (17, 25). The *uhpT-lacZ* operon fusion from this plasmid was transferred to phage λRZ5 (22), a gift from R. Zagursky, by recombination to form λRZ5-P_T*lac* (17). An analogous construct, λRZ5-P_T*lac*[Km], conferring kanamycin resistance, was prepared by cloning a Km cassette into the unique *Pst*I site of the β-lactamase gene in pRS415:P_T*RsaI* and then transferring the fusion to λRZ5 as described above. Lysogens containing λRZ5-P_T*lac*[Km] in strain RK6894 or λRZ5-P_T*lac* in strain MC4100 were selected by making use of the antibiotic resistance of the phage (24). To obtain Uhp$^+$ revertants, we plated 3×10^8 to 5×10^8 cells from overnight cultures of three independent isolates of strains carrying each chromosomal *uhp*::Km insertion or a deletion of *uhpBC* on minimal plates with Fru6P as the carbon source. The rich medium was L broth; the minimal medium was medium A with required supplements (19).

DNA manipulations. Routine recombinant DNA procedures were performed by use of standard methods or conditions specified by the product's manufacturer. Rapid plasmid

preparations were made by a modification of the alkaline lysis method (2), in which an ammonium acetate precipitation step replaced the phenol-chloroform extraction. Double-stranded DNA sequencing was performed on rapid plasmid preparations with the Pharmacia T7 sequencing kit. Oligonucleotide sequencing primers were purchased from Synthecell Corp. (Rockville, Md.).

Generation of Km cassette insertions. Partial digestions with restriction enzymes were performed in the presence of ethidium bromide (60 μg/ml) to favor the recovery of DNA molecules cut at a single site. The full-length linear fragments were isolated by agarose gel electrophoresis and ligated with the 1.4-kb Km determinant. This Km cassette was obtained by digestion of pUC-4K (Pharmacia, Piscataway, N.J.) with *Hin*cII for ligation into pMI29 fragments with blunt ends, with *Bam*HI for the *Sau*3A-treated plasmid, or with *Acc*I for the *Hpa*II-digested plasmid. The ligation mixtures were introduced by transformation into strain XL-1 Blue, and transformants resistant to ampicillin and kanamycin were selected. The oligonucleotides used as sequencing primers (5′-CAGAGATTTTGAGACACAAC and 5′-CGCT GACTTGACGGGACGGC) hybridize to opposite strands of the Km cassette, just inside the 200-bp inverted repeat segments at the ends of the cassette.

Transfer of mutations to the chromosome. Plasmids containing *uhp*::Km insertions were digested with *Bam*HI, and the full-length, linear DNA molecules were introduced into strain JC7623 by transformation as described by Winans et al. (32). The cassette was transferred at a low frequency to the chromosome by homologous recombination events in the flanking *uhp* regions. Plasmid pMI29 and derivatives carrying *uhp* deletions were transferred to strain RK9338 (*polA uhpT-lacZ*) by transformation and selection for the antibiotic resistance of the plasmid. In this strain, the plasmid cannot replicate autonomously (11), and ampicillin-resistant transformants in which the plasmid was integrated into the chromosome at a low frequency were isolated.

Assays of Uhp function and regulation. Uhp function was screened by growth on 0.2% Fru6P as carbon source, relative to growth on glucose at the same concentration. β-Galactosidase activities were determined on cultures grown to an optical density at 590 nm (OD_{590}) of ca. 0.1 with or without 0.34 mM Glu6P as inducer and treated with 0.01% sodium dodecyl sulfate (SDS)-chloroform (1:1) in polypropylene microtiter plates. The rate of hydrolysis of 3 mM *o*-nitrophenyl-β-D-galactopyranoside was determined from the change in the A_{415} and normalized for culture density (at 590 nm). Assays were performed with polystyrene microtiter plates and read with a spectrophotometric plate reader as described below. Each reported activity, $mOD_{415-590} \times minute^{-1} \times OD_{590}\ culture^{-1}$, is the average of at least three determinations on three independent isolates of each strain.

Cloning of *S. typhimurium* DNA. *S. typhimurium* genomic DNA was extracted from strain ST422 as previously described (24), partially digested with *Eco*RI, ligated into the *Eco*RI site of λgt11, and packaged in vitro with Gigapack II Plus extracts (Stratagene) in accordance with the manufacturer's instructions. The library was amplified by passage through *E. coli* Y1090r$^-$ and used to infect *E. coli* RK9332 (Δ*uhpA*) and RK9337 (Δ*uhpT*). λ transducing phages were recovered from Uhp$^+$ transductants by induction of liquid cultures at 42°C for 15 min and incubation at 37°C for 2 h.

Nucleotide sequence determination. Nucleotide sequencing was performed by the method of Sorge and Blinderman (27) with the ExoMeth sequencing kit from Stratagene. Cesium chloride-purified plasmid DNA was digested with pairs of restriction enzymes to generate 3′ and 5′ overhangs at one end of the region to be sequenced. For pKB24, *Kpn*I and *Sal*I were used to determine the sequence in one direction and *Pst*I and *Eco*RI were used for the opposite direction. For pKB23, *Kpn*I-*Xho*I and *Sst*I-*Bam*HI were used. The resulting linear DNA fragments were treated with exonuclease III for various times, and portions of the reaction mixture at each time were subjected to the dideoxy sequencing protocol with 5-methyl-dCTP. The reaction mixtures were digested with *Rsa*I prior to running of the sequencing gel to establish fixed endpoints for reading the sequence.

Construction of *uhp-phoA* fusions. Fusions of the *phoA* gene to regions of *uhp* were constructed in vitro by use of the *Pst*I sites at the ends of the Km insertions and cassettes containing the *phoA* gene bounded by polylinker sequences. Two *phoA* cassettes were used to allow fusion to *Pst*I sites in either of two translational reading frames. One cassette was derived from plasmid pSWFII (6) by cleavage with *Pst*I and partial digestion with *Xba*I. This cassette was cloned into *Pst*I-*Xba*I-cut pMI29::Km plasmids to replace the Km cassette and distal *uhp* sequences with *phoA* fused in-frame to *uhp* at positions corresponding to amino acid residues B-182, B-250, B-411, C-91, C-361, C-404, T-157, T-189, T-217, T-221, and T-366. The second cassette was constructed by subcloning the *Sma*I-*Sac*I, *phoA*-containing fragment of pSWFII into pKS$^+$ cut with *Sma*I-*Sac*I, which displaces the *phoA* coding frame by 1 base relative to the *Pst*I site. The *Pst*I-*Xba*I fragment from this construct (pKS-*phoA*) was used as described above to form fusions at *uhp* positions corresponding to amino acid residues B-87, B-151, B-240, B-345, C-224, C-241, T-249, T-332, and T-427.

Additional fusions were obtained by in vivo transposition of Tn*phoA* into an expression vector in which *uhpBC* was inserted downstream of the *lac* promoter, with amino acid 11 of LacZ fused to residue 189 of UhpA via the pUC8 polylinker. This plasmid was constructed by cloning the 2.8-kb *Pst*I-*Nsi*I fragment (nucleotides 666 to 3531) of *uhp* from plasmid pMI29::Km[A189] into vector pMLB1113 cut at the *Pst*I site of its polylinker by partial digestion. Plasmid pMLB1113 is a pBR322 derivative containing the *lacI*^{q} and *lacZ* genes, with the pUC8 polylinker inserted at codon 11 of *lacZ*, and was provided by Michael Berman. Products with the fragment in the proper orientation were identified by restriction mapping, and the plasmid was introduced into strain CC118. Transposition of Tn*phoA* into the plasmid was selected after introduction of F′ Tn*phoA* as described by Manoil and Beckwith (16), with some modifications. Light and dark blue colonies on LB plates containing 300 μg each of ampicillin and kanamycin per ml and 40 μg of 5-bromo-4-chloro-3-indolyl phosphate per ml were purified, and plasmid DNA was prepared. For plasmids carrying a Tn*phoA* insert, the sequence at the fusion junction was determined with a primer (5′-CGCCCTGAGCAGCCCGG) which anneals within the *phoA* gene. There were many fusions to the β-lactamase gene, and no low-activity fusions occurred in *uhp*.

High-activity fusions in *uhp* were transferred to plasmid pMI29 for comparison with in vitro-generated fusions. The *Pst*I-*Xba*I fragment containing the *phoA* gene from pKS-*phoA* was cloned into plasmid pMI29::Km[A189], replacing the Km cassette and 3′ *uhp* sequences with *phoA*. The *Pst*I-*Sph*I *uhp-phoA′* fragment of this construct was replaced with the *Pst*I-*Sph*I *uhp-phoA′* fragment of each Tn*phoA* fusion to generate plasmids with the same structure as the in vitro-generated fusions. The only differences between these constructs and the fusions constructed in vitro were four

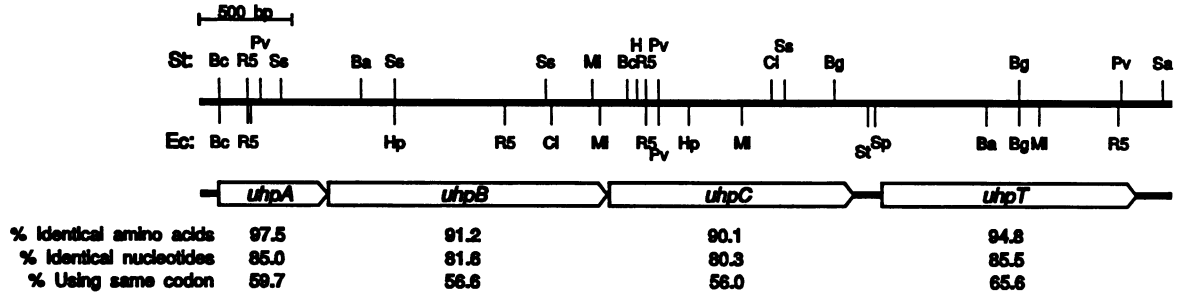

FIG. 1. Structure and comparison of the *uhp* regions in *E. coli* and *S. typhimurium*. Restriction maps of the *E. coli* (Ec) and *S. typhimurium* (St) *uhp* regions are presented on the upper line. Cleavage sites are abbreviated as follows: Bc, *Bcl*I; R5, *Eco*RV; Pv, *Pvu*I; Ba, *Bal*I; Ss, *Sst*II; Hp, *Hpa*I; Cl, *Cla*I; Ml, *Mlu*I; H, *Hin*dIII; Bg, *Bgl*II; Sa, *Sal*I. There were no sites for *Bam*HI, *Eco*RI, *Kpn*I, *Nco*I, *Pst*I, *Sst*I, *Sma*I, or *Xho*I. The location and the direction of the *uhp* open reading frames are presented on the middle line, and a summary of the percent identity in nucleotide and amino acid sequences is presented on the lower line.

amino acids inserted at UhpA residue 189. There were also differences in the linker region preceding *phoA* in constructs made with pSWFII, pKS-*phoA*, or Tn*phoA*.

Assays of alkaline phosphatase and β-lactamase activities. Alkaline phosphatase activity was measured by a modification of the procedure of Michaelis et al. (18). Isolates of strain CC118 carrying different *uhp-phoA* fusion plasmids were grown overnight in LB broth supplemented with 40 μg of ampicillin per ml, diluted 1:100 in the same medium, and grown at 37°C to an OD_{590} of ca. 0.2. Two 1.5-ml portions of each culture were sedimented and suspended in 0.15 ml of 2 M Tris-HCl (pH 8.0) for the alkaline phosphatase assay or 0.1 M Tris-HCl (pH 7.4) for the β-lactamase assay. The turbidity of a 1:10 dilution of each sample was measured, and 0.1 ml of the remaining material was lysed with 0.02 ml of 0.1% SDS-chloroform (1:1) in polypropylene microtiter plates. The enzymatic reaction was started by adding 0.05 ml of cell lysate to 0.05 ml of reaction mixture in wells of polystyrene microtiter plates. Changes in the absorbance were measured over 10 min at 37°C with a $THERMO_{max}$ microplate reader (Molecular Devices, Menlo Park, Calif.). For the alkaline phosphatase assay, the reaction mixture contained 2 M Tris (pH 8.0) and 4 mg of Sigma 104 phosphatase substrate (*p*-nitrophenyl phosphate) per ml and the increase in the $OD_{405-590}$ was measured. For the β-lactamase assay, the reaction mixture contained 0.1 M Tris (pH 7.4) and 4.0×10^{-5} M pyridine-2-azo-*p*-dimethylalanine cephalosporin (PADAC; Calbiochem, San Diego, Calif.) and the decrease in the $OD_{590-650}$ was measured (12). Both alkaline phosphatase and β-lactamase activities were corrected for cell density, and the expression of β-lactamase was used to correct alkaline phosphatase activity for slight differences in plasmid copy number.

Nucleotide sequence accession numbers. The GenBank accession numbers for the *E. coli* and *S. typhimurium uhp* sequences are M89479 and M89480, respectively.

RESULTS

Sequence of the *S. typhimurium uhp* region. The *S. typhimurium uhp* genes were cloned from a library of *Eco*RI-generated DNA fragments in λgt11 by selection for growth on Fru6P of transductants of *E. coli* Δ*uhpA* (RK9332) and Δ*uhpT* (RK9337) strains. One recombinant phage able to complement transposon insertion mutations in all four *E. coli uhp* genes was obtained. This phage carried an insert of a 6.3-kb *Eco*RI fragment, which was cloned into the *Eco*RI site of pKS^+ to yield plasmid pKB18, which complemented *uhpA* and *uhpT* mutations. Plasmid pKB18 was cut by *Cla*I into two fragments. The 2.6-kb fragment was cloned into the *Cla*I site of pKS^+ to yield pKB23, which carried part of *uhpC* and all of *uhpT*. The other fragment, of 6.6 kb, was ligated to a circular form, yielding pKB24, which had a 3.7-kb insert carrying *uhpA*, *uhpB*, and part of *uhpC*.

The nucleotide sequence of the 5,465-bp region was determined by the ExoMeth procedure of Sorge and Blinderman (27) for the generation of nested deletions from the ends of the insert. An average of 5.6 gel readings were obtained for each sequence character. The restriction map for the *S. typhimurium uhp* locus differed considerably from that for the *E. coli uhp* locus, although some sites were common to both (Fig. 1). The nucleotide sequences from the two organisms exhibited 81.5% overall identity and were readily aligned (Fig. 2). Four *uhp* open reading frames were present in the *S. typhimurium* sequence. The *uhpA* and *uhpT* genes were identical in size and encoded polypeptides having greater than 94% amino acid sequence identity with their *E. coli* homologs. However, the *uhpB* and *uhpC* coding regions were different in size and deduced polypeptide sequence from the published *E. coli* sequence (9).

Correction of the *E. coli uhp* sequence. During this work, several errors in the published *E. coli uhp* sequence were discovered in regions of gel compression. The changes that must be made in the published sequence (9) are deletion of a G at position 1762, insertion of a C at position 1823, change of CG to GC at position 2123, and insertion of a G at positions 2134, 2167, 2722, and 3271. These changes were corroborated by comparison with the location, length, and coding capacity of the open reading frames in the *S. typhimurium* sequence.

Comparison of the *uhp* regions. The corrections in the nucleotide sequence of the *uhp* region changed the deduced sequence of the *uhpB* and *uhpC* gene products and markedly reduced the length of the intergenic spaces. Several salient features of *uhp* gene organization are summarized here. An open reading frame precedes the *uhpA* coding region in *S. typhimurium* and ends at the same position as the *ilvN* gene in *E. coli* (residues 24 to 26 in Fig. 2), although the deduced sequence is different. These coding regions in both organisms are followed by a typical rho-independent terminator structure which is highly conserved and appears to overlap the *uhpABC* promoter. The amino acid sequence of UhpA is strongly conserved, with only five conservative substitutions among 196 codons (nucleotides 101 to 692). The termination codon of *uhpA* overlaps the initiation codon of *uhpB* in the sequence TGATG in *E. coli* and the sequence TAATG in *S.*


Sty
Eco
UhpA
UhpB
UhpT
UhpC

FIG. 2. Sequence of the *uhp* region. Aligned nucleotide sequences for the *uhp* regions from *E. coli* and *S. typhimurium* are presented. The amino acid sequences predicted for *S. typhimurium* are indicated above in single-letter code; the amino acids for *E. coli* are shown below only when they differ from those for *S. typhimurium*. Putative Shine-Dalgarno sequence residues are indicated by dots; termination codons are indicated by @; gaps inserted to maintain maximal alignment are indicated by dashes.

typhimurium. The *uhpB* coding region extends for 500 codons (nucleotides 691 to 2193). The UhpB amino acid sequences are 91% identical in the two organisms. The similarity of UhpA and UhpB to receiver and transmitter modules of sensory transduction pathways has been described (14, 28, 31), and the correction of two short regions in the published sequence at the carboxyl-terminal end of UhpB resulted in an increased match to other members of the protein kinase family.

The *uhpB* coding sequence is followed by a 9-bp intergenic region before the start of the 442-codon *uhpC* coding sequence (nucleotides 2203 to 3531). The *E. coli uhpC* gene begins with two tandem AUG codons, whereas in *S. typhimurium* the first of these codons is AUA. On the basis of the distance to the Shine-Dalgarno sequence, we suspect that both genes start at the conserved AUG codon at nucleotide 2203. UhpC of *S. typhimurium* is longer than UhpC of *E. coli* owing to the insertion of three amino acids just before the last residue. There is 90% amino acid sequence identity for UhpC from the two organisms.

The termination codon of *uhpC* overlaps part of the promoter region of *uhpT*. There is an intergenic segment of about 150 bp before the start of the *uhpT* coding region of 463 amino acids (nucleotides 3671 to 5063). The UhpT protein sequences are very strongly conserved, with 94.8% identity. UhpT and UhpC are related along their entire lengths, with 30.7% identical amino acids. About 11% of the residues are identical in all members of the family of membrane proteins, which includes UhpC and UhpT in both organisms, glycerol 3-phosphate transporter GlpT of *E. coli* (8), and phosphoglycerate transporter PgtP of *S. typhimurium* (10). The degree of amino acid identity shared by UhpC and UhpT is about the same as that among transporters UhpT, GlpT, and PgtP, between 27.8 and 33.3%. There were no large stretches of amino acid sequence in which UhpC differed from the transport proteins, which might represent a signalling domain.

Generation of insertion mutations in *uhp*. Insertions of a Km cassette in the *uhp* region were made to provide a set of precisely localized mutations and to allow the generation of in-frame deletions and *phoA* fusions. The 1.4-kb Km cassette was ligated to compatible ends on the linear form of plasmid pMI29 generated by digestion with limiting concentrations of *Alu*I, *Hae*III, *Hpa*II, *Rsa*I, *Sau*3A, or *Tha*I. Restriction enzyme analysis and nucleotide sequencing across both junctions identified the sites of insertion and verified the presence of a single Km cassette in a full-length *uhp*$^+$ plasmid. Several insertions contained short deletions on one side of the cassette as a result of cleavage at multiple, closely spaced sites on pMI29 and were not studied further.

The Uhp phenotypes conferred by 46 Km cassette insertions in *uhp* were determined following transfer to the chromosome and replacement of the chromosomal *uhp*$^+$ allele (Fig. 3). Two insertions just upstream of the *uhp* coding region retained Uhp function, as shown by growth on Fru6P as a carbon source. One Uhp$^+$ insertion was less than 100 bp upstream of the start of the *uhpA* reading frame and 30 bp from its putative −35 region. As expected from previous studies with transposon insertions, almost all of the Km insertions in the corrected *uhp* reading frames eliminated Uhp function. The only *uhp*::Km insertion that retained a Uhp$^+$ phenotype lay after amino acid 437 in UhpC,

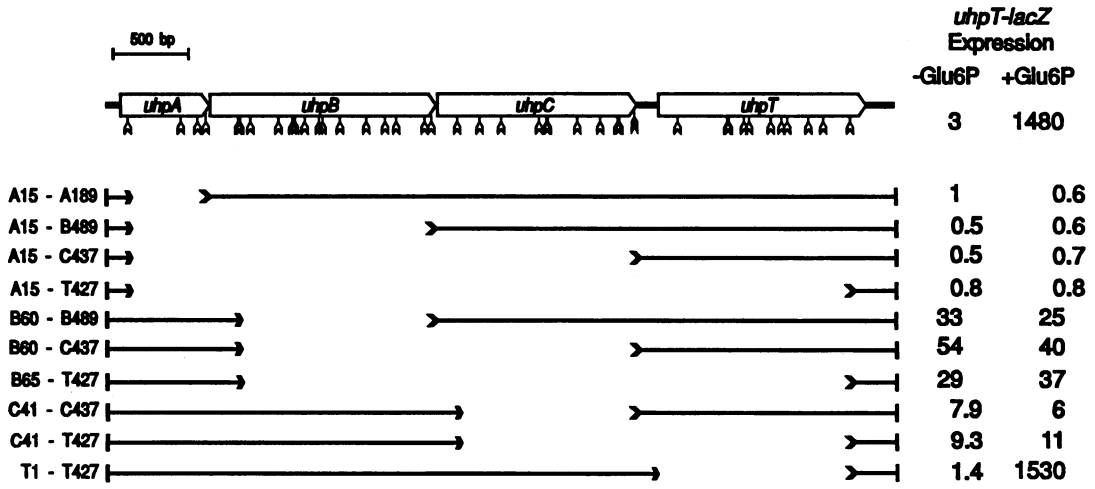

FIG. 3. Km cassette insertions and deletions in the *E. coli uhp* region. Arrowheads indicate the positions of Km cassette insertions, determined by nucleotide sequencing. The location of the insertion conferring a Uhp$^+$ phenotype is indicated by the filled arrowhead. The sequences remaining in the in-frame deletions are indicated below. Each deletion was derived by combining fragments of appropriate cassette insertions generated by cleavage at *Hin*dIII and *Bam*HI sites outside the *uhp* region and at the *Pst*I sites at the boundary of each insertion. The designation given to each in-frame deletion indicates the amino acids removed by that deletion: A, UhpA; B, UhpB; C, UhpC; T, UhpT. Deletions were transferred to the chromosome of strain RK9338 (λRZ5-P_T*lac*[Km]), and β-galactosidase activity expressed from the *uhpT-lacZ* reporter was measured after growth in the absence or presence of Glu6P. Each value is the mean of at least three determinations for three separate isolates and is given as $mOD_{415-590} \times minute^{-1} \times OD_{590}\ culture^{-1}$.

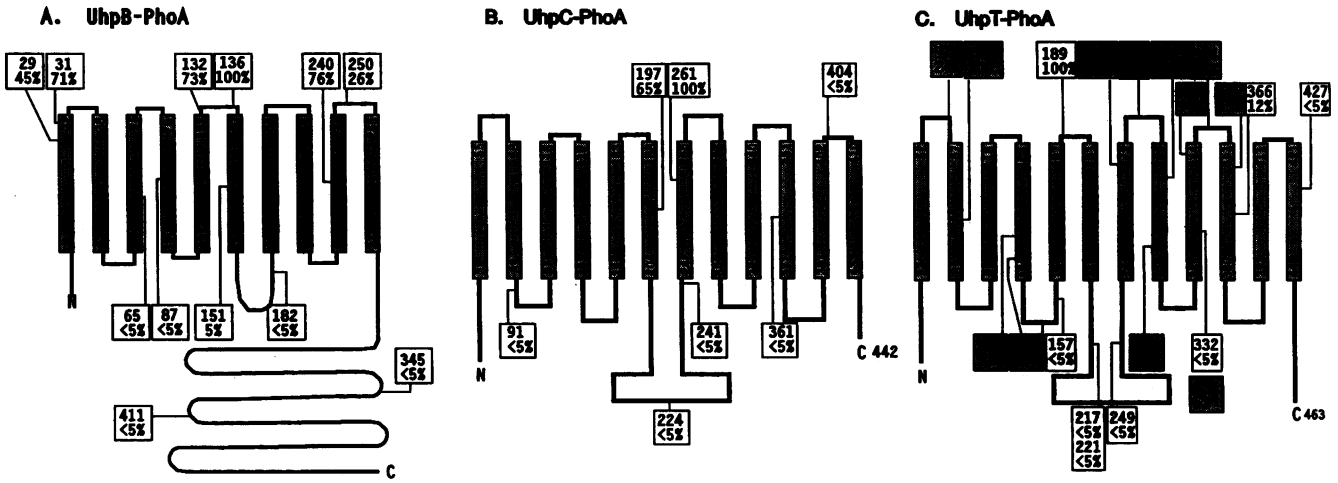

FIG. 4. Topological models based on the location and activity of *uhp-phoA* gene fusions. Two-dimensional topological models are presented for UhpB (A), UhpC (B), and UhpT (C). The shaded columns represent putative transmembrane segments, with the amino acid residues beginning and ending each segment indicated at each end. Heavy lines portray the cytoplasmic and periplasmic loops or termini of the polypeptides, oriented with the periplasmic space at the top. Thin lines connect the position of each fusion junction to a box which gives the last Uhp amino acid before the fusion junction (upper number) and the alkaline phosphatase/β-lactamase activity ratio relative to those of other fusions in the same gene (lower number). Shaded boxes in UhpT-PhoA (C) are previously reported fusions (15).

three codons from the end of that protein, indicating that the extreme carboxyl terminus of UhpC is not essential for its function.

Frequency of Uhp$^+$ revertants. The ability of the *uhp*::Km insertion mutants to yield Uhp$^+$ offspring was determined. Strains with insertions in *uhpA* or *uhpT* did not give rise to Uhp$^+$ progeny. The small colonies that arose on prolonged incubation yielded small colonies when restreaked on Fru6P medium. All of the mutants with insertions in *uhpB* or *uhpC* yielded Uhp$^+$ progeny at frequencies of ca. 10^{-7} per cell plated. All of these revertants retained the Km phenotype conferred by the original insertion mutation. Thus, it appears that the requirement for either *uhpB* or *uhpC* can be bypassed by second-site mutations, whereas the requirement for *uhpA* and *uhpT* cannot.

Regulatory phenotypes of null mutations. To test the effect of the loss of each *uhp* gene on the regulation of *uhpT* expression, we prepared null mutations in each *uhp* gene as in-frame deletions that removed most of the coding region without perturbing the reading frame. These deletions were prepared by use of the *Pst*I sites in the polylinker flanking each Km cassette. Digestion of pMI29::Km plasmids with *Pst*I and *Hin*dIII or with *Pst*I and *Bam*HI and exchange and ligation of the appropriate DNA fragments resulted in removal of the Km determinant and the desired portion of *uhp*. Each deletion was transferred to the chromosomal *uhp* locus, and its effect on *uhpT* expression was determined from the level of β-galactosidase encoded by a *uhpT-lacZ* fusion carried on a λRZ5 prophage. The wild-type induction ratio exhibited by this reporter system was at least 500 (Fig. 3).

The deletion of *uhpA* between amino acids 15 and 189 resulted in complete loss of *uhpT-lacZ* expression, as did longer deletions extending from early in *uhpA* to the distal ends of *uhpB*, *uhpC*, or *uhpT*. The levels of β-galactosidase were as low as those in uninduced cells. Thus, *uhpA* is absolutely required for *uhpT* transcription, regardless of whether the other *uhp* genes are present or absent. The in-frame deletion that removed most of *uhpB* conferred an elevated basal level of *uhpT* expression, about 1 to 3% the induced wild-type level. This expression was not changed by the presence of Glu6P. The same behavior was exhibited by strains with deletions of *uhpBC* and *uhpBCT*. The deletion of *uhpC* also resulted in an increased basal level of *uhpT-lacZ* expression that was not induced by the addition of Glu6P, although the increase in expression was smaller than that in the *uhpB* deletion. These results showed that UhpB and UhpC are required for the response to an inducer and that they also appear to act in a negative manner.

The deletion of *uhpT* had no apparent effect on the expression or regulation of the *uhpT-lacZ* reporter gene, confirming the previous conclusion that significant transport of Glu6P into the cell is not needed for the induction of *uhpT* expression (23).

Topology deduced from Uhp-PhoA fusions. Translational fusions of *uhpBCT* to the mature portion of *phoA* were generated with the restriction sites introduced on some of the Km cassettes. Additional fusions were made with Tn*phoA*. The alkaline phosphatase activities of strain CC118 carrying *uhp-phoA* constructs in plasmid pMI29 were measured relative to the β-lactamase levels to control for plasmid copy number. As the extent of the *uhp* sequence from all three genes was increased, there was a periodic alternation between fusions with substantial enzymatic activity and those with very low levels of enzymatic activity. These results are expected for fusions with proteins that span the cytoplasmic membrane multiple times. It is expected that high enzymatic activity will be displayed by fusions whose junction allows the PhoA moiety to be exported to the periplasmic space. Fusions that retain the PhoA moiety in the cytoplasm will show low enzymatic activity, as a result of the instability of the phosphatase domain, cleavage by proteases, or an inability to form disulfide bonds or to acquire the metal ions necessary for stability or activity (3, 15, 18).

The enzymatic activities of the 12 UhpB-PhoA, 7 UhpC-PhoA, and 9 UhpT-PhoA hybrids constructed in this study correlated with transmembrane topological models based on

hydropathy distributions and deductions from charge distributions of extramembranous loops (29) (Fig. 4). Figure 4 includes the results for 12 previously described UhpT-PhoA fusions (15). In almost all cases, the observed enzymatic activities fit the predicted model, although other possible topologies are not excluded. The few exceptions are discussed below. The deduced topologies of UhpC and UhpT are identical. UhpB is predicted to have 6 to 10 transmembrane segments which anchor the polar carboxyl-terminal half in the cytoplasm.

DISCUSSION

Cassette insertions in the *E. coli uhp* region helped define the boundaries and requirements for the genes in this locus. The phenotypes conferred by these precisely mapped insertions indicated that our previous reading frame assignment for *uhpC* was incorrect, as indicated by the Uhp^- phenotype of five insertions in sequences previously thought to be outside *uhpC* (9). The structure of the *uhp* locus has been clarified by correction of several errors in the published sequence and by an independent determination of the homologous region from *S. typhimurium*. Correction of the *E. coli* sequence revealed a higher-than-usual degree of relatedness between the *S. typhimurium* and *E. coli uhp* operons. The four *uhp* genes are arranged with minimal intergenic spacing, raising the possibility that translation of the *uhpABC* transcripts is coupled. The only *uhp*::Km insertion that was Uhp^+ was located three codons from the C terminus of UhpC. A further indication that the carboxyl terminus of UhpC is not critical to its function is the fact that this is the site of three additional amino acids in the *S. typhimurium* sequence. For UhpA or UhpB, insertions 7 or 11 codons, respectively, from the carboxyl terminus resulted in loss of function.

We had previously identified UhpC as a 20-kDa polypeptide expressed from a plasmid in which a DNA fragment encoding the distal portion of *uhpB* and intact *uhpC* was transcribed from a phage T7 late promoter (30). This polypeptide is instead likely to be the 168-amino-acid polypeptide encoded by the fusion of *ilvN* to the end of *uhpB*. The detection of UhpC is complicated by its low level of expression and its mobility on gel electrophoresis as a diffuse band with a size similar to that of UhpT.

Previous studies used transposon insertions, deletions, and overproduction of *uhp* gene products to address the role of individual *uhp* genes in sugar phosphate transport and its regulation (30, 31). These studies were limited by the imprecise localization of transposon insertion sites and deletion boundaries, by the polar effect that insertions can have on downstream genes, and by the difficulty in constructing strains which express pertinent gene products at normal levels. The in vitro-generated cassette insertions allowed the construction of in-frame deletions that had known endpoints and that removed the *uhp* genes individually or in combinations. These plasmid constructs were returned to the *E. coli* chromosome to generate strains expressing the remaining *uhp* gene products from their native promoter at a normal gene dosage. Using a *uhpT-lacZ* reporter to provide a more sensitive assay of gene expression than the measurement of transport activity, we obtained new information about the effect that the absence of one or more *uhp* gene products had on expression from the *uhpT* promoter.

As expected, all three *uhp* regulatory genes were required for the proper regulation of the *uhpT* gene (30). However, the loss of the *uhpB* and *uhpC* gene products resulted in a moderate increase in the basal expression of *uhpT* over the very low levels in uninduced cells. Thus, UhpB and UhpC might act in both positive and negative manners. One hypothesis is that, in the absence of induction, UhpB functions to maintain UhpA in an inactive state and thereby contributes to the extremely low basal level of expression in the Uhp system. Sequence homology suggests that UhpB may be a histidine-protein kinase and, in the presence of Glu6P, may phosphorylate and activate UhpA. In the absence of induction, it may act as a protein phosphatase to maintain UhpA in its unphosphorylated, inactive state. Such a system would reduce the effects of so-called "cross-talk" between Uhp and other two-component regulatory kinases.

The deletion of *uhpC* caused an increase in *uhpT-lac* expression similar to but smaller than that caused by the deletion of *uhpB*. Perhaps UhpC enhances UhpB phosphatase activity by the formation of a UhpBC complex. Although no direct evidence for an interaction between UhpB and UhpC has been presented, such an interaction may be critical to proper Uhp regulation. If UhpC contains the Glu6P-binding site relevant for regulation, transduction of the signal to UhpB through a protein-protein interaction is a probable mechanism. An unexpected result was the finding that both UhpB and UhpC could be bypassed by suppressor mutations, whereas we previously only had evidence that the UhpC function could be bypassed (13, 31).

The Km insertions also allowed the generation of translational fusions with topological reporter PhoA. We had previously described the use of transposon Tn*phoA* for the isolation of *uhpT-phoA* fusions (15). This approach was of limited value with the weakly expressed *uhp* regulatory genes because the level of PhoA activity in fusions to internal domains was so low that these fusions could not be reliably identified. In contrast, making *uhp-phoA* fusions at defined sites and in the proper reading frame allowed the application of this approach to these poorly expressed genes. Maintenance of the proper reading frame from the site of Km insertion required the use of several *phoA* cassettes, and we constructed fusions in two of the three possible frames. As was found with Tn*phoA* insertions and as is indicative of a transmembrane distribution, we found a periodic alternation of alkaline phosphatase activities with increasing lengths of the contributed UhpB, UhpC, or UhpT sequences.

The topology deduced for UhpT extended our previous results and agreed well with our previous model (15). The one site of discrepancy with the model for UhpT involves the longest fusion to residue 411, which had much lower activity than expected. Several explanations are possible. There may be only 10 transmembrane segments. On the other hand, the junction between UhpT411 and PhoA introduces several Arg residues, and this concentration of positive charge may interfere with the translocation of the preceding segment across the cytoplasmic membrane. Another possibility is that insertion of the 11th transmembrane segment, which has the lowest mean hydropathy value of the putative transmembrane segments, requires the presence of the 12th transmembrane segment, which is disrupted in the fusion. It is noteworthy that the same behavior is exhibited by a PhoA fusion to the analogous area of UhpC (UhpC404-PhoA).

The set of UhpC-PhoA fusions was not as extensive as that of UhpT-PhoA fusions and did not allow an independent assignment of all of the extramembranous loops. However, the activities of the fusions that were constructed were fully compatible with the UhpT topological model and provide new support for the cytoplasmic location of the central loop in both proteins.

The amino-terminal half of UhpB (residues 1 to 273) exhibits substantial hydrophobic character. A model predicting 10 transmembrane segments can be drawn on the basis of the hydropathy distribution and the distribution of charges in the extramembranous loops, according to the "interior-positive" rule of von Heijne (29). Two charged residues are included in the transmembrane region, with no net charge. This model is fully compatible with the properties of the UhpB-PhoA fusions, although other models cannot be excluded. The removal of putative transmembrane segments 3 and 4 or 7 and 8 or both from the membrane to cytoplasmic loops would also be consistent with the results presented. Nonetheless, these results indicate that UhpB contains 6, 8, or 10 transmembrane segments and that the carboxy-terminal polar half is cytoplasmic. Our results do not exclude the possibility that the amino terminus is periplasmic, although the absence of a signal sequence renders this possibility unlikely. This topology of UhpB supports models of *uhp* regulation which postulate that the membrane-embedded portions of UhpB and UhpC interact to transmit the signal presented by external Glu6P, while the kinase portion of UhpB is accessible to cytoplasmic UhpA.

ACKNOWLEDGMENTS

This work was supported by Public Health Service research grant GM38681 from the National Institute of General Medical Sciences. M.D.I. received postdoctoral training support from NRSA CA09109 from the National Cancer Institute.

We thank Tod Merkel for many helpful discussions and for supplying λRZ5-P_T*lac* and pRS415:P_T*Rsa*I. We are also grateful to Kate Jensen for excellent technical assistance.

REFERENCES

1. **Ambudkar, S. V., T. J. Larson, and P. C. Maloney.** 1986. Reconstitution of sugar phosphate transport systems of *Escherichia coli*. J. Biol. Chem. **261:**9083–9086.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
3. **Derman, A. I., and J. Beckwith.** 1991. *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. J. Bacteriol. **173:**7719–7722.
4. **Dietz, G. W.** 1976. The hexose phosphate transport system of *Escherichia coli*. Adv. Enzymol. **44:**237–259.
5. **Dietz, G. W., and L. A. Heppel.** 1971. Studies on the uptake of hexose phosphates. II. The induction of the glucose-6-phosphate transport system by exogenous but not by endogenously formed glucose-6-phosphate. J. Biol. Chem. **246:**2885–2890.
6. **Ehrmann, M., D. Boyd, and J. Beckwith.** 1990. Genetic analysis of membrane protein topology by a sandwich gene fusion approach. Proc. Natl. Acad. Sci. USA **87:**7574–7578.
7. **Eidels, L., P. D. Rick, N. P. Stimler, and M. J. Osborn.** 1974. Transport of D-arabinose-5-phosphate and D-sedoheptulose-7-phosphate by the hexose phosphate transport system of *Salmonella typhimurium*. J. Bacteriol. **119:**138–143.
8. **Eiglmeier, K., W. Boos, and S. T. Cole.** 1987. Nucleotide sequence and transcriptional startpoint of the *glpT* gene of *E. coli*: extensive sequence homology of the G-3-P transport protein with components of the H-6-P transport system. Mol. Microbiol. **1:**251–258.
9. **Friedrich, M. J., and R. J. Kadner.** 1987. Nucleotide sequence of the *uhp* region of *Escherichia coli*. J. Bacteriol. **169:**3556–3563.
10. **Goldrick, D., G.-Q. Yu, S.-Q. Jiang, and J.-S. Hong.** 1988. Nucleotide sequence and transcription start point of the phosphoglycerate transporter gene of *Salmonella typhimurium*. J. Bacteriol. **170:**3421–3426.
11. **Greener, L., and C. W. Hill.** 1980. Identification of a novel genetic element in *Escherichia coli* K-12. J. Bacteriol. **144:**312–321.
12. **Jones, R. N., H. W. Wilson, and W. J. Novick, Jr.** 1982. In vitro evaluation of pyridine-2-azo-*p*-dimethylaniline cephalosporin, a new diagnostic chromogenic reagent, and comparison with nitrocefin, cephacetrile, and other beta-lactam compounds. J. Clin. Microbiol. **15:**677–683.
13. **Kadner, R. J., and D. M. Shattuck-Eidens.** 1983. Genetic control of the hexose phosphate transport system of *Escherichia coli*: mapping of deletion and insertion mutations in the *uhp* region. J. Bacteriol. **155:**1052–1061.
14. **Kofoid, E. C., and J. S. Parkinson.** 1988. Transmitter and receiver modules in bacterial signaling proteins. Proc. Natl. Acad. Sci. USA **85:**4981–4985.
15. **Lloyd, A. D., and R. J. Kadner.** 1990. Topology of the *Escherichia coli uhpT* sugar-phosphate transporter analyzed by using Tn*phoA* fusions. J. Bacteriol. **172:**1688–1693.
16. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82:**8129–8133.
17. **Merkel, T. J., D. M. Nelson, C. L. Brauer, and R. J. Kadner.** 1992. Promoter elements required for positive control of transcription of the *Escherichia coli uhpT* gene. J. Bacteriol. **174:**2763–2770.
18. **Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith.** 1983. Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. J. Bacteriol. **154:**366–374.
19. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Oishi, M., and S. D. Cosloy.** 1972. The genetic and biochemical basis of the transformability of *Escherichia coli* K12. Biochem. Biophys. Res. Commun. **49:**1568–1572.
21. **Pogell, B. M., B. R. Maity, S. Frumkin, and S. Shapiro.** 1966. Induction of an active transport system for glucose-6-phosphate in *Escherichia coli*. Arch. Biochem. Biophys. **116:**406–415.
22. **Roland, K. L., C. Liu, and C. L. Turnbough.** 1988. Role of the ribosome in suppressing transcriptional termination at the *pyrBI* attenuator of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **85:**7149–7153.
23. **Shattuck-Eidens, D. M., and R. J. Kadner.** 1981. Exogenous induction of the *Escherichia coli* hexose phosphate transport system defined by *uhp-lac* operon fusions. J. Bacteriol. **148:**203–209.
24. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
26. **Sonna, L. A., S. V. Ambudkar, and P. C. Maloney.** 1988. The mechanism of glucose 6-phosphate transport by *Escherichia coli*. J. Biol. Chem. **263:**6625–6630.
27. **Sorge, J. A., and L. A. Blinderman.** 1989. ExoMeth sequencing of DNA: eliminating the need for subcloning and oligonucleotide primers. Proc. Natl. Acad. Sci. USA **86:**9208–9212.
28. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53:**450–490.
29. **von Heijne, G.** 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. EMBO J. **5:**3012–3027.
30. **Weston, L. A., and R. J. Kadner.** 1987. Identification of Uhp polypeptides and evidence for their role in exogenous induction of the sugar phosphate transport system of *Escherichia coli* K-12. J. Bacteriol. **169:**3546–3555.
31. **Weston, L. A., and R. J. Kadner.** 1988. Role of *uhp* genes in expression of the *Escherichia coli* sugar-phosphate transport system. J. Bacteriol. **170:**3375–3383.
32. **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.
33. **Winkler, H. H.** 1970. Compartmentation in the induction of the hexose-6-phosphate transport system of *Escherichia coli*. J. Bacteriol. **101:**470–475.