JOURNAL OF BACTERIOLOGY, May 1994, p. 2611–2618
0021-9193/94/$04.00+0

Vol. 176, No. 9

UTP:α-D-Glucose-1-Phosphate Uridylyltransferase of *Escherichia coli*: Isolation and DNA Sequence of the *galU* Gene and Purification of the Enzyme

AUDREY C. WEISSBORN,* QINGYUN LIU, MARILYNN K. RUMLEY, AND EUGENE P. KENNEDY

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received 25 October 1993/Accepted 25 February 1994

The *galU* gene of *Escherichia coli*, thought to encode the enzyme UTP:α-D-glucose-1-phosphate uridylyltransferase, had previously been mapped to the 27-min region of the chromosome (J. A. Shapiro, J. Bacteriol. 92:518–520, 1966). By complementation of the membrane-derived oligosaccharide biosynthetic defect of strains with a *galU* mutation, we have now identified a plasmid containing the *galU* gene and have determined the nucleotide sequence of this gene. The *galU* gene is located immediately downstream of the *hns* gene, and its open reading frame would be transcribed in the direction opposite that of the *hns* gene (i.e., clockwise on the *E. coli* chromosome). The nucleotide sequences of five *galU* mutations were also determined. The enzyme UTP:α-D-glucose-1-phosphate uridylyltransferase was purified from a strain containing the *galU* gene on a multicopy plasmid. The amino-terminal amino acid sequence (10 residues) of the purified enzyme was identical to the predicted amino acid sequence (after the initiating methionine) of the *galU*-encoded open reading frame. The functional enzyme appears to be a tetramer of the *galU* gene product.

UDP-glucose is of central importance in the synthesis of the components of the cell envelope of *Escherichia coli* and in both galactose and trehalose metabolism. The biosyntheses of lipopolysaccharide (6, 26), capsular polysaccharide (13), and membrane-derived oligosaccharides (24) all require UDP-glucose as a glucosyl donor. In addition, UDP-glucose is an essential intermediate for growth on galactose (for a review, see reference 9), growth on trehalose (4), and synthesis of trehalose (7). The enzyme UTP:α-D-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) and the *galU* gene that is thought to encode it were therefore the subjects of much interest in pioneering studies of the synthesis of cell surface carbohydrates in *E. coli* and of galactose metabolism. However, many important problems were left unsolved, including the molecular characterization of the uridylyltransferase, its relation to other enzymes that catalyze the synthesis of UDP-glucose in cell extracts, and the precise localization and nucleotide sequence of the *galU* gene. Further unresolved problems arise from the early work of Nakae and Nikaido and colleagues (17–19, 21), who described experiments suggesting that the *galF* gene of *Salmonella typhimurium* may encode an activity that modifies the chromatographic behavior of the uridylyltransferase, although the molecular basis of the observed phenomenon remained undefined.

The present study arose from the characterization (34) of a set of mutants blocked in the production of membrane-derived oligosaccharides (MDO) of *E. coli*. One type of mutant (class IV), whose mutation was designated *mdo-453*, synthesizes neither capsular polysaccharide nor MDO and also fails to grow on galactose. On the basis of these phenotypes and the genetic map position of the mutation near 27 min (the approximate position of *galU* [25]), the class IV mutants were considered to have mutations in the *galU* gene. Another type of mutant (class III), whose mutation was designated *mdoX452*, was more puzzling. Like the class IV mutant, this isolate was blocked in the production of both MDO and capsular polysaccharide and its mutation also mapped to the 27-min region of the chromosome. However, this strain was still capable of growth on galactose. These characteristics of the class III mutant suggested that there was another gene distinct from *galU* in the 27-min region that is needed for the biosynthesis of MDO and capsular polysaccharide. To clarify the molecular basis of these mutations, we have now accurately localized and sequenced the *galU* gene. Both the class III and class IV *mdo* mutations are in fact caused by lesions in *galU*. Further, we have definitively shown that the *galU* gene is the structural gene for UTP:α-D-glucose-1-phosphate uridylyltransferase by purifying the enzyme and showing the identity of the enzyme's amino-terminal amino acid sequence with the predicted amino acid sequence of the *galU* open reading frame.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1.

Strain construction. The *galU* alleles were moved into the desired strains by P1 transduction (16). Plasmids were introduced into the desired strains by the $CaCl_2$ transformation procedure as previously described (23). The *galU* gene with the Tn*5* insertion was transferred to the chromosome of strain JC7623 by homologous recombination (35) after the plasmid pGM7::Tn*5-16* was linearized with *Pvu*I (cutting in the *lacZ'* segment of the vector). Subsequently, this allele was moved by P1 transduction to the desired strains.

Assay of UTP:α-D-glucose-1-phosphate uridylyltransferase. Many earlier studies employed an assay in which the formation of UTP and glucose 1-phosphate was measured after the pyrophosphorolysis of UDP-glucose. However, pyrophosphorolysis is the reverse of the physiologically important reaction catalyzed in vivo. The presence of high levels of inorganic pyrophosphatase in crude extracts has the result that the reverse reaction is not linear with time or with the amount of extract added. In our hands, assays of this type could not be

* Corresponding author. Phone: (617) 432-1860. Fax: (617) 738-0516.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristic(s)[a]	Source, derivation, and/or reference
Strains		
AB1133	*galK*	B. Bachmann, CGSC 1133
AR216	As AB1133 but with *mdoX452*	P1(RZ60-10H) × AB1133
AR240	As AB1133 but with *mdo-453*	P1(RZ60-3A) × AB1133
AR286	As AB1133 but with *galU*::Tn*5-16*	P1(JC7623-16) × AB1133
CA10	*galU95*	B. Bachmann, CGSC 4973
CA198	*galU106*	B. Bachmann, CGSC 4983
JC7623	*recB recC sbcD*	35; via D. Siegele
JC7623-16	As JC7623 but with *galU*::Tn*5-16*	This study
JW380	*zch-506*::Tn*10*	B. Bachmann, CGSC 6094
QE35	Δ*rfbA*	B. Bachmann, CGSC 5951
QE35-10	Δ*rfbA mdoX452*	P1(RZ60-10H) × QE35
QE35-3	Δ*rfbA mdo-453*	P1(RZ60-3A) × QE35
QE35-16	Δ*rfbA galU*::Tn*5-16*	P1(JC7623-16) × QE35
RZ60-3A	*mdo-453*	34
RZ60-10H	*mdoX452*	34
SØ874	Δ*rfbA*	B. Bachmann, CGSC 5935
W4597	*galU51*	B. Bachmann, CGSC 5012
Plasmids		
pAR200	1.35-kb *Bam*HI-*Pvu*II fragment from pGM7 into pUC18, Cb^r	This study
pGM7	2.7-kb *Sal*I-*Sal*I insert with the *galU*, *hns*, and *tdk'* genes, Cm^r	15
pGM7::Tn*5-7*	pGM7 with Tn*5* in *hns*	15
pGM7::Tn*5-16*	pGM7 with Tn*5* in *galU*	15

[a] Abbreviations: Cb^r, carbenicillin resistance; Cm^r chloramphenicol resistance.

used for the reliable estimation of glucose 1-phosphate in crude extracts, although they may be suitable for the measurement of purified enzyme preparations.

In the following procedure for estimation of the uridylyltransferase, the synthesis of labeled UDP-glucose from UTP and [^{14}C]glucose 1-phosphate is measured by a simple procedure for separation of the labeled product from glucose 1-phosphate on small columns of DEAE-cellulose in Pasteur pipets.

Cells to be tested for the uridylyltransferase were grown in medium containing 1% (wt/vol) Bacto Peptone supplemented with 50 mM NaCl, 5 mM K_2HPO_4, 1 mM $MgSO_4$, and 10 mM glucose. The added glucose represses the synthesis of glucose-1-phosphatase (22) which otherwise interferes with the assay. The cells in late log phase were harvested by centrifugation. This and all subsequent steps were carried out at 0 to 4°C. The cells were washed with a volume of 0.1 M Tris HCl (pH 7.8) equal to 1/5 of the original culture volume, resuspended in a volume of a buffer containing 10 mM Tris HCl (pH 7.8), 5 mM 2-mercaptoethanol, and 1 mM dithiothreitol, and disrupted by sonic irradiation with an MSE Ultrasonic Disintegrator. Cellular debris was removed by centrifugation at 2,000 × *g* for 10 min. Streptomycin sulfate was added to the crude extract such that the ratio (wt/wt) of streptomycin sulfate to protein was 0.6. The precipitate was removed by centrifugation at 43,000 × *g* for 1 h. Glycerol at a final concentration of 25% (vol/vol) was added to stabilize the enzyme (10), and the preparation was stored at −20°C. The crude extract could be stored for 3 to 4 weeks at −20°C with little loss of uridylyltransferase activity.

The usual assay system in a final volume of 0.1 ml contained 50 mM Tris HCl buffer (pH 7.8), 10 mM $MgCl_2$, 0.3 mM [U-^{14}C]glucose 1-phosphate (Amersham) (specific activity, 4,000 cpm/nmol), 0.3 mM UTP, 2 mM KH_2PO_4, 0.1 U of inorganic pyrophosphatase (Sigma), and various amounts of the streptomycin-treated extract described above (usually about 1 to 10 μg of total protein). After a 20-min incubation at 37°C in an Eppendorf centrifuge tube, the reaction was stopped by the addition of 0.9 ml of 30% (vol/vol) ethanol. The precipitate was removed by centrifugation, and 0.9 ml of the supernatant solution was passed over a 0.5-ml DEAE-cellulose acetate column equilibrated in 30% ethanol and prepared in a Pasteur pipet. The column was washed with 1.1 ml of 30% ethanol and then with 7 ml of an acetate buffer composed of 0.05 M acetic acid and 0.05 M ammonium acetate (final pH of 4.6) in 30% ethanol. These washes, which eluted unutilized glucose 1-phosphate, were discarded. The labeled UDP-glucose was then eluted with 4 ml of an acetate buffer composed of 0.15 M acetic acid and 0.15 M ammonium acetate (final pH of 4.6) in 30% ethanol. One milliliter of the eluate was mixed with 10 ml of Liquiscint (National Diagnostics), and the radioactivity was measured in a Packard 1600 TR liquid scintillation counter. One unit of enzyme activity is defined as 1 μmol of UDP-glucose produced per min.

When the apparent K_m determinations for the uridylyltransferase were made, the following substrate concentrations, specific activities, and enzyme amounts were used. The determination of the K_m for glucose 1-phosphate used UTP and [U-^{14}C]glucose 1-phosphate (200,000 cpm/nmol) at concentrations of 0.3 or 0.5 mM and of 2 to 25 μM, respectively, and 3 × 10^{-6} to 7 × 10^{-6} U of enzyme. The determination of the K_m for UTP used UTP and [U-^{14}C]glucose 1-phosphate (5,000 to 20,000 cpm/nmol) at concentrations of 20 to 1,000 μM and of 0.3 mM, respectively, and 2 × 10^{-5} to 5 × 10^{-5} U of enzyme. The determination of the K_m for TTP used TTP and [U-^{14}C]glucose 1-phosphate, (3,000 to 10,000 cpm/nmol) at concentrations of 0.3 to 5 mM and of 0.3 mM, respectively, and 0.7 × 10^{-4} to 1.4 × 10^{-4} U of enzyme. When the apparent K_m determinations for the thymidylyltransferase were made, the substrate concentrations, specific activities, and enzyme amounts were as follows. The determination of the K_m for TTP used TTP and [U-^{14}C]glucose 1-phosphate (14,000 cpm/nmol) at concentrations of 15 to 250 μM and of 0.2 mM, respectively,

and 1.7×10^{-5} U of enzyme. The K_m determination for UTP used UTP and [U-^{14}C]glucose 1-phosphate (1,100 cpm/nmol) at concentrations of 0.3 to 5 mM and 0.5 mM, respectively, and 1.7×10^{-4} U of enzyme. The substrate utilization was 15% or less for each velocity determination.

Measurement of the synthesis of MDO. A rapid procedure for the determination of the synthesis of MDO labeled with [2-^{3}H]glycerol was used as previously described (34).

DNA sequencing of the *galU* gene. The *Bam*HI-*Sna*BI and *Sna*BI-*Pvu*II fragments from pGM7 (15) and the *Bam*HI-*Hpa*I (the *Hpa*I site is 187 bp from the end of the Tn*5* element) fragment from pGM7::Tn*5-16* (15) (see Fig. 1) were subcloned into pUC18 by conventional techniques (23) with enzymes purchased from New England Biolabs and Boehringer Mannheim Biochemicals. The sequences of both strands were determined from the subclones by using alkali-denatured DNA, the Sequenase 2.0 kit from United States Biochemical, α-^{35}S-dATP from Dupont, NEN Research Products, primers from Promega (24-mer forward sequencing and 22-mer reverse sequencing), and primers (22-mers) synthesized in our departmental facility on a Milligen model 7500 DNA synthesizer (Millipore Corp.). The original plasmid, pGM7, was used as a template to confirm the sequence through the *Sna*BI site.

DNA sequencing of the *galU* mutations. The DNA sequences of *galU* mutations were determined by using DNA fragments amplified by the PCR. For PCR amplification, one colony from each *galU* strain was suspended in 50 μl of H_2O, heated to 100°C for 3 min, and chilled to 0°C immediately. Each suspension was then mixed with 50 μl of 2× PCR buffer (2× PCR buffer is 40 mM Tris HCl [pH 8.3], 0.1 M KCl, 4 mM $MgCl_2$, 0.2 mM each deoxynucleoside triphosphate, 0.5 μM each primer, 2 μg of gelatin per ml) and 2 U of AmpliTaq (Perkin-Elmer Cetus). The upstream and downstream primers contained the sequence from bases 83 to 104 and from bases 1312 to 1291 of the *galU* gene, respectively. After the DNA fragments were amplified with a DNA Thermal Cycler model 480 (Perkin-Elmer Cetus) for 25 cycles (one cycle consisting of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C), the DNA fragments were purified with GeneClean (Bio 101) and used directly for sequencing. The DNA sequences of the fragments were determined by using the primers (22-mers) described in the previous paragraph and the Sequenase 2.0 kit as previously described (29), except that the primer annealing step and the labeling step were done at 0°C (28).

Purification of UTP:α-D-glucose-1-phosphate uridylyltransferase. The strain SØ874 containing the plasmid pAR200 was grown in the Bacto Peptone medium described above with the additional ingredients histidine (22 μg/ml), tryptophan (20 μg/ml), and carbenicillin (100 μg/ml). When the culture reached an optical density at 650 nm of 0.3, the cells were harvested by centrifugation at 4,200 × *g* for 20 min. This and all subsequent steps after the cell pellet was frozen and thawed were carried out at 0 to 4°C. The thawed cell pellet was resuspended in 1/100 of the original culture volume with a buffer containing 10 mM Tris HCl (pH 8.1), 5 mM 2-mercaptoethanol, and 1 mM dithiothreitol. The cells were broken open by two passages through an Aminco French press at 16,000 lb/in^2. Unbroken cells were removed by centrifugation at 2,500 × *g* for 10 min. Streptomycin sulfate was added to the supernatant in an amount equal to 0.6 times the amount (in milligrams) of protein. The precipitation was allowed to reach equilibrium while being stirred gently for 20 min. The precipitate and membrane fragments were then removed by centrifugation at 39,000 × *g* for 75 min. A sample of the supernatant with 24 mg of protein was then loaded onto a Reactive Green 5-Agarose (Sigma Chemical Co.) column (1 by 23.5 cm) equilibrated in a buffer composed of 50 mM Tris HCl (pH 8.1), 3 mM $MgCl_2$, 100 mM KCl, and 1 mM dithiothreitol (buffer A). Fractions of 1.5 ml were collected at a flow rate of 10 ml/h. Under these column conditions, the uridylyltransferase flows through the column, while 40% of the protein is bound to the column. The flowthrough fractions containing the bulk of the protein were pooled, and the protein, including the enzyme of interest, was precipitated by the addition of ammonium sulfate to a final concentration of 80% saturation. The precipitate was collected after 30 min by centrifugation at 12,000 × *g* for 15 min. The precipitate was resuspended in 5 ml of buffer A and loaded onto a Sephacryl S-200 (Pharmacia-LKB) column (1.1 by 92 cm) equilibrated in buffer A. One-milliliter fractions were collected at a flow rate of 10 ml/h. The four fractions with the highest specific activities were pooled, and glycerol was added to a final concentration of 25%. The final enzyme preparation was stored in the freezer. Two 24-mg batches of protein (from 350 ml of original culture volume) were purified through the two columns and combined for the final purified enzyme preparation as outlined in Table 4.

The molecular weight of the enzyme was estimated by chromatography on Sephacryl S-200. Blue Dextran and glucose were used to determine the void volume and included volume, respectively. Human immunoglobulin G (160,000), bovine serum albumin (66,000), ovalbumin (45,000), carbonic anhydrase (29,000), and chymotrypsinogen A (25,000) were used as molecular weight standards (all proteins from Sigma Chemical Co.).

Protein concentrations were determined by the microassay procedure described for use with the Bio-Rad Protein Assay Reagent (Bio-Rad Laboratories). Ovalbumin was the standard protein.

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis was done by the procedure of Laemmli (11) with a 12% acrylamide separating gel.

Amino acid sequencing of the amino terminus of UTP:α-D-glucose-1-phosphate uridylyltransferase. The final purified enzyme was prepared for sequencing by transfer to an Immobilon-P polyvinylidene difluoride membrane (Millipore Corp.) following SDS-polyacrylamide gel electrophoresis as previously described (14). Gas-phase protein sequencing was done on an Applied Biosystems 477A/120A instrument modified as described by Tempst and Riviere (30).

Nucleotide sequence accession number. The nucleotide sequence of the *galU* gene has been submitted to the GenBank data base and assigned accession number M98830.

RESULTS

Identification of a gene complementing *mdoX*, *mdo-453*, and *galU*. P1 transductional mapping of the *mdoX452* (class III) and *mdo-453* (class IV) mutations indicated that they were counterclockwise to the *hns* gene in the 27-min region of the chromosome and very closely linked to *hns* (34). Plasmid pGM7 (Table 1) of May et al. (15) was therefore tested for its ability to complement the MDO biosynthetic defect caused by the *mdoX452*, *mdo-453*, and *galU* mutations. This plasmid (Fig. 1) contains a 2.7-kb insert of *E. coli* DNA and is known to have the *hns* gene, an additional 1.3 kb of DNA downstream of *hns*, and, upstream, part of the thymidine kinase gene (3). As is seen in Table 2, introduction of the plasmid pGM7 allows complementation of the MDO defect in the *mdoX452* and *mdo-453* strains. In addition, MDO production in three other classically studied *galU* strains, W4597 (6), CA198 (25), and CA10 (25), could also be restored by plasmid pGM7 (Table 2). The complementing gene was not the *hns* gene, because a

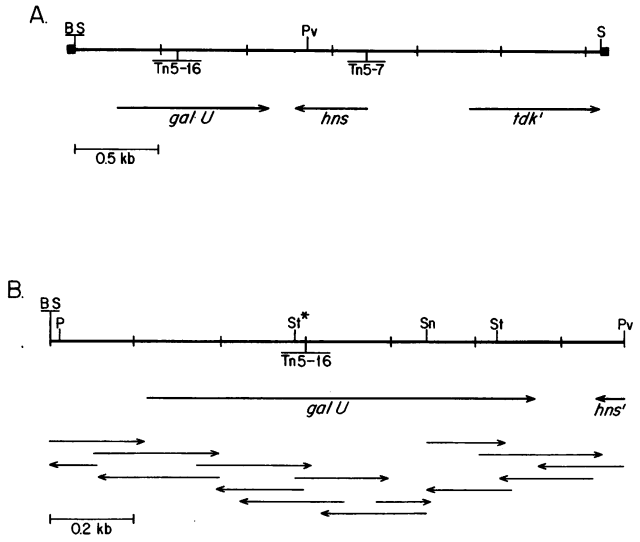

FIG. 1. Restriction map and sequencing strategy of the *galU* gene. (A) The *SalI-SalI* insert of plasmid pGM7 (15) is depicted. The locations of the two Tn*5* insertion elements in derivative plasmids (15) are also indicated. The extents of the *galU* and *hns* genes and part of the *tdk* gene on the insert are shown below the insert map. (B) The *SalI-PvuII* fragment from the map shown in panel A is depicted here with the restriction sites either used in subcloning fragments for sequencing or noted in previous references (15, 31). The *StuI* site on the left (asterisk) is cleaved only if the plasmid DNA in prepared from a Dcm$^-$ strain. The location of the Tn*5* insert in the *galU* gene is again shown. The extent of the *galU* gene and the portion of the *hns* gene are depicted immediately below the map. The arrows under the map indicate the directions and lengths of the regions that were sequenced. Abbreviations: B, *Bam*HI; S, *Sal*I; P, *Pst*I; St, *Stu*I; Sn, *Sna*BI; Pv, *Pvu*II.

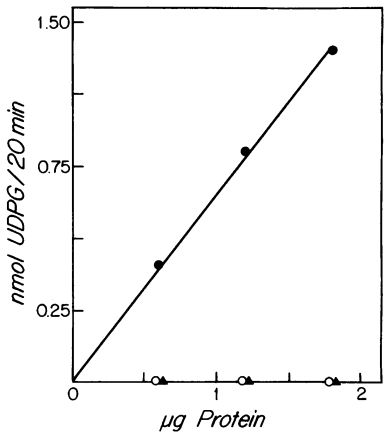

FIG. 2. Assay of UTP:α-D-glucose-1-phosphate uridylyltransferase in crude extracts of *E. coli*. Various amounts of cell extracts of QE35 (Δ*rfbA*), QE35-3 (Δ*rfbA mdo-453*), and QE35-16 (Δ*rfbA galU*::Tn*5-16*) were assayed for uridylyltransferase activity as described in Materials and Methods. Symbols: ●, QE35; ○, QE35-3; ▲, QE35-16.

derivative of pGM7 (pGM7::Tn*5-7*) with the Tn*5* element inserted in the *hns* gene (15) still corrected the MDO defect. However, a second derivative, pGM7::Tn*5-16* (15), that has an intact *hns* gene but has a Tn*5* element inserted downstream of *hns* at the position indicated in Fig. 1 failed to complement the defect in MDO synthesis in strains containing the *mdoX452*, *mdo-453*, and *galU* alleles. The apparent varied levels of complementation by the plasmids pGM7 and pGM7::Tn*5-7* reflect a characteristic of the bacterial strain instead of the plasmid. When grown under identical conditions, each bacterial strain (including strains containing *galU* after correction of the *galU* mutation by P1 transduction) seems to have its own level of MDO synthesis that varies up to sixfold among various strains (unpublished results).

TABLE 2. Complementation of MDO synthesis in *mdoX*, *mdo-453*, and *galU* strains

Strain	Allele	% of total cell cpm in MDO[a] in strain carrying:			
		No plasmid	pGM7	pGM7::Tn*5-7*	pGM7::Tn*5-16*
AB1133	None (wild type)	8.5	7.0	ND[b]	6.5
AR216	*mdoX452*	0.8	9.9	11.4	1.2
AR240	*mdo-453*	0.1	10.0	10.7	0.3
CA10	*galU95*	0.2	3.8	3.0	0.4
CA198	*galU106*	0.1	3.3	2.6	0.2
W4597	*galU51*	0.1	6.9	5.8	0.3

[a] Synthesis of MDO was determined as described previously (34).
[b] ND, not determined.

Synthesis of UDP-glucose in cell extracts of *galU* strains. When the synthesis of UDP-glucose from UTP (at a concentration of 5 mM in the assay) and [^{14}C]glucose 1-phosphate was initially measured in cell extracts of each of the *galU* mutants described above, levels of activity about 50 to 75% of the wild type were found. The question of whether all these mutants are somewhat leaky or whether the observed activity is catalyzed by another enzyme, such as TTP:α-D-glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24) therefore arose (21). To investigate this point, we constructed a chromosomal *galU* null mutation by moving the Tn*5-16* insertion from the plasmid pGM7::Tn*5-16* to the chromosome (see Materials and Methods). The *galU*::Tn*5-16* derivative of AB1133 that was obtained by P1 transduction, AR286, had a complete block in the production of MDO (data not shown). In addition, the accuracy of the homologous recombination was checked by moving the kanamycin-resistant (Mdo$^-$) phenotype by P1 transduction into a strain with the insertion element *zch-506*::Tn*10* (Table 1). The high frequency of kanamycin-resistant transductants that were now tetracycline sensitive (88%) was as expected from our previous P1 transductional mapping of the *mdoX452* and *mdo-453* mutations (34). When cell extracts of the *galU*::Tn*5-16* derivative (AR286) and the parent strain (AB1133) were compared, synthesis of UDP-glucose in the mutant strain was found to be 70% that of the wild type, indicating that the *galU* gene product accounts for only about 30% of the activity of wild-type crude extracts. The activity in the mutant extract, however, had about an 80-fold-higher K_m for UTP than for TTP, making it probable that the activity in the cell extracts from the strain with the *galU*::Tn*5-16* mutation is that of TTP:α-D-glucose-1-phosphate thymidylyltransferase. Further, as these experiments indicate, the thymidylyltransferase cannot function for the synthesis of significant amounts of UDP-glucose in vivo.

Confirmation of the presence of the interfering activity in the AB1133-derived strains was obtained when the *galU*::Tn*5-16* and *mdo-453* allele were moved into a strain, QE35, that has a deletion of the *rfb* region where the gene for the thymidylyltransferase is located (27). As seen in Fig. 2, there is essentially no enzyme activity present in the *galU*::Tn*5*

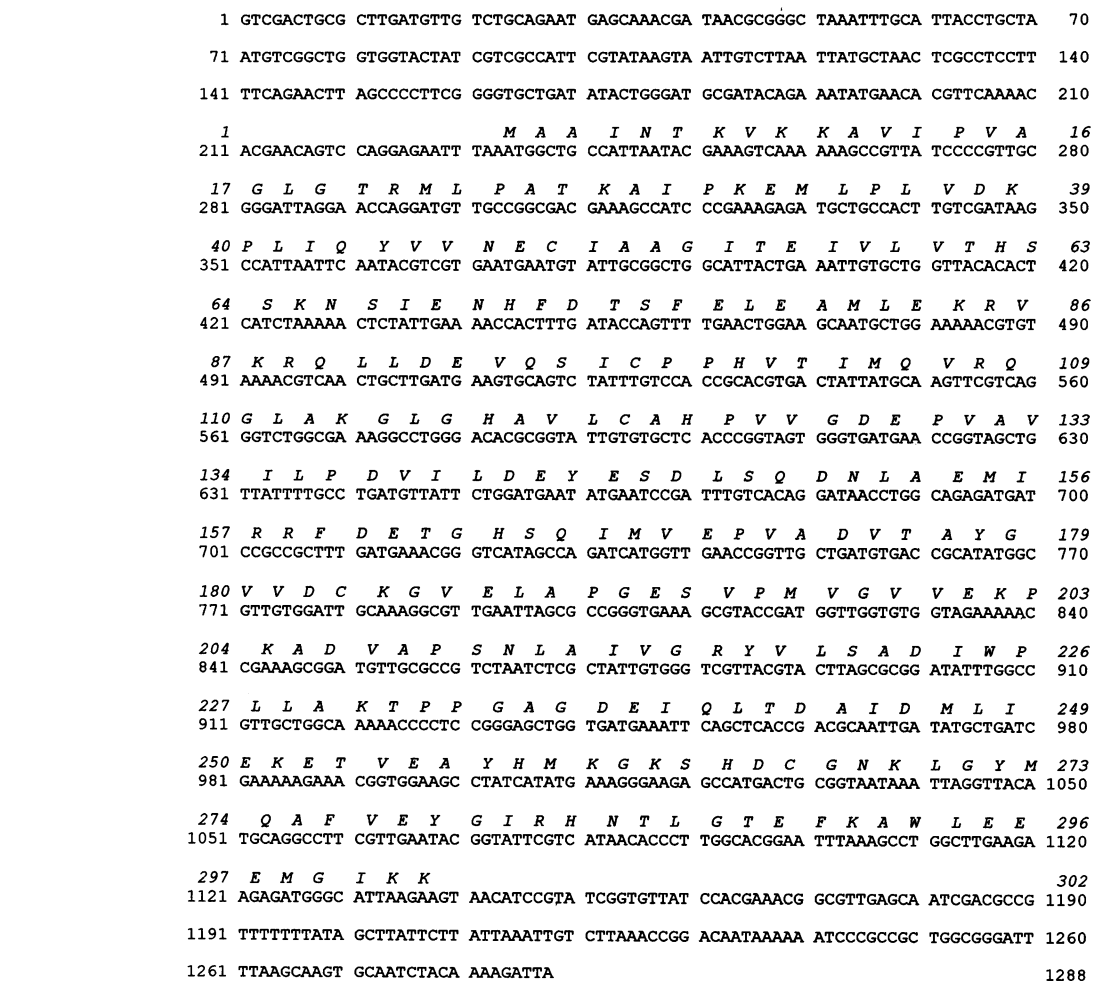

FIG. 3. Nucleotide sequence of the *galU* gene. The sequence of the 1,288-base segment of the antisense strand from the left *Sal*I site (Fig. 1) through the *galU* gene and up to and including the stop codon for the *hns* gene (its sense strand) is depicted. The derived amino acid sequence of the *galU* gene, extending from bases 234 to 1139, is shown above the nucleotide sequence in italic type in the single-letter code for amino acids.

and *mdo-453* derivatives of QE35, while high activity of the uridylyltransferase is present in extracts of QE35. Measurement of the K_m for TTP and UTP of the enzyme present in extracts of QE35 showed that this activity had about a 40-fold-higher K_m for TTP than for UTP, confirming the absence of thymidylyltransferase in this strain.

Overproduction of UTP:α-D-glucose-1-phosphate uridylyltransferase in strains carrying the *galU* gene on a multicopy plasmid. When the plasmid pGM7, a low-copy-number plasmid, was introduced into AB1133, the level of uridylyltransferase activity increased about ninefold (1.3 versus 0.14 nmol of UDPG · 20 $min^{-1} \cdot \mu g^{-1}$) over that of the strain without the plasmid (after correction for the interfering thymidylyltransferase). The *galU* gene was then subcloned as a *Bam*HI-*Pvu*II fragment (Fig. 1) into the high-copy-number plasmid pUC18 resulting in plasmid pAR200. When this plasmid was introduced into AR240, the activity of the uridylyltransferase increased over 400-fold (63.5 versus 0.14 nmol of UDPG · 20 $min^{-1} \cdot \mu g^{-1}$) over that normally found in the parental wild-type strain, AB1133.

Sequencing of the *galU* gene. The sequencing strategy is shown in Fig. 1. The nucleotide sequence (1,288 bases) of the antisense strand up to and including the stop codon for the *hns* gene (its sense strand) is shown in Fig. 3. Analysis of the DNA sequence using the DNA Strider program of Marck (12) showed only one open reading frame starting at position 234 and continuing through position 1139 with the stop codon, TAA, at positions 1140 to 1142. The orientation of this open reading frame is clockwise on the *E. coli* chromosome. The protein encoded here would have 302 amino acids, with a molecular mass of 32,921 daltons, in fair agreement with the value of 38,000 estimated by May et al. (15) to be the size of the protein disrupted by the insertion of Tn*5-16*. We find the Tn*5* insertion to be between nucleotides 598 and 599 in the codon for amino acid 122 of the derived protein.

Sequencing of *galU* mutations. The DNA sequence of the *galU* gene in the two MDO mutants and three classically utilized Gal mutants was determined after amplification of the *galU* gene by the PCR as described in Materials and Methods. The data presented in Table 3 show that two point mutations result in amino acid substitutions near the amino terminus of the protein, two mutations create aberrant stop codons, and one mutation, *mdoX452*, results in an amino acid substitution near the carboxyl terminus of the protein. This class III allele (*mdoX452*) allowed growth on galactose, although no MDO was produced by the bacteria. Presumably, this mutant enzyme

TABLE 3. DNA sequence changes of *galU* mutations[a]

Strain	Allele	DNA change	Codon	Amino acid change
RZ60-3A	*mdo-453*	280C→T[b]	16	A→V
RZ60-10H	*mdoX452*	976T→C	248	L→P
CA10	*galU95*	273C→T	14	P→S
CA198	*galU106*	498C→T	89	Q→(Ochre)
W4597	*galU51*	950T deleted	239	IQLTDAIDML . .→ISSPTQLIC (Umber)

[a] The DNA sequence of the *galU* gene in the listed strains was determined as described in Materials and Methods. The codon number and amino acid change refers to the predicted amino acid sequence as given in Fig. 3.
[b] 280C→T, the cytosine at position 280 was changed to thymine.

has sufficient uridylyltransferase activity remaining to produce the catalytic amount of UDPG needed for growth on galactose. We have not yet been able to demonstrate unambiguously residual enzyme activity in an *mdoX452* strain.

Purification of UTP:α-D-glucose-1-phosphate uridylyltransferase. The plasmid pAR200 was introduced into the strain SØ874, and the resultant strain was used as the source of material for purification. SØ874, like QE35, has a deletion of the *rfb* region (20), but SØ874 grows more rapidly than QE35. The use of an Rfb mutant ensured that the thymidylyltransferase would not interfere with the purification. In the SØ874 derivative, the presence of the plasmid, pAR200, results in a 40-fold increase in uridylyltransferase activity.

The purification scheme is outlined in Table 4 and described in Materials and Methods. Briefly, after preparing a cell extract and precipitating nucleic acids with streptomycin sulfate, the remaining soluble proteins were passed over a reactive dye column, Reactive Green 5-Agarose. About 40% of the protein binds to the column, while essentially all of the uridylyltransferase flows through. Concentration of the sample by ammonium sulfate precipitation results in a modest improvement in purity. Chromatography on a Sephacryl S-200 column resulted in a substantial improvement in purity, since the enzyme elutes as a protein with a molecular weight of 130,000. The purity of the enzyme can be seen in the SDS-polyacrylamide gel of Fig. 4, and the molecular weight of the enzyme subunit, designated by the arrow in lane 2, is about 38,000, as reported previously for the plasmid-encoded protein (15). Thus, the enzyme itself appears to be a tetramer of the *galU*-encoded protein.

Determination of the apparent K_m of this enzyme for the substrates glucose 1-phosphate and UTP gave the values 10 and 70 μM, respectively. These values are in fair agreement with those reported by Kamogawa and Kurahashi (10) in their earlier purification of this enzyme. The apparent K_m of this enzyme for TTP is 2 mM, about 30-fold higher than the K_m for UTP.

The amino-terminal amino acid sequence of the purified enzyme's subunit was determined after the protein with a molecular weight of 38,000 on the SDS-polyacrylamide gel was transferred to an Immobilon-P polyvinylidene difluoride membrane as described in Materials and Methods. The sequence obtained, AAINTKVKKA, is identical to the predicted amino acid sequence (Fig. 3) of the *galU*-encoded protein except for the initiating methionine that is missing in the enzyme subunit.

TABLE 4. Purification of UTP:α-D-glucose-1-phosphate uridylyltransferase

Purification step	Amt (mg) of protein	Activity (μmol/min)	Sp act (μmol/min/mg)	Purification (fold)	Yield (%)
Crude extract	142	30.4	0.21	1.0	100
Streptomycin sulfate	47	30.6	0.65	3.1	101
Reactive Green 5-Agarose	19	21.0	1.10	5.2	69
Ammonium sulfate	13	16.7	1.28	6.1	55
Sephacryl S-200	2.5	11.9	4.76	22.7	39

DISCUSSION

We have now provided definitive proof that the *galU* gene is the structural gene for the enzyme UTP:α-D-glucose-1-phosphate uridylyltransferase. The enzyme was purified, and the amino-terminal amino acid sequence of the enzyme's subunit was shown to be identical to the amino acid sequence of the open reading frame of the *galU* gene. The derived amino acid sequence was used to search known protein sequences by using the BLAST (1) network service of the National Center for Biotechnology Information. The derived sequence of the *galU* gene product proved to be identical to that of an unidentified open reading frame that was found by Ueguchi and Ito (31) to be immediately adjacent to the *hns* gene. The exact identity of their DNA sequence, as well as their amino acid sequence of the open reading frame, provides independent confirmation of the DNA sequence and of the position of the *galU* gene immediately adjacent and counterclockwise to the *hns* gene on the *E. coli* chromosome.

As has been reported previously (33), the derived amino acid sequences of two bacterial genes, the *Acetobacter xylinum celA* gene (5, 32) and the *Bacillus subtilis gtaB* gene (33), that are likely to encode uridylyltransferases are highly similar to each other and, also, to the unidentified open reading frame *rfb2.8* from *S. typhimurium* (8). Further, all three are highly similar to the now definitively identified *galU*-encoded protein, UTP:α-D-glucose-1-phosphate uridylyltransferase. The region

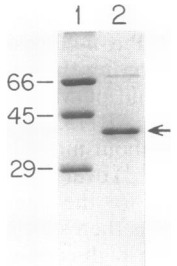

FIG. 4. SDS-polyacrylamide gel electrophoresis of purified uridylyltransferase. The molecular weight standards, bovine serum albumin (66,000), ovalbumin (45,000), and carbonic anhydrase (29,000), are shown in lane 1, and their molecular weights (in thousands) are shown to the left of the gel. The uridylyltransferase purified through the Sephacryl S-200 column was run in lane 2. The arrow to the right of the gel points to the protein with a molecular weight of 38,000 that was used for the amino-terminal amino acid sequence determination.

containing the highest concentration of identical amino acids among the four genes is the amino terminus of the protein; 27 of the first 54 residues of the *galU*-encoded protein are, in fact, identical in all four proteins. It is in this region that two of the point mutations determined here are found. One, *galU95*, changes an invariant proline to serine (residue 14 in the *galU*-encoded protein [Fig. 3]) and the other, *mdo-453*, changes an invariant alanine to valine (residue 16 in the *galU*-encoded protein). These mutations certainly suggest the importance of this region for the enzyme's catalytic activity. There is little similarity of these four proteins to the ATP:α-D-glucose-1-phosphate adenylyltransferase of *E. coli* (2). However, a motif, VEKP, near the reactive lysine 195 of the adenylyltransferase (2) is also present in the following uridylyltransferases: in the *galU*-derived sequence at residues 200 to 203, in the *celA*-derived sequence at residues 191 to 194, in the *gtaB*-derived sequence at residues 189 to 192, and in the *rfb2.8*-derived sequence at residues 196 to 199 with the conservative change of I for V.

In their earlier work, Nakae and Nikaido and colleagues reported (17–19, 21) that the product of the *galF* gene that is located in the *rfb* region of *S. typhimurium*'s chromosome caused changes in the chromatographic behavior of that bacterium's uridylyltransferase. They suggested that the *galF* gene product might be involved in some posttranslational modification or physical association with the *galU*-encoded enzyme. The *rfb2.8* gene is in the same region of the chromosome as the *galF* gene. In addition, the derived amino acid sequence of *rfb2.8* is highly similar to the amino acid sequence of the *galU* gene product; of 297 amino acid residues in the open reading frame of *rfb2.8*, 161 are identical to those of the *galU*-derived gene product and an additional 43 residues represent conservative changes. The close resemblance of the *rfb2.8* gene product to that of the *E. coli galU* gene product (now definitively shown to be a uridylyltransferase), the map location of *rfb2.8*, and the original findings of Nakae and Nikaido and colleagues strongly suggest that *rfb2.8* is the same gene as *galF* and that this gene encodes a second UTP:α-D-glucose-1-phosphate uridylyltransferase in *S. typhimurium* LT2. Our finding that the uridylyltransferase is a tetramer provides an explanation for the findings of Nakae and Nikaido and colleagues (17–19, 21). The varied and multiple peaks of activity they saw during DEAE-cellulose chromatography were likely assorted multimers of the *rfb2.8* (*galF*) and *galU* gene products, two highly similar proteins. In fact, they found only a single peak of activity when the *galF* gene was deleted (19). Earlier, Nikaido et al. (21) and Nakae (17) had indeed speculated that the multiple peaks of activity represented dimers of the *galF* gene product and the uridylyltransferase.

ACKNOWLEDGMENTS

We are grateful for the generous gift of plasmids by Erhard Bremer. We thank Jeff Himawan and Stan Tabor of the Charles Richardson laboratory for extensive help in the DNA sequencing and analysis of the sequence. We acknowledge the expert assistance of both Alex Nussbaum in the synthesis of the oligonucleotides and Melissa Saylor and John Rush of the HHMI/Harvard Medical School Biopolymers Facilities in the determination of the amino-terminal amino acid sequence.

This work was supported by grants GM19822 and GM22057 from the National Institutes of General Medical Sciences.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Baecker, P. A., C. E. Furlong, and J. Preiss.** 1983. Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* ADP-glucose synthetase as deduced from the nucleotide sequence of the *glgC* gene. J. Biol. Chem. **258:**5084–5088.
3. **Black, M. E., and D. E. Hruby.** 1991. Nucleotide sequence of the *Escherichia coli* thymidine kinase gene provides evidence for conservation of functional domains and quaternary structure. Mol. Microbiol. **5:**373–379.
4. **Boos, W., U. Ehmann, H. Forkl, W. Klein, M. Rimmele, and P. Postma.** 1990. Trehalose transport and metabolism in *Escherichia coli*. J. Bacteriol. **172:**3450–3461.
5. **Brede, G., E. Fjærvik, and S. Valla.** 1991. Nucleotide sequence and expression analysis of the *Acetobacter xylinum* uridine diphosphoglucose pyrophosphorylase gene. J. Bacteriol. **173:**7042–7045.
6. **Fukasawa, T., K. Jokura, and K. Kurahashi.** 1962. A new enzymic defect of galactose metabolism in *Escherichia coli* K-12 mutants. Biochem. Biophys. Res. Commun. **7:**121–125.
7. **Giæver, H. M., O. B. Styrvold, I. Kaasen, and A. R. Strøm.** 1988. Biochemical and genetic characterization of osmoregulatory trehalose synthesis in *Escherichia coli*. J. Bacteriol. **170:**2841–2849.
8. **Jiang, X.-M., B. Neal, F. Santiago, S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar *typhimurium* (strain LT2). Mol. Microbiol. **5:**695–713.
9. **Kalckar, H. M.** 1958. Uridinediphosphogalactose: metabolism, enzymology, and biology. Adv. Enzymol. **20:**111–134.
10. **Kamogawa, A., and K. Kurahashi.** 1965. Purification and properties of uridine diphosphate glucose pyrophosphorylase from *Escherichia coli* K12. J. Biochem. **57:**758–765.
11. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
12. **Marck, C.** 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16:**1829–1836.
13. **Markovitz, A.** 1977. Genetics and regulation of bacterial capsular polysaccharide biosynthesis and radiation sensitivity, p. 415–462. *In* I. Sutherland (ed.), Surface carbohydrates of the prokaryotic cell. Academic Press, New York.
14. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262:**10035–10038.
15. **May, G., P. Dersch, M. Haardt, A. Middendorf, and E. Bremer.** 1990. The *osmZ* (*bglY*) gene encodes the DNA-binding protein H-NS (Hla), a component of the *Escherichia coli* K12 nucleoid. Mol. Gen. Genet. **224:**81–90.
16. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 201–205. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Nakae, T.** 1971. Multiple molecular forms of uridine diphosphate glucose pyrophosphorylase from *Salmonella typhimurium*. III. Interconversion between various forms. J. Biol. Chem. **246:**4404–4411.
18. **Nakae, T., and H. Nikaido.** 1971. Multiple molecular forms of uridine diphosphate glucose pyrophosphorylase from *Salmonella typhimurium*. I. Catalytic properties of various forms. J. Biol. Chem. **246:**4386–4396.
19. **Nakae, T., and H. Nikaido.** 1971. Multiple molecular forms of uridine diphosphate glucose pyrophosphorylase from *Salmonella typhimurium*. II. Genetic determination of multiple forms. J. Biol. Chem. **246:**4397–4403.
20. **Neuhard, J., and E. Thomassen.** 1976. Altered deoxyribonucleotide pools in P2 eductants of *Escherichia coli* K-12 due to deletion of the *dcd* gene. J. Bacteriol. **126:**999–1001.
21. **Nikaido, H., M. Levinthal, K. Nikaido, and K. Nakane.** 1967. Extended deletions in the histidine-rough-B region of the *Salmonella* chromosome. Proc. Natl. Acad. Sci. USA **57:**1825–1832.
22. **Pradel, E., and P. L. Boquet.** 1989. Mapping of the *Escherichia coli* acid glucose-1-phosphatase gene *agp* and analysis of its expression in vivo by use of an *agp-phoA* protein fusion. J. Bacteriol. **171:**3511–3517.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Schulman, H., and E. P. Kennedy.** 1977. Identification of UDP-glucose as an intermediate in the biosynthesis of the membrane-

derived oligosaccharides of *Escherichia coli*. J. Biol. Chem. **252:** 6299–6303.

25. **Shapiro, J. A.** 1966. Chromosomal location of the gene determining uridine diphosphoglucose formation in *Escherichia coli* K-12. J. Bacteriol. **92:**518–520.
26. **Sundararajan, T. A., A. M. C. Rapin, and H. M. Kalckar.** 1962. Biochemical observations on *E. coli* mutants defective in uridine diphosphoglucose. Proc. Natl. Acad. Sci. USA **48:**2187–2193.
27. **Sunshine, M. G., and B. Kelly.** 1971. Extent of host deletions associated with bacteriophage P2-mediated eduction. J. Bacteriol. **108:**695–704.
28. **Tabor, S.** Unpublished results.
29. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:**4767–4771.
30. **Tempst, P., and L. Riviere.** 1989. Examination of automated polypeptide sequencing using standard phenyl isothiocyanate reagent and subpicomole high-performance liquid chromatographic analysis. Anal. Biochem. **183:**290–300.
31. **Ueguchi, C., and K. Ito.** 1992. Multicopy suppression: an approach to understanding intracellular functioning of the protein export system. J. Bacteriol. **174:**1454–1461.
32. **Valla, S., D. H. Coucheron, E. Fjærvik, J. Kjosbakken, H. Weinhouse, P. Ross, D. Amikam, and M. Benziman.** 1989. Cloning of a gene involved in cellulose biosynthesis in *Acetobacter xylinum*: complementation of cellulose-negative mutants by the UDPG pyrophosphorylase structural gene. Mol. Gen. Genet. **217:**26–30.
33. **Varón, D., S. A. Boylan, K. Okamoto, and C. W. Price.** 1993. *Bacillus subtilis gtaB* encodes UDP-glucose pyrophosphorylase and is controlled by stationary-phase transcription factor σ^B. J. Bacteriol. **175:**3964–3971.
34. **Weissborn, A. C., M. K. Rumley, and E. P. Kennedy.** 1992. Isolation and characterization of *Escherichia coli* mutants blocked in production of membrane-derived oligosaccharides. J. Bacteriol. **174:**4856–4859.
35. **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.