J. Membrane Biol. 164, 187–195 (1998)

The Journal of
Membrane
Biology

Identification of Two Essential Arginine Residues in UhpT, the Sugar Phosphate Antiporter of *Escherichia coli*

M.-C. Fann[1], A.H. Davies[1], A. Varadhachary[1], T. Kuroda[1,*], C. Sevier[1], T. Tsuchiya[2], P.C. Maloney[1]

[1]Department of Physiology, Johns Hopkins Medical School, Baltimore, MD, 21205 USA

[2]Department of Microbiology, Faculty of Pharmaceutical Sciences, Okayama University, Okayama, 700 Japan

Received: 29 January 1998/Revised: 13 April 1998

Abstract. Three lines of evidence indicate that arginine-46 (R46) and arginine-275 (R275) are essential to the function of UhpT, the Pi-linked antiport protein of *Escherichia coli.* A role for arginine was initially suggested by the sensitivity of UhpT to inhibition by 2,3-butanedione, an arginine-directed probe. Since the presence of substrate protected against this inhibition, this work further suggested that arginine(s) may lie at or near the UhpT active site. In other work, each UhpT arginine was examined individually by using site-directed mutagenesis to generate a cysteine or a lysine derivative. With two exceptions (R46, R275), all arginines could be replaced by either cysteine (10 of 14 residues) or lysine (12 of 14) without loss of function, implicating R46 and R275 as essential to UhpT function. This idea was strengthened by examining a multiple alignment of the eleven known UhpT-related proteins (⩾30% identity). That alignment showed R46 and R275 were two of the only three arginines strongly conserved in this group of proteins. Considered together, these different approaches lead us to conclude that UhpT and its relatives have only two arginine residues (R46, R275) whose presence is essential to function. Prior biochemical work had placed R275 at the external entrance to the translocation pathway, and a symmetry argument emerging from the multiple alignment suggests a similar position for R46. Accordingly, by virtue of their locations at the entrance to this pathway, we speculate that R46 and R275 function in establishing substrate specificity.

Key words: Phosphate-binding — Reconstitution — Membrane transporter protein — Butanedione

Introduction

In *Escherichia coli*, transport of hexose phosphates is mediated by the membrane protein, UhpT (for *u*ptake of *h*exose *p*hosphate), a chemiosmotic transporter enabling the accumulation of external sugar phosphate by an exchange with internal anions such as inorganic phosphate (Pi) [32, 33, 46]. Biochemical and genetic studies [23, 30, 48] indicate that UhpT has 12 transmembrane segments, presumed to be α-helices [51], and that it resembles in general structure a large number of other porters in prokaryotes and eukaryotes [21, 32, 34]. UhpT and its relatives in *E. coli* [9, 15, 23], *Salmonella typhimurium* [16] *Bacillus subtilis* [38] and *Haemophilus influenzae* [11] comprise a distinct cluster of bacterial anion exchange proteins within the Major Facilitator Superfamily (MFS) [34].

UhpT is the best-studied of the bacterial anion exchange transporters [32] and presently serves as an attractive model for identification of residues essential to transport by Pi-linked exchange. An important recent finding has been that TM7, the seventh transmembrane helix, lines the substrate translocation pathway [36, 50, 51], and for this reason one questions whether residues in TM7 participate in substrate recognition. In particular, we note that arginine-275 (R275) is located at the end of TM7, just inside the external mouth of the translocation pathway [36, 48], and that crystallographic study of soluble proteins shows arginine to be often involved in binding inorganic and organic phosphates [39; *see later*]. However, mutation of R275 yields nonfunctional product [48; *see below*], so the role of this residue could not be

* *Present address:* Department of Microbiology, Faculty of Pharmaceutical Sciences, Okayama University, Okayama, 700 Japan

Correspondence to: P.C. Maloney

addressed directly. For this reason, the present work has explored other ways to probe the relevance of arginine residues in UhpT function. Experiments using chemical modification now confirm the presence of critical arginine(s) in UhpT, and the results of mutagenesis show that only R46 and R275 among UhpT arginines are required for expression and function. These findings, together with early biochemical work, the proximity of R275 to the transport pathway and a presumption of symmetry in UhpT structure, lead to speculation that R46 and R275 are involved directly in interactions with substrate anions.

Materials and Methods

Expression Plasmids and Bacterial Strains

Escherichia coli strain RK5000, deleted for the chromosomal *uhp* region, was host for p261, an ampicillin-resistant phagemid carrying a 2 kb NsiI-BamHI fragment encoding UhpT under control of the lac promoter [15]. (RK5000 and p261 were from R.J. Kadner, University of Virginia.) Cells harboring p261 or its derivatives were grown overnight in LB medium (plus antibiotic), diluted 100-fold into fresh LB medium (plus antibiotic) and grown at 33°C until mid-exponential phase (OD_{650} of 0.6–1.0). Since RK5000 lacks the LacI gene, this leads to a 50- to 100-fold constitutive overexpression of phagemid-encoded UhpT.

Mutagenesis

Site-directed mutagenesis was performed by the method of Kunkel [26], using uracil-substituted ssDNA selected from phagemids grown in the *dut ung* strain, CJ236 [50]. The mutagenic oligonucleotides (Core Facility, Johns Hopkins Medical School) contained diagnostic restriction sites for preliminary identification of putative mutants; in all instances, mutagenesis was verified by DNA sequencing using the dideoxy chain termination method [43].

Reconstitution

As described earlier [1], membranes obtained by osmotic lysis of lysozyme-sensitized cells [47] were incubated for 20–30 min on ice with 40 mM potassium phosphate, 20% glycerol, 0.2% *E. coli* phospholipid, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride and 1.2% octylglucoside. Insoluble material was removed by centrifugation in an Eppendorf refrigerated microcentrifuge (15,000 × *g* for 15 min) prior to storage of the crude detergent extract at −70°C.

For reconstitution, a portion of the crude detergent extract was mixed, in a final volume of 1 ml, with bath-sonicated liposomes (5 mg *E. coli* phospholipid) and sufficient detergent and buffer to maintain the concentrations noted above. After 20 min on ice, proteoliposomes were formed at room temperature by 16-fold dilution into 100 mM potassium phosphate (pH 7), followed by a 20-min incubation, also at room temperature [1]. Proteoliposomes were pelleted in the cold (145,500 × *g* for 60 min) and resuspended at 2–6 mg protein per ml in Buffer A (75 mM potassium sulfate, 20 mM MOPS/K, 2.5 mM magnesium sulfate).

Modification of arginine was achieved by incubating proteoliposomes (0.2 to 0.4 mg protein/ml) with 50 mM 2,3-butanedione at pH 8, in the presence or absence of protective substrate, at 25°C and in a solution containing (in mM): 60 Tris/sulfate, 50 potassium borate, 10 potassium sulfate, 1.5 magnesium sulfate. The reaction was terminated at the specified times by 10-fold dilution of aliquots into a chilled stop solution (pH 7) lacking butanedione (in mM): 100 MOPS/K, 50 potassium borate, 15 potassium sulfate, 2.5 magnesium sulfate; quenched samples were kept on ice until assay.

Assays of Transport

UhpT function was monitored by transport of [^{14}C]G6P into either intact cells or Pi-loaded proteoliposomes. For the former assay, cells were washed and resuspended in assay Buffer B (in mM): 250 potassium chloride, 20 MOPS/K (pH 7), 1 magnesium sulfate) at a final density of 0.3 mg protein/ml. On equilibration at 23°C, 50 μM [^{14}C]G6P was added, and aliquots were removed after 1 min for filtration and washing using presoaked Millipore HAWP filters (0.45 μm pore size). Assays were in triplicate, and the reported values reflect the averages of three or more independent experiments. To monitor transport after reconstitution, proteoliposomes (*above*) were spotted in duplicate onto 0.22 μm GSTF Millipore filters and washed with 5-ml Buffer A. The vacuum was discontinued, and transport was initiated by layering 0.25 ml Buffer A containing 50 μM [^{14}C]G6P over the immobilized proteoliposomes; after 3 min the reaction was stopped by vacuum filtration, followed by two rinses with 5-ml Buffer A [1].

Electrophoresis and Western Blotting

SDS-PAGE using 12.5% acrylamide [27] was performed without preheating of samples. After transfer to nitrocellulose [36, 41], protein was visualized using a rabbit polyclonal antiserum (1/2,500) reactive to a peptide derived from the UhpT C-terminus [36]. Western blots were developed with chemiluminescence (Amersham); quantitation of scanned gels was performed with NIH Image v1.59.

Chemicals

Calbiochem-Novabiochem was the source of n-octyl-β-D-glucopyranoside; butanedione was from Aldrich Chemical. *E. coli* phospholipid was purified from crude material provided by Avanti Polar Lipid [1]. D-1-[^{14}C]G6P was from DuPont-New England Nuclear.

Sequence Alignments

Amino acid sequences are specified in the single letter code, and when discussed in the text, they are written in PROSITE pattern format [4]. Pairwise and multiple alignments were performed with ClustalW1.7 (default settings) from the Baylor College of Medicine Search Launcher (http://kiwi.imgen.bcm.tmc.edu:8088/search-launcher/launcher.html) for eleven sequences drawn from public databases. Most sequences were found at full-length (or nearly so), but in three cases only partial information was available: 32 residues for GlpT of *H. influenzae;* 72 residues for GlpT of *H. actinomycetemcomitans;* 215 residues for UhpT of *L. lactococcus.* In these instances, upstream and downstream DNA sequences were also examined to search for errors that may have led to inadvertent mistranslation. We found evidence of this for GlpT of *H. influenzae* [11], where an artefactual truncation appeared likely, due a region of high GC content; the full-length 480 residue sequence was recovered by taking this into account. A similar truncation may have occurred for PgtP of *S. typhimurium*, whose reported amino acid sequence terminates after eleven transmembrane

helices [16], rather than the twelve found in other examples; for purposes of multiple alignment we extended the PgtP sequence to include TM12 (*see later*). As well, because the UhpT and UhpC proteins of *E. coli* so closely resemble their homologs in *S. typhimurium* (*ca.* 94% identity) [23], we excluded the latter to avoid bias in multiple alignment.

To search for internal duplications in UhpT-related proteins, we used Block Maker (http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html) to analyze a set containing the N- and C-terminal halves of the sequences described above; the partial sequences from *H. actinomycetemcomitans* and *L. lactis* (*above*) were assigned to the N- or C-terminal halves according to their position in the multiple alignment of complete sequences (*see later*). Block Maker utilizes two separate programs (MOTIF and GIBBS), each based on different assumptions and principles, to find highly conserved ungapped segments in a sequence set [22]. Strong evidence for the presence of a conserved block is provided when these two programs identify the same segment [22].

A signature sequence for UhpT-related proteins was identified by repetitive sampling of the Swiss Protein database using strings of conserved residues as recommended by PatternSearch (http://www.genome.ad.jp/SIT/MOTIF.html).

Abbreviations

The abbreviations and trivial names used are: Pi, inorganic phosphate; OG or octylglucoside, n-octyl-β-D-glucopyranoside; G6P, glucose 6-phosphate.

Results

An Arginine-specific Reagent Inhibits UhpT

Our first experiments used chemical modification as a generic tool to ask whether UhpT arginine(s) might play some functional role. Such selective modification of arginine is usually achieved by treatment with 2,3-butanedione, 1,2-cyclohexanedione or phenylglyoxal, each of which has adjacent carbonyl groups spaced appropriately for interaction with the guanido group of arginine [7, 8, 12, 35, 40]. From among these potential probes, we selected 2,3-butanedione, whose interaction with arginine is stabilized in the presence of borate (*see* Fig. 1, inset).

Incubation of Pi-loaded proteoliposomes with 2,3-butanedione, in the presence of borate, led to prompt decay of UhpT function as judged by tests of glucose 6-phosphate transport (Fig. 1, closed symbols); this inactivation required the presence of borate (*not shown*), which itself had minimal effect (Fig. 1, open symbols). This behavior suggested inhibition of UhpT occurred following derivatization of arginine, and the experience with soluble proteins indicates that such functional deficit might involve modification of one or a few residues [39]. We also examined this response in the presence of substrate, since protection by substrate could indicate that the susceptible residue(s) is(are) at or near the sugar

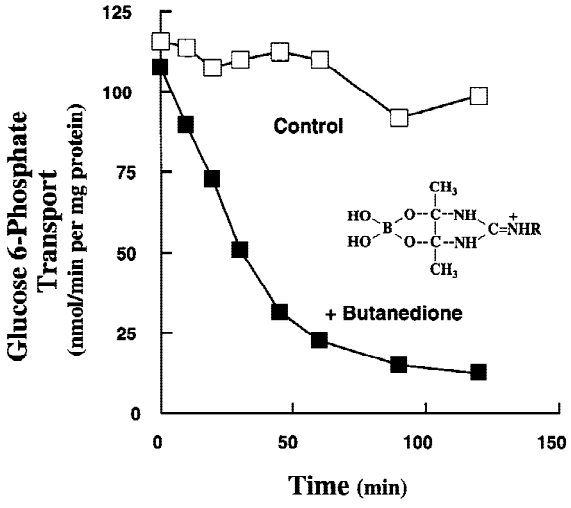

Fig. 1. 2,3-Butanedione inhibition of UhpT. Phosphate-loaded proteoliposomes were placed for the indicated times in the presence (closed symbols) or absence (open symbols) of 50-mM butanedione, in borate-containing buffer at pH 8 (*see* Materials and Methods). The reaction was quenched by dilution of aliquots into a chilled stop buffer at pH 7 lacking butanedione. Subsequently, duplicate samples of each aliquot were placed on Millipore filters for assays of glucose 6-phosphate transport. *Inset:* Structure of the borate-stabilized complex of arginine and 2,3-butanedione.

phosphate binding site [e.g., 13, 24, 39]. Indeed, as noted in Table 1, the presence of saturating levels of external 2-deoxyglucose 6-phosphate (2DG6P) largely prevented inhibition by 2,3-butanedione. By contrast, the same concentration of galactose 6-phosphate, a poor substrate for UhpT [33], had no significant effect on the action of this arginine-directed reagent. This last observation rules out nonspecific actions of 2,3-butanedione on proteoliposomes and instead points to a specific action on UhpT itself.

Using these same experimental conditions (i.e., Table 1), we also analyzed proteoliposomes containing GlpT, the Pi-linked glycerol 3-phosphate antiporter of *E. coli* [2]. In this instance, 2,3-butanedione inhibition was more pronounced (19 ± 11% residual activity with butanedione alone), and substrate protection was correspondingly diminished (57 ± 8% residual activity in the presence of both butanedione and 20 mM glycerol 3-phosphate). Inhibition by 2,3-butanedione was also observed in two trials using proteoliposomes containing PgtP, the phosphoenolpyruvate transporter of *Salmonella typhimurium* [16, 33], but substrate protection could not be documented (*not shown*). Together with the observations described above (Fig. 1, Table 1), these findings show that chemical modification of arginine residue(s) leads to inactivation of Pi-linked antiporters, and suggests that in at least two cases (UhpT and GlpT), the susceptible arginine(s) is (are) shielded from attack as a result of conformational changes accompanying substrate binding and transport.

Table 1. Butanedione inhibition of UhpT[1]

Additions	Activity Remaining[2]
None	(1.00)
Butanedione	0.49 ± 0.08
Butanedione *plus* 2-deoxyglucose 6-phosphate	0.89 ± 0.07
Butanedione *plus* galactose 6-phosphate	0.41

[1] Pi-loaded proteoliposomes were placed at pH 8 with 50 mM 2,3-butanedione in the presence or absence of 20 mM protective substrate, as indicated. After incubation for 45 min, proteoliposomes were diluted 10-fold into a chilled stop solution at pH 7 lacking butanedione (*see* Materials and Methods), and residual UhpT function was monitored by transport of [^{14}C]G6P.
[2] Mean values from two or more independent experiments; if appropriate, the mean is given ± SE.

Mutagenesis Identifies Critical Arginines in UhpT

The findings from chemical modification suggested that the active site of UhpT has one or a few nearby arginines, but no particular residue(s) could be identified in this way. To provide specificity, we therefore examined UhpT with methods of higher resolution, first noting the sensitivity of function to mutagenesis of arginines and then surveying the degree to which UhpT arginines are conserved in related proteins; as summarized below, both avenues of study gave similar answers.

In the first of these alternative approaches, site-directed mutagenesis was used to replace arginines, individually, with either cysteine or lysine. In this way, each target position was assigned to either of three categories. On the one hand, retention of normal activity by an R → C mutant was taken as evidence that the presence of arginine has no bearing on protein function. If, on the other hand, an R → C mutation disrupted function but the R → K substitution did not, it was presumed that arginine satisfies a general requirement for positive charge. Finally, if both R → C and R → K derivatives were defective, we concluded that arginine itself is specifically required at the particular position. (There was no instance in which an R → C variant retained activity while the R → K mutant did not.)

This approach showed that most of the fourteen UhpT arginines are not essential to function (Table 2). Thus, ⩾60% wild-type specific activity was retained for ten examples in which cysteine replaced arginine; in the four remaining cases, residual specific activities ranging from 1% to 12% indicated some significant role for the wild-type residue. When the lysine replacements were analyzed (Table 2), derivatives of the ten ''nonessential'' arginines continued to show wild-type function, as might be expected. In addition, two cases which had proven negative in the earlier test (the R20C and R325C variants) were now recovered with wild-type or near wild-type activity (45 and 80% parental activity, respectively).

Table 2. UhpT function of arginine mutants

Target Residue[1]	Relative Activity[2]	
	Arginine → Cysteine	Arginine → Lysine
Arginine-9	0.99 ± 0.07	0.58 ± 0.03
Arginine-20	→ 0.12 ± 0.01	0.45 ± 0.03
Arginine-21	1.26 ± 0.02	1.23 ± 0.05
Arginine-46	→ 0.01 ± 0.01	→ 0.01 ± 0.01
Arginine-154	0.84 ± 0.07	1.21 ± 0.01
Arginine-155	0.75 ± 0.03	1.08 ± 0.04
Arginine-157	0.72 ± 0.04	0.58 ± 0.02
Arginine-211	0.83 ± 0.05	1.04 ± 0.03
Arginine-275	→ 0.01 ± 0.02	→ 0.01 ± 0.01
Arginine-325	→ 0.02 ± 0.01	0.79 ± 0.03
Arginine-326	0.60 ± 0.07	0.46 ± 0.03
Arginine-449	1.04 ± 0.09	1.11 ± 0.01
Arginine-452	0.83 ± 0.07	0.90 ± 0.03
Arginine-453	0.88 ± 0.01	1.13 ± 0.03
Control, wild type[3]	1.03 ± 0.07	1.15 ± 0.06
Control, no insert[3]	0.01 ± 0.001	

[1] Residues in bold face indicate those identified as highly conserved by multiple alignment of UhpT and its relatives (*see* Fig. 3).
[2] Mean values ± SE relative to the parental wild type protein assayed in parallel; from 3–5 independent experiments. Arrows show instances of low specific activity.
[3] These data include assays of wild type-plasmids recovered during mutagenesis and assays of RK5000 carrying the expression plasmid without a UhpT insert.

Thus, among the fourteen UhpT arginines, the approach based on mutagenesis scored only two residues (R46, R275) as essential to overall function.

It seemed likely that the UhpT variants having parental levels of transport would also show parental levels of expression, and this presumption was verified in a Western blot analysis (*not shown*) of the twelve lysine derivatives with normal or near normal function (cysteine replacements were not tested). But because of their low levels of function, we directly assessed protein expression for both the cysteine and lysine substitutions for R46 and R275. Such work showed that the R46C and R275C substitutions gave normal yields of UhpT (90%–110%), while the R46K and R275K derivatives were present in considerably reduced amounts (17 and 4%, respectively) (Fig. 2). This showed that cysteine substitution at these two positions yields nonfunctional product, a conclusion also reached for the lysine derivatives, despite their low level of expression, since the functional assay was of sufficient sensitivity. (Reduced expression of the lysine derivatives may in part reflect an increased charge density, since the positive charge of arginine is distributed throughout the terminal guanido group.)

Conserved Arginines in UhpT

Mutational studies identified two positions (nos. 46 and 275) at which arginine itself appears necessary for UhpT

expression and/or activity, and two other positions (nos. 20 and 325) at which positive charge is required. To evaluate the broader significance of such observations, we correlated these findings with the distribution of arginine residues in proteins related to UhpT. A search of sequences available as of August, 1997, revealed eleven examples with significant homology with UhpT (≥30% identity) (Fig. 3). These examples, from six different eubacterial genera (Bacillus, Escherichia, Haemophilis, Lactobacillus, Salmonella, Shigella), include presumptive antiporters for hexose 6-phosphate (termed UhpT), glycerol 3-phosphate (GlpT) or phosphoglycerates (PgtP), as well as receptors for glucose 6-phosphate (termed UhpC). By accepting a less restrictive definition of homology (≥20% identity), more distantly related sequences were also identified, in both prokaryotes and eukaryotes, but we chose the stricter criterion so as to establish a set of sequences that could be assigned a unique signature (*below*).

Multiple alignment of UhpT-related sequences showed that 49 residues in UhpT (excluding the initial methionine)—about 10% of the total—are conserved in 7 or more of the nine sequences considered. These highly conserved residues are distributed in a nonuniform fashion (Fig. 3), showing a clear tendency to cluster, and, in the N-terminal half of the protein, to lie within or near the borders of the transmembrane segments identified by reporter gene fusions in UhpT and GlpT [9, 23, 30] (the boundaries of TM11 and TM12 are not well-delineated by such work).

Among the ten arginines classified as nonessential by mutagenesis (Table 2), only R157 is strongly conserved (Fig. 3), indicating that it may be of evolutionary rather than mechanistic significance. Such a conclusion is consistent with the fact that R157 lies within a signature sequence (R-G-x(5)-W-N-x(2)-H-N-G) that uniquely defines this family of transporters and receptors. R20, which had been classified as critical because of its basic nature, has no counterpart among relatives of UhpT, but R325, the other required basic residue, does show partial conservation (4 of 7 aligned residues at this position are arginine or lysine). This is not unexpected, since the sequence between G316 and R325 reflects a duplicated motif widely distributed within the MFS (*see below*). Mutagenesis had also identified two arginines (R46, R275) as absolutely essential to UhpT. Emphasizing the broader relevance of this conclusion, these two residues are present in all sequences examined (Fig. 3). Conservation of this sort supports the idea that R46 and R275 are specifically required for the proper assembly and/or catalytic function of UhpT.

Multiple alignment of UhpT-related proteins was also the basis of a useful additional finding. In a search for internal duplications specific to UhpT-related proteins, we applied the MOTIF and GIBBS algorithms [22]

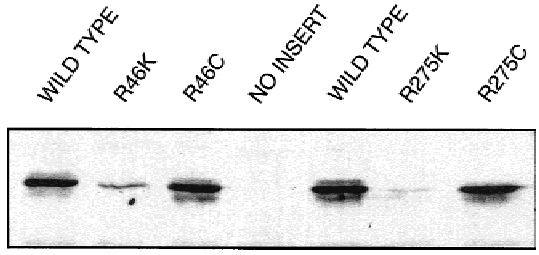

Fig. 2. Expression of R46 and R275 substitution mutants. Washed cells harboring plasmids encoding the indicated proteins were processed for SDS-PAGE (10 μg total cell protein per lane). After transfer to nitrocellulose, Western blots were developed using chemiluminescence; scanned images were quantitated using NIH Image v1.59.

to a sequence set comprised of the N- and C-terminal halves of the sequences noted above (Fig. 3). Four candidate blocks were identified, three by MOTIF and two from GIBBS, but there was overlap only in one case. Thus, a 24-residue consensus emerged from the MOTIF search (T-L-F-E-I-A-Y-G-V-G-K-L-L-C-G-W-L-S-D-R-A-N-P-R), while a comparable 21-residue block came from the GIBBS routine (F-E-I-G-Y-G-V-G-K-F-I-M-G-W-L-S-D-R-A-N-P). The near identity of these segments is considered highly significant, since the two programs are based on different assumptions and operating principles [22]. The idea that these segments reflect an authentic duplication is further strengthened by the finding that at their distal ends, one finds the motif identified by Griffiths et al. [18] as a duplication within the MFS (G-x-(3)-D-[RK]-x-G-[x]-R-[RK]). We conclude that UhpT and its relatives contain a 22-residue internal duplication whose consensus takes the form: F-E-I-[AG]-Y-G-V-G-K-x(3)-G-W-L-S-D-R-A-N-P-R. The position of this duplication is indicated in Fig. 3 as the sequence blocks surrounded by a dotted perimeter.

By reducing the level of homology required for inclusion (to ≥20% identity) the alignment of Fig. 3 was expanded to include a glucarate transporter in *Bacillus subtilis* (GLUC_BACSU), a glycerol 3-phosphate transporter in *Arabidopsis thaliana* (GLPT_ARATH) and in *Caenorhabditis elegans*, several proteins related to the sugar phosphate receptor, UhpC (UHPC_T11G6.3, UHPC_T10C6.J, UHPC_T11G6.4). This enlargement (*not shown*) reinforces the significance of the mutagenesis of UhpT (Table 2), since R46 is present in all instances (14/14), and R275 in most (8/12) (in three other cases, this positive charge may be represented by the lysine at position *n*-3). Moreover, R157, which seems of interest in (at least) a taxonomic sense, continues to be present at high frequency (13/14). There is, nevertheless, a clear heterogeneity within this larger set—for example, it was not possible to identify a signature sequence in this case; nor did the methods used to identify an internal duplication yield a consistent solution. Despite such divergence, the continued significance of R46,

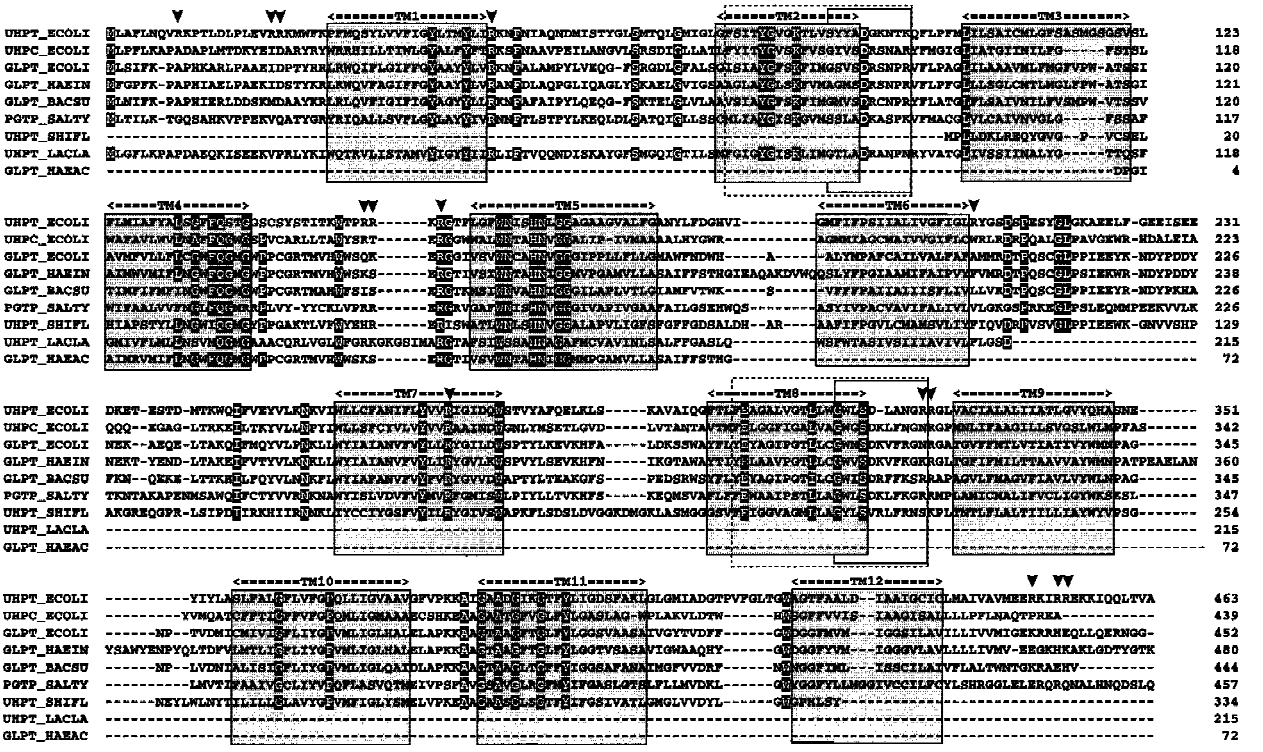

Fig. 3. Multiple alignment of UhpT-related sequences. Using sequences of ⩾30% identity with UhpT, multiple alignment was performed using ClustalW1.7, without endgap penalty (*see* Materials and Methods). UhpT arginine residues are indicated by inverted arrowheads. The shaded boxes give transemembrane helices as identified by hydropathy analysis and by PhoA and LacZ gene fusions in UhpT and GlpT [9, 23, 50], modified as follows: the extracellular boundary of TM8 has been shifted distally (by four residues) relative to earlier models [50] to allow for access of trypsin to lysine-294 [36]; the extracellular boundary of TM12 has been shifted distally to allow for the apparent insertion of six residues in the external loop connecting UhpT TM11 and TM12. The dotted perimeter shows the internal duplication identified by the MOTIF and GIBBS algorithms [*cf.* 22]; the consensus sequence for this duplication is given in the text. The smaller, solid perimeter gives the location of the internal duplication identified by Griffiths et al. [18] in members of the MFS. Italicized residues for PGTP_SALTY indicate residues encoded by sequences 3′ of the reported stop codon (*see* Materials and Methods).

R157 and R275 highlights their likely relevance in this lineage, and this finding prompts one final comment. The genomic sequence of *Mycobacterium genitalium* has a putative sugar phosphate transporter broadly related to UhpT (19% identity) [14]. Yet this protein, UHPT_MYCOG, has neither the two essential arginines (R46, R275), basic residues where positive charge may be required (R20, R325), nor the arginine found to be an index of evolutionary relatedness (R157). For these reasons, it seems likely that the *M. genitalium* protein functions other than as a sugar phosphate transporter or receptor.

Discussion

Identification of Essential Arginines

This work had its origin in the crystallographic observation that enzymes and receptors acting on phosphate or organic phosphates use the amino group(s) of arginine (or, in one case, lysine) as a ligand for phosphate oxygen(s). Examples in which arginine is at the active center include the periplasmic phosphate-binding protein [28] and alkaline phosphatase [24] of *E. coli*, along with various enzymes of intermediary metabolism [5, 10, 19, 24, 29, 37, 44, 45]. In most of these instances, a role for arginine had been inferred, prior to crystallography, from the effects of agents such as 2,3-butanedione [reviewed by 39]; and in some of these cases, it is now known that the arginine modified by such probes is the one located in the substrate binding pocket [10, 24, 25, 39].

With these precedents in mind, we theorized that arginine or lysine might play a similar role in UhpT, and in testing the idea we focused specifically on arginine, the more frequently found as a phosphate ligand [39]. Using 2,3-butanedione as a probe, we asked first for chemical evidence of critical arginines in UhpT. Based on a positive response in these experiments (Fig. 1, Table 1), we then used site-directed mutagenesis to examine each UhpT arginine directly. In this second phase, the approach was based on results from the survey of Bordo

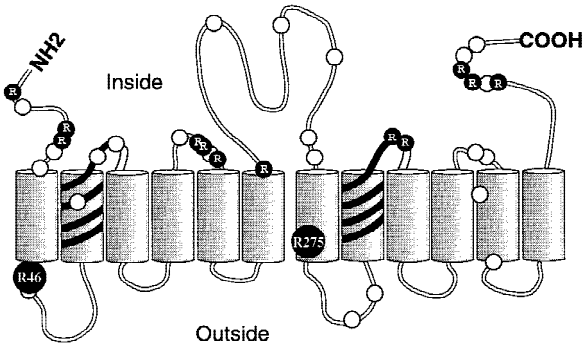

Fig. 4. UhpT topology. Topology has been slightly modified from that given earlier [50], as noted in the legend to Fig. 3. Solid circles show arginine residues; enlarged solid circles indicate the essential arginines, R46 and R275. Open circles give the locations of UhpT lysines. The internal duplication beginning in TM2 and in TM8 is also indicated.

& Argos [6], who examined homologous proteins of known structure and found that arginine is most often replaced by lysine, but never by cysteine. For this reason, when function was retained in an R → C variant, we presumed no specific need for arginine, a response characteristic of most positions occupied by arginine in UhpT. Correspondingly, when function was lost in an R → K derivative, we presumed arginine was specifically required, a conclusion reached in only two cases (Table 2). To strengthen the experimental work, we also assessed the conservation of arginines in proteins related to UhpT (Fig. 3). Together, these independent approaches document that two arginines (R46 and R275) play critical roles in UhpT function. While no specific function can be as yet assigned to these residues, sufficient information is presently available to allow reasonable speculation on this point.

Elements of Symmetry in UhpT

Theoretical arguments [*see* 17, 31, 32] and experimental results [*see* 21, 42] suggest that a dimer of identical subunits, each with six transmembrane helices, served as the ancestral form of transporters such as UhpT, and that contemporary versions arose in a gene duplication and fusion event generating a covalent heterodimer. Accordingly, one might expect present day examples to exhibit elements of structural or sequence homology between TM1 and TM7, TM2 and TM8, TM3 and TM9, and so on. In fact, this idea is compatible with several observations made for UhpT, the most significant of which is an internal duplication (consensus: F-E-I-[AG]-Y-G-V-G-K-x(3)-G-W-L-S-D-R-A-N-P-R) encompassing much of TM2 and TM8 and their cytoplasmic extensions (Fig. 3, dotted perimeters; *see also* Fig. 4). Because this sequence also includes the motif noted by Griffiths et al. [18] as duplicated in the MFS, this region has been the target of site-directed mutagenesis in both the lactose [17], and tetracycline [49] transporters. Neither that work nor the present study links this region with any specific function, but by incorporating this motif as part of a larger duplication in UhpT-related proteins, we strengthen the idea of an evolutionary link between the N- and C-terminal halves of such transporters.

The placement of an internal duplication in TM2 and TM8 (*above*) is consistent with the idea that the N- and C-terminal halves of UhpT share (at least) an evolutionary history. So, too, is the distribution of conserved and nonconserved residues. Thus, in UhpT-related proteins, helices TM1-2, TM4-5, TM7-8, and TM10-11 are enriched for conserved and/or hydrophilic residues, while TM3, TM6, TM9, TM12 have either no conserved residues or but a single hydrophobic example (Fig. 3). Hydrophobicity calculations for the MFS as a whole give a similar stratification of helices, and for this reason Goswitz & Brooker [17] were led to develop a general model in which these more hydrophobic segments (TM3, TM6, TM9, TM12) form a scaffold supporting the remaining, more hydrophilc segments. The latter, in turn, surround and thereby define the translocation pathway. This is in accord with past suggestions concerning UhpT [36, 50, 51], and with arguments presented below.

We note as well that the essential arginines in UhpT—R46 and R275—are found at or near the external surfaces of TM1 and TM7, respectively, a registration in agreement with the view that the N- and C-terminal halves of UhpT might display pseudo-twofold symmetry reflecting a common origin for the two halves of the molecule.

Role(s) of R46 and R275

R46 and its neighbors on TM1 have not before been examined by mutagenesis, but considerable effort has been directed to a study of residues in and around TM7, the locus of R275 [36, 50, 51]. In particular, cysteine-scanning mutagenesis provides strong evidence that TM7 is an amphipathic alpha helix [51] whose hydrophilic surface, defined by its several polar residues (including R275), lines the translocation pathway. It is feasible, therefore, that electrostatic interaction enables the positively charged R275 to take part in recognition of the anionic sugar phosphate substrate. There are other positive residues in this region, but it appears that only R275 (or R46) could function in this way. Thus, while UhpT arginines and lysines are found mainly at the cytoplasmic border (Fig. 4), as in most membrane proteins [20], there are six positively charged residues near the UhpT external surface (R46, K47, R275, K291, K294, K404) (Fig. 4). However, K47 is not strongly conserved (Fig. 3), and K291, K294 and K404 can each be replaced by cysteine without affecting function [36; M.-C. Fann and P.C.

Maloney, *in preparation*]. Accordingly, we conclude that R46 and R275 are not only uniquely required arginines, but are also uniquely required among the basic residues at the UhpT external surface.

One cannot assign explicit function to R275, since its cysteine and lysine derivatives are nonfunctional, so the following speculation relies heavily on inference. Given this precaution, however, we would emphasize the following: (i) that substrate protection of 2,3-butanedione inhibition (Table 1) suggests arginine(s) is located near the UhpT active site (i.e., the translocation pathway); (ii) that the functions fulfilled by R46 and R275 require arginine specifically; and (iii) that there is a considerable literature implicating arginine as an active site ligand for phosphate (*above*). These facts are then coupled with the postulate that R275 itself faces the translocation pathway. This latter seems highly likely, since proximal (V273) and distal (1276) residues on the same helical surface lie within the central core of the pathway [48], where they are presumed to be alternately accessible to both membrane surface [*cf.*, 47]. Together, these findings strongly imply that R275 interacts directly with substrates transported by UhpT. A related line of reasoning rationalizes the essential nature of R46. Thus, studies of exchange stoichiometry in *L. lactis* [3] showed that the UhpT active site has an unusual bifunctional character, in that it can accept either two monovalent substrate anions or a single divalent anion. Supposing that R275 aids in the binding of one substrate anion (one negative charge), the argument concerning symmetry (*above*) suggests the correlative function (binding of the second negative charge) might be assigned to R46, which would lie in an equivalent position in the covalent heterodimer. Such inferences are necessarily indirect. Even so, there is considerable supportive evidence that at least R275 is found in a position where it may be expected to influence substrate specificity (see above). For this reason, we suggest that by virtue of their locations at or near one end of the translocation pathway, R46 and R275 may function in establishing substrate specificity.

These studies were supported by United States Public Health Service Grant GM24195 from the National Institutes of Health. We thank Dr. R.J. Kadner for the generous give of strain RK5000 and phagemid p261.

References

1. Ambudkar, S.V., Maloney, P.C. 1986. Bacterial anion exchange: Use of osmolytes for solubilization and reconstitution of Pi-linked antiport from *Streptococcus lactis. J. Biol. Chem.* **261:**10079–10086
2. Ambudkar, S.V., Larson, T.J., Maloney, P.C. 1986. Reconstitution of sugar phosphate transport systems of *Escherichia coli. J. Biol. Chem.* **261:**9083–9086
3. Ambudkar, S.V., Sonna, L.A., Maloney, P.C. 1986. Variable stoichiometry of phosphate-linked anion exchange in *Streptococcus lactis:* Implications for the mechanism of sugar phosphate transport by bacteria. *Proc. Natl. Acad. Sci. USA* **83:**280–284
4. Bairoch, A., Bucher, P., Hofmann, K. 1997. The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25:**217–221
5. Berger, S.A., Evans, P.R. 1990. Active-site mutants altering the cooperativity of *E. coli* phosphofructokinase. *Nature* **343:**575–576
6. Bordo, D., Argos, P. 1991. Suggestions for ''safe'' residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* **217:**721–729
7. Branlant, G., Tritsch, D., Biellmann, J-F. 1981. Evidence for the presence of anion-recognition sites in pig-liver aldehyde reductase. Modification by phenylglyoxal and *p*-carboxyphenylglyoxal of an arginyl residue located close to the substrate-binding site. *Eur. J. Biochem.* **116:**505–512
8. Cheng, K.-C., Nowak, T. 1989. Arginine residues at the active site of avian liver phosphoenolpyruvate carboxykinase. *J. Biol. Chem.* **264:**3317–3324
9. Eiglmeier, K., Boos, W., Cole, S.T. 1987. Nucleotide sequence and transcriptional startpoint of the *glpT* gene of *Escherichia coli:* extensive sequence homology of the glycerol-3-phosphate transport protein with components of the hexose-6-phosphate transport system. *Molec. Microbiol.* **1:**251–258
10. Eklund, H., Samama, J.P., Jones, T.A. 1984. Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. *Biochemistry* **23:**5982–5996
11. Fleischmann, R.D., Adams, M.D., White, O. 1995. Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269:**496–512
12. Fortin, A.F., Hauber, J.M., Kantrowitz, E.R. 1981. Comparison of the essential arginine residue in *Escherichia coli* ornithine and aspartate transcarbamylases. *Biochim. Biophys. Acta* **662:**8–14
13. Fox, C.F., Kennedy, E.P. 1965. Specific labeling and partial purification of the M protein, a component of the beta-galactoside transport system of *Escherichia coli. Proc. Natl. Acad. Sci. USA* **51:**891–896
14. Fraser, C.M. 1995. The minimal gene complement of *Mycoplamsa genitalium. Science* **270:**397–403
15. Friedrich, M.J., Kadner, R.J. 1987. Nucleotide sequence of the *uhp* region of *Escherichia coli. J. Bacteriol.* **169:**3556–3563
16. Goldrick, D., Yu, G.Q., Jiang, S.Q., Hong, J.S. 1988. Nucleotide sequence and transcription start point of the phosphoglycerate transporter gene of *Salmonella typhimurium. J. Bacteriol.* **170:**3421–3426
17. Goswitz, V.C., Brooker, R.J. 1995. Structural features of the uniporter-symporter-antiporter superfamily. *Protein Science* **4:**534–537
18. Griffith, J.K., Baker, M.E., Rouch, D.A., Page, M.G.P., Skurray, R.A., Paulsen, I.T., Chater, K.F., Baldwin, S.A., Henderson, P.J.F. 1992. Membrane transport proteins: implications of sequence comparisons. *Curr. Opin. Cell Biol.* **4:**684–695
19. Hardy, L.W., Finer-Moore, J.S., Montfort, W.R., Jones, M.O., Santi, D.V., Stroud, R.M. 1987. Atomic structure of thymidylate synthase: target for rational drug design. *Science* **235:**448–455
20. von Heijne, G. 1992. Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225:**487–494
21. Henderson, P.J.F. 1993. The 12-transmembrane helix transporters. *Curr. Opin. Cell Biol.* **5:**708–721
22. Henikoff, S., Henikoff, J.G., Alford, W.J., Pietrokovski, S. 1995. Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* **163:**GC17–26
23. Island M.D., Wei, B.Y., Kadner, R.J. 1992. Structure and function of the *uhp* genes for the sugar phosphate transport system in *Esch-*

erichia coli and *Salmonella typhimurium. J. Bacteriol.* **174:**2754–2762
24. Johanson, R.A., Henkin, J. 1985. Affinity labeling of dihydrofolate reductase with an antifolate glyoxal. *J. Biol. Chem.* **260:**1465–1474
25. Kim, E.E., Wyckoff, H.W. 1990. Structure of alkaline phosphatases. *Clin. Chem. Acta* **186:**175–188
26. Kunkel, T.A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* **82:**488–492
27. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:**680–685
28. Luecke, H., Quiocho, F.A. 1990. High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature* **347:**402–406
29. Lundqvist, T., Schneider, G. 1989. Crystal structure of the complex of ribulose-1,5-bisphosphate carboxylase and a transition state analogue, 2-carboxy-D-ara-binitol 1,5-bisphosphate. *J. Biol. Chem.* **264:**7078–7083
30. Lloyd, A.D., Kadner, R.J. 1990. Topology of the *Escherichia coli* UhpT sugar phosphate transporter analyzed by using TnphoA fusions. *J. Bacteriol.* **172:**1688–1693
31. Maloney, P.C. 1990. Resolution and reconstitution of anion exchange reactions. *Phil. Trans. R. Soc. Lond. B* **326:**437–454
32. Maloney, P.C. 1994. Bacterial transporters. *Curr. Opin. Cell Biol.* **6:**571–582
33. Maloney, P.C., Ambudkar, S.V., Anantharam, V., Sonna, L.A., Varadhachary, A. 1990. Anion exchange mechanisms in bacteria. *Microbiol. Rev.* **54:**1–17
34. Marger, S.D., Saier, M.H., Jr. 1993. A major superfamily of transmembrane facilitators that catalyze uniport, symport and antiport. *Trends Biochem. Sci.* **18:**13–20
35. Marshall, M., Cohen, P.P. 1980. Evidence for an exceptionally reactive arginyl residue at the binding site for carbamyl phosphate in bovine ornithine transcarbamylase. *J. Biol. Chem.* **255:**7301–7305
36. Matos, M., Fann, M.-C., Yan, R.-T., Maloney, P.C. 1996. Enzymatic and biochemical probes of residues external to the translocation pathway of UhpT, the sugar phosphate carrier of *Escherichia coli. J. Biol. Chem.* **271:**18571–19575
37. McPherson, A., Brayer, G.D., Morrison, R.D. 1986. Crystal structure of RNase A complexed with dp(A)4. *J. Mol. Biol.* **189:**305–327
38. Nilsson, R.P., Beijer, L., Rutbert, B. 1994. The glpT and glpQ genes of the glycerol regulon in *Bacillus subtilis. Microbiol.* **140:**723–730
39. Riordan, J.F. 1979. Arginyl residues and anion binding sites in proteins. *Molec. Cell Biochem.* **26:**71–92
40. Riordan, J.F., McElvany, K.D., Borders, Jr., C.L. 1977. Arginyl residues: anion recognition sites in enzymes. *Science* **195:**884–886
41. Ruan, Z.-S., Anantharam, V., Crawford, I.T., Ambudkar, S.V., Rhee, S.Y., Allison, M.J., Maloney, P.C. 1992. Identification, purification and reconstitution of the oxalate: formate antiport protein of *Oxalobacter formigenes. J. Biol. Chem.* **267:**10537–10543
42. Rubin, R.A., Levy, S.B., Heinrikson, R.L., Kezdy, F.J. 1990. Gene duplication in the evolution of the two complementing domains of Gram-negative bacterial tetracycline efflux proteins. *Gene* **87:**7–13
43. Sanger, F., Nicklen, S., Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:**5463–5467
44. Stec, B., Lebioda, L. 1990. Refined structure of yeast apo-enolase at 2.25 A resolution. *J. Mol. Biol.* **211:**235–248
45. Skarzynski T., Moody, P.C., Wonacott, A.J. 1987. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 A resolution. *J. Mol. Biol.* **193:**171–187
46. Sonna, L.A., Ambudkar, S.V., Maloney, P.C. 1988. The mechanism of glucose 6-phosphate transport by *Escherichia coli. J. Biol. Chem.* **263:**6625–6630
47. Varadhachary, A., Maloney, P.C. 1990. A rapid method for reconstitution of bacterial membrane proteins. *Molec. Microbiol.* **4:**1407–1411
48. Weston, L.A., Kadner, R.J. 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
49. Yamaguchi, A., Kimura, T., Someya, Y., Sawai, T. 1993. Metal-tetracycline/H+ antiporter of *Escherichia coli* encoded by transposon Tn10. The structural resemblance and functional difference in the role of the duplicated sequence motif between hydrophobic segments 2 and 3 and segments 8 and 9. *J. Biol. Chem.* **268:**6496–6504
50. Yan, R.-T., Maloney, P.C. 1993. Identification of a residue within the translocation pathway of a membrane carrier. *Cell* **75:**37–44
51. Yan, R.-T., Maloney, P.C. 1995. Residues seen by substrate moving through a membrane carrier. *Proc. Natl. Acad. Sci. USA* **92:**5973–5976