A water-soluble aldose sugar dehydrogenase (Asd) has been purified for the first time from *Escherichia coli*. The enzyme is able to act upon a broad range of aldose sugars, encompassing hexoses, pentoses, disaccharides, and trisaccharides, and is able to oxidize glucose to gluconolactone with subsequent hydrolysis to gluconic acid. The enzyme shows the ability to bind pyrroloquinoline quinone (PQQ) in the presence of Ca^{2+} in a manner that is proportional to its catalytic activity. The x-ray structure has been determined in the apo-form and as the PQQ-bound active holoenzyme. The β-propeller fold of this protein is conserved between *E. coli* Asd and *Acinetobacter calcoaceticus* soluble glucose dehydrogenase (sGdh), with major structural differences lying in loop and surface-exposed regions. Many of the residues involved in binding the cofactor are conserved between the two enzymes, but significant differences exist in residues likely to contact substrates. PQQ is bound in a large cleft in the protein surface and is uniquely solvent-accessible compared with other PQQ enzymes. The exposed and charged nature of the active site and the activity profile of this enzyme indicate possible factors that underlie a low affinity for glucose but generic broad substrate specificity for aldose sugars. These structural and catalytic properties of the enzymes have led us to propose that *E. coli* Asd provides a prototype structure for a new subgroup of PQQ-dependent soluble dehydrogenases that is distinct from the *A. calcoaceticus* sGdh subgroup.

Proteins that require the prosthetic group pyrroloquinoline quinone (PQQ) for catalysis form the largest and best-characterized subclass of quinoproteins, catalyzing the conversion of a wide range of different alcohol- and aldehyde-containing compounds to their corresponding aldehydes/ketones and acids as the first step in the catabolism of these compounds. On the basis of amino acid sequence and three-dimensional structure, these proteins can be divided into two classes. One class, of which methanol dehydrogenase is the classic example, exhibits an eight-bladed β-propeller structure of ~600 amino acids. The second class includes the structurally defined soluble glucose dehydrogenase (sGdh) from *Acinetobacter calcoaceticus*, which is a 454-residue protein with a six-bladed β-propeller domain (1). Putative homologues of sGdh have been identified in phylogenetically diverse prokaryotic genera spanning Bacteria and Archaea (1). A role for PQQ-dependent enzymes in eukaryotes has recently been suggested but is still a matter for debate (2–5).

Until now, the *A. calcoaceticus* sGdh was the only protein of the second class to have been extensively characterized. This enzyme functions as a dimer of identical 50-kDa subunits (6). Each monomer binds three calcium ions and one PQQ cofactor molecule (7, 8). One of the calcium ions is needed for activation of PQQ, and the others are required for functional dimerization of the protein (8, 9). The protein oxidizes a wide variety of mono- and disaccharides as well as the trisaccharide maltotriose to their corresponding lactones (10, 11). The enzyme has a β-propeller fold comprising six four-stranded, antiparallel β-sheets (12). PQQ is bound in a wide solvent-accessible active site in the center of the molecule, near the pseudo-6-fold rotation symmetry axis (8, 13). Structures of complexes with the substrate glucose and competitive inhibitor methylhydrazine have resolved the substrate-binding site and led to elucidation of the catalytic mechanism of the reductive half-cycle (the substrate oxidation process) (8, 13).

The *A. calcoaceticus* sGdh has been successfully incorporated into successful diagnostic systems to monitor glucose levels in the blood of diabetics (14–16). This system has the advantage of being extremely rapid due to the high catalytic activity of sGdh and of being oxygen-independent. The latter feature sets it apart from systems based on the flavoprotein glucose oxidase, for which the oxidative half-cycle of the protein involves an obligatory reaction with oxygen (17). In many cases the bacterial and archaeal homologues of sGdh identified previously (1) have rather low amino acid sequence identity with the *A. calcoaceticus* sGdh and thus their specificity and activity toward glucose and maltose cannot be predicted. Determination of the three-dimensional structures of these proteins and comparison with *A. calcoaceticus* sGdh may allow identification of factors involved in determining the substrate specificity of these proteins and suggest potential protein engineering targets to improve the current diagnostic systems.
Here we report the first expression of the yliI (accession code NP415358) gene of *Escherichia coli* and purification of its gene product. From bioinformatic analysis, the gene product is predicted to be a PQQ-binding protein that has a low 18% identity with the *A. calcoaceticus* sGdh. Our biochemical analysis shows this enzyme to have a low affinity for both glucose and maltose but a generic promiscuity toward mono-, di-, and trisaccharide aldose sugars. Structural determination of YliI reveals that the PQQ-binding site lies in a shallow, solvent-exposed cleft on the surface of the protein, providing structural insight into the poor substrate selectivity. We have analyzed the primary sequences of the putative *A. calcoaceticus* sGdh homologues available in the data base and suggest that functional evolutionary divergence of these proteins has generated distinct subtypes. The significant catalytic and structural differences between the enzymes lead us to propose that the *E. coli* yliI gene product provides the first structure for a new subtype of soluble PQQ-dependent dehydrogenases, with substrate selectivity and structural features that distinguish them from *A. calcoaceticus* sGdh enzyme type and which we propose to call soluble aldose sugar dehydrogenase (Asd).

**MATERIALS AND METHODS**

**Bacterial Strains, Growth Media, and Recombinant DNA Techniques**—*E. coli* cells were grown at 37 °C in LB liquid cultures or agar plates (18). Kanamycin was used at 25 mg/liter when appropriate. Plasmid DNA was isolated using Qiagen miniprep (Qiagen) kits. Chromosomal DNA was isolated using DNeasy Tissue kits (Qiagen). *E. coli* cells were transformed using standard methods. PCR was carried out using Pfu (Promega) or Taq (Roche Applied Science) DNA polymerases. Restriction enzymes were purchased from Promega. All constructs obtained were verified by DNA sequence analysis.

**Construction of the YliI Expression Plasmid**—The yliI expression plasmid pET M-11 yliI was constructed by amplifying a DNA fragment containing the yliI gene (without the first twenty codons encoding the original peptide) from *E. coli* K12 genomic DNA using PCR with these primers: forward, 5’-CATGCGCATGCTCCGCAACGGTAAATGTCGAA-3’; reverse, 5’-CGCGGATCCCTAAATGCGTGGCATAACTTTAAGTAAATTC-3’. The product was digested with Ncol and BamHI restriction enzymes and ligated into the pET M-11 vector (a gift from Arie Geerlof and Gunter Stier, European Molecular Biology Laboratory, Heidelberg, Germany) cut with the same enzymes and treated with calf intestinal phosphatase. The plasmid obtained was transformed into competent DH5α cells and selected on LB plates containing kanamycin. Positive clones were confirmed by DNA sequence analysis.

**Expression of E. coli yliI**—The pET M-11 yliI construct was transformed into *E. coli* BL21 (DE3) cells for production of recombinant YliI protein. Single colonies were used to inoculate 5 ml of LB medium containing kanamycin at 37 °C for 16 h. This culture was used to inoculate 500 ml of LB medium containing kanamycin. Cultures were grown until late log phase (A600 nm = 0.8) and then cooled to 25 °C prior to the induction of protein expression with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cells were maintained at 25 °C for 5 h and then harvested by centrifugation (3000 × g, 20 min, 4 °C).

For production of selenomethionine-containing *E. coli* YliI, the *E. coli* selenium auxotroph strain B834 (DE3) was transformed with the plasmid pET M-11 yliI. The transformed B834 (DE3) cells were initially grown at 37 °C on M9 minimal medium, supplemented with 50 μg/ml methionine. When the cells reached an A600 of 1.0, the culture was centrifuged at 6000 × g for 10 min at 4 °C. Cells were resuspended in M9 minimal medium without methionine and incubated with shaking at 37 °C. After 6 h, seleno-L-methionine was added to a final concentration of 50 μg/ml, and the culture was incubated with shaking at 37 °C for 30 min. Overexpression of yliI was then induced by cooling the culture to room temperature and adding IPTG to a final concentration of 1 mM. The culture was incubated with shaking at 25 °C for a further 10 h. Cells were harvested by centrifugation at 7000 × g for 15 min at 4 °C. The incorporation of selenomethionine into purified protein was verified by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

**Purification of E. coli YliI**—Harvested cells were washed and resuspended in 20 mM Tris-HCl, 150 mM NaCl, pH 7.5. Cell suspension was incubated on ice with DNase for 15 min then sonicated while on ice. Cell debris was pelleted by centrifugation (35,000 × g, 30 min, 4 °C). The soluble fraction obtained after centrifugation was applied to, and eluted from, a hand-poured Ni2+-nitrilotriacetic acid metal chelation affinity column (Amersham Biosciences) according to the manufacturer’s instructions. Fractions containing His-tagged YliI were pooled and dialyzed into 20 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA, 1 mM dithiothreitol, pH 8.0. The hexahistidine tag was removed by cleavage with tobacco etch virus protease. The cleaved protein was concentrated using an ultrafiltration cell (Amicon) and applied to a S200 Sephadex gel filtration column equilibrated with 20 mM HEPES, 100 mM NaCl, pH 7.5. Fractions containing pure YliI were pooled and concentrated using an ultrafiltration cell (Amicon). Protein expression and purification was monitored by SDS-PAGE using 12% acrylamide Tris-glycine gels. After electrophoresis, the gels were stained with Coomassie Blue. Protein concentrations of extracts and purified recombinant YliI were determined according to the method of Bradford using bovine serum albumin as a standard.

**Enzyme Reconstitution with PQQ**—In order to reconstitute apo-YliI with PQQ, purified protein (5.45 mg/ml) was incubated with a 10-fold molar excess of PQQ in 20 mM HEPES, 100 mM NaCl, 1 mM CaCl2, pH 7.5, at 4 °C for 16 h. Unbound PQQ was removed by passing the mixture over a Sephadex G-25 column (PD-10; Amersham Biosciences). This step was also used to exchange the protein into 20 mM potassium phosphate, pH 7. To determine the effect that PQQ binding had on YliI activity toward glucose, an assay was set up as follows. 8.6 μl of apo-YliI (5.8 mg/ml), 5 μl of 10 mM CaCl2, and 0–25 μl of 50 mM PQQ were prepared to a final volume of 50 μl with 50 mM potassium phosphate buffered to pH 7. Solutions were incubated for 1 h at 4 °C, and 10 μl was used to reduce 2,6-dichloroindophenol (DCIP) in the presence of 100 mM D-glucose according to the standard assay described below.

**Structure of a Soluble Aldose Sugar Dehydrogenase**

---

**OCTOBER 13, 2006 • VOLUME 281 • NUMBER 41**

**JOURNAL OF BIOLOGICAL CHEMISTRY**

30651
**Structure of a Soluble Aldose Sugar Dehydrogenase**

The Aldose Sugar Dehydrogenase Assay—Enzyme activity was measured by observing the reduction at 25 °C of DCIP at 600 nm in a UV-visible spectrophotometer (U-3310; Hitachi). The reaction volume was 1 ml, typically containing 50 μM DCIP, 100 mM substrate, 50 mM potassium phosphate buffer, pH 7, and 1–8 μg of holo-YliI. Reduction of DCIP was followed over 150 s, initial rates being estimated from a linear reduction in absorbance at 600 nm (E600 (DCIP) = 21 mm−1 cm−1), typically over the first 20 s of observation. Sugar substrates were prepared 24 h in advance to allow equilibration of anomers. Enzyme activities are expressed in units mg⁻¹, where 1 unit of enzyme activity represents the reaction of 1 μmol of sugar substrate/min. Kinetic constants were estimated by a nonlinear least squares method using MicroMath Scientist. The pH optimum of the holo-YliI was determined in buffers of 50 mM potassium phosphate and 50 mM Bis-Tris propane over the pH ranges 6.0–8.0 and 7.0–9.5, respectively.

Analysis of Reaction Products by HPLC—Reaction mixtures totaling 1 ml were prepared containing 790 μl of 1 mM DCIP, 100 μl of 1 mM α-glucose, 100 μl of 50 mM potassium phosphate buffer, pH 7, and 10 μl of reconstituted E. coli YliI (3.45 mg/ml). Control reactions without enzyme and without enzyme or glucose but with 79 μl of 10 mM gluconic acid were also prepared. Reaction mixtures were incubated with shaking at 25 °C for 2 h. The reaction was deemed to have reached completion when the mixture became colorless (i.e., all of the DCIP was reduced). DCIP was then removed by four successive extractions with a 1:2-0.001 solution of chloroform/methanol/hydrochloric acid. 10-μl samples were loaded onto an Ion Pac AS4A SC running at 1 ml/min with an isocratic gradient of 1% 150 mM NaOH. Peaks were detected by ED40 conductivity.

Crystallization of YliI—A newly developed seeding protocol to obtain well diffraction crystals of soluble PQQ-dependent enzymes from a number of sources has been established. The Yli protein was crystallized at 16 °C using the hanging drop vapor diffusion method. A protein solution of 3 mg/ml was mixed with an equal volume of precipitant solution (17–22% PEG 6000, 1 mM CaCl2, 120 mM NaCl, 100 mM HEPES, pH 7.0, or 17–22% PEG 6000, 1 mM CaCl2, 120 mM NaCl, 100 mM CHES, pH 9.2) to reach a total volume of 2 μl, to which 1 μl of a seed stock solution was added. The drops were equilibrated against a 1-mL reservoir of precipitant solution.

Soaking of YliI Crystals with PQQ—Crystals were transferred to a stabilizing solution comprising 30% (w/v) PEG 6000, 1 mM CaCl2, 120 mM NaCl, 100 mM HEPES, pH 7.0, or 25% PEG 6000, 1 mM CaCl2, 120 mM NaCl, 100 mM CHES, pH 9.2, containing 1 mM PQQ. Crystals were incubated in this solution for up to 12 h and were subsequently flash-frozen using liquid nitrogen in a solution containing 30% (w/v) PEG 6000, 20% (v/v) ethylene glycol, 1 mM CaCl2, 120 mM NaCl, and 100 mM HEPES, pH 7.0, or 30% (w/v) PEG 6000, 20% ethylene glycol, 1 mM CaCl2, 120 mM NaCl, and 100 mM CHES, pH 9.2.

X-ray Diffraction Data Collection and Analysis—X-ray diffraction data sets were collected to a resolution of 1.5 Å at beam line ID29 of the European Synchrotron Radiation Facility (Grenoble, France). Data were processed and reduced using MOSFLM (20) and programs from the CCP4 suite (21). 5% of the data were set aside to follow progress of refinement using the Rfree factor (22). The final round of refinement was carried out using all of the data.

Structure Determination and Refinement—The structure of apo-YliI from E. coli was determined by single wavelength anomalous dispersion using the anomalous signal of six incorporated selenium atoms per YliI molecule. Identification of the selenium atom sites and subsequent phasing was carried out using programs from the CCP4 suite (21). The obtained phases allowed for 91% of all amino acids in the asymmetric unit of the apoenzyme to be built automatically by ARP/wARP (23). The structure of holo-YliI was determined by molecular replacement using Molrep using a model of the apoenzyme taken from an early stage in the refinement of that model. The models were completed using alternating rounds of manual model building using Coot (24) and automated refinement using REFMAC (25) and ARP (26). As refinement of the protein models approached completion, phases were calculated from the model. Water molecules were incorporated using ARP, and, in the latter stages of refinement, the models were further improved by applying individual atomic anisotropic B-factor refinement using REFMAC. Throughout refinement, progress was monitored with the aid of an Rfree value calculated with 5% of the data. Upon completion of the models, a final round of refinement using all of the data were used to give a single crystallographic R factor. The stereochemistry and geometry of the final model were judged with PROCHECK (27).

**RESULTS AND DISCUSSION**

Expression, Purification, and Biochemical Characterization of E. coli YliI—The E. coli YliI was successfully cloned and expressed in E. coli BL21 (DE3). SDS-PAGE revealed a
strong IPTG-inducible band at around 40,000 Da (Fig. 1). The observed molecular mass is consistent with the predicted recombinant protein mass of ~42,000 Da, which decreased to ~39,000 Da after cleavage of the His tag (Fig. 1).

A typical purification resulted in pure recombinant YliI as judged by SDS-PAGE with a yield of up to 30 mg of pure protein/liter of culture (Fig. 1). The identity of the pure protein was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

The UV-visible spectrum of apo-YliI is typical of that of a cofactorless enzyme and has a maximum at ~280 nm (Fig. 2). Upon reconstitution with PQQ, a second broader peak corresponding to bound PQQ can be observed, with a maximum at 353 nm (Fig. 2). D-Glucose was selected as a substrate with which to characterize the dependence of activity on PQQ and pH. In the presence of increasing amounts of glucose, the UV-visible spectrum of YliI showed an increase in absorbance corresponding to reduction of PQQ (Fig. 2). The dominant peak at 353 nm increased in size with a new maximum at 338 nm. The activity of apo-YliI toward glucose increased in a linear manner with respect to added PQQ, reaching a plateau at a value of 1 (Fig. 2, inset). From this we deduce that one molecule of PQQ binds per molecule of YliI, similar to that shown with sGdh from *A. calcoaceticus* (28). CaCl$_2$ was present in these reconstitution assays, and inclusion of the calcium chelator EDTA (1 mM) during reconstitution with PQQ resulted in no activity toward glucose, suggesting that YliI is also a Ca$^{2+}$-binding protein and that the cation is important for PQQ binding. Once PQQ is bound, the holoenzyme is extremely stable, activity being essentially unchanged after 3 months of storage at 4 °C. The observed oxidation of DCIP to DCIPH$_2$ at 600 nm was shown to be linear with respect to concentration of YliI (data not shown). The products of the D-glucose-dependent enzyme assay were assessed by HPLC. The chromatogram profile was similar to that seen when gluconic acid was applied to the HPLC column (data not shown) and showed no overlap with the chromatograms produced by any of the reaction constituents. This substantiates the theory that *E. coli* YliI is capable of acting as a D-glucose dehydrogenase, with the reaction product being gluconolactone that is rapidly hydrolyzed in solution to yield gluconic acid.

The activity of YliI toward glucose was measured over a range of pH using two different buffers, 50 mM potassium phosphate and 50 mM Bis-Tris propane. The activity in the presence of Bis-Tris propane was substantially higher than that observed in potassium phosphate at similar pH. Activity assays were therefore performed in Bis-Tris propane buffer. The activity of YliI toward glucose showed an optimum at pH 8.75 (Fig. 3).
obervation is in contrast to that observed with the sGdh from *A. calcoaceticus*, which has a pH optimum of 7 in potassium phosphate buffer (10). Comparison of the pH profile of YliI for D-glucose with that of maltose reveals substantial differences. In the presence of maltose, the pH activity profile has a maximum of pH 8.0 (Fig. 3). Varying the concentration of D-glucose, maltose, or maltotriose revealed that activity was related to substrate concentration in a manner that approximated Michaelis-Menten kinetics. At pH 8.75, the apparent kinetic parameters for D-glucose were $K_{cat} = 3360 \text{s}^{-1}$ and $K_m = 400 \text{mM}$ (Table 1). This is 1 order of magnitude lower than that for *A. calcoaceticus* sGdh (Table 1) and leads to *E. coli* YliI having a much lower catalytic efficiency for D-glucose than *A. calcoaceticus* sGdh (Table 1). The $K_m$ values of *E. coli* YliI for the disaccharide maltose or trisaccharide maltotriose were lower than for glucose, leading to a higher catalytic efficiency for these substrates compared with D-glucose (Table 1). In the case of maltose, this affinity and catalytic efficiency were still much lower than published for *A. calcoaceticus* (Table 1).

**Determination of the X-ray Crystal Structure of E. coli YliI**—The selenomethionine-labeled apoprotein of *E. coli* YliI crystallized in two different space groups, $P_2_1$ with two molecules in the asymmetric unit and $P_2_12_12_1$ with one molecule in the asymmetric unit. Holoenzyme crystals were obtained by soaking crystals of the apoenzyme with a molecule in the asymmetric unit. Holoenzyme crystals were comprised residues 3–350, one molecule of PQQ, 2 calcium ions, 13 water molecules, one phosphate molecule, and 25 molecules of ethylene glycol from the cryocoolant bound in the intermolecular spaces.

Both structures have good overall geometry as judged by PROCHECK, with no residues lying in disallowed regions of the Ramachandran plot (29). 88.7% of residues lie in the most favored regions, with a further 10.9% in additional allowed regions. Two residues consistently lie in the generously allowed region of the Ramachandran plot in both structures: Pro$^{215}$ and Leu$^{66}$. Both of these residues are in the cis-peptide conformation. This allows the backbone carbonyls of Gly$^{214}$ and Pro$^{215}$ to act as ligands to the active site calcium. The backbone conformation of Leu$^{66}$ allows it to hydrogen-bond to the OGl of Thr$^{72}$ via its carbonyl group, further stabilizing the 2A-2B strand structure. This may be important in maintaining potential hydrogen bonding partners to the substrate, such as Gln$^{197}$, whose corresponding residue in *A. calcoaceticus* sGdh, Gln$^{196}$, is involved in interactions with the glucose $\text{O}_2$ hydroxyl.

**Overall Structure and PQQ Binding of E. coli YliI**—The overall fold of *E. coli* YliI in both the apoenzyme and the active holoenzyme forms consists of the $\beta$-propeller fold characteristic of this family containing six four-stranded antiparallel $\beta$-sheets (Fig. 4). The PQQ-binding site lies in a shallow, solvent-exposed cleft on the surface of the protein close to the pseudosymmetry axis generated by the six $\beta$-sheets. In contrast to *A. calcoaceticus* sGdh, the *E. coli* YliI enzyme crystallizes as a monomer, and this was also confirmed for its native state in solution as determined by gel filtration and dynamic light scattering measurements (data not shown). Each monomer binds two calcium ions, which compares to three in *A. calcoaceticus* sGdh. In both proteins, one of these calcium ions lies in the PQQ binding pocket and will be discussed later, and another is sandwiched between two of the six strands that make up the propeller fold. It has been shown previously that the propeller fold has a general preference for 7-fold pseudosymmetry (reviewed in Ref. 32). Given that (i) the *A. calcoaceticus* sGdh and *E. coli* YliI both contain six blades, (ii) the calcium site is sandwiched between two of the six sheets and thereby part of the structural core, and (iii) four sheets of the *A. calcoaceticus* sGdh and *E. coli* YliI structures can be superimposed with low root mean square deviation values on seven-bladed propellers, such as nitrous oxide reductase (30) and quinohemoprotein amine dehydrogenase (31), it seems possible that this conserved calcium site serves to stabilize the propeller structure by allowing six sheets to adopt a stable configuration similar to those adopted by seven-bladed propellers. Moreover, one of the calcium ligands is a conserved glutamate (Glu$^{259}$) in strand 4C, which may add to the stability of this $\beta$-sheet. The third calcium binding site in *A. calcoaceticus* also lies in the long loop between strands 4C and 4D, stabilizing the loop in the vicinity of the dimerization interface. This region of the loop is significantly shorter in *E. coli* YliI and thus does not require the structural stabilization provided by calcium binding (Fig. 4).

The largest, albeit still small, structural change upon binding of PQQ to the enzyme is the rotation of the side chain of His$^{213}$ to form hydrophobic stacking interactions with PQQ. The corresponding residue in *A. calcoaceticus* sGdh, Gln$^{246}$, cannot make such interactions. These hydrophobic contacts might increase the affinity of the enzyme for PQQ and thus increase enzyme stability. There are minor rearrangements of side chains in the active site to facilitate hydrogen bonding interactions between the protein and PQQ (Fig. 5). The side chain of Tyr$^{241}$ rotates toward PQQ to form hydrogen bonds between the side chain hydroxyl and the O5, O7A, and N6 of PQQ. The side chain of Gln$^{197}$ is also seen to rotate facilitating additional hydrogen bond formation with PQQ. The structures of the apo-
and holoenzyme can be overlaid with a root mean square deviation of 0.34 across the main chain atoms.

The SigmaA-weighted $2F_o - F_c$ electron density of the PQQ cofactor is well defined, and the quality of the electron density maps allows us to ascertain that the conformation of PQQ is planar (Fig. 5). It was shown that the conformation of oxidized PQQ is not completely planar in the *Acinetobacter* sGdh (13), whereas the likely reduced state is planar (8). In the *E. coli* YliI structure, it is clear that PQQ is planar, but it is not possible to be certain of the oxidation state of the enzyme, given that the solvent contains maltose and one turnover might have taken place. All protein residues in direct contact with the cofactor have well defined electron density. Comparison of the *E. coli* YliI structures with those from *A. calcoaceticus* sGdh (Fig. 6) shows that the geometry of all known participants in the reductive half-reaction (i.e. PQQ, calcium, His$^{127}$ (His$^{144}$), and Arg$^{194}$ (Arg$^{228}$)) are very similar in the two structures. The

| TABLE 2 | Data collection and refinement statistics for crystals of *E. coli* YliI (Asd) |
|----------|-------------------------------------------------------------------------|
| **Apoenzyme** | **Holoenzyme** |
| Beam line | ESRF Beam Line ID29 | ESRF Beam Line ID29 |
| Wavelength (Å) | 0.9184 | 0.9184 |
| Resolution (Å) | 1.5 | 1.5 |
| Space group | $P_2_1$ | $P_2_1$ |
| Cell dimensions | $a = 40.83$, $b = 113.48$, $c = 75.37$, $\beta = 90.03$ | $a = 58.15$, $b = 75.32$, $c = 76.60$ |
| Completeness (%) | 94.2 (63.7) | 90.6 (54.5) |
| Redundancy (%) | 3.8 (2.3) | 4.2 (3.0) |
| $R_{min}$ | 0.057 (0.178) | 0.011 (0.208) |
| Refinement resolution range (Å) | 7.96-1.5 | 19.18-1.5 |
| No. of protein atoms | 5965 | 3088 |
| No. of calcium ions | 4 | 2 |
| No. of PQQ atoms | 0 | 1 |
| No. of water molecules | 1085 | 462 |
| $R$-factor (%) | 14.8 | 16.0 |
| $R_{free}$ factor (%) | 18.2 | 13.0 |
| Final $R$-factor using all data (%) | 15.0 | 13.2 |
| Root mean square deviations from ideality for bond lengths (Å) | 0.01 | 0.007 |
| Root mean square deviations from ideality for bond angles (degrees) | 1.221 | 1.231 |

**FIGURE 4.** The overall structure of *E. coli* YliI (Asd). A, schematic representation of the overall structure of *E. coli* YliI. “Blades” of the propeller structure are labeled 1–6. B, overlay of the loop region between strands 4C and 4D from *E. coli* YliI (Asd) (yellow) and *A. calcoaceticus* sGdh (red), which is involved in dimerization in *A. calcoaceticus* sGdh. Calcium ions are shown as spheres.

**FIGURE 5.** Stick representation of the PQQ-binding site of *E. coli* YliI (Asd) in both the apoenzyme and holoenzyme forms. The PQQ cofactor is shown in yellow with surrounding SigmaA-weighted $2F_o - F_c$ electron density. The hydrogen bonds to the protein from PQQ and the active site calcium ion (red sphere) are shown as dashed red lines. Cyan, apoenzyme; pink, holoenzyme.
charge or hydrogen bonding potential of other key residues, such as Glu\textsuperscript{146} (Asp\textsuperscript{163}), is maintained. This suggests that the catalytic mechanism for aldose sugar oxidation may be similar.

**Substrate Specificity of E. coli YliI**—The three-dimensional structure of *E. coli* YliI as described here and the substrate specificity profile of this enzyme have allowed us to understand more about the determinants of substrate specificity in this family of enzymes. Although the key components of the reaction mechanism and a number of other residues with important structural roles in *A. calcoaceticus* sGdh are conserved in *E. coli* YliI, not all of the amino acids involved in substrate binding are present, and indeed novel active site features are apparent. The long loop in *A. calcoaceticus* sGdh between strands 3B and 3C, which contains helix α2, is significantly shorter in *E. coli* YliI. This loop contains residues involved in hydrophobic interactions with glucose, such as Leu\textsuperscript{169}, and in hydrogen bonds to glucose, such as Gln\textsuperscript{168}. The hydrophobic residues Tyr\textsuperscript{343} and Trp\textsuperscript{346}, which are involved in hydrophobic interactions with glucose in *A. calcoaceticus* sGdh, are also absent in *E. coli* YliI due to the dramatic difference in structure of the loop between strands 4D and 5A. The significantly shorter loop regions in *E. coli* YliI generate a very open active site structure with the entire surface of PQQ exposed to solvent (Fig. 7). This imposes little steric hindrance to substrate binding but reduces the interactions that can potentially be made between substrate and enzyme to yield tight binding, which is reflected by the extremely high $K_m$ values obtained for glucose and the larger substrates maltose and maltotriose (Table 1). This also prevented us from obtaining structures of *E. coli* YliI with substrate bound at the active site, despite extensive soaking trials or crystallization in the presence of substrates. Indeed, one substrate, maltose, was used as a cryoprotectant, yet maltose molecules did not bind to the active site but rather at other sites on the protein surface.

The reduced capacity for tight interaction with the glucose and maltose in *E. coli* YliI as compared with the *A. calcoaceticus* sGdh (Table 1) led us to explore the relative activity compared with α-glucose of *E. coli* YliI toward a broader range of aldose sugar substrates (Table 3). Consideration of the chemical nature of the substrates investigated and making the assumption that the substrates bind in a similar orientation to glucose
in the *A. calcoaceticus* sGdh allowed the following conclusions to be made. *E. coli* YliI, like the *A. calcoaceticus* sGdh, acts upon the C1-OH of the sugar ring, as shown by the virtual abolition of activity when this group is substituted. This is apparent in the difference in rates between glucose 1-phosphate and glucose 6-phosphate. The YliI enzyme also shows little or no activity toward linear sugars or sugar alcohols as displayed by rates observed with xylitol, inositol, sorbose, and mannitol. The C2-OH functional group is not required for productive binding, as shown by an increase in rate with 2-deoxyglucose and lyxose as substrates. This is not the case for *A. calcoaceticus* sGdh, which exhibits a 110-fold decrease in rate for *A. calcoaceticus* sGdh (11). The amino acids involved in interactions with the C2-OH of glucose in *A. calcoaceticus* sGdh are completely conserved in *E. coli* YliI, with the only difference being the distance of Gln62 (Gln76 in *A. calcoaceticus*) from the putative site of glucose binding. In the superposition, this side chain is positioned at only 1.7 Å from the C2-OH. This implies that it might have to rotate away from the substrate in order to accommodate glucose and most aldose sugar substrates but not 2-deoxyglucose, consistent with the notion that the hydroxyl at the substrate C2 position is not important for binding. In *A. calcoaceticus*, the inability to form productive hydrogen bonds with C2 of 2-deoxyglucose reduces its relative activity (10). Larger substitutions at the C2 position result in a loss of activity of the *E. coli* YliI, as indicated by rates observed with glucosamine and N-acetyl glucosamine (Table 3). This illustrates that subtle differences between the protein structures can lead to dramatic changes in reactivity toward certain substrates.

The trend of increasing observed velocity from D-glucose through D-fucose to D-arabinose indicates that the C5 hydroxymethyl extension is not necessary for catalysis, and indeed, its absence enhances activity. According to our superposition of the *E. coli* YliI and *A. calcoaceticus* sGdh structures, this extension should not make any productive interactions with the protein or PQQ, and its absence may allow hydrophobic interaction with the *E. coli*-specific Tyr241. Similar rates observed with D- and L-arabinose indicate no preference for different enantiomers. This is in contrast with sGdh from *A. calcoaceticus*, which showed an absolute preference for L-arabinose over D-arabinose. This may be due to the loss of hydrogen bonds with the C2-OH and the proximity of Leu169, which is absent in *E. coli* YliI.

Comparing monosaccharides and disaccharides, maltose (4-O-β-D-glucopyranosyl-D-glucose), cellobiose (4-O-β-D-glucopyranosyl-D-glucose), and melibiose (6-O-α-D-galactopyranosyl-D-glucose) all show rates significantly faster than that with glucose. Maltotriose (O-α-D-glucopyranosyl(1→4)-O-α-D-glucopyranosyl(1→4)-D-glucose), a trisaccharide, also showed an elevated rate when compared with glucose. This pattern is not observed in *A. calcoaceticus* sGdh and may reflect the more solvent-exposed active site that provides less steric hindrance to these more bulky sugars.

### Table 3

Activity profiles of *E. coli* YliI (Asd) toward different substrates

| Group and name | pH 7 | pH 8.75 |
|---------------|------|---------|
| **Simple sugars** | | |
| D-Glucose | 76 | 1 | 1260 | 1 |
| D-Galactose | 192 | 2.5 | 2930 | 2.3 |
| D-Fructose | 44 | 0.6 | 480 | 0.4 |
| D-Arabinose | 295 | 3.9 | 4420 | 3.5 |
| L-Arabinose | 252 | 3.3 | 4710 | 3.3 |
| D-Fucose | 232 | 3.0 | 3950 | 3.1 |
| D-Mannose | 133 | 1.8 | 1040 | 0.8 |
| D-Lyxose | 261 | 3.4 | 2640 | 2.1 |
| D-Xylose | 140 | 1.8 | 2130 | 1.7 |
| D-Ribose | ND | ND | 2520 | 2.0 |
| **Linear sugars/alcohols** | | |
| Xylitol | ND | ND | 32 | 0 |
| myo-Inositol | 7 | 0.1 | 16 | 0 |
| L-Sorbose | 31 | 0.4 | 330 | 0.3 |
| Mannitol | ND | ND | 41 | 0 |
| Methanol | 0 | 0 | 0 | 0 |
| Ethanol | 0 | 0 | 0 | 0 |
| **Substituted glucose** | | |
| 2-Deoxy-glucose | 279 | 3.7 | 3620 | 2.9 |
| Glucosamine | ND | ND | 100 | 0.1 |
| N-acetylglucosamine | ND | ND | 590 | 0.5 |
| Glucose 1-phosphate | 9 | 0.1 | 0 | 0 |
| Glucose 6-phosphate | 69 | 0.9 | 660 | 0.5 |
| **Disaccharide** | | |
| Maltose | 130 | 1.7 | 2240 | 1.8 |
| α-Lactose | 165 | 2.2 | 720 | 0.6 |
| D-Sucrose | 11 | 0.1 | 24 | 0 |
| D-Cellobiose | 290 | 3.8 | 2910 | 2.3 |
| Melibiose | 154 | 2.0 | 2030 | 1.6 |
| **Polysaccharide** | | |
| Maltotriose | 143 | 1.9 | 1530 | 1.2 |
Structure of a Soluble Aldose Sugar Dehydrogenase

CONCLUSIONS

In this study, we have characterized a novel PQQ-dependent enzyme from *E. coli* in terms of its three-dimensional structure, cofactor binding features, and substrate specificity. A large number of other homologues of this protein exist in a phylogenetically wide range of prokaryotes and remain as yet unstudied (1). Phylogenetic analysis of the primary structure of putative homologues available in the data base using BLAST (32), ClustalW (33) and TreeView (19) suggests distinct evolutionary subclasses that cannot be correlated with 16 S ribosomal evolutionary relationships (supplemental Fig. 1). In terms of their primary structure, *E. coli* Yiii and *A. calcoaceticus* sGdh are only distantly related within this group. The sGdh from *A. calcoaceticus* shows significantly greater identity (40%) with the 50-kDa PQQ-binding domain of soluble l-sorbosone dehydrogenase from *Ketogulonicigenium vulgare* (34, 35) than it does with *E. coli* Yiii (18%). The elongated loop regions found in *A. calcoaceticus* sGdh are also present in *K. vulgare* l-sorbosone dehydrogenase but are not found in any other homologues identified so far. This underlies a difference size between the sGdh/-s-sorbosone dehydrogenase subgroup (~50 kDa) compared with the *E. coli* Yiii subgroup (~40 kDa). The key residues believed to be involved in the reaction mechanism, however, remain conserved across all of the homologues. Significant differences remain between these other homologues, not least in the residues potentially binding to substrate molecules in the active site. Given that the physiological role of all of these proteins is unknown, it is conceivable that they have different functions in these very different organisms. The structure and enzymatic properties of the *E. coli* PQQ-dependent enzyme have provided insight into what appears to be the common overall structure of this family of soluble quinoproteins but have identified a new subtype with broad substrate specificity toward aldose sugars and unique structural features. Since it has a low affinity for glucose and it is not at all clear that this enzyme has a physiological substrate, we feel it would be misleading to refer to this enzyme as a soluble glucose dehydrogenase and propose that it is henceforth given the more generic name aldose sugar dehydrogenase (Asd).

The *E. coli* Asd (Yiii) is predicted to be part of a periplasmic electron transfer system that can serve to input electrons into the respiratory network. It represents a second example of a PQQ-dependent dehydrogenase that can input sugar-derived electrons into the *E. coli* respiratory network. The existence of the first system, an integral membrane glucose-ubiquinone oxidoreductase, has been recognized for some time. This enzyme has a *Km* for d-glucose of about 4 mM (36), 2 orders of magnitude lower than that reported here for the soluble Asd. The paradox of expressing genes encoding two different PQQ enzymes is that *E. coli* lacks the genetic information required to synthesize the PQQ cofactor. It is, however, chemotactic toward PQ (41) and thus could use environmental PQ released by other organisms to activate periplasmic (Asd) or periplasmic facing (membrane glucose-ubiquinone oxidoreductase) apoenzymes. Many of the environmental niches that *E. coli* and other enteric bacteria with Asd can occupy will have aldose sugars present, and thus this sGdh could provide for bioenergetic options in such environments that further highlight the flexibility of its respiratory network. There has been significant progress in structurally defining this respiratory network in recent years with structures of nitrite reductase, membrane-bound nitrate reductase, formate dehydrogenase, fumarate reductase, succinate dehydrogenase, and cytochrome *bo* oxidase all emerging (37–43). The present study expands this molecular definition of the *E. coli* respirome through the determination of the structure of a hitherto unstudied respiratory enzyme.

Acknowledgments—We thank Arrie Geerlof and Gunter Stier (EMBL, Heidelberg) for the gift of the pET M-11 vector, Dr. Charles Brearley for help with HPLC analysis, Ann Reilly for production of samples of *E. coli* Asd (Yiii), Dr. Frank Sargent for helpful comments on the manuscript, and Professor Robert Eisenthal (University of Bath) for helpful discussions.

REFERENCES

1. Oubrie, A. (2003) Biochim. Biophys. Acta 1647, 143–151
2. Kasahara, T., and Kato, T. (2005) Nature 433, E11–E12
3. Kasahara, T., and Kato, T. (2003) Nature 422, 832
4. Felton, L. M., and Anthony, C. (2005) Nature 433, E10–E12
5. Rucker, R., Storms, D., Sheets, A., Tschaparian, E., and Fascetti, A. (2005) Nature 433, E10–E12
6. Olsbothorn, A. J. J., and Duine, J. A. (1996) Arch. Biochem. Biophys. 336, 42–48
7. Olsbothorn, A. J. J., Otsuki, T., and Duine, J. A. (1997) Eur. J. Biochem. 247, 659–665
8. Oubrie, A., Rozeboom, H. J., Kalk, K. H., Olsthoorn, A. J. J., Duine, J. A., and Dijkstra, B. W. (1999) EMBO J. 18, 5187–5194
9. Olsbothorn, A. J. J., and Duine, J. A. (1998) Eur. J. Biochem. 255, 255–261
10. Olsbothorn, A. J. J., and Duine, J. A. (1998) Biochemistry 37, 13854–13861
11. Igarashi, S., Hirokawa, T., and Sode, K. (2004) Biomol. Eng. 21, 81–89
12. Oubrie, A., Rozeboom, H. J., Kalk, K. H., Duine, J. A., and Dijkstra, B. W. (1999) J. Mol. Biol. 289, 319–333
13. Oubrie, A., Rozeboom, H. J., and Dijkstra, B. W. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 11787–11791
14. Hoenes, J., and Unkrig, V. (1996) U. S. Patent 5,484,708
15. Tang, Z., Louie, R. F., Lee, J. H., Lee, D. M., Miller, E. E., and Kost, G. J. (2001) Crit. Care Med. 29, 1062–1070
16. Solnica, B., Naskalski, J. W., and Sieradzki, J. (2003) Clin. Chim. Acta 331, 29–35
17. Pazur, J. H., and Kleppe, K. (1964) Biochemistry 3, 578–583
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY
19. Lesley, A. G. W. (1992) *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 26
20. Collaborative Computational Project 4 (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
21. Brunger, A. T. (1997) Macromol. Crystallogr. 277, 366–396
22. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 57, 1445–1450
23. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 2126–2132
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* 53, 240–255
25. Lamzin, V. S., and Wilson, K. S. (1997) Methods Enzymol. 277, 269–305
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291
27. Doktor, P., Frank Izn, J., and Duine, J. A. (1986) *Biochem. J.* 239, 163–167
28. Ramachandran, G. N., and Sasiexekhanar, V. (1968) *Adv. Protein Chem.* 23, 283–438
30. Haltia, T., Brown, K., Tegoni, M., Cambillau, C., Saraste, M., Mattila, K., and Djinovic-Carugo, K. (2003) Biochem. J. 369, 77–88
31. Satoh, A., Kim, J. K., Miyahara, I., Devreese, B., Vandenbergh, I., Hacisalihoglu, A., Okajima, T., Kuroda, S., Adachi, O., Duine, J. A., Van Beeumen, J., Tanizawa, K., and Hirotsu, K. (2002) J. Biol. Chem. 277, 2830–2834
32. Altschul, S. F., Madden, T. L., Schiffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
33. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) Nucleic Acids Res. 31, 3497–3500
34. Miyazaki, T., Sugisawa, T., and Hoshino, T. (2006) Appl. Environ. Microbiol. 72, 1487–1495
35. Sugisawa, T., Miyazaki, T., and Hoshino, T. (2005) Biosci. Biotechnol. Biochem. 69, 659–662
36. Dewanti, A. R., and Duine, J. A. (1998) Biochemistry 37, 6810–6818
37. Jormakka, M., Richardson, D., Byrne, B., and Iwata, S. (2004) Structure 12, 95–104
38. Bertero, M. G., Rothery, R. A., Palak, M., Hou, C., Lim, D., Blasco, F., Weiner, J. H., and Strynadka, N. C. (2003) Nat. Struct. Biol. 10, 681–687
39. Bamford, V. A., Angove, H. C., Seward, H. E., Thomson, A. J., Cole, J. A., Butt, J. N., Hemmings, A. M., and Richardson, D. J. (2002) Biochemistry 41, 2921–2931
40. Jormakka, M., Tornroth, S., Byrne, B., and Iwata, S. (2002) Science 295, 1863–1868
41. Iverson, T. M., Luna-Chavez, C., Cecchini, G., and Rees, D. C. (1999) Science 284, 1961–1966
42. Yankovskaya, V., Horsefield, R., Tornroth, S., Luna-Chavez, C., Miyoshi, H., Leger, C., Byrne, B., Cecchini, G., and Iwata, S. (2003) Science 299, 700–704
43. Abramson, J., Riistama, S., Larsson, G., Jasaitis, A., Svensson-Ek, M., Laakkonen, L., Puustinen, A., Iwata, S., and Wikstrom, M. (2000) Nat. Struct. Biol. 7, 910–917