

## Identification of the *yqhE* and *yafB* Genes Encoding Two 2,5-Diketo-D-Gluconate Reductases in *Escherichia coli*

DO-YOUNG YUM, BONG-YONG LEE, AND JAE-GU PAN\*

Bioprocess Engineering Division, Korea Research Institute of Bioscience and Biotechnology (KRIBB),  
Yusong, Taejeon 305-600, Korea

Received 5 April 1999/Accepted 26 May 1999

The identification of a gene (*viaE*) encoding 2-ketoaldonate reductase (2KR) in our previous work led to the hypothesis that *Escherichia coli* has other ketogluconate reductases including 2,5-diketo-D-gluconate reductase (25DKGR) and to study of the related ketogluconate metabolism. By using the deduced amino acid sequences of 5-diketo-D-gluconate reductase (5KDGR) of *Gluconobacter oxydans* and 25DKGR of *Corynebacterium* sp., protein databases were screened to detect homologous proteins. Among the proteins of *E. coli*, an oxidoreductase encoded by *yjgU* and having 56% similarity to 5KDGR of *G. oxydans* and two hypothetical oxidoreductases encoded by *yqhE* and *yafB* and having 49.8 and 42% similarity, respectively, to 25DKGR of *Corynebacterium* sp. were detected. Recently, the *yjgU* gene was identified as encoding 5KDGR and renamed *idnO* (C. Bausch, N. Peekhaus, C. Utz, T. Blais, E. Murray, T. Lowary, and T. Conway, J. Bacteriol. 180:3704–3710, 1998). The pathways involved in the metabolism of ketogluconate by *E. coli* have been predicted by biochemical analysis of purified enzymes and chemical analysis of the pathway intermediates. The gene products of *yqhE* and *yafB* were identified as 25DKGR-A, and 25DKGR-B, respectively, catalyzing the reduction of 25DKG to 2-keto-L-gulonate (2KLG). The native 25DKGR-A, 25DKGR-B, and 5KDGR had apparent molecular weights of about 30,000, 30,000, and 54,000, respectively. In sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels, all three enzymes showed protein bands with a molecular weight of about 29,000, which indicated that 25DKGR-A, 25DKGR-B, and 5KDGR may exist as monomeric, monomeric, and dimeric proteins, respectively. The optimum pHs for reduction were 7.5, 7.0, and 8.0, respectively. The 5KDGR was active with NADH, whereas 25DKGR-A and 25DKGR-B were active with NADPH as a preferred electron donor. 25DKG can be converted to 5KDG by 2KR, which is then reduced to D-gluconate by 5KDGR. The pathways were compared with those of *Erwinia* sp. and *Corynebacterium* sp. A BLAST search of published and incomplete microbial genome sequences revealed that the ketogluconate reductases and their related metabolism may be widespread in many species.

Incomplete oxidation of glucose to ketogluconates, 2-keto-D-gluconate (2KDG), 5-keto-D-gluconate (5KDG), and 2,5-diketo-D-gluconate (25DKG), in a variety of microorganisms has been shown to proceed via membrane-bound dehydrogenases linked to the electron transport chain (4, 33). Some ketogluconates can be used as the source of carbon and energy for many bacteria. The metabolic pathways involved in the use of such ketogluconates have been studied for *Corynebacterium* sp. (32) and *Erwinia* spp. (5, 35). The pathways in the two species are quite similar, except that, in *Erwinia* sp., 25DKG is converted to 5KDG but in *Corynebacterium* sp., 25DKG is converted to 2KDG before being converted to D-gluconate. The sequential conversion of ketogluconates to D-gluconate is mediated by soluble NAD(P)H-dependent reductases. In those bacteria, D-gluconate is phosphorylated to 6-phosphogluconate and further metabolized through the pentose phosphate pathway. In our previous work, it was found that the decrease of 2KDG produced from D-gluconate in the cultivation of recombinant *Escherichia coli* harboring the cloned membrane-bound gluconate dehydrogenase gene (38) was due to 2-ketoaldonate reductase (2KR) as the cytosolic enzyme responsible for conversion of 2KDG to D-gluconate (37). We also identified a gene (*viaE*) encoding 2KR and found that

2KR catalyzes the reduction of 25DKG to 5KDG and of 2-keto-L-gulonate (2KLG) to L-idonate (IA), as well as of 2KDG to D-gluconate (37). This result suggested strongly that the other ketogluconate reductases known in other bacteria, 5-keto-D-gluconate reductase (5KDGR) and 2,5-diketo-D-gluconate reductase (25DKGR), and the related ketogluconate metabolism could also exist in *E. coli*.

Several ketogluconate reductases, 5KDGR, 2KDGR, 2KR, and 25DKGR, have been purified and characterized (1–3, 27, 34, 36) elsewhere, and the genes encoding 5KDGR (21) and 25DKGR (6, 17, 18) have been cloned elsewhere. To find ketogluconate reductases in *E. coli*, we searched homologous proteins in the protein databases with the amino acid sequences of 5KDGR of *Gluconobacter oxydans* (21) and 25DKGR of *Corynebacterium* sp. (6). As a result, our attention was drawn to two hypothetical oxidoreductases, encoded by *yqhE* and *yafB*, showing similarity to 25DKGR. A hypothetical oxidoreductase, encoded by *yjgU*, showing similarity to 5KDGR of *G. oxydans* was also found; however, the *yjgU* gene was found to encode 5KDGR and renamed *idnO* recently (8). 5KDGR of *E. coli* was identified by sequence analysis of the GntII (subsidiary) system for gluconate metabolism. The GntII system encodes a pathway for catabolism of IA, in which IA is converted to 5KDG by IA dehydrogenase before being reduced to D-gluconate by 5KDGR. D-Gluconate is phosphorylated to 6-phosphogluconate and further metabolized through the Entner-Doudoroff pathway.

In this report, we will show that the proteins encoded by

\* Corresponding author. Mailing address: Bioprocess Engineering Division, Korea Research Institute of Bioscience and Biotechnology (KRIBB), P.O. Box 115, Yusong, Taejeon 305-600, Korea. Phone: 82-42-860-4483. Fax: 82-42-860-4594. E-mail: jgpan@kribb4680.kribb.re.kr.

*yqhE* and *yafB* are 25DKGR-A and 25DKGR-B, catalyzing the reduction of 25DKG to 2KLG, and the pathways involved in the metabolism of ketogluconates by *E. coli* are similar to those involved in the metabolism of ketogluconates by *Erwinia* and *Corynebacterium* spp. 25DKGR is an important enzyme in the bioconversion of 2KLG, a key intermediate in the biosynthesis of ascorbic acid (vitamin C) by the glucose pathway (17, 27, 31, 32, 34). Besides the application of these 25DKGRs, the understanding of ketogluconate metabolism should lead to further development of an *E. coli* host strain that lacks the catabolic pathway of ketogluconates produced when a membrane-bound gluconate dehydrogenase (38), a 2KDG dehydrogenase, and a 25DKGR are expressed for the conversion (5, 6, 17) of glucose to 2KLG.

## MATERIALS AND METHODS

**Bacterial strains and media.** *E. coli* W3110 [ $F^-$  IN(*rrnD-rrnE*)], DH5 $\alpha$  [ $F^-$  *endA1 hsdR17(r<sub>K</sub> m<sub>K</sub>^+)* *supE44 thi-1 recA1 gyrA relA1 Δ(argF-lac)U169 deoR Φ80lacZΔM15*], and ER2566 [ $F^-$   $\lambda^-$  *fluA2 (lon) ompT lacZ::T7 gene 1 gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::mini-Tn10)2 R(zgb-210::Tn10)1 (Tet<sup>r</sup>) endA1 (dcm)*] (13) were used as host strains. All strains were grown in Luria-Bertani (LB) medium (1% Bacto Tryptone, 0.5% yeast extract, 0.5% NaCl) or M9 minimal medium (26) with 2KDG, 5KDG, 25DKG, or 2KLG as the carbon source. For antibiotic selection, the concentration of 100  $\mu$ g/ml (ampicillin) was used.

**Preparation of cell extracts.** *E. coli* cells were grown at 37°C for 18 h, harvested by centrifugation, and washed with 100 mM sodium phosphate buffer (pH 6.0). The cell paste was resuspended in the same buffer containing 0.1 mM phenylmethylsulfonyl fluoride (PMSF) to 1/10 of the original culture volume, and the cells were broken by sonication. The cell debris and unbroken cells were removed by centrifugation at 22,000  $\times g$  for 20 min.

**Enzyme assay and determination of D-glucose, D-gluconate, 2KDG, 5KDG, 2KLG, 25DKG, and IA.** The ketogluconate reductase activities were assayed as described previously (27) with 20 mM substrate and 0.1 mM NADH or NADPH. The reaction was monitored for the initial decrease in absorbance at 340 nm ( $\epsilon = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ ). One unit of activity corresponds to the production of 1  $\mu$ mol of NADP<sup>+</sup> or NAD<sup>+</sup> per min. The protein concentration of each sample was determined with the bicinchoninic acid protein assay kit (Pierce). Glucose was determined with a glucose analyzer (model 2300 STAT; Yellow Springs Instrument Co.). D-Gluconate, 2KDG, 5KDG, 2KLG, 25DKG, and IA in the reactions were determined by high-pressure liquid chromatography with an HPLC-87C column (Bio-Rad) at 30°C at a flow rate of 0.5 ml of 0.008 N H<sub>2</sub>SO<sub>4</sub> per min as the eluent.

**DNA preparation, manipulation, and sequence analysis.** Total DNA from *E. coli* was prepared by using Qiagen Genomic Tips. DNAs of the vector plasmids were prepared by a rapid alkaline lysis procedure (9), with a QIAprep Spin Miniprep kit (Qiagen). General DNA manipulations were carried out as described by Maniatis et al. (26). DNA sequencing of both strands was performed with an ABI 373 automated sequencer with dye-labelled terminators (Applied Biosystems Division of Perkin-Elmer). Oligonucleotides were synthesized by Bioneer (Chungwon, Korea). Sequence analysis, comparisons, and CLUSTAL alignments were performed, in part with LASERGENE (DNASTAR Inc., Madison, Wis.). Comparisons were also performed via BLAST and FASTA analysis of the SWISS-PROT and Protein Information Resource sequences. The PROSITE database was used for motif searches.

**PCR cloning of the *yqhE*, *yafB*, and *yjgU* genes.** Design of the primers was based on the published *yqhE*, *yafB*, and *yjgU* nucleotide sequences (GenBank accession no. AE000383, AE000129, and AE000497, respectively) (10). Two PCR primers used for *yqhE* were 5'-TGAACGCGTCTAGAACATCACTG-3' and 5'-CTTCCGGCTCTAGATGATGATGT-3', those used for *yafB* were 5'-TCGCGGGTCTAGACCCGTCCTG-3' and 5'-GTGTTTGTCCGCTCTAGATGATCGACA-3', and those used for *yjgU* were 5'-AATACCTGTCTAGAGGTCCTCGT-3' and 5'-ATACCATTGTCTAGAGTGCAGG-3' (the *Xba*I restriction site is underlined, and the point-mutated position is shown in boldface). The primers contained several point mutations introducing an *Xba*I site at the 5' and 3' ends. All PCRs were carried out in a GeneAmp PCR 2400 system (Perkin-Elmer) with 30 cycles of denaturation for 30 s at 95°C, annealing for 30 s at 65°C, and extension for 1 min at 72°C, followed by a 3-min extension period at 72°C. The resulting PCR products were cloned into plasmid pUC19 cleaved with *Xba*I to produce plasmids pYQHE, pYAFB, and pYJGU. The plasmids were sequenced to confirm that the sequences of the inserts were identical to those of the *yqhE*, *yafB*, and *yjgU* genes.

**Purification of 25DKGR-A.** 25DKGR-A was isolated by using the IMPACT (intein-mediated purification with an affinity chitin-binding tag) one-step protein purification system (New England Biolabs). The DNA fragment containing the coding region of the *yqhE* gene was synthesized by using two primers, 5'-AGG AACATATGGCTAATCAACCGTTATTAAG-3' and 5'-GAGAAGCTCTCCGAGCCCGGAAGTGGTCAGGATC-3' (the *Nde*I and *Sap*I restriction

sites are underlined, and the point-mutated position is shown in boldface). In the system, cloning into the *Sap*I site leaves no additional residues attached to the C terminus of the target protein. After being digested with *Nde*I and *Sap*I, the fragment was cloned into the vector pTYB1 (New England Biolabs), which contains an isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG)-inducible T7 promoter. The resulting plasmid, pTY-YQH, containing the *yqhE* gene fused to the intein gene at the 3' end, was transformed into *E. coli* ER2566. For the purification of the recombinant protein, the resulting *E. coli* strain, ER2566(pTY-YQH), was grown in LB medium to an optical density at 600 nm of 0.5. Expression of the proteins was induced by the addition of 1 mM IPTG for 4 h. The cells were then resuspended in cell lysis buffer (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 0.1 mM EDTA, and 0.1% Triton X-100) and disrupted by sonication. Cell debris was removed by centrifugation, and the supernatant was loaded onto the chitin columns. The columns were washed two times with cell lysis buffer. The native proteins were cleaved and eluted with cleavage buffer (20 mM Tris-HCl [pH 8.0], 50 mM NaCl, 0.1 mM EDTA, and 30 mM dithiothreitol).

**Purification of 25DKGR-B.** Overproduction of 25DKGR-B was achieved by constructing plasmid pKK-YAF. To construct pKK-YAF, the *yafB* gene was amplified by PCR with two primers, 5'-TAAAAGAGGGAATTCATGGCTATCCC-3' and 5'-GCTGTCAGAGAAGCTTAATCCCAT-3' (the *Eco*RI and *Hind*III restriction sites are underlined, and the point-mutated position is shown in boldface), introducing an *Eco*RI and a *Hind*III site to *yafB*. The resulting DNA fragment was ligated to the *Eco*RI and *Hind*III cloning sites of plasmid pKK223-3 (Pharmacia), yielding plasmid pKK-YAF. For protein purification, the *E. coli* DH5 $\alpha$ (pKK-YAF) cells were harvested after 4 h of induction with IPTG. The cells were harvested by centrifugation, washed two times with 20 mM Tris-HCl buffer (pH 8.3), resuspended in the same buffer with 0.1 mM phenylmethylsulfonyl fluoride and streptomycin (0.1 mg/ml), and disrupted with an ultrasonic oscillator. Cell debris was removed by centrifugation at 15,000  $\times g$  for 30 min. The resulting supernatant as a crude enzyme solution was used for the purification of the reductase. The dialyzed sample was put on a DEAE-Toyopearl column (2.5 by 4 cm) that had been equilibrated with 20 mM Tris-HCl buffer (pH 8.3). The column was washed with the same buffer, and proteins were eluted with a linear gradient of 0.1 to 0.5 M NaCl in the same buffer at a flow rate of 1.19 ml/min. Fractions that contained 25DKGR-B activity were pooled and dialyzed overnight against 20 mM Tris-HCl buffer (pH 7.0). The dialyzed sample was placed on a Blue-Sepharose CL-6B column (1.5 by 7 cm) equilibrated with 20 mM Tris-HCl buffer (pH 7.0). The column was washed with the same buffer, and proteins were eluted with a linear gradient of 0 to 0.3 M NaCl in the same buffer. The active fractions eluted from the affinity column were combined, concentrated with Centriprep 10 (Amicon) to 2.0 ml, and put on a Bio-Sil SEC-250 (Bio-Rad) gel filtration column (0.78 by 30 cm) equilibrated with 20 mM Tris-HCl buffer (pH 7.0) containing 0.15 M NaCl. The active fractions were pooled and stored at 4°C.

**Purification of N-terminal His<sub>6</sub>-tagged 5KDGR.** Overproduction of His<sub>6</sub>-tagged 5KDGR was achieved by constructing plasmid pQE-YJG. To construct pQE-YJG, the *yjgU* gene was amplified by PCR with two primers, 5'-GAATAAGGATCCGAACGATCTATTTCACTGGCAGGAA-3' and 5'-GTAGGGGGGAAGCTTAAACAGCCAC-3' (the *Bam*HI and *Hind*III restriction sites are underlined, and the point-mutated position is shown in boldface), introducing a *Bam*HI and a *Hind*III site to *yjgU*. The resulting DNA fragment was ligated to the *Bam*HI and *Hind*III cloning sites of plasmid pQE31 (Qiagen), yielding plasmid pQE-YJG. In this plasmid, *yjgU* is fused N-terminally in frame to the six-His tag-coding region of plasmid pQE31, leading to expression under the control of the IPTG-inducible promoter. Overproduction of His<sub>6</sub>-tagged 5KDGR was achieved in *E. coli* DH5 $\alpha$ (pQE-YJG) after IPTG (1 mM) induction in LB medium (containing 100  $\mu$ g of ampicillin per ml) for at least 1 h. For protein purification, the cells were harvested after 4 h of induction with IPTG. The His<sub>6</sub>-tagged fusion protein was purified from recombinant *E. coli* cells with Ni-nitrilotriacetic acid (NTA) resin (Qiagen). Centrifugations were carried out at 4°C, and column chromatographies were carried out at room temperature. For His<sub>6</sub>-tagged 5KDGR purification, the cell pellet from a 500-ml culture of *E. coli* DH5 $\alpha$ (pQE-YJG) was suspended in 20 ml of 50 mM sodium phosphate (pH 8.0)-0.3 M NaCl and sonicated on ice. The resulting cell lysate was centrifuged at 16,000  $\times g$ , and the supernatant was passed directly over a column containing 1.6 ml of Ni-NTA resin (Qiagen). After the column was washed with 30 ml of 50 mM sodium phosphate buffer (pH 8.0) containing 0.3 M NaCl-10% glycerol, the N-terminal His<sub>6</sub>-tagged 5KDGR was eluted with 4 ml of 50 mM Na-citrate buffer (pH 6.0). Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was done by Laemmli's method (23).

**Chemicals.** Calcium 25DKG and sodium 2KLG were obtained from Dong-A Pharmaceutical Company (Seoul, Korea). Calcium 2KDG and potassium 5KDG were purchased from Sigma Chemical Company.

## RESULTS AND DISCUSSION

**Identification of *yqhE* and *yafB* genes encoding 25DKGR-A and 25DKGR-B.** In the previous work, we identified a gene (*yiaE*) encoding 2KR catalyzing the reduction of 25DKG to 5KDG and of 2KLG to IA, as well as of 2KDG to D-gluconate

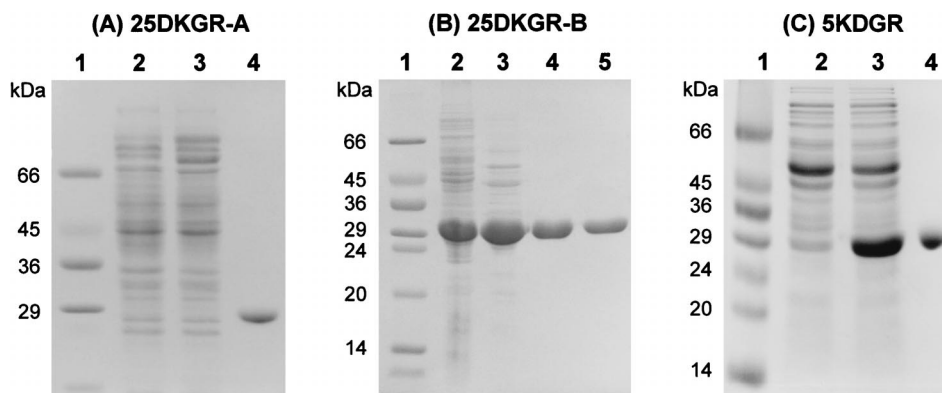


FIG. 1. SDS-PAGE monitoring of purification of 25DKGR-A (A), 25DKGR-B (B), and His<sub>6</sub>-tagged 5KDGR (C) from *E. coli* ER2566(pTY-YQH), DH5 $\alpha$ (pKK-YAF), and DH5 $\alpha$ (pQE-YJG), as described in Materials and Methods. (A) Lane 1, molecular mass markers (Sigma); lane 2, noninduced cells; lane 3, cells induced with IPTG; lane 4, eluate from chitin column. (B) Lane 1, molecular mass markers; lane 2, cells induced with IPTG; lanes 3 to 5, active fractions after DEAE-Toyopearl chromatography (lane 3), Blue-Sepharose CL-6B (lane 4), and Bio-Sil SEC-250 gel filtration (lane 5). (C) Lane 1, molecular mass markers; lane 2, noninduced cells; lane 3, cells induced with IPTG; lane 4, eluate from Ni-NTA column.

(37), which suggested strongly that the other ketogluconate reductases, 5KDGR and 25DKGR, and the related ketogluconate metabolism could also exist in *E. coli*. To find the possible ketogluconate reductases in *E. coli*, we searched homologous proteins in the protein databases by running FASTA searches with the amino acid sequences of 5KDGR (SP: P50199) of *G. oxydans* (21) and 25DKGR (PIR:I40838) of *Corynebacterium* sp. (6) as query sequences. As a result, our attention was drawn to two hypothetical oxidoreductases (SP: Q46857 and SP:P30863), encoded by *yqhE* and *yafB*, having 49.8 and 42% similarity, respectively, to 25DKGR and an oxidoreductase (SP:P39345), encoded by *yjgU*, showing 56% similarity to 5KDGR. Of the three genes, the *yjgU* gene was found to encode 5KDGR and was renamed *idnO* recently (8). To clarify the functions of the hypothetical oxidoreductases encoded by *yqhE* and *yafB*, we attempted to clone these putative genes. We amplified each of these genes by PCR under conditions that minimized errors and cloned them in the plasmid vector pUC19. The reductase activities in crude cell extracts of several different clones of each construct were determined by using 2KDG, 5KDG, 2KLG, or 25DKG as a substrate in the presence of NADPH or NADH. The presence of pYQHE and pYAFB containing *yqhE* and *yafB* genes, respectively, increased specific 25DKGR activity by about 10-fold, and the reaction products were identified as 2KLG. This result suggested that both *yqhE* and *yafB* genes encoded 25DKGRs, which were named 25DKGR-A and 25DKGR-B, respectively. As expected, the clone pYJGU containing the *idnO* gene also increased 5KDGR-specific activity by about 10-fold. The reductases in *E. coli* were purified further to confirm the substrate specificity.

**Purification of 25DKGR-A, 25DKGR-B, and His<sub>6</sub>-tagged 5KDGR and their substrate specificity.** 25DKGR-A was purified by using the IMPACT (intein-mediated purification with an affinity chitin-binding tag) one-step protein purification system. The N-terminal His<sub>6</sub>-tagged 5KDGR was isolated by using metal-chelate affinity chromatography on a Ni-NTA column instead of using the IMPACT system, since Leu at the C terminus of 5KDGR decreases *in vitro* cleavage with dithiothreitol at 4°C, resulting in a decrease in the yield of mature protein. Since the N-terminal or C-terminal His-tagged 25DKGR-B was catalytically inactive and the intein-mediated purification procedure was also inapplicable, 25DKGR-B was purified to homogeneity by column chromatographies on

DEAE-Toyopearl, Blue-Sepharose CL-6B, and Bio-Sil SEC-250 from the crude cell extracts of the *E. coli* clone. Upon SDS-PAGE, the purified proteins yielded a single band when stained with Coomassie blue (Fig. 1). 2KDG, 25DKG, 2KLG, 5KDG, D-fructose, and L-sorbose were examined as substrates. Consistent with earlier results with crude cell extracts, purified 25DKGR-A and 25DKGR-B were highly specific for 25DKG in the presence of NADPH, and 2KDG, 2KLG, 5KDG, D-fructose, and L-sorbose were inactive when assayed at pH 6.0 in the presence of NADH or NADPH. In the enzymatic reduction of 25DKG, the reaction products with purified 25DKGR-A and 25DKGR-B were identified as exclusively 2KLG. His<sub>6</sub>-tagged 5KDGR was also highly specific for 5KDG, and the reaction product was identified as D-gluconate. The other substrates, 2KDG, 25DKG, 2KLG, D-fructose, and L-sorbose, showed no reduction activity with NADH or NADPH. The His<sub>6</sub>-tagged 5KDGR was active with either NADPH (three-fold lesser extent than NADH) or NADH, whereas 25DKGR-A and 25DKGR-B were active with only NADPH as a preferred electron donor. This result indicates that 25DKGRs and 5KDGR of *E. coli* have similar substrate specificities with 25DKGRs of *Corynebacterium* sp. (27, 34) and 5KDGR of *G. oxydans* (1, 3, 21), respectively. The native 25DKGR-A, 25DKGR-B, and His<sub>6</sub>-tagged 5KDGR had apparent molecular weights of about 30,000, 30,000, and 54,000, respectively, as determined by molecular exclusion chromatography with Bio-Sil SEC-250 (Bio-Rad). In SDS-PAGE, all three enzymes showed protein bands with a molecular weight of about 29,000 (Fig. 1). The molecular weights (31,003, 29,401, and 27,562) of 25DKGR-A, 25DKGR-B, and 5KDGR, respectively, calculated by deduced amino acid sequences were in good agreement with the molecular masses obtained by SDS-PAGE. These data suggest that 25DKGR-A, 25DKGR-B, and His<sub>6</sub>-tagged 5KDGR are monomeric, monomeric, and dimeric proteins, respectively. The optimum pHs for reduction were 7.5, 7.0, and 8.0, respectively.

**Comparison of ketogluconate reductases.** The ketogluconate reductases for which the amino acid sequences are known were compared with those of *E. coli* including 2KR (37). The alignments of amino acid sequences are compared in Table 1. The reductases having the same substrate specificity showed relatively high homology. The PROSITE database was used for motif searches. As a result, the signatures were identified as the aldo- and ketoreductase family for 25DKGR-A

TABLE 1. Statistical comparison of ketogluconate reductases<sup>a</sup>

Reductase (sp.)	% Identity with reductase (sp.):						
	2KR ( <i>Erwinia herbicola</i> )	25DKGR-A ( <i>E. coli</i> )	25DKGR-A ( <i>Corynebacterium</i> sp.)	25DKGR-B ( <i>Corynebacterium</i> sp.)	25DKGR-B ( <i>E. coli</i> )	5KDGR ( <i>E. coli</i> )	5KDGR ( <i>G. oxydans</i> )
2KR ( <i>E. coli</i> )	69.1	10.9	11.2	10.5	12.7	11.8	12.9
2KR ( <i>Erwinia herbicola</i> )		12.0	11.6	12.7	12.0	12.2	13.7
25DKGR-A ( <i>E. coli</i> )			47.4	40.1	31.8	10.2	11.7
25DKGR-A ( <i>Corynebacterium</i> sp.)				37.7	37.1	12.1	12.9
25DKGR-B ( <i>Corynebacterium</i> sp.)					28.8	11.0	12.1
25DKGR-B ( <i>E. coli</i> )						12.6	11.3
5KDGR ( <i>E. coli</i> )							55.5

<sup>a</sup> Accession numbers are as follows: 2KR of *E. coli* (SP:P37666), 25DKGR-A and -B of *E. coli* (SP:Q46857 and P30863, respectively), 25DKGR-A and -B of *Corynebacterium* sp. (PIR: I40838 and A45961, respectively), 5KDGR of *E. coli* (GB:AAC77223.1), and 5KDGR of *G. oxydans* (SP:P50199). Values presented in the table represent percent identities.

and 25DKG-B, the short-chain dehydrogenase-reductase family for 5KDGR, and the D-isomer-specific 2-hydroxyacid dehydrogenase family for 2KR (37). The alignment indicates conserved motifs that may be involved in substrate binding or catalysis, including signature motifs. The aldo- and ketoreductase family has three consensus patterns located in the N-terminal, central, and C-terminal sections, and the ( $\alpha/\beta$ ) eight-barrel fold provides a common scaffold for NAD(P)(H)-dependent catalytic activity, with substrate specificity determined by variation of loops on the C-terminal side of the barrel (19).

#### Proposed pathways for ketogluconate metabolism in *E. coli*.

The pathways involved in the metabolism of ketogluconates by *Erwinia* sp. (35) and *Corynebacterium* sp. (32) have been investigated by use of a combination of enzyme assays and isolation of mutants. The presence of the ketogluconate reductase genes (*yiaE*, *yqhE*, *yafB*, and *idnO*) and the enzymatic activities described above suggest the pathway shown in Fig. 2 as the catabolic pathway for ketogluconates in *E. coli*. In *Erwinia* sp.,

25DKG is converted to 5KDG, but in *Corynebacterium* sp., 25DKG is converted to 2KDG before being converted to D-gluconate. Interestingly, in *E. coli*, 25DKG can be catabolized by sequential reductions to D-gluconate via 5KDG. Then the resulting D-gluconate from ketogluconates is metabolized via the Entner-Doudoroff pathway, with the pentose phosphate pathway playing a secondary role (16). The other possible catabolic pathway, of 25DKG to D-gluconate via 2KLG, IA, and 5KDG, needs the idonate dehydrogenase (5KR[I]) catalyzing the oxidation of IA to 5KDG, which is known in *Erwinia* and *Corynebacterium* spp. Recently, 5KDGR of *E. coli* was identified by sequence analysis of the GntII system for gluconate metabolism (8). The GntII system encodes a pathway for catabolism of IA, in which IA is converted to 5KDG by IA dehydrogenase before being reduced to D-gluconate by 5KDGR.

Even though the genes encoding 25DKGR-A and 25DKGR-B exist in *E. coli*, the W3110 strain showed poor growth on 2KDG (37), 5KDG, 25DKG, or 2KLG as the sole carbon source (data not shown), which may have been due to poor uptake of ketogluconates into the cell or redox imbalance. This assumption is consistent with experiments with acetic acid bacteria having NAD(P)H-dependent ketogluconate reductases, which showed slow growth on 2KDG or 5KDG (30). It has been suggested elsewhere that the ketogluconate reduction reactions in acetic acid bacteria may be important in maintaining redox balance rather than in providing a carbon source (35). *E. coli* grows well on IA, which is converted to D-gluconate by redox-coupled interconversion via intermediate 5KDG (8). The regulation of the sugar acid regulons allows *E. coli* to cometabolize the sugar acids, even in the presence of glucose (28). Therefore, the ketogluconate metabolism may provide *E. coli* with the metabolic advantage necessary for it to compete with other bacteria by maintaining redox balance. Because *E. coli* shows poor growth on 25DKG as the sole carbon source, cultivation in minimal medium containing both glucose (2 g/liter) and 25DKG (6 g/liter) was tried (Fig. 3). 25DKG was added into the medium after 4 h of cultivation with glucose. The level of 25DKG in the medium decreased after the exhaustion of glucose, and the cell optical density was reincreased after 18 h of cultivation. This observation indicates that the ketogluconate metabolism in *E. coli* may serve as a subsidiary metabolism in certain ecological environments.

The GntII system encoding a pathway for catabolism of IA in *E. coli* consists of an operon, which encodes IA dehydrogenase, 5KDGR, an IA transporter, and an IA regulatory protein, respectively, in which the IA regulatory protein acts as a positive regulator, with IA and 5KDG serving as the inducers

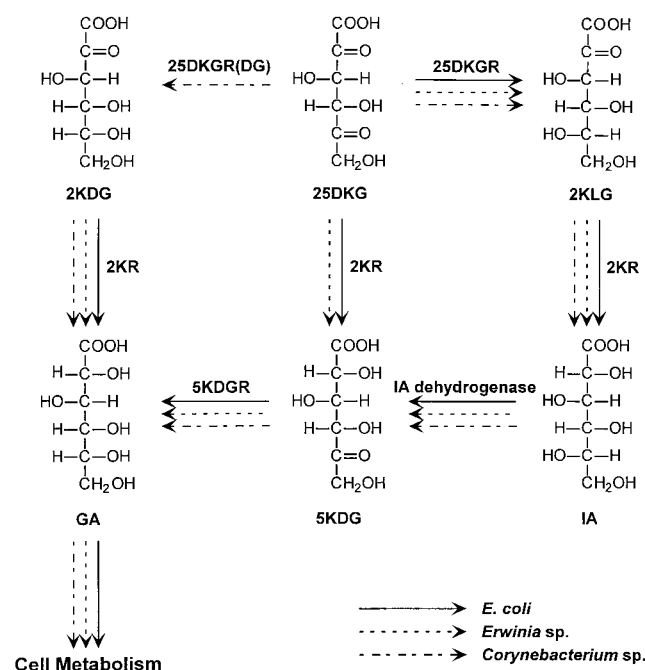


FIG. 2. Possible pathways for ketogluconate metabolism in *E. coli*. The pathways were compared with those of *Erwinia* sp. and *Corynebacterium* sp.

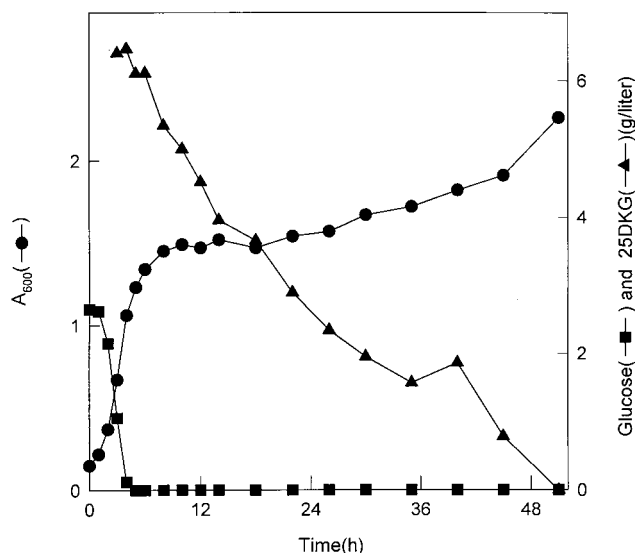


FIG. 3. Growth of *E. coli* in M9 minimal medium. *E. coli* W3110 was grown on M9 minimal medium containing glucose (2 g/liter). After 4 h of cultivation, 25DKG (6 g/liter) was added to the medium.

of the pathway (8). In the analysis of the genes in the vicinity of the genes encoding 2KR, 25DKGR-A, and 25DKGR-B, two genes, *yqhC* and *yafC*, encoding hypothetical transcriptional regulators adjacent to *yqhE* and *yafB* were found. Further, the *yqhD* gene encoding a hypothetical oxidoreductase was found to be located between the *yqhE* and *yqhC* genes. Two putative regulatory proteins may play a role in the regulation of *yqhE* and *yafB* gene expressions. Therefore, further studies should be focused on the identification of *yafC*, *yqhC*, and *yqhD* genes.

**BLAST search of incomplete microbial genomes.** To investigate the possibility that other bacteria could also have genes encoding 2KR, 5KDGR, 25DKGR-A, and 25DKGR-B, amino acid sequences encoded by *yiaE*, *idnO*, *yqhE*, and *yafB* genes were subjected to a homology search with the tBLASTn program of the National Center for Biotechnology Information database of 26 incomplete microbial genome sequences of *Actinobacillus actinomycetemcomitans*, *Bordetella pertussis*, *Borrelia burgdorferi*, *Campylobacter jejuni*, *Chlamydia trachomatis*, *Chlorobium tepidum*, *Clostridium acetobutylicum*, *Deinococcus radiodurans*, *Enterococcus faecalis*, *Helicobacter pylori*, *Methanobacterium thermoautotrophicum*, *Methanococcus jannaschii*, *Mycobacterium tuberculosis* CSU 93, *Mycobacterium tuberculosis* H37Rv, *Neisseria gonorrhoeae*, *Neisseria meningitidis* MC58, *Neisseria meningitidis* serogroup Ai, *Pseudomonas aeruginosa*, *Porphyromonas gingivalis* W83, *Pyrococcus furiosus*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Thermotoga maritima*, *Vibrio cholerae*, and *Yersinia pestis*. Even if these microbial genomes are not yet complete, several open reading frames (ORFs) encoding proteins showing high homologies with *E. coli* ketogluconate reductases were found in several strains. ORFs encoding proteins showing more than 30% identity with *E. coli* 2KR (8 ORFs), 5KDGR (14 ORFs), 25DKGR-A (14 ORFs), and 25DKGR-B (15 ORFs) were found. Especially in *Y. pestis* and *P. aeruginosa*, ORFs encoding putative proteins showing homologies with all four reductases were found.

The range of bacterial genome sequences available has grown rapidly and is likely to continue expanding. A search of published genome sequences, as well as partially completed

genome sequences, indicates that putative ketogluconate reductases are present in several of these organisms. This is not surprising given that ketogluconates can serve as the sole source of carbon and energy for various bacteria. It has also been reported elsewhere that a large number of enterobacterial strains have the glucose oxidation pathway (11) and the ability to use 2KDG as a source of carbon and energy in a defined medium (12). It has also been postulated elsewhere that sugar acid metabolism is an important aspect of the ecology of *E. coli* (28). Although generally considered to be restricted to a few bacteria, the ketogluconate metabolism may be widespread in many species. Until now, complete genome sequences were available for only 16 microorganisms. More than 50 additional microbial genomes are scheduled to be completely sequenced by the year 2000 (14). The availability of the complete genome sequence should facilitate identification of how the metabolism is distributed in microorganisms. It may be possible that exploring the ketogluconate metabolism could provide a useful tool for classification and identification of microbial strains.

**Application.** Current work concerns the study of the genes encoding enzymes involved in ketogluconate metabolism in *E. coli* by using a direct strategy for finding and characterizing unidentified genes from the *E. coli* genomic database. This finding should lead to further development of an *E. coli* host strain that lacks the catabolic pathway of ketogluconates produced when a membrane-bound gluconate dehydrogenase (38), 2KDG dehydrogenase, and 25DKGR are expressed for the conversion of glucose to 2KLG, a key intermediate in the biosynthesis of ascorbic acid (vitamin C). A process for the bioconversion of an ordinary microbial metabolite such as glucose into 2KLG with a single recombinant strain by combining the relevant traits of both *Erwinia* sp. and *Corynebacterium* sp. into a single organism has been tried (5, 6, 17). However, even when a 2KR gene of *Erwinia* sp. was disrupted, the production of an undesirable by-product, IA, could not be prevented, which was found to be due to the unexpected additional pathway of metabolism of ketogluconates by another 2KR (5). *E. coli* could be useful as a recombinant host strain for the production of 2KLG, because it does not grow well on ketogluconates. If needed, an *E. coli* mutant deficient in 2KR activity (37) could be used as a host strain.

25DKGR is an important enzyme in the bioconversion of 2KLG by the glucose pathway since the rate-limiting step may be the reduction of 25DKG to 2KLG by 25DKGR. Until now, the 25DKGR gene from *Corynebacterium* sp. has been used, and several studies have been carried out to increase the temperature stability of the enzyme (24, 29). Recently, the three-dimensional structure of 25DKGR was also determined to aid the elucidation of the structural determinants of specificity, catalysis, and stability for the enzyme (20). Though we do not know yet how stable and efficient two *E. coli* 25DKGRs are in the application and it is possible that their thermal stability might not prove to be satisfactory, two enzymes could be useful in enzyme improvement, considering that directed evolution with family shuffling of enzymes has been shown to be an effective method in recent studies (7, 15, 22, 25).

#### ACKNOWLEDGMENTS

We thank E. S. Choi for critical reading of the manuscript.

This investigation was supported by grant HS1810 from the Ministry of Science and Technology of Korea (MOST) and Korea Microbial Technology Inc. (KOMITECH).

#### REFERENCES

- Adachi, O., E. Shinagawa, K. Matsushita, and M. Ameyama. 1979. Crystallization and properties of 5-keto-D-gluconate reductase from *Gluconobacter*

- suboxydans*. Agric. Biol. Chem. **43**:75–83.
2. Ameyama, M., and O. Adachi. 1982. 2-Keto-D-gluconate reductase from acetic acid bacteria. Methods Enzymol. **89**:203–209.
  3. Ameyama, M., and O. Adachi. 1982. 5-Keto-D-gluconate reductase from *Gluconobacter suboxydans*. Methods Enzymol. **89**:198–202.
  4. Ameyama, M., K. Matsushita, E. Shinagawa, and O. Adachi. 1987. Sugar-oxidizing respiratory chain of *Gluconobacter suboxydans*. Evidence for a branched respiratory chain and characterization of respiratory chain-linked cytochromes. Agric. Biol. Chem. **51**:2943–2950.
  5. Anderson, S., R. A. Lazarus, H. I. Miller, and R. K. Stafford. July 1991. U.S. patent 5,032,514.
  6. Anderson, S., C. B. Marks, R. Lazarus, J. Miller, K. Stafford, J. Seymour, D. Light, W. Rastetter, and D. Estell. 1985. Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified *Erwinia herbicola*. Science **230**:144–149.
  7. Arnold, F. H., and J. C. Moore. 1997. Optimizing industrial enzymes by directed evolution. Adv. Biochem. Eng. Biotechnol. **58**:1–14.
  8. Bausch, C., N. Peekhaus, C. Utz, T. Blais, E. Murray, T. Lowary, and T. Conway. 1998. Sequence analysis of the GntII (subsidiary) system for gluconate metabolism reveals a novel pathway for L-idenic acid catabolism in *Escherichia coli*. J. Bacteriol. **180**:3704–3710.
  9. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
  10. Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277**:1453–1462.
  11. Bouver, O. M., P. Lenormand, and P. A. Grimont. 1989. Taxonomic diversity of the D-glucose oxidation pathway in the *Enterobacteriaceae*. Intl. J. Syst. Bacteriol. **39**:61–67.
  12. Buisiere, J., G. Brault, and L. Le Minor. 1981. Utilization and fermentation of 2-ketogluconate by *Enterobacteriaceae*. Ann. Microbiol. (Paris) **132**:191–195.
  13. Chong, S., F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus, and M. Q. Xu. 1997. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene **192**:271–281.
  14. Clayton, R. A., O. White, and C. M. Fraser. 1998. Findings emerging from complete microbial genome sequences. Curr. Opin. Microbiol. **1**:562–571.
  15. Crameri, A., S. A. Raillard, E. Bermudez, and W. P. Stemmer. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature **391**:288–291.
  16. Fraenkel, D. G., and S. R. Levisohn. 1967. Glucose and gluconate metabolism in an *Escherichia coli* mutant lacking phosphoglucose isomerase. J. Bacteriol. **93**:1571–1578.
  17. Grindley, J. F., M. A. Payton, H. van de Pol, and K. G. Hardy. 1988. Conversion of glucose to 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a recombinant strain of *Erwinia citreus*. Appl. Environ. Microbiol. **54**:1770–1775.
  18. Hardy, K., H. van de Pol, J. Grindley, and M. A. Payton. July 1990. U.S. patent 4,945,052.
  19. Jez, J. M., M. J. Bennett, B. P. Schlegel, M. Lewis, and T. M. Penning. 1997. Comparative anatomy of the aldo-keto reductase superfamily. Biochem. J. **326**:625–636.
  20. Khurana, S., D. B. Powers, S. Anderson, and M. Blaber. 1998. Crystal structure of 2,5-diketo-D-gluconic acid reductase A complexed with NADPH at 2.1-Å resolution. Proc. Natl. Acad. Sci. USA **95**:6768–6773.
  21. Klasen, R., S. Bringer Meyer, and H. Sahn. 1995. Biochemical characterization and sequence analysis of the gluconate:NADP 5-oxidoreductase gene from *Gluconobacter oxydans*. J. Bacteriol. **177**:2637–2643.
  22. Kuchner, O., and F. H. Arnold. 1997. Directed evolution of enzyme catalysts. Trends Biotechnol. **15**:523–530.
  23. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
  24. Lazarus, R. A., M. Hurle, S. Anderson, and D. B. Powers. September 1992. U.S. patent 5,376,544.
  25. MacBeath, G., P. Kast, and D. Hilvert. 1998. Redesigning enzyme topology by directed evolution. Science **279**:1958–1961.
  26. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  27. Miller, J. V., D. A. Estell, and R. A. Lazarus. 1987. Purification and characterization of 2,5-diketo-D-gluconate reductase from *Corynebacterium* sp. J. Biol. Chem. **262**:9016–9020.
  28. Peekhaus, N., and T. Conway. 1998. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J. Bacteriol. **180**:3495–3502.
  29. Powers, D. B., and S. Anderson. July 1997. Patent WO 97/25432.
  30. Shinagawa, E., T. Chiyonobu, K. Matsushita, O. Adachi, and M. Ameyama. 1978. Distribution of gluconate dehydrogenase and ketogluconate reductases in aerobic bacteria. Agric. Biol. Chem. **42**:1055–1057.
  31. Sonoyama, T., B. Kageyama, and S. Yagi. 1987. Distribution of microorganisms capable of reducing 2,5-diketo-D-gluconate to 2-keto-L-gulonate. Agric. Biol. Chem. **51**:2003–2004.
  32. Sonoyama, T., B. Kageyama, S. Yagi, and K. Mitsushima. 1987. Biochemical aspects of 2-keto-L-gulonate accumulation from 2,5-diketo-D-gluconate by *Corynebacterium* sp. and its mutants. Agric. Biol. Chem. **51**:3039–3047.
  33. Sonoyama, T., B. Kageyama, S. Yagi, and K. Mitsushima. 1988. Facultatively anaerobic bacteria showing high productivities of 2,5-diketo-D-gluconate from D-glucose. Agric. Biol. Chem. **52**:667–674.
  34. Sonoyama, T., and K. Kobayashi. 1987. Purification and properties of two 2,5-diketo-D-gluconate-reductases from a mutant strain derived from *Corynebacterium* sp. J. Ferment. Technol. **65**:311–317.
  35. Truesdell, S. J., J. C. Sims, P. A. Boerman, J. L. Seymour, and R. A. Lazarus. 1991. Pathways for metabolism of ketoaldonic acids in an *Erwinia* sp. J. Bacteriol. **173**:6651–6656.
  36. Yum, D.-Y., S.-S. Bae, and J.-G. Pan. 1998. Purification and characterization of the 2-ketoaldonate reductase from *Brevibacterium ketosoreductum* ATCC 21914. Biosci. Biotechnol. Biochem. **62**:154–156.
  37. Yum, D.-Y., B.-Y. Lee, D.-H. Hahm, and J.-G. Pan. 1998. The *yiaE* gene, located at 80.1 minutes on the *Escherichia coli* chromosome, encodes a 2-ketoaldonate reductase. J. Bacteriol. **180**:5984–5988.
  38. Yum, D.-Y., Y.-P. Lee, and J.-G. Pan. 1997. Cloning and expression of a gene cluster encoding three subunits of membrane-bound gluconate dehydrogenase from *Erwinia cypripedii* ATCC 29267 in *Escherichia coli*. J. Bacteriol. **179**:6566–6572.