Identification and Characterization of L-Arabonate Dehydratase, L-2-Keto-3-deoxyarabonate Dehydratase, and L-Arabinolactonase Involved in an Alternative Pathway of L-Arabinose Metabolism

NOVEL EVOLUTIONARY INSIGHT INTO SUGAR METABOLISM*

Seiya Watanabe,†§, Naoko Shimada,§ Kunihiko Tajima,§ Tsutomu Kodaki,§ and Keisuke Makino†¶

From the †Faculty of Engineering, Kyoto University, Kyoto-daigaku-katsura, Saiya-ku, Kyoto 615-8530, Japan, the §Institute of Advanced Energy, Kyoto University, Gakoyo, Uji, Kyoto 611-0011, Japan, the ¶Department of Biomolecular Engineering, Kyoto Institute of Technology, Matsugasaki, Saiyo-ku, Kyoto 606-8585, Japan, and the §CREST, JST (Japan Science and Technology Agency), Gakoyo, Uji, Kyoto 611-0011, Japan

Azospirillum brasilense possesses an alternative pathway of L-arabinose metabolism, different from the known bacterial and fungal pathways. In the preceding articles, we identified and characterized L-arabinose-1-dehydrogenase and α-ketogluutaric semialdehyde dehydrogenase, which catalyzes the first and final reaction steps in this pathway, respectively (Watanabe, S., Kodaki, T., and Makino, K. (2006) J. Biol. Chem. 281, 28876–28888). We here report the remaining three enzymes, L-arabonate dehydratase, L-2-keto-3-deoxyarabonate (L-KDA) dehydratase, and L-arabinolactonase. N-terminal amino acid sequences of L-arabonate dehydratase and L-KDA dehydratase purified from A. brasilense cells corresponded to those of AraC and AraD genes, which form a single transcriptional unit together with the L-arabinose-1-dehydrogenase gene. Furthermore, the L-arabinolactonase gene (AraB) was also identified as a component of the gene cluster. Genetic characterization of the alternative L-arabinose pathway suggested a significant evolutionary relationship with the known sugar metabolic pathways, including the Entner-Doudoroff (ED) pathway and the several modified versions. L-arabonate dehydratase belongs to the ILVD/EDD family and spectrophotometric and electron paramagnetic resonance analysis revealed it to contain a [4Fe-4S]2+ cluster. Site-directed mutagenesis identified three cysteine ligands essential for cluster coordination. L-KDA dehydratase was sequentially similar to DHDPS/NAL family proteins. d-2-Keto-3-deoxygluconate aldolase, a member of the DHDPS/NAL family, catalyzes the equivalent reaction to L-KDA aldolase involved in another alternative L-arabinose pathway, probably associating a unique evolutionary event between the two alternative L-arabinose pathways by mutation(s) of a common ancestral enzyme. Site-directed mutagenesis revealed a unique catalytic amino acid residue in L-KDA dehydratase, which may be a candidate for such a natural mutation.

Many bacteria are able to grow with L-arabinose as the sole carbon and energy source, and the bacterial pathway of L-arabinose metabolism has been extensively investigated. Many bacteria including Escherichia coli depend on protein products of the araBAD operon, which contains araB (ribulokinase, EC 2.7.1.16), araA (L-arabinose isomerase, EC 5.3.1.4) and araD (L-ribulose-phosphate 4-epimerase, EC 5.1.3.4) to convert L-arabinose to D-xylulose 5-phosphate through L-ribulose and L-ribulose 5-phosphate (1). In a recently characterized fungal pathway (2–4), L-arabinose is also converted to D-xylulose 5-phosphate but through different intermediates by two reductases, two dehydrogenases, and a kinase. On the other hand, it is believed that a hypothetical pathway of L-arabinose metabolism is operative in some bacteria (5–13) (Fig. 1A). In this pathway, L-arabinose is oxidized to L-arabinon-γ-lactone by NAD(P)⁺-dependent dehydrogenase. The lactone is cleaved by a lactonase to L-arabonate, followed by two successive dehydrogenation reactions forming L-2-keto-3-deoxyarabonate (L-KDA)2 and α-ketogluutaric semialdehyde (αKGSα). The last step is the NAD(P)⁺-dependent dehydrogenation of αKGSα to α-ketoglutaric acid (referred to as the first pathway). Alternatively, L-KDA is cleaved through an aldolase reaction to glycolaldehyde and pyruvate (referred to as the second pathway). It is note-

* This work was supported by the Center of Excellence (COE) program for the “Establishment of COE on Sustainable Energy System”, a Grant-in-aid for Scientific Research, and Grants for Regional Science and Technology Promotion “Kyoto Nanotechnology Cluster” project from the Ministry of Education, Science, Sports and Culture, Japan. This work was also supported by CREST and “Research for Promoting Technological Seeds” of the Japan Science and Technology Agency, CREST, JST (Japan Science and Technology Agency), Gakoyo, Uji, Kyoto 611-0011 Japan. Tel.: 81-774-38-3524; E-mail: kmak@iae.kyoto-u.ac.jp.

† To whom correspondence should be addressed: Inst. of Advanced Energy, Kyoto University, Gakoyo, Uji, Kyoto 611-0011 Japan. Tel.: 81-774-38-3517; Fax: 81-774-38-3524; E-mail: kmak@iae.kyoto-u.ac.jp.

‡ Faculty of Engineering, Kyoto University, Kyotodaigaku-katsura, Saikyo-ku, Kyoto 615-8530, Japan, the ¶Department of Biomolecular Engineering, Kyoto Institute of Technology, Matsugasaki, Saiyo-ku, Kyoto 606-8585, Japan, and the §CREST, JST (Japan Science and Technology Agency), Gakoyo, Uji, Kyoto 611-0011, Japan.

§¶ This work was supported by the Center of Excellence (COE) program for the “Establishment of COE on Sustainable Energy System”, a Grant-in-aid for Scientific Research, and Grants for Regional Science and Technology Promotion “Kyoto Nanotechnology Cluster” project from the Ministry of Education, Science, Sports and Culture, Japan. This work was also supported by CREST and “Research for Promoting Technological Seeds” of the Japan Science and Technology Agency, CREST, JST (Japan Science and Technology Agency), Gakoyo, Uji, Kyoto 611-0011 Japan. Tel.: 81-774-38-3524; E-mail: kmak@iae.kyoto-u.ac.jp.

The nucleotide sequence(s) reported in this paper has been submitted to the DDBJ/GenBank(TM)/EBI Data Bank with accession number(s) AB241136.

1 To whom correspondence should be addressed: Inst. of Advanced Energy, Kyoto University, Gakoyo, Uji, Kyoto 611-0011 Japan. Tel.: 81-774-38-3517; Fax: 81-774-38-3524; E-mail: kmak@iae.kyoto-u.ac.jp.

2 The abbreviations used are: L-KDA, L-2-keto-3-deoxyarabonate; αKGSα, α-ketoglutaric semialdehyde; ED pathway, Entner-Doudoroff pathway; EDD, 6-phosphogluconate dehydrogenase; KDPG, D-2-keto-3-deoxypyruvate; GDH, glucose dehydrogenase; KDG, D-2-keto-3-deoxyglyconate; DD pathway, De Ley-Doudoroff pathway; HPLC, high performance liquid chromatography; M, molecular mass; ILVD, dihydroxyacid dehydratase; WT, wild-type enzyme; EPR, electron paramagnetic resonance; DHDPS, dihydrodipicolinate synthase; NAL, N-acetylneuraminate lyase; ORF, open reading frame.
A Novel Bacterial L-Arabinose Pathway

A. brasiliense (first pathway). L-KDA is also converted to pyruvate and glycolaldehyde in several bacteria (second pathway). Bold enzymes are the focus of this study. Comparison between the alternative L-arabinose pathway (red), non- and semi-phosphorylative ED pathways (orange), D-gluconate pathway (blue), ED pathway (light-green), DD pathway (violet), D-galactarate/D-glucarate pathway (gray). In these pathways, sugar is metabolized through the participation of dehydrogenase, lactonase, dehydratase, and/or aldolase. In this study, we show that enzymes indicated with a bold EC number are related evolutionally.

**FIGURE 1.** A, alternative pathway proposed by Novick and Tyler (12) for *A. brasiliense* (first pathway). L-KDA is also converted to pyruvate and glycolaldehyde in several bacteria (second pathway). Bold enzymes are the focus of this study. B, comparison between the alternative L-arabinose pathway (red), non- and semi-phosphorylative ED pathways (orange), D-gluconate pathway (blue), ED pathway (light-green), DD pathway (violet), D-galactarate/D-glucarate pathway (gray). In these pathways, sugar is metabolized through the participation of dehydrogenase, lactonase, dehydratase, and/or aldolase. In this study, we show that enzymes indicated with a bold EC number are related evolutionally.
worthy that the alternative pathway of bacterial L-arabinose metabolism (red-colored route in Fig. 1B), in particular the second pathway, seems equivalent to the Entner-Doudoroff (ED) pathway (14).

Concerning the metabolism of sugars related to L-arabinose, the ED pathway is a general route for D-glucose catabolism in bacteria together with glycolysis (green-colored route in Fig. 1B). This pathway involves: (i) the initial phosphorylation of glucose to glucose 6-phosphate either by a glucokinase or by the action of a phosphoenolpyruvate-dependent phosphotransferase system, (ii) the oxidation to 6-phosphogluconate by glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and 6-phosphogluconolactonase (EC 3.1.1.31), (iii) dehydration to D-2-keto-3-deoxygluconate (KDG) by 6-phosphogluconate dehydrogenase (EDD, EC 4.2.1.12), and (iv) cleavage of the characteristic KDPG intermediate by KDPG aldolase (EC 4.1.2.14), yielding glyceraldehyde 3-phosphate and pyruvate. Glyceraldehyde 3-phosphate is further metabolized via the glycolytic pathway, yielding a second molecule of pyruvate. Although the ED pathway seems to be restricted to bacteria, modifications have been identified in all three kingdoms of life: Eukarya, Eubacteria, and Archaea (15). The “non-phosphorylative ED pathway” (orange-colored route in Fig. 1B) is a modified version of the ED pathway and consists of (i) the oxidation of glucose to gluconate via glucose dehydrogenase (GDH), (ii) the conversion of gluconate by D-gluconate dehydrogenase (EC 4.2.1.39) into D-2-keto-3-deoxygluconate (KDG), (iii) the aldol cleavage of KDG by KDG aldolase (EC 4.1.2.20), forming glyceraldehyde and pyruvate. This pathway is operative in (hyper)thermophilic archaea (16) and a fungi of Aspergillus niger (17). Recently, a set of metabolic genes was identified in archaean Sulfolobus solfataricus (18–20), revealing no evolutionary relationship between the ED pathway and non-phosphorylative ED pathway despite similar schematic sugar conversions (16). Another ED modification is the so-called “semi-phosphorylative ED pathway” (21), in which KDG is subsequently phosphorylated by KDG kinase to form KDPG. Oxidation of D-glucose to D-gluconate is also the initial step in the “D-gluconate pathway” of bacteria (22), yeast (23) and A. niger (24) (blue-colored route in Fig. 1B). In contrast with the semi-phosphorylative ED modification, D-gluconate is phosphorylated by gluconokinase, and the produced 6-phosphogluconate enters the ED pathway and/or pentose-phosphate pathway. Similarly, an alternative pathway different from the well-known “Leloir pathway” for D-galactose metabolism (25) is known in some bacteria. In the so-called “De Ley-Doudoroff (DD) pathway” (26), D-galactose is metabolized through analogous conversion to the semi-phosphorylative ED pathway, to produce the same end products as the ED pathway (violet-colored route in Fig. 1B). Analogous to the non-phosphorylative ED pathway, the DD pathway involves no phosphorylation in Aspergillus fungi (27). Few metabolic genes involved in the DD pathway have been identified.

Because no gene encoding the enzyme involved in the alternative pathways of l-arabinose metabolism has been identified, we are in the process of enzymatically and genetically characterizing the alternative l-arabinose pathway of Azospirillum brasilense, in which the first pathway is operative (12), and recently identified l-arabinose 1-dehydrogenase (EC 1.1.1.46) and αKGSA dehydrogenase (EC 1.2.1.26) catalyzing the first and final-step reactions, respectively (28, 29). In this study, we focused on the remaining enzymes, l-arabinolactonase (EC 3.1.1.15), l-arabonate dehydrogenase (EC 4.2.1.25) and l-KDA dehydrogenase (EC 4.2.1.43). Interestingly, these metabolic enzymes involved in the alternative l-arabinose pathway are evolutionally related to equivalent enzymes involved in the ED pathway, non-phosphorylative ED pathway, D-gluconate pathway, or DD pathway. Furthermore, unique catalytic insights into l-arabonate dehydrogenase and l-KDA dehydrogenase are discussed.

**EXPERIMENTAL PROCEDURES**

**General Procedures**—Basic recombinant DNA techniques were performed as described by Sambrook et al. (30). PCR was carried out using a PCR Thermal Cycler PERSONAL (TaKaRa) for 30 cycles in a 50-μl reaction mixture containing 1.25 units of Ex Taq DNA polymerase (TaKaRa), appropriate primers (10 pmol) and template DNA under the following conditions: denaturation at 98 °C for 10 s, annealing at 50 °C for 30 s and extension at 72 °C for time periods calculated as an extension rate of 1 kbp/min. DNA sequencing was carried out using the Dual CyDye™ Terminator Sequencing Kit (Veritas) and appropriate primers with Long-Read Tower, UBC DNA sequencer (Amersham Biosciences). Protein concentrations were determined by the method of Lowry et al. (31) with bovine serum albumin as the standard. SDS-PAGE was performed as described by Laemmli (32). Non-denaturing PAGE was performed by omitting SDS and 2-mercaptopethanol from the solution used in SDS-PAGE. High performance liquid chromatography (HPLC) was performed using a Multi-Station LC-8020 model II system (TOSOH). Samples were applied at 35 °C to an Aminex HPX-87H Organic Analysis column (300 × 7.8 mm, Bio-Rad) linked to an RID-8020 refractive index detector (TOSOH) and eluted with 5 mM H2SO4 at a flow rate of 0.6 ml/min. If necessary, 12% (w/v) trichloroacetic acid was added to samples (0.1 volume) to remove proteins. After filtration, 100 μl of this solution was then analyzed.

**Assay of l-Arabonate Dehydrogenase**—L-Arabonate (K⁺ salt) was prepared by the hypiodite-in-methanol oxidation of L-arabinose (33). l-Arabonate dehydrogenase activity was determined chemically using the semicarbazide method (34). Briefly, the total reaction mixture of 400 μl was incubated at 30 °C in 50 mM HEPES, pH 7.2, containing 10 mM MgCl₂ and 10 mM L-arabonate and a small volume of L-arabinose dehydrogenase. After 10 min, the enzyme reaction was stopped by adding 100 μl of 2 M HCl. To this solution, 300 μl of semicarbazide solution consisting of 1.0% (w/v) semicarbazide hydrochloride and 1.5% (w/v) sodium acetate was then added, and the mixture was incubated at 30 °C for 10 min. Finally, the reaction mixture was diluted with 500 μl of distilled water and its absorbance was measured at 250 nm (molar absorption coefficient ε = 571 M⁻¹·cm⁻¹). D-Arabonate, D-xylonate, D-gluconate, and D-galactonate were prepared from the corresponding sugar as a K⁺ salt by the same procedure as L-arabonate.
A Novel Bacterial L-Arabinose Pathway

Assay of L-KDA Dehydratase—L-KDA dehydratase was assayed spectrophotometrically in the coupling system with L-arabonate dehydratase and αKGSA dehydrogenase. The recombinant αKGSA dehydrogenase of A. brasilienensis was prepared by the method described previously (29) (see lane 5 in Fig. 2C). The assay was performed in 50 mM HEPES (pH 7.2) buffer containing 10 mM L-arabonate, 1.5 mM NADH. After the addition of each 1 unit of purified L-arabonate dehydratase and αKGSA dehydrogenase, the mixture was incubated for 10 min at 25 °C. No change in absorbance at 340 nm was found in this stage. The reaction was started by the addition of a small amount of L-KDA dehydratase. Recombinant L-arabinose dehydratase mutants. Lane 1, native enzyme purified from A. brasilienensis cells; lane 2, wild-type; lane 3, C47A; lane 4, C56A; lane 5, C124A; lane 6, C197A; lane 7, C441A. 5 μg of purified protein was applied on 12% (w/v) gel.

Assay of L-Arabinolactonase—L-Arabinolactonase was identified as a continuous reaction with L-arabinose 1-dehydrogenase. Recombinant L-arabinose 1-dehydrogenase was prepared by the method described previously (see lane 1 in Fig. 2C) (28). The reaction mixture contained 10 mM L-arabinose, 1 mM MgCl₂, 1 unit of purified L-arabinose 1-dehydrogenase, an appropriate amount of L-arabinolactonase, and 1 mM NADH in 20 mM potassium phosphate (pH 7.2). The reaction was performed at 30 °C for the appropriate time and stopped by adding 12% (v/v) trichloroacetic acid. After filtration, the amount of L-arabonate formed in the reaction mixture was determined by HPLC. Pure L-arabinolactone was chemically synthesized by boiling L-arabonate (K⁺ salt) in 0.2 M HCl for 5 min (34).

Purification of L-Arabinonate Dehydratase from A. brasilienensis—A. brasilienensis ATCC 29145 was cultivated aerobically for 48 h at 30 °C in a synthetic medium containing 37 mM L-arabinose as described previously (28). The grown cells were harvested by centrifugation at 30,000 × g for 20 min, washed with 50 mM Bisic shield buffer, pH 7.4, containing 10 mM MgCl₂ (referred to as Buffer A) and stored at −35 °C until use. The washed cells were suspended in Buffer A, disrupted by sonication for 20 min with appropriate intervals on ice using ASTRASON® Ultrasonic Liquid Processor XL2020 (Misonix Inc.) and then centrifuged at 108,000 × g for 20 min at 4 °C to obtain cell-free extracts.

All purification steps were performed below 4 °C. The cell-free extracts were fractionated between 50–60% saturation of (NH₄)₂SO₄. The precipitate was dissolved in a small volume of Buffer A, and the solution was then dialyzed overnight against a large volume of Buffer A containing 1.3 M (NH₄)₂SO₄. All chromatography was carried out using an AKTA purifier system (Amersham Biosciences). After insoluble materials were removed by centrifugation, the supernatant was applied to a HiPrep 16/10 Butyl FF column (1.6 × 10 cm, Amersham Biosciences) equilibrated with buffer A containing 1.3 M (NH₄)₂SO₄ and washed with the same buffer. Proteins were eluted using a reversed linear gradient of 1.3–0 M (NH₄)₂SO₄ in Buffer A (300 ml). Fractions with high enzymatic activity were pooled and dialyzed overnight against a large volume of Buffer A. The enzyme solution was loaded onto a column of HiPrep 16/10 DEAE FF (1.6 × 10 cm, Amersham Biosciences) equilibrated with Buffer A, and washed thoroughly with the same buffer. The column was developed with 200 ml of a linear gradient of 0–1 M NaCl in Buffer A. Fractions containing L-arabinonate dehydratase activity were combined and dialyzed overnight against a large volume of 10 mM potassium phosphate, pH 7.4, containing 10 mM MgCl₂ (referred to as Buffer B). The enzyme solution was applied to a column of CHT Ceramic Hydroxyapatite Type I (1.6 × 5 cm, Bio-Rad) equilibrated with 10 mM potassium phosphate, pH 7.4, containing 2 mM MgCl₂ (referred to as Buffer C).
The column was washed thoroughly with the same buffer and developed with 200 ml of a linear gradient of 0.01—0.3 M potassium phosphate buffer in Buffer C. Fractions with high enzymatic activity were combined and concentrated by ultrafiltration with Centriflo YM-30 (Millipore) at 18,000 × g for ~2 h. The enzyme solution was loaded onto a HiLoad 16/60 Superdex 200 pg (1.6 × 60 cm, Amersham Biosciences) column equilibrated with Buffer A. Highly active fractions were pooled, concentrated and dialyzed against Buffer A containing 50% (v/v) glycerol, and stored at −35 °C until use (referred to as Sample I).

**Purification of L-KDA Dehydratase from A. brasilense—** Procedures for the preparation of cell-free extracts and (NH₄)₂SO₄ fractionation were similar to those described in the purification of L-arabonate dehydratase except that 20 mM potassium phosphate buffer, pH 7.2, (referred to as Buffer D) was used instead of Buffer A. Subsequently, Butyl FF chromatography was performed using a reversed linear gradient of 1.3–0 M (NH₄)₂SO₄ in Buffer D. Fractions containing L-KDA dehydratase activity were combined and dialyzed overnight against Buffer D. The dialyzed enzyme preparation was applied to a column of DEAE FF equilibrated with Buffer D and washed thoroughly with the same buffer. The column was developed with 200 ml of a linear gradient 0–0.5 M NaCl in Buffer D. The active fractions were combined and dialyzed overnight against 10 mM potassium phosphate buffer, pH 7.2 (referred to as Buffer E). The enzyme solution was applied to a CHT Ceramic Hydroxyapatite Type I column equilibrated with Buffer E and washed thoroughly with the same buffer. Proteins were eluted using a linear gradient of 0.01–0.3 M potassium phosphate in Buffer E (200 ml). Fractions with high enzymatic activity were combined, concentrated and loaded onto a HiLoad 16/60 Superdex 200 pg column equilibrated with Buffer D. Proteins in fractions containing high L-KDA dehydratase activity were analyzed by SDS-PAGE, and fractions containing a single protein were collected, concentrated, dialyzed against Buffer D containing 50% (v/v) glycerol, and stored at −35 °C until use.

**Determination of N-terminal Sequence of Native Enzyme—** As there was a little contamination in Sample I of the preparation of L-arabonate dehydratase, in-gel staining assay (35) was carried out to identify a polypeptide corresponding to this enzyme. Sample I was run on native-PAGE with 6% (w/v) gel that was then soaked in 50 mM HEPES, pH 7.2, containing 10 mM MgCl₂ and 10 mM L-arabonate. After incubation for 1 h at 30 °C, the gel was further soaked in the dark in 1 M NaOH containing 2,3,5-triphenyltetrazolium chloride. Dehydratase activity appeared as a pink band. The active band that appeared in staining was cut out and broken in 50 mM Tris-HCl (pH 8.0) containing 1% (w/v) SDS. The protein was extracted from the gel by vigorous mixing overnight. A 4-fold volume of acetone chilled at −35 °C was then added to the extracts. After cooling for 1 h at −35 °C, the mixture was centrifuged at 39,120 × g for 15 min at 4 °C. The sample was dissolved in a small volume of SDS-PAGE sample buffer (500 mM Tris-HCl, pH 6.8, containing 5% (w/v) SDS, 10% (v/v) glycerol, 0.25% (w/v) bromphenol blue and 5% (v/v) 2-mercaptoethanol) and separated by SDS-PAGE with 10% (v/v) gel. This procedure revealed that a polypeptide with a molecular weight (Mᵣ) of ~62 kDa in Sample I corresponded to L-arabonate dehydratase. Therefore, Sample II was used to determine the N-terminal sequence of L-arabonate dehydratase. With L-KDA dehydratase, the purified enzyme was directly used to determine the N-terminal sequence.

After SDS-PAGE, protein on the gel was transferred to Hybond™-P (Amersham Biosciences) at 3 mA/cm² for 0.5 h in transfer buffer (10 mM cyclohexylaminopropane sulfonic acid, pH 11, containing 10% (v/v) methanol) with a horizontal electrophoretical blotting system (model AE-7500, Atto). After staining and destaining the protein, the area of the membrane corresponding to the L-arabonate dehydratase or L-KDA dehydratase band was excised and analyzed with a Procise™ 492 HT protein sequencer (Applied Biosystems).

**Cloning of L-Arabinose Operon from A. brasilense—** A. brasilense genomic DNA was isolated as described previously (28). To clone the L-arabinose operon of *A. brasilense*, four synthetic oligonucleotide primers (UP-S, UP-AS, DOWN-S, and DOWN-AS) were designed (Table 1); UP-S and DOWN-AS were derived from the genome sequence of *Burkholderia cepacia* strain R18149, and UP-AS and DOWN-S were derived from that of *A. brasilense*. Genomic PCR was carried out separately using a pair of UP-S/UP-AS and DOWN-S/DOWN-AS as primers and 300 ng of *A. brasilense* genomic DNA as a template. Single PCR products of ~2.7-kbp and ~4.3-kbp were purified, cloned into a pGEM®-T vector (Promega) (referred to as pGEM1 and pGEM2, respectively) and sequenced. We designated the deduced open-reading frame (ORF) in the l-arabinose operon, as described in Fig. 3, in which a functional l-arabinose 1-dehydrogenase gene (*AraA*) is contained (28).

**Functional Expression and Purification of Recombinant His₆-tagged Proteins—** *AraB*, *AraC*, and *AraD* were amplified by PCR using primers containing appropriate restriction enzyme sites at the 5'- and 3'-ends and *A. brasilense* genome DNA as a template (Table 1). Each amplified DNA fragment was introduced into BamHI-PstI sites (for *AraB* and *AraD*) or BamHI-KpnI (for the *AraC* gene) in pQE-80L (Qiagen), a plasmid vector for conferring N-terminal His₆ tag on expressed proteins, to obtain pHi₆*AraB*, pHi₆*AraC* and pHi₆*AraD*, respectively. *E. coli* DH5α harboring the expression plasmid for His₆-tagged enzymes was grown at 37 °C to a turbidity of 0.6 at 600 nm in Super broth medium (pH 7.0, 12 g of tryptone, 24 g of yeast extract, 5 ml of glycerol, 3.81 g of KH₂PO₄, and 12.5 g of K₂HPO₄ per liter) containing 50 mg/liter ampicillin. After the addition of 1 mM of isopropyl-β-d-thiogalactopyranoside, the culture was further grown for 6 h to induce the expression of His₆-tagged protein. Cells were harvested and resuspended in Buffer X (pH 7.0, 50 mM sodium phosphate containing 300 mM NaCl and 10 mM imidazole). The cells were then disrupted by sonication, and the solution was centrifuged. The supernatant was loaded onto a Ni-NTA Superflow column (Qiagen) equilibrated with Buffer X, which is linked to an AKTA purifier system. The column was washed with Buffer Y (pH 7.0, Buffer X containing 10% (v/v) glycerol and 50 mM imidazole instead of 10 mM imidazole). The enzymes were then eluted with Buffer Z (pH 7.0, Buffer Y containing 250 mM imidazole instead of 50 mM imidazole). For the purification of *AraB* and *AraC* proteins, 10 mM MgCl₂ was added to the buffer system. The eluant was concentrated and dialyzed against Buffer A containing 50%
(v/v) glycerol (for AraB and AraC proteins) or Buffer D containing 50% (v/v) glycerol (for AraD protein), and stored at −35 °C until use.

For Western blot analysis, purified His6-tagged proteins were separated by SDS-PAGE, and the proteins on the gels were transferred onto a nitrocellulose membrane (Hybond™-ECL; Amersham Biosciences). Western blot analysis was carried out using the ECL™ Western blotting Analysis System (Amersham Biosciences) and RGS®His horseradish peroxidase antibody, a horseradish peroxidase-fused mouse monoclonal antibody against Arg-Gly-Ser-(His)6 in the N-terminal additional body, by subsequent sequencing in both directions. Mutant proteins were expressed and purified by the same procedures as each wild-type (WT) enzyme.

Iron-Sulfur Cluster Analysis of L-Arabinose Dehydratase—To estimate the nature of the iron-sulfur cluster in L-arabinose dehydratase, electron paramagnetic resonance (EPR) analysis was carried out. The His6-tagged recombinant protein sample (40 mg/ml) was prepared in 50 mM HEPES, pH 7.2, containing 10 mM MgCl2 and 50% (v/v) glycerol. EPR measurements were performed for L-arabinose dehydratase reduced anaerobically with 10-fold excess sodium dithionite (Na2S2O4). EPR spectra were recorded using a JEOL TE-300 X-band spectrometer operating with 100 kHz field modulation. Temperature-dependent EPR measurements were performed in the range of 4.2–30 K, using an LTR-3 liquid helium cryostat (Air Products). The g-values were evaluated based on the g-value of Li-salt of tetracyanoquinodimethane (2.0025) as an external standard. The magnetic field strength of EPR spectra were calibrated using a NMR field meter (Echo Electronics) for each measurement.

Amino Acid Sequence Alignment and Phylogenetic Analysis—Protein sequences were analyzed using the Protein-BLAST and ClustalW program distributed by DDBJ (DNA Data Bank of Japan). The phylogenetic tree was produced using the TreeView 1.6.1 program.

RESULTS

L-Arabinose Dehydratase from A. brasilienise—Although there has been no enzymological report of L-arabinose dehydratase, information about other dehydratases may be helpful to characterize this enzyme. In particular, D-glucosonate dehydratase and D-galactonate dehydratase (EC 4.2.1.6) utilize non-phosphorylated sugar acid as a substrate similar to L-arabinose dehydratase and belong to an enolase superfamily (20, 36, 37). Because the activity of L-arabinose dehydratase in the cell-free extract was dependent on several metal ions such as Mg2+ sim-
ilar to general dehydratases including D-gluconate dehydratase and D-galactonate dehydratase (data not shown), the enzyme was partially purified by (NH₄)₂SO₄ fractionation and four chromatographic steps with a buffer system containing Mg²⁺ (Sample I). A typical result of purification is summarized in Table 2. Enzymes in the final fractionation showed specific activity of 2.07 units/mg protein. Sample I was further subjected to non-denaturing PAGE. Sample II was prepared from a single active band detected by “in-gel staining assay” using L-arabonate as the substrate and was shown to consist of a single polypeptide with an approximate molecular weight (Mr) of 62 kDa (lane 6 in Fig. 3A). Sample I was eluted as an almost single symmetric peak of protein of Mr 128 kDa from the column of gel filtration, suggesting that L-arabonate dehydratase forms a homotypic dimer. The N-terminal amino acid sequence up to 15 amino acid residues of L-arabonate dehydratase was determined with Sample II to be ATKPRLRSTQXFGXN (X was not determined).

L-KDA Dehydratase from A. brasiliense—As described above, we were successful in preparing native L-arabonate dehydratase from A. brasiliense. Furthermore, we recently isolated αKGSA dehydrogenase from A. brasiliense, which catalyzes the conversion of αKGSA to α-ketoglutarate in the presence of NAD(P)⁺ (29). When cell-free extracts (containing a low concentration of L-KDA dehydratase) were incubated with L-arabonate and NAD⁺ in the presence of L-arabonate dehydratase and αKGSA dehydrogenase, a significant increase in absorbance at 340 nm because of the appearance of NADH was found, suggesting that L-arabonate was sequentially converted to α-KDA and αKGSA. This assay allowed us to probe many protein fractions during the purification of L-KDA dehydratase, which occurred homogeneously by SDS-PAGE through (NH₄)₂SO₄ fractionation and four chromatographic steps. The purified enzyme showed specific activity of 19.8 units/mg protein. A typical result of purification is summarized in Table 3. SDS-PAGE revealed only one subunit with an apparent Mr value of 36 kDa. Native Mr was estimated to be ~84.8 ± 3.0 kDa by gel filtration, suggesting that L-KDA dehydratase is dimeric. The N-terminal amino acid sequence up to 20 amino acid residues was as follows TSSSTPRHRGIFPVVPTTFA.

Candidate of L-Arabonate Dehydratase and L-KDA Dehydratase Genes—Surprisingly, the N-terminal amino acid sequence of L-arabonate dehydratase was similar to ILVD/EDD family proteins but not enolase family proteins including D-gluconate dehydratase and D-galactonate dehydratase. In particular, it had high homology with a putative ILVD/EDD protein from B. cepacia strain R18194 (NCBI accession number ZP_00218289): MSATKPRLRSTQWFGTN (underlined letters

FIGURE 3. Novel L-arabinose operon of B. thailandensis (left) and A. brasiliense and B. cepacia strain R18149 (right). Same colored ORF indicates each homologous gene. Italic letters are putative functional genes annotated in the genome sequence of B. cepacia strain R18149. Previously, Regions 1 and 2 on the genome of A. brasiliense were isolated, and AraA and AraE genes were assigned to L-arabinose 1-dehydrogenase and αKGSA dehydrogenase, respectively (28, 29). In this study, Regions 3 and 4 on the genome of A. brasiliense were amplified by genomic PCR, and AraA, AraC, and AraD genes were assigned to L-arabinolactonase, L-arabonate dehydratase, and L-KDA dehydratase, respectively.

TABLE 2
Summary of L-arabonate dehydratase purification from A. brasiliense

| Step                  | Total protein | Total activity | Specific activity | Yield | Purification |
|-----------------------|---------------|----------------|------------------|-------|--------------|
| Cell-free extract     | 1759          | 44             | 0.03             | 100   | 1            |
| (NH₄)₂SO₄ fractionation| 1432          | 39             | 0.03             | 89    | 1            |
| HiPrep 16/10 Butyl FF | 143           | 14             | 0.10             | 32    | 3.3          |
| HiPrep 16/10 DEAE FF  | 48            | 7.7            | 0.16             | 18    | 5.3          |
| CHT ceramic hydroxyapatite | 2.7       | 2.5            | 0.93             | 5.7   | 31           |
| HiLoad 16/60 Superdex 200 pg | 0.48 | 1.0            | 2.07             | 2.3   | 69           |
A Novel Bacterial L-Arabinose Pathway

TABLE 3
Summary of L-KDA dehydratase purification from A. brasiliense

| Step                      | Total protein (mg) | Total activity (units) | Specific activity (units/mg protein) | Yield (%) | Purification fold |
|---------------------------|--------------------|------------------------|-------------------------------------|-----------|-------------------|
| Cell-free extract         | 1174               | 53.8                   | 0.046                               | 100       | 1                 |
| (NH₄)₂SO₄ fractionation   | 650                | 32.0                   | 0.049                               | 59        | 3.11              |
| HiPrep 16/10 Butyl FF     | 47                 | 22.6                   | 0.48                                | 42        | 10.4              |
| HiPrep 16/10 DEAE FF      | 6.5                | 18.0                   | 2.8                                 | 33        | 60.9              |
| CHT ceramic hydroxypatite| 1.1                | 16.7                   | 15.4                                | 31        | 335               |
| HiLoad 16/60 Superdex 200 pg | 0.59          | 19.8                   | 31.2                                | 21        | 430               |

indicate the same amino acid residues as L-arabonate dehydratase. On the other hand, a putative DHDS/NAL protein of B. cepacia strain R18194 (NCBI accession number ZP_00218290) had high homology with the N-terminal amino acid sequence of L-KDA dehydratase of A. brasiliense: 1MTSSSTPRYGIFFPVPTTFA (underlined letters indicate the same amino acid residues as the homolog of A. brasiliense). Enzyme reaction product. Typical results are summarized in Fig. 4A. When L-arabonate was incubated with AraC protein, the enzyme reaction product, Compound I, appeared as a single peak with a retention time of 10.3 min. When AraC and AraD proteins were present, L-arabonate was converted to a single enzyme product, Compound II, with a slightly later retention time (10.4 min) than Compound I, suggesting that AraD protein converted Compound I to Compound II. It is noteworthy that the elution profile of Compound II was different from those of authentic glycolaldehyde and pyruvate, potential reaction products by L-KDA aldolase, considering that AraD protein was sequentially assigned as L-KDA aldolase rather than L-KDA dehydratase as described below. Finally, when L-arabonate was incubated with AraC, AraD and AraE (aKGS dehydrogenase) proteins in the presence of NAD⁺, the enzyme product, Compound III, corresponded to authentic α-ketogluvurate. These results indicated that AraC and AraD proteins function as L-arabonate dehydratase and L-KDA dehydratase, respectively.

Functional Expression of AraC and AraD Genes in E. coli—AraC and AraD proteins were overexpressed in E. coli cells by induction with isopropyl-β-d-thiogalactopyranoside as a His₆-tagged enzyme and purified homogeneously with a nickel-chelating affinity column (Fig. 2C). The mobility of the recombinant enzyme in SDS-PAGE was slightly later than that of the native enzyme purified from A. brasiliense cells (lanes 1 and 2 in Fig. 2, D and E), because of the additional 13 amino acid residues including His₆-tag at the N-terminal, which was also confirmed by Western blot analysis with anti-His₆-tag antibody (data not shown). The specific activity of purified recombinant AraC (9.4 unit/mg protein) was 4.5-fold higher than that of the native enzyme, probably because of partial inactivation by oxidation of the [4Fe-4S]²⁺ cluster in the native enzyme during purification (see “Discussion”). The purified recombinant AraD protein showed similar specific activity to the native enzyme (33.3 unit/mg protein).

Identification of Reaction Product by AraC and AraD Proteins—HPLC analysis was performed to estimate the enzyme reaction product. Typical results are summarized in Fig. 4A. When L-arabonate was incubated with AraC protein, the enzyme reaction product, Compound I, appeared as a single peak with a retention time of 10.3 min. When AraC and AraD proteins were present, L-arabonate was converted to a single enzyme product, Compound II, with a slightly later retention time (10.4 min) than Compound I, suggesting that AraD protein converted Compound I to Compound II. It is noteworthy that the elution profile of Compound II was different from those of authentic glycolaldehyde and pyruvate, potential reaction products by L-KDA aldolase, considering that AraD protein was sequentially assigned as L-KDA aldolase rather than L-KDA dehydratase as described below. Finally, when L-arabonate was incubated with AraC, AraD and AraE (aKGS dehydrogenase) proteins in the presence of NAD⁺, the enzyme product, Compound III, corresponded to authentic α-ketogluvurate. These results indicated that AraC and AraD proteins function as L-arabonate dehydratase and L-KDA dehydratase, respectively.

Amino Acid Sequence Analysis of L-Arabinate Dehydratase—As expected from the N-terminal amino acid sequence, L-arabonate dehydratase was a novel member of the ILVD/EDD protein family including the archetype dihydroxyacid dehydratase (ILVD, EC 4.2.1.9) and phosphogluconate dehydratase (EDD, EC 4.2.1.12) (Fig. 5A). Two consensus segments have been proposed as conserved motifs for this protein family, which were also conserved in L-arabonate dehydratase with a few modifications: Cys¹²⁴-Asp-Thr-Pro-Ala-Leu-Leu-Met-Gly¹³⁴ and Ile²²²-Thr-Asp-Ala-Arg-Met-Ser-Gly-[–]-Thr³⁸⁰ (underlined letters indicate the same amino acid residues as the proposed motifs and [–] is a sequential gap) (regions X and Z in Fig. 5A, respectively). It is interesting that EDD catalyzes the equivalent step in the ED pathway as L-arabonate dehydratase: the dehydration of phosphogluconate into KDGP (Fig. 1B). In the phylogenetic tree of the ILVD/EDD family, ILVD and EDD formed a single subclass, and L-arabonate dehydratase was closely related to archaeal and eukaryotic ILVDs (Fig. 5C).

Substrate Specificity and Cofactor of L-Arabinate Dehydratase—Various hexonates/pentonates were tested as substrates for dehydratase in addition to L-arabonate (Fig. 5B). Significant activity was observed with L-arabonate and D-xylonate,
but very little activity was observed with D-galactonate, D-arabonate, and D-gluconate. D-Xylonate, a C4 epimer of L-arabonate, was a better substrate than D-galactonate with the same configuration as L-arabonate at C2, C3, and C4. These results indicated that the enzyme was significantly specific for L-arabonate and the L-arabonate-specific configuration at C2 and C3 was important for substrate specificity, probably reflecting the characterization of dehydration by the α,β elimination of water and ketonization of the intermediate enol. This is reported to be located between the function(s) of several ILVD/EDD proteins and metal ions, particularly, Mg$^{2+}$. Similar behavior was found in L-arabonate dehydratase, and the optimal Mg$^{2+}$ concentration was about 1–10 mM (inset in Fig. 5B).

Identification of Catalytic Amino Acid Residues of L-Arabonate Dehydratase—It is known that enzymes in the ILVD/EDD protein family contain an [Fe-S] cluster in the active center (38–41). Purified L-arabonate dehydratase (the concentrated recombinant enzyme) was brown in color (Fig. 6B), and the UV-visible absorption spectrum consisted of a protein band centered at 280 nm with pronounced shoulders at ~330 and 400 nm arising from S→Fe(III) charger transfer bands localized on the [Fe-S] cluster(s) (Fig. 6C). On the other hand, the addition of sodium dithionite (Na$_2$S$_2$O$_4$) resulted in absorption bleaching, and absorbance changes were consistent with complete one-electron reduction to the paramagnetic [4Fe-4S]$^{2+}$ redox state. The EPR spectrum observed for the reduced form of the enzyme revealed a complex line shape, consisting of at least four paramagnetic species, as illustrated in Fig. 6D. A broad signal at $g = 1.91$ was assigned to $S = 1/2$ resonance derived from [4Fe-4S]$^{2+}$ (38, 42). In addition, the relatively sharp signal at $g = 2.06$ was characteristic of the $S = 1/2$ resonance of the [3Fe-4S]$^{2+}$ cluster (38), which was expected to be formed by a reaction occurring between excess sodium dithionite and the [4Fe-4S]$^{2+}$ cluster. In fact, a set of weak EPR signals ($g = 10.4, 3.3$) because of ferric high-spin species with rhombic geometry was also recognized. In the lower magnetic field, broad and weak EPR signals at $g = 4.6, 4.7, 3$, and 2 overlapping with $S = 1/2$ resonance are attributed to the γ-anisotropy of $S = 3/2, S = 3/2$, and $S = 5/2$ species of the cluster (38, 42). Wholly, the EPR property of the reduced form is analogous to that reported for most ILVD/EDD proteins containing [4Fe-4S]$^{2+}$ cluster (cf. Fig. 4 in Ref. 38) but not that reported for ILVD of spinach (cf. Fig. 5 in Ref. 42), in which a [2Fe-2S] cluster is contained. These results clearly suggested that the [4Fe-4S]$^{2+}$ cluster was found in L-arabonate dehydratase as well as most...
FIGURE 5. A, multiple sequence alignment of deduced amino acid sequences of L-arabonate dehydratase (AraC) from *A. brasiliense* and several ILVD/EDD proteins. Letters with asterisks indicate the positions of amino acid residues of L-arabonate dehydratase substituted by site-directed mutagenesis in this study. Regions X and Z are consensus segments of the ILVD/EDD family. Region Y in *E. coli* ILVD contains the Cys-X2-Cys-X2-Cys motif characteristic of \([4Fe-4S]_2\) cluster-containing enzymes. GenBank™ accession numbers for protein sequences are in Fig. 5C. B, substrate specificity of L-arabonate dehydratase. The indicated substrates were added to the standard assay solution instead of L-arabonate. Values are the means ± S.D., n = 3. Inset figure indicates the effect of Mg2⁺ on enzyme activity. The stocked enzyme sample was gel-filtrated on HiLoad 16/60 Superdex 200 pg of column with 50 mM HEPES, pH 7.2. The standard assay solution was used except that the Mg2⁺ concentrations varied as indicated. Values are the means ± S.D., n = 3. C, phylogenetic tree of ILVD/EDD proteins including L-arabonate dehydratase (left). GenBank™ accession numbers are as follows: *Corynebacterium glutamicum*, CAB57218; *Streptomyces coelicolor*, AAC35795; *E. coli* (ILVD), P05791; *Caulobacter crescentus*, P55186; *Neurospora crassa*, CAD70774; *Saccharomyces cerevisiae*, AAA34568; *Lactococcus lactis*, AAB81918; *S. solfataricus*, NP_344419; *Pyrococcus horikoshii*, P31961; *E. coli* (EDD), AAB59053; *Z. mobilis*, P21909. Right figure indicates the partial alignment of amino acid sequences around the site-directed mutated region of L-arabonate dehydratase with other ILVD/EDD proteins. Bold letter enzymes have been known experimentally to include the \([4Fe-4S]_2\) cluster. Cysteine residues in black boxes coordinate with the \([4Fe-4S]_2\) cluster in L-arabonate dehydratase from *A. brasiliense*. Regions X and Y correspond to the same regions in Fig. 5A.

A Novel Bacterial L-Arabinose Pathway

33530 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 281 • NUMBER 44 • NOVEMBER 3, 2006
ILVD/EDD proteins. It has been assumed that three iron atoms in the cluster of ILVD/EDD proteins are coordinated with three (unidentified) cysteine residues from protein and the remainder is exposed to solvent (38, 39) (Fig. 6A). The amino acid sequence of L-arabonate dehydratase contains nine cysteine residues, among which Cys124 and Cys197 are conserved completely in these ILVD/EDD proteins and an equivalent cysteine residue at the 56 position is found frequently in other ILVD/EDD proteins. Therefore, we constructed five mutants with cysteine substituted for alanine, C47A, C56A, C124A, C197A, and C441A, as described in “Experimental procedures.” They were overexpressed in E. coli cells as a His6-tagged enzyme, purified homogeneously with the same procedures as the WT enzyme (Fig. 2D) and characterized (Fig. 6, B and C). Under standard assay conditions, brown-colored C47A and C441A mutants were significantly active and UV-visible absorption spectra were similar to those of the WT. On the other hand, C124A and C197A mutants were non-colored enzymes and the UV visible absorption spectra contained no characteristic element of [4Fe-4S]2+ localized in the WT. The C56A mutant was also an inactive form, but partially shared similar behavior to the WT in color and UV-visible absorption spectra. These results strongly suggested that Cys56, Cys124, and Cys197 were involved in enzyme catalysis, in particular, the coordination of [4Fe-4S]2+. This is the first report to clearly identify the active site(s) in ILVD/EDD proteins.

Amino Acid Sequence Analysis of L-KDA Dehydratase—As expected from the N-terminal amino acid sequence, L-KDA dehydratase belongs to the DHDPS/NAL protein family, including the archetypal dihydrodipicolinate synthase (DHDPS, EC 4.2.1.52), N-acetylneuraminic lyase (NAL, EC 4.1.3.3), KDG aldolase, d-5-keto-4-deoxyglucarate dehydratase (decarboxylating) (EC 4.2.1.41), trans-o-hydroxybenzylidenepyruvate hydrolase-aldolase and trans-2'-carboxybenzalpyruvate hydratase-aldolase (EC 4.1.2.34) (Fig. 7A). Phylogenetic analysis revealed that L-KDA dehydratase has a poor relationship to any subclasses of the other members (Fig. 7B), probably causing unique enzyme catalysis, as described below. The DHDPS/NAL family includes two interesting enzymes, KDG aldolase and d-5-keto-4-deoxyglucarate dehydratase. The former enzyme is involved in the non-phosphorylative ED pathway and catalyzes an equivalent reaction to “L-KDA dehydratase-aldolase (EC 4.1.2.18)” in the second alternative pathway of L-arabinose metabolism but not “dehydratase” (Fig. 1B) and the latter enzyme produces αKGSAn but not pyruvate similar to L-KDA dehydratase (Figs. 1B and 7C).

Identification of Catalytic Amino Acid Residue of L-KDA Dehydratase—Recently, a crystallographic study of KDG aldolase (19) revealed that the overall structural fold and fundamental catalytic mechanism(s) including Tyr130 and Lys156 are conserved, compared with those found in other DHDPS/NAL enzymes (44–46). On the other hand, Tyr130–Lys156 correspond to Gln143–Lys171 in L-KDA dehydratase, indicating that this enzyme is the first example with no functional tyrosine residue in DHDPS/NAL family enzymes (asterisks in Fig. 7A). To obtain insight into the catalytic mechanism of L-KDA dehydratase, Gln143 was substituted with Asn, Glu, Ser, Thr, and Tyr by site-directed mutagenesis as described under “Experimental Procedures,” to construct Q143N, Q143E, Q143S, Q143T, and Q143Y mutants, respectively. They were overexpressed in
E. coli cells as a His₆-tagged enzyme and purified with the same procedures as the wild-type enzyme (Fig. 2E). No L-KDA dehydratase activity was found in any constructed mutants. Because Q143S, Q143T, and Q143Y mimicked dratase activity was found in any constructed mutants. Because catalyzes the ring cleaving of D-gluconate in the D-gluconate pathway, and this reaction is analogous to that of L-arabinolactonase (Fig. 1). The nucleotide sequence was submitted to GenBank™ with GenBank™ accession numbers for protein sequences are in Fig. 7. A consensus segment of the DHDPS/NAL family is shown as a gray-colored boxes in Fig. 2A. The number on each branch indicates the bootstrap value.

Identification of L-Arabinolactonase Gene—Isolation of the downstream region of the AraA gene by genomic PCR of A. brasilienensis revealed that A. brasilienensis possesses a complete equivalent l-arabinose opeon to B. cepacia strain R18149 (Fig. 3). The nucleotide sequence was submitted to GenBank™ with accession number AB241136. The AraX, AraY, and AraZ genes were homologous with bacterial putative ATP-binding cassette (ABC)-type sugar transporter components consisting of a periplasmic substrate-binding protein, an ATP-binding complex (ATPase), and a membrane-spanning protein (permease). A gluconolactonase (EC 3.1.1.17) from Zymomonas mobilis was the only characterized enzyme with similarity to the (putative) AraB protein (~24% of identity) (48, 49). This enzyme catalyzes the ring cleaving of d-glucono-γ-lactone to d-gluconate in the d-gluconate pathway, and this reaction is analogous to that of l-arabinolactonase (Fig. 1B); therefore, the AraB gene is thought to be a candidate for l-arabinolactonase. AraB protein was expressed in E. coli as a His₆-tagged enzyme and purified homogeneously by similar experimental procedures as for AraC and AraD proteins (lane 2 in Fig. 2C). The known sugar lactonase assays with a lactone are not reliable, conveniently because lactone is hydrolyzed spontaneously to produce the corresponding sugar acid. We measured potential l-arabinolactonase activity in AraB protein by an alternative method, which consisted of a coupling reaction with NADP⁺-dependent l-arabinose 1-dehydrogenase (28) and determined the reaction product by HPLC. This is summarized in Fig. 4B. The retention time of authentic l-arabinino-γ-lactone was slightly earlier than that of l-arabinose (data not shown). When l-arabinose was incubated with l-arabinose 1-dehydrogenase in the presence of NADP⁺, l-arabinose was only slightly present in the reaction mixture due to spontaneous lactone hydrolysis. On the other hand, when AraB protein co-existed in the same reaction mixture, substantial formation of l-arabonate was observed, indicating that AraB functions as l-arabinolactonase. Since l-arabinose 1-dehydrogenase also functions as d-galactose 1-dehydrogenase (EC 1.1.1.48 (120)) in vitro (28), d-galactose was added to the reaction mixture instead of l-arabinose. AraB protein converted d-galactonolactone to d-galactonate similarly to l-arabinolactone (data not shown).

DISCUSSION

A Novel Bacterial l-Arabinose Operon—In this and preceding (28) studies, we showed that A. brasilienensis possesses a novel gene cluster containing l-arabinose 1-dehydrogenase (AraA), l-arabinolactonase (AraB), l-arabonate dehydratase (AraC), and l-KDA dehydratase (AraD) genes together with sugar transporter genes (AraX, Y and Z) (Fig. 3). The “novel l-arabinose operon (AraCDAXYZB)” is different from the well-known bacterial l-arabinose operon as described in the “Introduction.” A similar operon is also found in the genome of B. cepacia strain R18149. Furthermore, we previously revealed that the αKGSA dehydrogenase gene (AraE) is located on the genome of A. brasilienensis (probably also B. cepacia strain R18149) separately from the gene cluster (29). On the other hand, Burkholderia thailandensis, a near relative bacterium of B. cepacia strain R18149, possesses an additional aldehyde dehydrogenase-like gene in the equivalent gene cluster (Fig. 3). The gene product is thought to be a candidate for AraE and this type of operon is named AraCEDAXYZB. Although each homolog to AraA–AraE genes is observed in many bacterial genomes, almost all form a partial and/or no gene cluster (data not shown), in particular several Burkholderia species despite their close relationships, as mentioned by Moore et al. (50). This suggests that the l-arabinose operon diversified very quickly during bacterial evolution. Wide-range genome transportation is thought to occur at a very early evolutionary stage in cases of A. brasilienensis and B. cepacia strain R18149, considering the very high similarity between their nucleotide sequences (~95% identity).

Evolutionary Insight into Several Sugar Metabolic Pathways—One of the most significant findings reported in this paper is the complete identification of five metabolic genes involved in the alternative l-arabinose pathway (the first pathway). We will discuss the evolutionary relationships between the alternative l-arabinose pathway and other sugar metabolic pathways including the ED, non-phosphorylative ED, d-glucanate, and DD pathways at the molecular level (Fig. 1B).

Conversion of l-Arabinose to l-Arabonate—In the non-phosphorylative ED pathway and most of the bacterial d-glucanate pathway, d-glucose is directly oxidized to d-glucanate by GDH: the GDH involved in the former pathway is an NAD(P)⁺-dependent enzyme (EC 1.1.1.118 (119)) (51, 52), while that in the latter pathway is a pyrroloquinoline-quinone (PQQ)-dependent enzyme (EC 1.1.5.2) (53). In this regard, the conversion of l-arabinose is analogous to the ED pathway, the d-glucanate pathway with a few bacteria, and the DD pathway, in which a lactone is included as an intermediate. We previously showed that l-arabinose 1-dehydrogenase is similar to glucose-fructose oxidoreductase (GFOR; EC 1.1.99.28) and d-galactose 1-dehydrogenase involved in the two later pathways, respectively (28).

FIGURE 7. A. A multiple sequence alignment of deduced amino acid sequences of l-KDA dehydratase (AraD) from A. brasilienensis and several DHDPs/NAL proteins. GenBank™ accession numbers for protein sequences are in Fig. 7B. A consensus segment of the DHDPs/NAL family is shown as a line on the sequence. Letters in gray-colored boxes indicate amino acid residues forming a hydrogen bond with the substrate. Lysine residue with an asterisk is also involved in the catalysis of DHDPs/NAL proteins, while the residue is substituted to glutamine in l-KDA dehydratase, β phylogenetic relationships among NAL (subclass I), DHDPs (subclass II), d-5-keto-4-deoxyglucurionate dehydratase (subclass III), trans-α-hydroxybenzylidenepyrurate hydrolyase-aldolase, and trans-2′-carboxybenzalpyruvate hydratase-aldolase (subclass IV), and KDG aldolase (subclass V). The number on each branch indicates the bootstrap value. C, schematic reactions catalyzed by DHDPs/NAL proteins. The cleaved carbon atoms are shown in red and orange. l-KDA dehydratase belongs to this family, while the enzyme reaction contains no cleavage of the C=C or C≡C bond. Although it is unclear whether l-KDA aldolase is a potential member of this family, the enzyme reaction is equivalent to those of this enzyme family (see text).
Therefore, it seems reasonable that \( l \)-arabinolactonase is similar to bacterial gluconolactonase but not 6-phosphogluconolactonase (involved in the ED pathway). On the other hand, \( Z. \) mobilis is the only characterized bacteria with the \( d \)-gluco-
nate pathway involving GFOR and gluconolactonase. Because \( d \)-galactonolactone is active for \( l \)-arabinolactonase, the \( d \)-gal-
 lactonolactonase involved in the DD pathway may be homolo-
gous with these lactonases.

Although lactone studies have recently received much attention in diverse fields such as pathology, applied microbiology, and prebiotic biology, there are only two reports of lactone-hydrolyzing enzymes at the molecular level: gluconolactonase of \( Z. \) mobilis and lactonohydrolase of a fungus, \textit{Fusarium oxysporum}. Recently, Kobayashi et al. (54) proposed that these lact-
tone-hydrolyzing enzymes form a unique superfamily consisting of C-O cleavage enzymes and P-O cleaving enzymes. Further biochemical research including crystallographic analysis would be a major advantage to identify the active sites and reaction mechanism.

\textit{Conversion of L-Arabonate to L-KDA—L-Arabonate dehy-
 dratase} belongs to the ILVD/EDD family, which is different from the enolase superfamily including \( d \)-gluconate dehy-
 dratase and \( d \)-galactonate dehydratase (20, 55). Because EDD, an equivalent enzyme involved in the ED pathway to \( l \)-arabonate dehydratase, is a member of this protein family, this phylogenetic relationship seems reasonable.

The metabolic promiscuity of sugar carbon sources is often found in microorganisms. For example, \textit{S. sulfolobus} metabolizes \( d \)-glucose and \( d \)-galactose through the non-phosphorylative ED pathway (18–20). Similarly, we previously doubted that the alternative \( l \)-arabinose pathway is involved in \( d \)-galactose metabolism as the DD pathway, because \( l \)-arabinose \( 1 \)-deh-
ydrogenase also functions as \( d \)-galactose \( 1 \)-dehydrogenase in \textit{E. coli} (28). In this study, we showed that \( d \)-galactonolactone is an additional active substrate for \( l \)-arabinolactonase, while \( l \)-arabinate dehydratase has no activity with \( d \)-galactonate (Fig. 6A). Thus, these enzymological data reveal that the physi-
ological role of this pathway is limited in \( l \)-arabinose metabol-
ism, confirming the previous genetic analysis; disruption of the \textit{AraA} gene (equal inactivation of the whole \( l \)-arabinose operon) leads to \( l \)-arabinose-negative and \( d \)-galactose-positive pheno-
types (28).

It is interesting that ILVD of archaeal \( S. sulfolobus \) also dis-
plays high activity toward \( d \)-gluconate in addition to dihy-
droxyisovalerate, a natural substrate (56). Thus, this Archaea
possesses different enzymes with dehydration activity for \( d \)-gluconate, although it is unclear whether ILVD is involved in the non-phosphorylative ED pathway. As described below, a \([4Fe-4S]\)y\textsuperscript{2+} cluster is contained in ILVD/EDD proteins but not in enolase superfamily proteins. There is a line of reasoning that the \([Fe-S]\) cluster is an evolutionary relic left over the days when the levels of oxygen in the atmosphere were much lower than now (57, 58). It follows from that line of reasoning that some biological functions originally performed by \([Fe-S]\) clusters may have been abandoned during evolution and that some functions currently assumed by \([Fe-S]\) clusters will eventually be replaced by more efficient metal-free systems. If we follow this hypo-
thesis, the ancestor of \( d \)-gluconate dehydratase might be a \([4Fe-
4S]\)^{2+} cluster-containing enzyme similar to EDD, ILVD and \( l \)-arabinonate dehydratase.

\textit{Conversion of L-KDA into aKGSa—The metabolic fate of \( l \)-KDA to \( \alpha \)-ketoglutarate via aKGSa is unique in the first alternative \( l \)-arabinose pathway, because 2-keto-3-deoxypen-
tonate/hexonate is cleaved through an aldolase reaction to aldehyde and pyruvate in the all known other sugar metabolic pathways similar to the second alternative \( l \)-arabinose pathway (Fig. 1B). \( l \)-KDA dehydratase is a unique member of the DHDP/NAL family because the enzyme reaction consists of only dehydration without cleavage of the C=C or C=O bond (Fig. 7C). Because KDG aldolase in particular belongs to this protein family, the \textit{AraD} gene would be assigned as \( l \)-KDA aldolase rather than \( l \)-KDA dehydratase without the enzymo-
logical data in this study. These insights suggest one evolutional hypothesis for the two alternative \( l \)-arabinose pathways, as described below. On the other hand, KDPG aldolase involved in the ED pathway belongs to the KDPG/KHG aldolase family (59), which is not similar to the DHDP/NAL family.

Overall, there is significant mosaicism between the alterna-
tive \( l \)-arabinose pathway and other sugar metabolic pathways, in contrast with no evolutional relationship between the ED pathway and non-phosphorylative ED pathway. According to the so-called “patchwork” evolution model, metabolism evolved by the recruitment of relatively inefficient enzymes of broad specificity that could react with a wide range of chemically related substrates (60–62). In this regard, “evolutionary relics” seem to remain in the alternative \( l \)-arabinose pathway.

\textit{Novel Coordination of [4Fe-4S]\textsuperscript{2+} Cluster in \( l \)-Arabonate Dehydratase—In many [Fe-S] cluster-containing proteins, some cysteine residues essential for the cluster exist in the Cys-
X\textsubscript{2}-Cys motif or its extended form, CYS-X\textsubscript{2}-CYS-X (\( X \) represents any other amino acid) (39, 63). The amino acid sequence of \textit{E. coli} ILVD also contains such a motif (Cys\textsuperscript{189}Pro-
Thr-Cys\textsuperscript{192}Gly-Ser-Cys\textsuperscript{195}, in which three cysteine resi-
dues are thought to be candidates for ligands to the \([4Fe-4S]\)^{2+} cluster (38) (Fig. 5, A and C). On the other hand, \( l \)-arabinonate dehydratase of \textit{A. brasilienise} (this study) and EDDs of \textit{E. coli} (39) and \textit{Z. mobilis} (40) contain the \([4Fe-4S]\)^{2+} cluster but pos-
sess no such a motif in the equivalent region or the whole region. Therefore, among these three proteins at least, it is not helpful to use this motif(s) as a guide to estimate ligands to the \([4Fe-4S]\)^{2+} cluster. Alternatively, multiple amino acid sequence alignment indicates two strictly conserved cysteine residues in ILVD/EDD proteins, which structurally correspond to Cys\textsuperscript{124} and Cys\textsuperscript{197} in \( l \)-arabinose dehydratase. A consensus Cys-
Asp-Lys-X\textsubscript{3}-Pro-[Gly/Ala]-X\textsubscript{3}-[Gly/Ala] segment including the for-
mer cysteine residue (underlined letter) is known to be a char-
acteristic conserved motif in ILVD/EDD proteins. On the other
hand, the latter cysteine residue is a structural equivalent to
Cys\textsuperscript{195} in \textit{E. coli} ILVD as described above. C124A and C197A mutants share no behavior with \([4Fe-4S]\)^{2+} cluster-containing proteins in color and UV-visible absorption spectra (Fig. 6, B and C), indicating that inactivation is due to complete elimina-
tion of the \([4Fe-4S]\)^{2+} cluster. Therefore, it seems probable that the equivalent two cysteine residues in ILVD/EDD proteins are commonly involved in catalysis similar to \( l \)-arabinose dehy-
dratase: these residues should be ligated directly to the \([4Fe-

4S$^{2+}$ cluster. On the other hand, the C56A mutant is also inactive but shares partial behavior with [Fe-S] cluster-containing proteins, in contrast to C124A and C197A mutants (Fig. 6, B and C). There is a possibility that this mutant contains other classes of [Fe-S] cluster including [2Fe-2S], linear and cubic [3Fe-4S], derived from oxidation of the [4Fe-4S]$^{2+}$ cluster. It is noteworthy that ILVD/EDD proteins containing the [4Fe-4S]$^{2+}$ cluster do not necessarily possess the equivalent cysteine residue to Cys$^{56}$ in L-arabonate dehydratase (Fig. 5C). Therefore, the identification of three cysteine ligands to the [4Fe-4S]$_2$ cluster forms a single transcriptional unit on the A. brasiiliense genome, indicating a role in the alternative L-arabinose pathway as follows: L-Arabonate dehydratase is a novel member of the ILVD/EDD family, which is characteristic of a [4Fe-4S]$^{2+}$ cluster among the ILVD/EDD family. One interesting contrast is that L-KDA dehydratase is rather assigned as an aldol-cleavage enzyme for L-KDA in the DHDPS/NAL family. Site-directed mutagenesis first revealed that three cysteine residues are involved in catalysis as ligands to the cluster among the ILVD/EDD family. One interesting contradiction is that L-KDA dehydratase is rather assigned as an aldol-cleavage enzyme for L-KDA in the DHDPS/NAL family. Site-

Here, we propose different evolutional hypotheses of two alternative pathways of L-arabinose metabolism. The first is that their metabolic evolution depended on a common ancestor with aldol reaction and dehydration activity for L-KDA, supporting that L-KDA dehydratase is assigned sequentially as L-KDA aldolase as described above. Furthermore, pyruvate prevents the inactivation of L-KDA dehydratase by NaBH$_4$ in the presence of a substrate similar to other DHDPS/NAL proteins, assuming that the enzyme possesses significant affinity with pyruvate (and glycolaldehyde?) (65). An alternative hypothesis is that there is no evolutionary relationship between L-KDA dehydratase and L-KDA aldolase. Dahms and Anderson (8, 66–69) reported that, in pseudomonad MSU-1, D-fucose is enzymatically converted into pyruvate and lactaldehyde through the participation of dehydrogenase, lactonase, dehydratase and aldolase. Indeed, these enzymes seem to be also involved in the second alternative L-arabinose pathway in this bacteria: D-2-keto-3-deoxyxylonate aldolase is consistent with L-KDA aldolase, which is the only enzymological report of L-KDA aldolase (8, 69). The enzyme clearly requires a divalent cation, including Mn$^{2+}$, Co$^{2+}$, and Mg$^{2+}$, for activity, whose behavior is characteristic of KDPG/KHG aldolase family proteins rather than the DHDPS/NAL family including L-KDA dehydratase. Furthermore, as shown in Fig. 7, B and C, the DHDPS/NAL family includes D-5-keto-4-deoxyglucarate dehydratase, by which KGSA is also produced from D-5-keto-4-deoxyglucarate (47, 70). Indeed, D-5-keto-4-deoxyglucarate is alternatively converted into pyruvate and tartronate semialdehyde by D-5-keto-4-deoxyglucarate aldolase, similar to L-KDA. Recently, crystallographic study revealed that this enzyme belongs to the KDPG/KHG aldolase family but not the DHDPS/NAL family (43). This seems to support the hypothesis that L-KDA dehydratase and L-KDA aldolase evolved independently from different ancestors. Identification of the native L-KDA aldolase gene and further site-directed mutagenesis of L-KDA dehydratase would help to further elucidate this evolutionary insight.

CONCLUSION

Genetic and enzymatic analysis revealed several interesting insights into the alternative L-arabinose pathway as follows: Four metabolic genes except the aKGSA dehydrogenase gene form a single transcriptional unit on the A. brasiiliense genome, which is different from the well-known bacterial L-arabinose operon.

Metabolic enzymes involved in the alternative L-arabinose pathway are related evolutionally to equivalent enzymes involved in the ED pathway, non-phosphorylative ED pathway, D-glucuronate pathway and/or DD pathway, indicating that this pathway evolved step-by-step from several ancestral metabolic pathways.

L-Arabanate dehydratase is a novel member of the ILVD/EDD family, which is characteristic of a [4Fe-4S]$^{2+}$ cluster in the active center. Site-directed mutagenesis first revealed that three cysteine residues are involved in catalysis as ligands to the cluster among the ILVD/EDD family. One interesting contradiction is that L-KDA dehydratase is rather assigned as an aldol-cleavage enzyme for L-KDA in the DHDPS/NAL family. Site-
A Novel Bacterial L-Arabinose Pathway

directed mutagenesis revealed a unique amino acid residue involved in catalysis.

Acknowledgments—We thank Dr. Yasuhiro Takada and Tomohiro Hirose, Hokkaido University, for help in the determination of amino acid sequences.

REFERENCES

1. Stryer, L. (1988). *Biochemistry*, 3rd Ed., pp. 805–807, W. H. Freeman and Company, New York.
2. Richard, P., Londesborough, J., Putkonen, M., Kalkkinen, N., and Penttila, M. (2001). *J. Biol. Chem.* 276, 40631–40637.
3. Richard, P., Putkonen, M., Vaananen, R., Londesborough, J., and Penttila, M. (2002) *Biochemistry* 41, 6432–6437.
4. Verho, R., Putkonen, M., Londesborough, J., Penttila, M., and Richard, P. (2004) *J. Biol. Chem.* 279, 14746–14751.
5. Welmberg, R., and Doudoroff, M. (1955). *J. Biol. Chem.* 217, 607–624.
6. Welmberg, R. (1961) *J. Biol. Chem.* 236, 629–635.
7. Dagley, S., and Trudgill, P. W. (1965).
8. Dahms, A. S., and Anderson, R. L. (1969).
9. Macchiarulo, A., Nobeli, I., and Thornton, J. M. (2004).
10. Duncan, M. J. (1979).
11. MacGee, J., and Doudoroff, M. (1954). *J. Biol. Chem.* 210, 617–626.
12. Gabriel, O., and Wang, S. F. (1969) *Anal. Biochem.* 27, 545–554.
13. Donald, A., Sibley, D., Lyons, D. E., and Dahms, A. S. (1979) *J. Biol. Chem.* 254, 2132–2137.
14. Entner, N., and Doudoroff, M. (1952).
15. Duncan, M. J. (1904).
16. Flint, D. H., Empette, M. H., Finnegnan, M. G., Fu, W., and Johnson, M. K. (1993) *J. Biol. Chem.* 268, 14732–14742.
17. Elzainy, T. A., Hassan, M. M., and Allam, A. M. (1973).
18. Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D., and Hough, D. W., Danson, M. J., and Taylor, G. L. (2004).
19. Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lamble, H. J., Hough, D. W., Danson, M. J., and Taylor, G. L. (2004) *J. Biol. Chem.* 279, 43886–43892.
20. Lambre, H. J., Milburn, C. C., Taylor, G. L., Hough, D. W., and Danson, M. J. (2004) *FEBS Lett.* 576, 133–136.
21. Ahmed, H., Ettema, T. J., Tjaden, B., Geerling, A. C., van der Oost, J., Siebers, B. (2005) *Biochem. J.* 390, 529–540.
22. Otani, M., Ibara, N., Umezawa, C., and Sano, K. (1986) *J. Bacteriol.* 167, 148–152.
23. Tsai, C. S., Shi, J. L., and Ye, H. G. (1995) *Arch. Biochem. Biophys.* 316, 163–168.
24. Muller, H. M. (1986) *Arch. Microbiol.* 144, 151–157.
25. Leibor, L. F. (1951) *Arch. Biochem.* 33, 186–190.
26. De Ley, J., and Doudoroff, M. (1957) *J. Biol. Chem.* 227, 745–757.
27. Elshtaei, M., Abdel-Fattah OM. (2001) *Enzyme Microb. Technol.* 29, 76–83.
28. Watanabe, S., Kodaki, T., and Makino, K. (2006) *J. Biol. Chem.* 281, 2612–2623.
29. Watanabe, S., Kodaki, T., and Makino, K. (2006) *J. Biol. Chem.* 281, 28876–28888.
30. Sambrook, J., Fritsch, E. F., and Maniatis, T. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd Ed., Cold Spring Harbor, NY.