

Cloning and Characterization of the *Zymobacter palmae* Pyruvate Decarboxylase Gene (*pdc*) and Comparison to Bacterial Homologues†

Krishnan Chandra Raj, Lee A. Talarico, Lonnie O. Ingram, and Julie A. Maupin-Furlow*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611-0700

Received 28 January 2002/Accepted 1 March 2002

Pyruvate decarboxylase (PDC) is the key enzyme in all homo-ethanol fermentations. Although widely distributed among plants, yeasts, and fungi, PDC is absent in animals and rare in bacteria (established for only three organisms). Genes encoding the three known bacterial *pdc* genes have been previously described and expressed as active recombinant proteins. The *pdc* gene from *Zymomonas mobilis* has been used to engineer ethanol-producing biocatalysts for use in industry. In this paper, we describe a new bacterial *pdc* gene from *Zymobacter palmae*. The pattern of codon usage for this gene appears quite similar to that for *Escherichia coli* genes. In *E. coli* recombinants, the *Z. palmae* PDC represented approximately 1/3 of the soluble protein. Biochemical and kinetic properties of the *Z. palmae* enzyme were compared to purified PDCs from three other bacteria. Of the four bacterial PDCs, the *Z. palmae* enzyme exhibited the highest specific activity (130 U mg of protein⁻¹) and the lowest *K_m* for pyruvate (0.24 mM). Differences in biochemical properties, thermal stability, and codon usage may offer unique advantages for the development of new biocatalysts for fuel ethanol production.

Pyruvate decarboxylase (PDC) is the key enzyme for all homo-fermentative ethanol pathways. This enzyme catalyzes the nonoxidative decarboxylation of pyruvate to acetaldehyde using Mg²⁺ and thiamine pyrophosphate (TPP) as cofactors. Acetaldehyde is reduced to ethanol by alcohol dehydrogenase (ADH) during NADH oxidation. ADH enzymes are widely distributed throughout nature (41). In contrast, PDC is common to only plants, yeasts, and fungi; it is absent in animals and rare in prokaryotes (25). PDC has been established by cloning and purification for only three bacteria, *Zymomonas mobilis* (6, 7, 11, 19, 33, 34), *Sarcina ventriculi* (28, 45), and *Acetobacter pasteurianus* (40). Though cloned from plants and fungi (25, 38), only bacterial PDC enzymes are produced at high levels as active recombinant products (6, 10, 11, 34, 40, 45). Recombinant *Z. mobilis* PDC has been used for the metabolic engineering of ethanol pathways in many organisms (4, 8, 14, 16, 22, 44).

An ethanologenic bacterium, *Zymobacter palmae*, has been reported which appears to contain a PDC enzyme (37). *Z. palmae* was originally isolated from palm sap (37) and produces ethanol as a primary fermentation product from a variety of hexose sugars and saccharides (20, 36). This bacterium is distinct from other ethanol-fermenting bacteria including those belonging to the genera *Zymomonas*. *Zymomonas* is well known for production of ethanol from plant sap in tropical areas; however, its fermentable carbohydrates are limited to glucose, fructose, and saccharose (22). In addition to substrate utilization, *Zymobacter* (55.8 ± 0.4 mol% G+C) differs dramatically from *Zymomonas* (48.5 ± 0.5 mol% G+C) in ge-

nome composition. This is consistent with the other phenotypic differences that have been observed, including the type of ubiquinone synthesized and peritrichous versus polar flagellation (37).

In this paper, we have established the presence of PDC in *Z. palmae* by cloning the gene and by purification of the enzyme (native and recombinant). Biochemical and kinetic properties of this gene and the encoded enzyme were compared to those of the three other bacterial homologues. Differences in codon usage and kinetic properties among these may provide a useful guide for the development of future biocatalysts for fuel ethanol production from renewable biomass.

MATERIALS AND METHODS

Strains and media. *Z. palmae* strain T109 (IAM 14233, ATCC 51623) was cultivated in ATCC 1956 MY medium (10 g of yeast extract, 20 g of maltose, 2 g of KH₂PO₄, and 5 g of NaCl per liter) at 26°C (200 rpm). *A. pasteurianus* strain NCIB8618 (ATCC 12874) was grown with aeration at 25°C in minimal medium (pH 5.0 to 5.5) supplemented with 2% (vol/vol) D-L-lactate as described previously (3, 13), with the addition of antifoam A (0.4 ml per 10 liters). *Escherichia coli* strains ER1648 F⁻ *fluA2Δ(lacZ)* *r1 supE44 trp31 mcrA1272::Tn10(Tet^r) his-1 rpsL104* (Str^r) *xyl-7 ml-2 metB1 Δ(mcrC-mrr)102::Tn10(Tet^r)* (New England Biolabs, Beverly, Mass.), DH5α F⁻ *recA1 endA1 hsdR17* (r_k⁻ m_k⁺) *supE44 thi-1 gyrA relA1* (Life Technologies, Rockville, Md.), BL21-CodonPlus-RIL F⁻ *ompT hsdS* (r_B⁻ m_B⁻) *dcm⁺ Tet^r gal λ* (DE3) *endA Hte [argU ileY leuW Cam^r]* (an *E. coli* B strain) (Stratagene, LaJolla, Calif.), and Rosetta (DE3) F⁻ *ompT hsdS_B* (r_B⁻ m_B⁻) *gal dem lacYI* (pRARE) (Novagen, Madison, Wis.) were used for recombinant DNA experiments. *E. coli* strains were grown at 37°C (200 rpm) in Luria-Bertani medium. Medium was supplemented with 2% (wt/vol) glucose and antibiotics, including ampicillin (100 mg per liter), kanamycin (100 mg per liter), and chloramphenicol (30 mg per liter), as needed.

DNA isolation and cloning of the *Z. palmae pdc* gene. Plasmid DNA was isolated using a Quantum Prep Plasmid Miniprep kit from Bio-Rad (Hercules, Calif.). DNA fragments were eluted from 0.8% SeaKem GTG agarose (FMC Bioproducts, Rockland, Maine) gels with 1× TAE buffer (40 mM Tris acetate, 2 mM EDTA, pH 8.5) using the QIAquick gel extraction kit from Qiagen (Valencia, Calif.). *Z. palmae* genomic DNA was isolated using the method described by Harwood and Cutting (17).

A degenerate oligonucleotide, 5'-ATGTAYACNGTNGGNATGTAYTNG CNGAR-3', was synthesized (Sigma Genosys, The Woodlands, Tex.) based on

* Corresponding author. Mailing address: Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611-0700. Phone: (352) 392-4095. Fax: (352) 392-5922. E-mail: jmaupin@ufl.edu.

† Journal Series R-08718 of the Florida Agricultural Experiment Station.

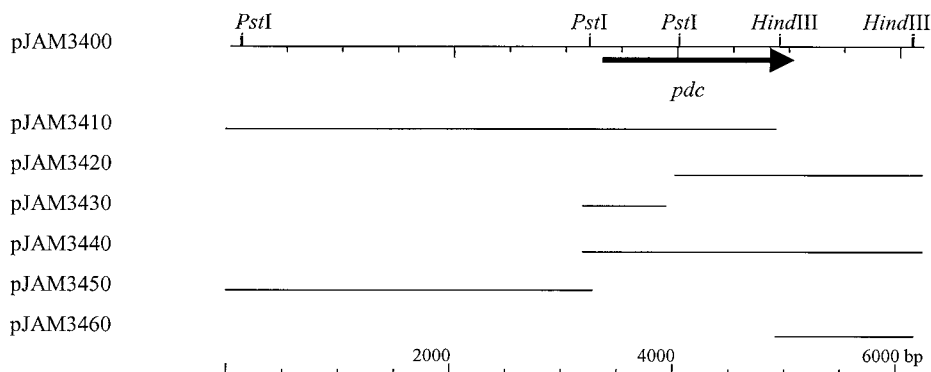


FIG. 1. Cloning of the *Z. palmarum* *pdc* gene. Plasmid pJAM3400 carries a 6-kb *Bam*HI fragment of *Z. palmarum* genomic DNA, blunt ended and ligated into plasmid vector pLITMUS28. The remaining plasmids were derived from pJAM3400 and were used for DNA sequence analysis of a 2.9-kb region that included the *pdc* gene. Plasmid pJAM3440 was used to produce the *Z. palmarum* PDC protein in recombinant *E. coli*. The arrowhead indicates the direction of *pdc* gene transcription.

the N-terminal amino acid sequence of PDC purified from *Z. palmarum* (R is A or G; N is A, C, G, or T; Y is C or T). The oligonucleotide was labeled at the 3' end using terminal transferase with digoxigenin-11-dUTP as recommended by the supplier (Roche Molecular Biochemicals, Indianapolis, Ind.) and used to probe *Z. palmarum* genomic DNA. Based on Southern hybridization, a partial library of 5.5- to 6.5-kb *Bam*HI genomic fragments was generated using plasmid vector pLITMUS28 (New England BioLabs). Colonies were screened with the labeled oligonucleotide. This allowed the isolation of plasmid pJAM3400 containing the complete *pdc* gene on a 6.0-kb *Bam*HI fragment (Fig. 1).

Analysis of DNA and protein sequences. Plasmid pJAM3400 was subcloned as depicted in Fig. 1. This enabled the sequencing of both DNA strands of a 2.9-kb region, including the *pdc* gene, by the Sanger dideoxy method (42) using a LICOR sequencer (DNA Sequencing Facility, Department of Microbiology and Cell Science, University of Florida).

Genpro 5.0 (Hoefler Scientific, San Francisco, Calif.), ClustalW, version 1.81 (46), Treeview, version 1.5 (39), and MultiAln (12) were used for DNA and protein sequence alignments and comparisons. Deduced amino acid sequences were compared to protein sequences in the GenBank, EMBL, and SwissProt databases at the National Center for Biotechnology Information (Bethesda, Md.) using the BLAST network server.

Purification of PDC proteins. An overview of the PDC purifications is presented in Table 1. All procedures were performed at room temperature and all purification buffers contained 1 mM TPP and 1 mM MgCl₂ unless otherwise indicated. *S. ventriculi* PDC was purified from recombinant *E. coli* BL21-Codon-Plus-RIL(pJAM419) and Rosetta (DE3)(pRARE, pJAM419) as previously described (45). *Zymomonas mobilis* PDC was purified from recombinant *E. coli* DH5 α (pLOI276) (11) using thermal treatment, Q-Sepharose, and Superdex 200 procedures as described below for the *Z. palmarum* PDC, with the exception that the *Zymomonas mobilis* PDC was eluted from Q-Sepharose at 0.22 to 0.32 M NaCl.

(i) **Purification of *Z. palmarum* PDC.** *Z. palmarum* and *E. coli* DH5 α (pJAM3440) cells were grown to stationary phase and harvested by centrifugation at 10,000 \times g (10 min, 4°C). Cells (12 to 15 g [wet weight]) were resuspended in 6 volumes of 50 mM sodium phosphate buffer at pH 6.5 containing 1 mM TPP and 1 mM MgCl₂ (buffer A) and passed through a chilled French pressure cell at 20,000 lb/in². Cell debris was removed by centrifugation at 16,000 \times g (20 min, 4°C). Cell lysate was heated to 60°C for 30 min, chilled on ice for 15 min, and centrifuged at 16,000 \times g (30 min at 4°C) to remove denatured proteins. The supernatant was applied to a HiLoad Q-Sepharose 26/10 column (Amersham Pharmacia Biotech, Piscataway, N.J.) equilibrated in buffer A. The column was washed with 200 ml of buffer A and a 400-ml linear 0 to 1 M NaCl gradient was applied at 4.0 ml per min. The peak of PDC activity eluted between 0.4 and 0.5 M NaCl and was pooled. The recombinant *Z. palmarum* PDC from *E. coli* was purified to apparent homogeneity by injecting aliquots (0.5 ml) onto a Superdex 200 10/30 column (Amersham Pharmacia Biotech) equilibrated with buffer A containing 150 mM NaCl and 10% glycerol at 0.2 ml per min.

Purification of native PDC from *Z. palmarum* required additional steps. Heat-treated lysate was applied to a hydroxyapatite CHT5-1 column (Bio-Rad) equilibrated with 10 mM sodium phosphate buffer at pH 6.5 containing 1 mM TPP and 1 mM MgCl₂ (buffer B). The column was washed with 30 ml of buffer B and

a 30-ml linear 10 mM to 1 M sodium phosphate gradient was applied at 0.5 ml per min. The peak of PDC activity eluted at 0.45 to 0.5 M sodium phosphate and was pooled. Aliquots were further purified using a Superdex 200 10/30 column as described above. Solid ammonium sulfate was added to the combined active fractions to 1 M and then loaded onto a phenyl Sepharose 6 Fast Flow (low sub) column (Amersham Pharmacia Biotech) equilibrated with buffer A containing 1 M ammonium sulfate. After a 15-ml wash, the column was developed with a 15-ml decreasing linear gradient of 1 to 0 M ammonium sulfate at 0.5 ml per min. The peak of native PDC activity was eluted between 0.55 and 0.3 M salt.

TABLE 1. Purification of bacterial PDC proteins

PDC protein or purification step ^a	Sp act (U mg ⁻¹)	Purification (fold)
ZpaPDC		
Lysate	0.13	1
Q-Sepharose	1.3	10
(Phenyl Sepharose)	66	510
ZpaPDC-R		
Lysate	32	1
Q-Sepharose	92	3
(Superdex 200)	100	4
ApaPDC		
Lysate	0.62	1
Q-Sepharose	12	18
(Hydroxyapatite)	71	110
ApaPDC-R		
Lysate	6.7	1
Q-Sepharose	58	9
(Superdex 200)	92	14
ZmoPDC-R		
Lysate	6.2	1
Q-Sepharose	57	9
(Superdex 200)	72	12
SvePDC-R		
Lysate	8.7	1
Q-Sepharose	33	2
(Superdex 200)	37	6

^a PDC-R, PDC purified from recombinant *E. coli*. SvePDC-R represents protein purified from *E. coli* Rosetta (DE3)(pRARE, pJAM419). SvePDC-R activity was not detected in the cell lysate of *E. coli* BL21-CodonPlus-RIL(pJAM419). Parentheses indicate the final purification step. Q-Sepharose and Superdex 200 were common purification steps for all of the bacterial PDCs.

(ii) **Purification of *A. pasteurianus* PDC.** PDC was purified from *A. pasteurianus* and *E. coli* ER1648(pJAM304) using heat treatment, Q-Sepharose, Superdex 200, and hydroxyapatite as described for the *Z. palmae* PDC, with the following modifications. Cells were passed through a French pressure cell at 30,000 lb/in². The PDC activity, which was eluted at 0.17 to 0.26 M NaCl from the Q-Sepharose column, was pooled and concentrated by dialysis against PEG8000 prior to loading onto the Superdex 200 column.

The *Z. palmae* and *A. pasteurianus* PDCs were stored for up to 1 month in 10% glycerol under liquid nitrogen without loss of activity. The *S. ventriculi* and *Z. mobilis* PDCs were stored at 4°C.

PDC assay. Protein concentration was determined using Bradford reagent with bovine serum albumin as the standard (Bio-Rad). PDC activity was measured by the coupled assay with baker's yeast ADH (Sigma-Aldrich, St. Louis, Mo.) as previously described (11). The reaction mixture contained 0.15 mM NADH, 5 mM MgCl₂, 0.1 mM TPP, 5 mM pyruvate, and 10 U of ADH in 50 mM sodium citrate buffer at pH 5.0 and was measured at 25°C unless otherwise indicated. One unit of enzyme activity is defined as the amount of enzyme that generates 1 μmol of acetaldehyde per min.

Molecular mass and amino acid sequence determination of the PDC protein. Subunit molecular masses and enzyme purity were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 12% polyacrylamide gels stained with Coomassie blue R-250 (26). Molecular size standards for SDS-PAGE were as follows: phosphorylase b (97.4 kDa), serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa), and lysozyme (14.4 kDa). For the determination of native molecular mass, samples were applied to a Superdex 200 10/30 column equilibrated in 50 mM sodium phosphate buffer at pH 6.5 with 150 mM NaCl. Molecular mass standards included serum albumin (66 kDa), ADH (150 kDa), α-amylase (200 kDa), apoferritin (443 kDa), and thyroglobulin (669 kDa).

The N-terminal sequences of the purified PDC proteins were determined after SDS-PAGE and electroblotting of the proteins onto a polyvinylidene difluoride membrane (PVDF-PLUS) (Micron Separations Inc., Westborough, Mass.). Sequences were determined by automated Edman degradation at the Protein Chemistry Core Facility of the University of Florida Interdisciplinary Center for Biotechnology Research.

Apoenzyme preparation. Purified PDC (0.75 mg per 0.5 ml) was diluted with 1.5 ml of 50 mM sodium phosphate buffer at pH 9.0 and immediately concentrated eightfold using a Centricon YM30 concentrator (Millipore, Bedford, Mass.). After adjusting the volume to 0.5 ml, the protein was incubated at 25°C for 45 min. PDC was purified from unbound cofactors by application to a Superdex 200 10/30 column equilibrated with 50 mM sodium phosphate buffer at pH 9.0 with 10% glycerol and 150 mM NaCl. Immediately after elution, the apoenzyme was diluted fivefold into 50 mM sodium citrate buffer at pH 5.0, stored at 4°C, and used within 16 h for reconstitution assays. TPP was measured by fluorescence after oxidation to thiochrome diphosphate using K₃Fe(CN)₆ in 15% NaOH (15). Excitation and emission wavelengths were 375 and 430 nm, respectively. For reconstitution, apoenzyme (755 ng) was diluted into 1 ml of 50 mM sodium citrate buffer at pH 5.0 with 1 mM TPP and/or 1 mM MgCl₂. The mixture was assayed for PDC activity in the same buffer.

Materials. Biochemicals were purchased from Sigma-Aldrich. Other organic and inorganic analytical grade chemicals were from Fisher Scientific (Atlanta, Ga.). Restriction endonucleases and DNA-modifying enzymes were from New England Biolabs. Digoxigenin-11-dUTP (2'-deoxyuridine-5'-triphosphate coupled by an 11-atom spacer to digoxigenin), alkaline phosphatase-conjugated antibody raised against digoxigenin, and nylon membranes for colony and plaque hybridizations were from Roche Molecular Biochemicals. Positively charged nylon membranes for Southern hybridization were from Ambion (Austin, Tex.).

Nucleotide sequence accession number. The sequence for the *pdC* gene from pJAM3400 was deposited under GenBank accession number AF474145.

RESULTS AND DISCUSSION

Pyruvate decarboxylase (*pdC*) operon of *Z. palmae*. A 6.0-kb *Bam*HI genomic DNA fragment that hybridized to a *pdC*-specific probe was isolated from *Z. palmae* (Fig. 1). A 1.7-kb *Pst*I-to-*Bam*HI subclone of this fragment (pJAM3440) was found to be required for PDC activity in recombinant *E. coli*. Based on DNA sequence, a 1,668-bp open reading frame (ORF) (55 mol% G+C content) was identified in this region that encoded a putative PDC protein. A canonical Shine-Dal-

garno sequence (GGAGG) was 10 bp upstream of the translation start codon (ATG). In addition, a putative -35 and -10 promoter (TTcAct-N₁₋₇-atTAAT, where N is any nucleotide and uppercase letters match the bacterial promoter consensus) was located 16 to 44 bp upstream of the start codon (18). Downstream (21 bp) of the translation stop codon was a 62-bp sequence predicted to form two stem-loop structures followed by a 16-bp AT-rich region, consistent with ρ-independent transcription termination (18). Thus, all four bacterial *pdC* genes, including those from *Z. mobilis*, *S. ventriculi*, and *A. pasteurianus*, are predicted to be independently transcribed from a monocistronic operon (11, 40, 45).

The deduced PDC protein of *Z. palmae* (ZpaPDC) contained 556 amino acids (including the N-terminal methionine) with an anhydrous molecular mass of 60,113 Da. This is similar to the three other bacterial PDCs, which range from 552 to 568 amino acids and 59,830 to 61,809 Da. The ZpaPDC had a calculated pI of 4.93 analogous to the pI determined experimentally (4.87 to 5.3) for the *Z. mobilis* PDC (7, 33). The relatively high alanine content of ZpaPDC (13.1%) was comparable to that of the gram-negative *Z. mobilis* and *A. pasteurianus* PDCs (ZmoPDC and ApaPDC) but almost twofold higher than that of the gram-positive *S. ventriculi* PDC (SvePDC) (6.9%). Multiple amino acid sequence alignment (Fig. 2) and dendrogram cluster analysis (data not shown) revealed that ZpaPDC was most closely related to ApaPDC (72% identity) and highly related to ZmoPDC (62% identity). All three of the gram-negative PDCs clustered in sequence similarity with the plant PDCs (e.g., *Zea mays* PDC, 39 to 40% identity). In contrast, the gram-positive SvePDC was only 30 to 31% identical to the gram-negative PDCs and instead grouped with the majority of filamentous fungi and yeast PDCs (e.g., *Saccharomyces cerevisiae* PDC1, 38% identity). The N-terminal extension common to PDCs of the plant kingdom was absent from all four of the bacterial PDCs. Residues involved in TPP and Mg²⁺ binding, based on the crystal structures of the *Z. mobilis* and yeast PDCs, were conserved in all four bacterial PDCs. In contrast, residues of the yeast PDC1 (Tyr157 and Arg224) that are likely to be involved in substrate activation, based on binding to the pyruvate analog pyruvamide (29), were only noted in the bacterial SvePDC and not in the remaining three bacterial PDCs.

Codon usage. There is considerable interest in using bacterial PDCs to engineer ethanol pathways in a wide variety of organisms (4, 14, 16, 22). Compatible patterns of codon usage are particularly important for functional expression in heterologous systems. A simple tabulation of G+C content can serve as an initial guide. G+C contents for these four *pdC* genes range from 60% for *A. pasteurianus* to 31% for *S. ventriculi*. The average G+C content for *E. coli* ORFs is 52%, identical to that for *Z. mobilis pdC* (52%) and similar to that for *Z. palmae pdC* (55%). For most amino acids, patterns of codon usage for these two organisms were quite similar to each other and to *E. coli*. The *Z. palmae pdC* was functionally expressed at high levels (approximately 1/3 of the protein based on activity) in recombinant *E. coli*. The *Z. mobilis pdC* and *A. pasteurianus pdC* were expressed at lower levels (7 to 9% of the protein based on activity) and *S. ventriculi pdC* was poorly expressed in recombinant *E. coli* (less than 0.3% of the protein based on activity) (45). To produce larger amounts of functional recom-

ZpaPDC	1	-----	MYTVGMYLAERLAC
ApePDC	1	-----	TYTVGMYLAERLVC
ZmoPDC	1	-----	SYTVGMYLAERLVC
ZmaPDC	1	METLLAGNPANGVAKPTCNGVGVLPVANSHTIATPAAAAATLAPAGAT	GRRLAERLVC
ScePDC1	1	-----	MSEITCKYLFERLKC
SvePDC	1	-----	KTIKTYLKRKLC
ZpaPDC	15	IGLKHFFAVAGDYNLVLLD	LLLNKDMEOVYCCNELNCGSAAGYARARGAAAAVTFSSV
ApePDC	16	IGLKHFFAVAGDYNLVLLD	LLLNKDMKQVYCCNELNCGSAAGYARNSGAAAAVTFSSV
ZmoPDC	16	IGLKHFFAVAGDYNLVLLD	LLLNKDMEOVYCCNELNCGSAAGYARARGAAAAVTFSSV
ZmaPDC	61	IGASDVFAVPGDNLTLDDYL	AEPCITLVGCCNELNAGAAAGYARSRGVACAVTFSSV
ScePDC1	17	INNTVTEIPGDNLTLDDK	YVEVEGMRWAGNANELNARAAAGYARITGVCITLTFSSV
SvePDC	16	INLHMEVEVPGDNLGLDY	EDSKDLENVGSCCNELNAGAAAGYARIRGCVLITLTFSSV
ZpaPDC	75	GALSAINATGGAYAENLPVILISG	PNNDYGTGHLHHTIGTIDNYOLEMVAHITCAR
ApePDC	76	GALSAINATGGAYAENLPVILISG	PNNDYGTGHLHHTIGTIDNYOLEMVAHITCAR
ZmoPDC	76	GALSAINATGGAYAENLPVILISG	PNNNDHAGHLHHAATIDNYOLEMVAHITCAR
ZmaPDC	121	GELSVNATGAYENLPVITV	PNNDYGTGHLHHTIGLDPDQILCFQDTITCYQ
ScePDC1	77	GELSAINATGAYENLPVITV	PNNDYGTGHLHHTIGLDPDQVFRMSAHTIETT
SvePDC	76	GELSAINATGAYENLPVITV	HISGVESALVQONRHHHTGADTFERMFETTEFQ
ZpaPDC	135	ESTVSAQAPAKIDHVIRTALRE	KPAYLETACNTA--GAECVPGPNSLRLEVDQT
ApePDC	136	ESTVSAQAPAKIDHVIRTALRE	KPAYLETACNTA--SEPCVPGPNSLRLEVDQT
ZmoPDC	136	ESTVSAQAPAKIDHVIRTALRE	KPAYLETACNTA--SMPCVPGPNSLRLEVDQT
ZmaPDC	181	AETNNLQARECHDTATATALRE	SKPVYLS--CNAGLSHETFSRDPVPMFSPRLSKA
ScePDC1	137	AETNNLQARECHDTATATALRE	SKPVYLS--CNAGLSHETFSRDPVPMFSPRLSKA
SvePDC	136	SEITLQYNAAREEDVLE	YIYQLEGYELPVITVS--KELDEIKPNTLMRSLEK
ZpaPDC	193	SITAAVDAAVWIDRQNVVMLV	SGSKLRAAAAEKQAVLADRLGCAVTMAAARKGFPEP
ApePDC	194	SLKAAVDAAVWIDRQNVVMLV	SGSKLRAAALAAATETLADRLGCAVTMAAARKGFPEP
ZmoPDC	194	SLNAAVETKFAVNRDKVA	LVGSKLRAAGAEAVKFDALGGAVATMAAARKGFPEP
ZmaPDC	241	NLEAVFAAAFLNKAVKPV	GGPKRVARARBAFAADASGYPFAVMAARKGLVFPFH
ScePDC1	196	SEKEVDTITVLAKDAKPV	LADCCSRHDVKAETKKLIDLTQFPFAVTFMKGKGISDQ
SvePDC	192	ILEKFDVVKMLASSKGOH	LADYELRAAEKELEFTAKIFVNTIETAVTS
ZpaPDC	253	HPNDRGLYWGEVSEGAQELVENADA	ILCEVFNQYTVGNSVWPKGDNVADRVIT
ApePDC	254	HAGRGLYWGEVSNFGVQELVET	SDALCEVFNQYSTVGVSGVPEKGNVADPDRVT
ZmoPDC	254	NPLICRSGEVSYPGVKTKKADAA	AEVFNQYSTVGVWEDLEPKKVAAPRSVY
ZmaPDC	301	HPHGLYWGAVSTFCAEVEADAVLEA	EFNDYSVGLSLRREKAVVQDPDRVY
ScePDC1	256	HPHGLYWGAVSTFCAEVEADAVLEA	ELSDNTGSSYSYKTKNVEFHSDLEK
SvePDC	252	NEMBAGLSSEETS	DLVKELECKADILLELVKSIDTACRYNKOVKLEGLTDCR
ZpaPDC	313	FQSEFGSLSTAAALAKAPSRP	----ATTQGTQVVLGTEAEENAPLTNDETRC
ApePDC	314	VDGRAVDCFTLAEALAKAPSRP	----ASAQKSSVETCSLTATSDEAGLTNDEIVRH
ZmoPDC	314	VNCRFPSTHLKDLTRLAQKVS	TGALDFFKSLNAELKKAAPDPSAPVFNARCR
ZmaPDC	361	VVDGPAFCCLMPEFLRALAKRL	ENTTAYDNYRRIFFVDRPFPNKENEPLRVVAEHI
ScePDC1	316	FRNATFEGQKFFVLOKLLINIADAAG	---YKPVAVRTPANAIVPASTPLKQVWVNC
SvePDC	312	RETITGTYKQVAKALTAKTFHN	-----DVKVERAEVKEFVPTDARLTDTRYRQ
ZpaPDC	369	FOITLTSITVAETGDSWFNASRMP	FCGARVELMOWGHIGWSVPAFGNAGSPEP
ApePDC	370	FNALLTSNTTVAETGDSWFNASRMP	LAGARVELMOWGHIGWSVPAFGNAGSPEP
ZmoPDC	374	VEALLTSNTTVAETGDSWFNASRMP	ENGARVEYEMOWGHIGWSVPAFGNAGSPEP
ZmaPDC	421	IKNLLGDAVAETGDSWFNC	---LPECCGYEFOMCGSIGWSVGAITCYAOKAR
ScePDC1	374	HGNLQEGDVVAETGDSWAGT	QTFENNTYGISQLWCSIGTCATLCAAFABEED
SvePDC	366	VEALKPNQVVEETETS	SGACNMFEFGVFGQGSWMSIGTPAVLGTHTDKSR
ZpaPDC	428	---RHIMVGDGGSFOLTAQVAQMR	EPVLIIFLNNRGYVIEAIIH--DGPYNNIKNW
ApePDC	428	---CHIMVGDGGSFOLTAQVAQMR	EPVLIIFLNNRGYVIEAIIH--DGPYNNIKNW
ZmoPDC	433	---RNIMVGDGGSFOLTAQVAQMR	RLKLEPVIIFLNNRGYVIEAIIH--DGPYNNIKNW
ZmaPDC	480	---RVTACVGDGGSFOLTAQVSTMR	RCGQKSLIIFLNNRGYVIEAIIH--DGPYNNIKNW
ScePDC1	434	PKKRVILFVGDGGSFOLTVQEVSTMR	GLKPYEFTVNNQGYTIEKIHGPKQVNLIQW
SvePDC	425	---RNILVGDGGSFOLTVQEVSTMR	QKINTVFTVNNQGYTIEKIHGPKQVNLIQW
ZpaPDC	483	NYAGLELVFND	---GHGELGLKATGEELEGATKKAAL--DNRRGPTLIECNTAQDCTP
ApePDC	483	DYAGLELVFN	---GHGELGLKATPKELTEALARAK--ANTRGPTLIECNTAQDCTP
ZmoPDC	488	DYAGLELVFN	NGGYSGACKGLKAKTCELAEAIKVAL--ANTGPTLIECFGTDCTP
ZmaPDC	535	DYTCGLNAIHN	---SGNCWTQKVRTEEELKEALATVTGAKKCLCELEVAHHTDTRK
ScePDC1	494	DHLSLPTGEGK	---YETHVATTQWDLKTDQKSFNSKIRIIEVTPVDDAQ
SvePDC	482	CYAELEKTLATER	---IQPTCKVTEKELAAVEEIN--KCTCIAFVWVDDMDAPK
ZpaPDC	538	TLVWGRVVAATNSRKEFOA	---
ApePDC	538	MLVWGRVVAATNSRKEFOA	---
ZmoPDC	547	ELVWGRVVAATNSRKEVNLKLL	---
ZmaPDC	591	ELVWGRVVAATNSRKEVNPQ	---
ScePDC1	548	NLVWGRVVAATNSRKEVNPQ	---
SvePDC	538	SLRQVSLVSNMNY	---

FIG. 2. Multiple amino acid sequence alignment of the deduced PDC protein of *Z. palmarum* aligned with selected PDC proteins. Functionally conserved (black highlight) and semiconserved (gray highlight) amino acid residues are indicated. Dashes indicate gaps introduced in alignment. ▼, residues within 0.4 nm of the Mg²⁺ and TPP binding sites of the yeast and *Z. mobilis* PDC proteins; ■, yeast PDC1 residues forming hydrogen bonds with pyruvamide. Double underlined residues are conserved among TPP-dependent enzymes. Abbreviations and GenBank or SwissProt accession numbers: Zpa, *Z. palmarum*, this study; Apa, *A. pasteurianus*, AR368435; Sce, *S. cerevisiae*, P06169; Sve, *S. ventriculi*, AF354297; Zma, *Z. mays*, P28516; Zmo, *Z. mobilis*, P06672.

binant *S. ventriculi* PDC, the use of a modified *E. coli* host containing additional tRNA genes for rare codons such as AGA (arginine), GGA (glycine), AUA (isoleucine), and CUA (leucine) proved essential for high-level protein production (almost 1/4 of the protein based on activity) (Table 2) (45). The pattern of codon usage in *S. ventriculi pdc* may prove particularly useful for the genetic engineering of an ethanol pathway in low-G+C gram-positive bacteria as illustrated by the comparison to *Bacillus subtilis* ORFs (Table 2). In contrast, *A. pasteurianus pdc* may be the best choice for engineering a homo-ethanol pathway into high-G+C cellulolytic bacteria such as *Thermobifida* sp. (23) or *Cellulomonas* sp. (35).

Biochemical properties. An overview of the purification of the four bacterial PDC proteins is presented in Table 1. ZpaPDC was purified to homogeneity from *Z. palmae* and recombinant *E. coli*. For comparison, the three other bacterial PDCs were purified from recombinant *E. coli*. PDC was also purified from *A. pasteurianus*. In general, purification of the native PDCs required additional steps (i.e., phenyl Sepharose and hydroxyapatite chromatography). This was due to the 10- to 250-fold lower levels of PDC activity of native lysate than of that of recombinant cell lysate. All of the PDC proteins had subunit molecular masses of 55 to 60 kDa, as estimated by reducing SDS-PAGE, which were consistent with the masses calculated from the deduced protein sequences.

The N-terminal amino acid sequence of PDC purified from *Z. palmae* (MYTVGMYLAE) was identical to the sequence deduced from the gene and included the N-terminal methionine. Previous studies have demonstrated that the PDC purified from *S. ventriculi* also retains the N-terminal methionine which is amino to a lysine residue (28, 45). In contrast, the N-terminal sequence of PDC purified from *A. pasteurianus* (TYTVGMYLAERL) lacked an N-terminal methionine, suggesting cleavage by a native methionine aminopeptidase. The PDC purified from *Z. mobilis* also appears to be cleaved to expose an N-terminal serine (34). These results are consistent with the highly conserved substrate specificity of methionine aminopeptidases which is dictated by the P₁' residue (5). The N-terminal methionine is only removed if the second residue is small and uncharged. This substrate preference is opposite to the "N-end rule" for protein degradation (47).

Thermostability and temperature optima. Previous studies have revealed that the ZmoPDC is thermostable after heating cell lysate to 60°C in the presence of cofactors (TPP and Mg²⁺) (11). To further address this phenomenon, activity of the purified bacterial PDCs was analyzed at 25°C after exposure to various temperatures in the presence of saturating cofactors (Fig. 3). All three gram-negative PDCs were relatively thermostable and retained 60 to 100% activity after heating to 60°C for 30 min. In contrast, the SvePDC was completely inactivated after exposure to temperatures of 50°C or higher. Although amino acid composition cannot be used as a universal guide to thermostability (30, 31, 48), it is interesting that the three thermostable PDCs contained higher levels of alanine and cysteine and lower levels of phenylalanine. These changes are consistent with an increase in thermostability based on compositional comparisons of homologous enzymes (2, 31). Increased levels of alanine may stabilize gram-negative PDC proteins at high temperature by facilitating the formation of helices and a more compact protein core.

Further analysis of the thermostable PDCs from the gram-negative bacteria revealed an over 2.5-fold increase in activity when assayed at their temperature optimum of 60°C compared to room temperature. Based on Arrhenius plots, these enzymes were estimated to have activation enthalpy in the range of 12.6 to 14.2 kJ · mol⁻¹ and entropy values of -92.8 to -98.2 J · mol⁻¹ · K⁻¹.

Quaternary structure. The ZmoPDC associates as a tetramer of 240 kDa even after removal of the cofactors TPP and Mg²⁺ (15). Likewise, the SvePDC and ZpaPDC associated as tetramers of 240 kDa even after cofactor dissociation (data not shown). Interestingly, the ApaPDC formed both tetramers and octamers (224 and 447 kDa) of similar specific activity and dissociated into dimers (120 kDa) after cofactor extraction. Activity and tetrameric configuration of the ApaPDC apoenzyme was fully restored after addition of Mg²⁺ and TPP, with half-saturation constants of 3.1 and 5.8 μM and Hill constants of 1.17 and 1.22, respectively. The tetrameric configuration observed for all four bacterial PDCs is consistent with the quaternary structure of the majority of eukaryotic PDCs. However, similar to ApaPDC, higher-order structures have been observed for the PDC from the filamentous fungus *Neurospora crassa*, which associates as a homopolymeric filament of 8 to 10 nm (1). Likewise, in plants (maize, pea, and wheat germ) PDC forms complexes up to 1 MDa (24, 27, 32). The dissociation of ApaPDC into dimers after removal of the cofactors is also consistent with eukaryotic PDCs (25).

pH dependence and kinetic constants. Significant differences in pH optima were observed for the bacterial PDC proteins. The ZpaPDC was most active at pH 5.5 to 6.0 (Fig. 4), similar to ZmoPDC, which has a pH optimum of 6.0 (34). The pH optimum of ApaPDC, however, was significantly lower than that observed for the other two gram-negative PDCs and ranged from 5.0 to 5.5 (Fig. 4). The pH optimum for the gram-positive SvePDC (pH 6.3 to 7.6) has been shown to be much broader and more neutral than the gram-negative PDCs (28). This suggests that there are differences in the conformation and/or composition of charged residues at or near the active site.

The eukaryotic PDCs and bacterial SvePDC exhibit positive cooperative kinetics in the presence of the substrate pyruvate (25, 28). The ZpaPDC and ApaPDC, however, exhibited normal Michaelis-Menten kinetics, with K_m values of 0.2 to 0.4 mM for pyruvate and k_{cat} values of 20,500 to 30,500 min⁻¹ (pH optimum, 25°C) (Table 3). The lack of allosteric control with respect to pyruvate which was observed for these gram-negative PDC enzymes is similar to that established for the highly related PDC from *Z. mobilis* (7, 9, 34). As a further comparison, the kinetic constants of all four bacterial PDC proteins were determined at their pH optimum as well as at pH 7.0 (Table 3). This approach was chosen to analyze the bacterial PDC enzymes to more closely reflect the neutral cytosol of recombinant hosts used previously for high-level ethanol production (e.g., *E. coli*, *Erwinia* spp.). These modest pH changes had only a slight effect on the maximum rate of the reaction for all four bacterial PDC proteins. In contrast, pH significantly influenced the affinity of the gram-negative PDCs for pyruvate. Most notably, an over 13-fold increase in K_m was observed when ApaPDC was shifted from its pH optimum to neutral conditions.

TABLE 2. Codon usage of bacterial *pdC* genes compared to the *E. coli* K-12 and *B. subtilis* genomes^a

Amino acid	Codon ^b	Frequency of codon per thousand bases (% codon usage per amino acid)					
		<i>Z. palmae pdC</i>	<i>A. pasteurianus pdC</i>	<i>Z. mobilis pdC</i>	<i>S. ventriculi pdC</i>	<i>E. coli</i> K-12 genome	<i>B. subtilis</i> genome
Ala	GCA	41.3 (32)	17.9 (14)	31.6 (21)	32.5 (47)	20.1 (21)	21.7 (28)
	GCC	39.5 (30)	73.5 (56)	31.6 (21)	0 (0)	25.5 (27)	15.8 (21)
	GCG	3.6 (03)	30.5 (23)	10.5 (07)	1.8 (03)	33.6 (36)	20.2 (26)
	GCU	46.7 (36)	9.0 (07)	75.6 (51)	34.4 (50)	15.3 (16)	19.0 (25)
Arg	AGA	0 (0)	0 (0)	0 (0)	39.8 (100)	2.1 (04)	10.6 (26)
	AGG	0 (0)	1.8 (04)	0 (0)	0 (0)	1.2 (02)	3.9 (09)
	CGC	23.3 (50)	28.7 (64)	14.1 (47)	0 (0)	22.0 (40)	8.6 (21)
	CGG	0 (0)	5.4 (12)	1.8 (06)	0 (0)	5.4 (10)	6.5 (16)
	CGU	23.3 (50)	9.0 (20)	14.1 (47)	0 (0)	20.9 (38)	7.6 (18)
	CGA	0 (0)	0 (0)	0 (0)	0 (0)	3.6 (07)	4.1 (10)
	Asn	AAC	41.3 (77)	32.3 (67)	49.2 (82)	19.9 (42)	21.7 (55)
Asp	AAU	12.6 (23)	16.1 (33)	10.5 (18)	27.1 (58)	17.7 (45)	22.1 (56)
	GAC	37.7 (78)	32.3 (67)	22.8 (54)	3.6 (07)	19.1 (37)	18.8 (36)
Cys	GAU	10.8 (22)	16.1 (33)	19.3 (46)	47.0 (93)	32.1 (63)	33.0 (64)
	UGC	10.8 (60)	32.3 (67)	10.5 (85)	1.8 (20)	6.5 (56)	4.3 (55)
Gln	UGU	7.2 (40)	16.1 (33)	1.8 (15)	7.2 (80)	5.2 (44)	3.6 (45)
	CAA	9.0 (25)	3.6 (11)	0 (0)	28.9 (100)	15.3 (35)	19.8 (51)
	CAG	26.9 (75)	28.7 (89)	17.6 (100)	0 (0)	28.8 (65)	18.7 (49)
Glu	GAA	61.0 (87)	43.0 (80)	63.3 (92)	85.0 (94)	39.4 (69)	48.9 (68)
	GAG	9.0 (13)	10.8 (20)	5.3 (08)	5.4 (06)	17.8 (31)	23.1 (32)
Gly	GGA	1.8 (02)	3.6 (05)	1.8 (02)	50.6 (72)	8.0 (11)	21.7 (31)
	GGC	35.9 (44)	64.5 (82)	19.3 (24)	0 (0)	29.6 (40)	23.4 (34)
	GGG	0 (0)	3.6 (05)	0 (0)	0 (0)	11.1 (15)	11.2 (16)
His	GGU	44.9 (54)	7.2 (09)	59.8 (74)	19.9 (28)	24.7 (34)	12.8 (19)
	CAC	14.4 (67)	10.8 (43)	12.3 (58)	3.6 (22)	9.7 (43)	7.5 (33)
	CAU	7.2 (33)	14.3 (57)	8.8 (42)	12.7 (78)	12.9 (57)	15.2 (67)
Ile	AUA	0 (0)	0 (0)	0 (0)	39.8 (56)	4.4 (07)	9.3 (13)
	AUC	55.7 (89)	34.1 (66)	36.9 (78)	9.0 (13)	25.1 (42)	27.0 (37)
	AUU	7.2 (11)	17.9 (34)	10.5 (22)	21.7 (31)	30.3 (51)	36.8 (50)
Leu	CUA	7.2 (09)	0 (0)	0 (0)	7.2 (09)	3.9 (04)	4.9 (05)
	CUC	5.4 (07)	9.0 (08)	19.3 (22)	0 (0)	11.1 (10)	10.8 (11)
	CUG	59.2 (73)	69.9 (64)	33.4 (38)	0 (0)	52.6 (50)	23.1 (24)
	CUU	3.6 (04)	10.8 (10)	15.8 (18)	14.5 (18)	11.0 (10)	23.0 (24)
	UUA	0 (0)	0 (0)	1.8 (02)	59.7 (73)	13.9 (13)	19.1 (20)
	UUC	5.4 (07)	19.3 (18)	17.6 (20)	0 (0)	13.7 (13)	15.3 (16)
Lys	AAA	26.9 (94)	12.5 (35)	35.1 (56)	63.3 (92)	33.6 (76)	49.1 (70)
	AAG	1.8 (06)	23.3 (65)	28.1 (44)	5.4 (08)	10.3 (24)	20.8 (30)
Met	AUG	26.9 (100)	26.9 (100)	21.1 (100)	27.1 (100)	27.9 (100)	26.9 (100)
	UUC	23.3 (87)	19.7 (78)	28.1 (89)	18.1 (38)	16.6 (43)	14.2 (32)
Phe	UUU	3.6 (13)	5.4 (22)	3.5 (11)	28.9 (62)	22.3 (57)	30.2 (68)
	CCA	1.8 (04)	1.8 (04)	5.3 (11)	18.1 (67)	8.4 (19)	7.0 (19)
	CCC	1.8 (04)	21.5 (50)	3.5 (07)	0 (0)	5.5 (12)	3.3 (09)
	CCG	30.5 (74)	14.3 (33)	28.1 (59)	1.8 (07)	23.2 (53)	16.1 (44)
Ser	CCU	7.2 (17)	5.4 (13)	10.5 (22)	7.2 (27)	7.0 (16)	10.6 (29)
	AGC	12.6 (26)	19.7 (35)	15.8 (37)	12.7 (20)	16.1 (28)	14.2 (23)
	AGU	1.8 (04)	0 (0)	5.3 (13)	12.7 (20)	8.8 (15)	6.6 (10)
	UCA	3.6 (07)	7.2 (13)	1.8 (04)	32.5 (50)	7.2 (12)	14.8 (23)
	UCC	9.0 (19)	19.7 (35)	14.1 (33)	0 (0)	8.6 (15)	8.1 (13)
	UCG	1.8 (04)	7.2 (13)	0 (0)	0 (0)	8.9 (15)	6.4 (10)
	UCU	19.7 (41)	1.8 (03)	5.3 (13)	7.2 (11)	8.5 (15)	12.9 (21)
Thr	ACA	5.4 (09)	10.8 (16)	0 (0)	34.4 (53)	7.1 (13)	22.3 (41)
	ACC	19.7 (34)	30.5 (45)	28.1 (62)	0 (0)	23.4 (43)	8.6 (16)
	ACG	10.8 (19)	23.3 (34)	10.5 (23)	0 (0)	14.4 (27)	14.6 (27)
	ACU	21.5 (38)	3.6 (05)	7.0 (15)	30.7 (47)	9.0 (17)	8.7 (16)
Trp	UGG	16.2 (100)	12.5 (100)	12.3 (100)	5.4 (100)	15.2 (100)	10.3 (100)
Tyr	UAC	18.0 (59)	14.3 (47)	12.3 (32)	7.2 (18)	12.2 (43)	12.0 (35)
	UAU	12.6 (41)	16.1 (53)	26.4 (68)	32.5 (82)	16.2 (57)	22.6 (65)
Val	GUA	14.4 (20)	7.2 (10)	0 (0)	36.2 (46)	10.9 (15)	13.4 (20)
	GUC	34.1 (49)	23.3 (32)	31.6 (40)	0 (0)	15.3 (22)	17.3 (26)
	GUG	3.6 (05)	26.9 (38)	5.3 (07)	0 (0)	26.4 (37)	17.7 (26)
	GUU	18.0 (26)	14.3 (20)	42.2 (53)	43.4 (54)	18.3 (26)	19.2 (28)
Coding DNA mol% G+C		55	60	52	31	52	44

^a GenBank accession numbers are as follows: *A. pasteurianus pdC*, AF368435; *Z. palmae pdC*, AF474145; *Z. mobilis*, M15393; *S. ventriculi*, AF354297.^b Stop codons are not included.

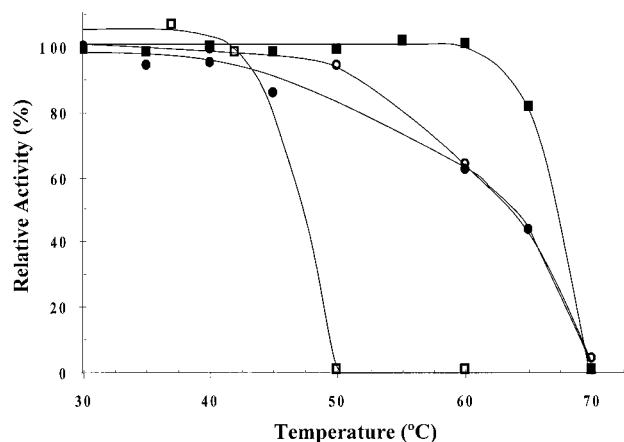


FIG. 3. Thermostability of bacterial PDC enzymes. Recombinant ZmoPDC (●), ZpaPDC (■), ApaPDC (○), and SvePDC (□) proteins were preincubated at the temperatures indicated in 50 mM sodium citrate buffer at pH 5.0 with 1 mM TPP and 1 mM MgCl₂ for 30 min, cooled to 0°C, and assayed for residual activity at 25°C in the same buffer. 100% is the activity after preincubation at 0°C. SvePDC was purified from recombinant *E. coli* BL21-CodonPlus-RIL(pJAM419).

These results suggest that all three gram-negative PDCs have an amino acid residue(s) that enables the enzyme to more efficiently bind pyruvate when protonated (at the pH optimum) than when deprotonated (neutral pH). However, the protonated state of this residue does not influence the overall rate of decarboxylation. Since the changes in K_m for the gram-negative PDCs were observed for pH values of 5 to 7, it is possible that the residue involved in modulating substrate binding for all three enzymes is a histidine. The pK_a for free histidine is 6.04, whereas the pK_a for pyruvate as well as the other ionizable amino acid residues (i.e., Asp, Glu, Lys, and Arg) does not fall within this pH range. Protonation of one or more histidines may be necessary to facilitate substrate binding by forming an ion pair with the carboxyl group of pyruvate or by making the active site more substrate accessible. It is also possible that a

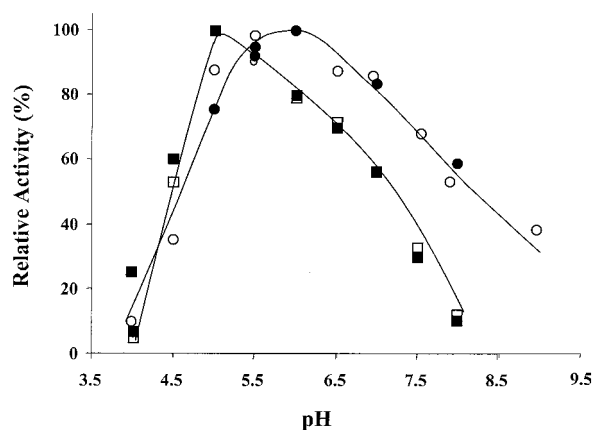


FIG. 4. Effect of pH on activity of recombinant bacterial PDC enzymes ZpaPDC (●) and ApaPDC (■) and native enzymes ZpaPDC (○) and ApaPDC (□). Activity was measured at 25°C in 50 mM sodium citrate buffer from pH 4.0 to 5.0 and in 50 mM sodium phosphate buffer from pH 5.5 to 8.0. 100% is the activity at optimal pH.

TABLE 3. Kinetic parameters of recombinant bacterial PDC enzymes

Enzyme	Type of kinetics ^a	pH optimum	K_m (mM [pH])	V_{max} (U per mg of protein [pH])
ZpaPDC	N	5.5–6.0	0.24 (6.0) 0.71 (7.0)	130 (6.0) 140 (7.0)
ApaPDC	N	5.0–5.5	0.39 (5.0) 5.1 (7.0)	97 (5.0) 79 (7.0)
ZmoPDC	N	6.0	0.43 (6.0) 0.94 (7.0)	100 (6.0) 78 (7.0)
SvePDC ^b	S	6.3–6.7	5.7 (6.5) 4.0 (7.0)	45 (6.5) 35 (7.0)

^a N, normal Michaelis-Menten kinetics; S, sigmoidal kinetics.

^b SvePDC was purified from *E. coli* Rosetta (DE3) (pRARE, pJAM419).

different residue(s) is involved in modulating substrate binding; however, its pK_a is modified by the protein environment.

The PDC enzyme of *Z. mobilis* has been under intense biochemical and structural investigation. The effect of pH on k_{cat}/K_m compared to k_{cat} has been well documented for this enzyme by titration curves, with pK_a values estimated at 6.23 to 6.45 for the residue involved in modulating substrate binding (21, 43). The deprotonation of His113 has been suggested to lead to conformational changes that result in a flexible loop (residues 105 to 112) to close over the active site during catalysis (43). Both His113 and His114 are conserved in all PDC proteins that have been characterized. Thus, His113 is a logical candidate for the observed pH-dependent changes in K_m observed for all three gram-negative PDCs. However, five crucial residues with ionizable side chains protruding into the active site of *Z. mobilis* PDC (Asp27, Glu50, His113, His114, and Glu473) were modified to residues that were nonionizable or had altered pK_a values (21). The pH dependence of k_{cat}/K_m was relatively unaffected by these modifications, suggesting that these residues, including His113, may not be involved. Interestingly, in this same study modification of His113 to a residue with significantly higher pK_a value (i.e., Arg and Lys) increased the affinity of ZmoPDC for pyruvate over 20-fold. This result reinforces the existing notion that the positively charged form of His113 keeps the active site open for substrate binding.

Conclusions. This study establishes the presence of a functional PDC gene in *Z. palmarum*. Together with ADH, this PDC is presumed to function as the primary ethanol pathway in *Z. palmarum*. Based on primary amino acid sequence, PDC from *Z. palmarum* was highly related to the two other PDCs from the gram-negative bacteria and to plants. The four known bacterial PDC enzymes were purified and compared. All respective genes were expressed as active recombinant enzymes in *E. coli*. Efficient functional expression of *S. ventriculi* pdc required additional tRNA genes. Purified *Z. palmarum* PDC exhibited the highest specific activity and lowest K_m for pyruvate. Only the PDCs from gram-negative bacteria (*Z. palmarum*, *Z. mobilis*, and *A. pasteurianus*) were thermostable, which is useful during purification. Thermostability was correlated with higher levels of alanine and cysteine and lower levels of phenylalanine. The three gram-negative PDCs displayed Michaelis-Menten kinetics and a relatively high affinity for pyruvate (0.2 to 0.4 mM) at

optimal pH. In contrast, *S. ventriculi* PDC exhibited allosteric activation by pyruvate (K_m of 4 to 6 mM) similar to most eukaryotic PDCs. Patterns of codon usage differed widely among the four *pdc* genes, with G+C contents ranging from 60% for *A. pasteurianus pdc* to 31% for *S. ventriculi pdc*. Differences in kinetic properties, thermal stability, and codon usage among these *pdc* genes may provide particular advantages for the metabolic engineering of ethanol pathways in diverse recombinant hosts.

ACKNOWLEDGMENTS

We thank K. T. Shanmugam for assistance with the research and the manuscript. We also thank Francis Davis and Jack Shelton for DNA sequencing.

This work was supported in part by the USDOE National Renewable Energy Laboratory (ZDH-9-29009-04), the USDOE Energy Biosciences Program (FG02-96ER20222), and the Florida Agricultural Experiment Station.

REFERENCES

- Alvarez, M. E., A. L. Rosa, E. D. Temporini, A. Wolstenholme, G. Panzetta, L. Patrito, and H. J. Maccioni. 1993. The 59-kDa polypeptide constituent of 8–10-nm cytoplasmic filaments in *Neurospora crassa* is a pyruvate decarboxylase. *Gene* **130**:253–258.
- Argos, P., M. G. Rossmann, U. M. Grau, H. Zuber, G. Frank, and J. D. Tratschin. 1979. Thermal stability and protein structure. *Biochemistry* **18**:5698–5703.
- Atkinson, D. E. 1956. Hydrogen metabolism in *Acetobacter peroxydans*. *J. Bacteriol.* **72**:189–194.
- Barbosa, M. F. S., and L. O. Ingram. 1994. Expression of the *Zymomonas mobilis* alcohol dehydrogenase II (*adhB*) and pyruvate decarboxylase (*pdc*) genes in *Bacillus*. *Curr. Microbiol.* **28**:279–282.
- Bradshaw, R. A., W. W. Brickey, and K. W. Walker. 1998. N-terminal processing: the methionine aminopeptidase and N^α-acetyl transferase families. *Trends Biochem. Sci.* **23**:263–267.
- Braü, B., and H. Sahn. 1986. Cloning and expression of the structural gene for pyruvate decarboxylase of *Zymomonas mobilis* in *Escherichia coli*. *Arch. Microbiol.* **144**:296–301.
- Bringer-Meyer, S., K.-L. Schimz, and H. Sahn. 1986. Pyruvate decarboxylase from *Zymomonas mobilis*. Isolation and partial characterization. *Arch. Microbiol.* **146**:105–110.
- Bucher, M., R. Braendle, and C. Kuhlemeier. 1994. Ethanol fermentation in transgenic tobacco expressing *Zymomonas mobilis* pyruvate decarboxylase. *EMBO J.* **13**:2755–2763.
- Candy, J. M., and R. G. Duggleby. 1998. Structure and properties of pyruvate decarboxylase and site-directed mutagenesis of the *Zymomonas mobilis* enzyme. *Biochim. Biophys. Acta* **1385**:323–338.
- Candy, J. M., R. G. Duggleby, and J. S. Mattick. 1991. Expression of active yeast pyruvate decarboxylase in *Escherichia coli*. *J. Gen. Microbiol.* **137**:2811–2815.
- Conway, T., Y. A. Osman, J. I. Konnan, E. M. Hoffmann, and L. O. Ingram. 1987. Promoter and nucleotide sequences of the *Zymomonas mobilis* pyruvate decarboxylase. *J. Bacteriol.* **169**:949–954.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**:10881–10890.
- DeLey, J. 1958. Studies on the metabolism of *Acetobacter peroxydans*. Part I. General properties and taxonomic position of the species. *Antonie Leeuwenhoek J. Microbiol. Serol.* **24**:281–297.
- Deng, M. D., and J. R. Coleman. 1999. Ethanol synthesis by genetic engineering in cyanobacteria. *Appl. Environ. Microbiol.* **65**:523–528.
- Diefenbach, R. J., and R. G. Duggleby. 1991. Pyruvate decarboxylase from *Zymomonas mobilis*. Structure and re-activation of apoenzyme by the cofactors thiamin diphosphate and magnesium ion. *Biochem. J.* **276**:439–445.
- Gold, R. S., M. M. Meagher, S. Tong, R. W. Hutkins, and T. Conway. 1996. Cloning and expression of the *Zymomonas mobilis* "production of ethanol" genes in *Lactobacillus casei*. *Curr. Microbiol.* **33**:256–260.
- Harwood, C. R., and S. M. Cutting. 1990. *Molecular biological methods for Bacillus*. Wiley Interscience, New York, N.Y.
- Hénaut, A., and A. Danchin. 1996. Analysis and predictions from *Escherichia coli* sequences, or *E. coli* in silico, p. 2047–2066. In F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella* cellular and molecular biology. ASM Press, Washington, D.C.
- Hoppner, T. C., and H. W. Doelle. 1983. Purification and kinetic characteristics of pyruvate decarboxylase and ethanol dehydrogenase from *Zymomonas mobilis* in relation to ethanol production. *Eur. J. Appl. Microbiol. Biotechnol.* **17**:152–157.
- Horn, S. J., I. M. Aasen, and K. Østgaard. 2000. Production of ethanol from mannitol by *Zymobacter palmae*. *J. Ind. Microbiol. Biotechnol.* **24**:51–57.
- Huang, C. Y., A. K. Chang, P. F. Nixon, and R. G. Duggleby. 2001. Site-directed mutagenesis of the ionizable groups in the active site of *Zymomonas mobilis* pyruvate decarboxylase: effect on activity and pH dependence. *Eur. J. Biochem.* **268**:3558–3565.
- Ingram, L. O., H. C. Aldrich, A. C. Borges, T. B. Causey, A. Martinez, F. Morales, A. Saleh, S. A. Underwood, L. P. Yomano, S. W. York, J. Zaldivar, and S. Zhou. 1999. Enteric bacterial catalysts for fuel ethanol production. *Biotechnol. Prog.* **15**:855–866.
- Irwin, D. C., S. Zhang, and D. B. Wilson. 2000. Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. *Eur. J. Biochem.* **267**:4988–4997.
- Kenworthy, P., and D. D. Davies. 1976. Kinetic aspects of regulation of pyruvic decarboxylase. *Phytochemistry* **15**:279–282.
- König, S. 1998. Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. *Biochim. Biophys. Acta* **1385**:271–286.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
- Lee, T. C., and P. J. Langston-Unkefer. 1985. Pyruvate decarboxylase from *Zea mays* L. I. Purification and partial characterization from mature kernels and anaerobically treated roots. *Plant Physiol.* **79**:242–247.
- Lowe, S. E., and J. G. Zeikus. 1992. Purification and characterization of pyruvate decarboxylase from *Sarcina ventriculi*. *J. Gen. Microbiol.* **138**:803–807.
- Lu, G., D. Dobritzsch, S. Baumann, G. Schneider, and S. König. 2000. The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study. *Eur. J. Biochem.* **267**:861–868.
- Matthews, B. W. 1993. Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* **62**:139–160.
- Mozhaev, V. V., and K. Martinek. 1984. Structure-stability relationships in proteins: new approaches to stabilizing enzymes. *Enzyme Microb. Technol.* **6**:50–59.
- Mücke, U., S. König, and G. Hübner. 1995. Purification and characterisation of pyruvate decarboxylase from pea seeds (*Pisum sativum* cv. Miko). *Biol. Chem. Hoppe-Seyler* **376**:111–117.
- Neale, A. D., R. K. Scopes, R. E. Wettenhall, and N. J. Hoogenraad. 1987. Nucleotide sequence of the pyruvate decarboxylase gene from *Zymomonas mobilis*. *Nucleic Acids Res.* **15**:1753–1761.
- Neale, A. D., R. K. Scopes, R. E. Wettenhall, and N. J. Hoogenraad. 1987. Pyruvate decarboxylase of *Zymomonas mobilis*: isolation, properties, and genetic expression in *Escherichia coli*. *J. Bacteriol.* **169**:1024–1028.
- Notenboom, V., C. Birsan, R. A. Warren, S. G. Withers, and D. R. Rose. 1998. Exploring the cellulose/xylan specificity of the beta-1,4-glycanase *cex* from *Cellulomonas fimi* through crystallography and mutation. *Biochemistry* **37**:4751–4758.
- Okamoto, T., H. Taguchi, K. Nakamura, and H. Ikenaga. 1994. Production of ethanol from maltose by *Zymobacter palmae* fermentation. *Biosci. Biotechnol. Biochem.* **58**:1328–1329.
- Okamoto, T., H. Taguchi, K. Nakamura, H. Ikenaga, H. Kuraishi, and K. Yamasato. 1993. *Zymobacter palmae* gen. nov., sp. nov., a new ethanol-fermenting peritrichous bacterium isolated from palm sap. *Arch. Microbiol.* **160**:333–337.
- Or, E., J. Baybik, A. Sadka, and A. Ogdovitch. 2000. Fermentative metabolism in grape berries: isolation and characterization of pyruvate decarboxylase cDNA and analysis of its expression throughout berry development. *Plant Sci.* **156**:151–158.
- Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**:357–358.
- Raj, K. C., L. O. Ingram, and J. A. Maupin-Furlow. 2001. Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactic acid by *Acetobacter pasteurianus*. *Arch. Microbiol.* **176**:443–451.
- Reid, M. F., and C. A. Fewson. 1994. Molecular characterization of microbial alcohol dehydrogenases. *Crit. Rev. Microbiol.* **20**:13–56.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- Schenk, G., F. J. Leeper, R. England, P. F. Nixon, and R. G. Duggleby. 1997. The role of His113 and His114 in pyruvate decarboxylase from *Zymomonas mobilis*. *Eur. J. Biochem.* **248**:63–71.
- Tadege, M., R. Brandle, and C. Kuhlemeier. 1998. Anoxia tolerance in tobacco roots: effect of overexpression of pyruvate decarboxylase. *Plant J.* **14**:327–335.
- Talarico, L. A., L. O. Ingram, and J. A. Maupin-Furlow. 2001. Production of the gram-positive *Sarcina ventriculi* pyruvate decarboxylase in *Escherichia coli*. *Microbiology* **147**:2425–2435.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
- Tobias, J. W., T. E. Shrader, G. Rocap, and A. Varshavsky. 1991. The N-end rule in bacteria. *Science* **254**:1374–1377.
- Vieille, C., and G. J. Zeikus. 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**:1–43.