Production of the Gram-positive *Sarcina ventriculi* pyruvate decarboxylase in *Escherichia coli*

Lee A. Talarico, Lonnie O. Ingram and Julie A. Maupin-Furlow

Author for correspondence: Julie A. Maupin-Furlow. Tel: +1 352 392 4095. Fax: +1 352 392 5922.
e-mail: jmaupin@ufl.edu

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

***Sarcina ventriculi* grows in a remarkable range of mesophilic environments from pH 2 to pH 10. During growth in acidic environments, where acetate is toxic, expression of pyruvate decarboxylase (PDC) serves to direct the flow of pyruvate into ethanol during fermentation. PDC is rare in bacteria and absent in animals, although it is widely distributed in the plant kingdom. The *pdc* gene from *S. ventriculi* is the first to be cloned and characterized from a Gram-positive bacterium. In *Escherichia coli*, the recombinant *pdc* gene from *S. ventriculi* was poorly expressed due to differences in codon usage that are typical of low-G+C organisms. Expression was improved by the addition of supplemental codon genes and this facilitated the 136-fold purification of the recombinant enzyme as a homo-tetramer of 58 kDa subunits. Unlike *Zymomonas mobilis* PDC, which exhibits Michaelis–Menten kinetics, *S. ventriculi* PDC is activated by pyruvate and exhibits sigmoidal kinetics similar to fungal and higher plant PDCs. Amino acid residues involved in the allosteric site for pyruvate in fungal PDCs were conserved in *S. ventriculi* PDC, consistent with a conservation of mechanism. Cluster analysis of deduced amino acid sequences confirmed that *S. ventriculi* PDC is quite distant from *Z. mobilis* PDC and plant PDCs. *S. ventriculi* PDC appears to have diverged very early from a common ancestor which included most fungal PDCs and eubacterial indole-3-pyruvate decarboxylases. These results suggest that the *S. ventriculi pdc* gene is quite ancient in origin, in contrast to the *Z. mobilis pdc*, which may have originated by horizontal transfer from higher plants.**

Keywords: pyruvate metabolism, thiamin pyrophosphate-dependent enzymes, alcohol dehydrogenase, ethanol fermentation, allosteric regulation

INTRODUCTION

Metabolic engineering of bacteria for high-level ethanol production incorporates the use of genes encoding pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH) to channel the flow of pyruvate to ethanol (Ingram *et al.*, 1999). There has been success in engineering Gram-negative bacteria for ethanol production using the *pdc* and *adh* genes from *Zymomonas*

mobilis. However, similar approaches in engineering Gram-positive bacteria have been limited, possibly due to differences in transcription, translation or protein degradation between Gram-negative and Gram-positive bacteria (Barbosa & Ingram, 1994; Gold *et al.*, 1996). Thus, the isolation of *pdc* operons from Gram-positive bacteria, such as *Sarcina ventriculi*, is likely to broaden the diversity of bacteria which can be modified for high-level ethanol production from biomass under conditions such as low pH, high temperature and high salt.

PDC (EC 4.1.1.1) serves as the key enzyme in all homo-ethanol fermentations. This enzyme catalyses the non-oxidative decarboxylation of pyruvate to acetaldehyde and carbon dioxide using Mg^{2+} and thiamin pyrophosphate (TPP) as cofactors. Acetaldehyde is subsequently reduced to ethanol by alcohol dehydrogenase

Abbreviations: ADH, alcohol dehydrogenase; E1, decarboxylase component of pyruvate dehydrogenase; IPD, indole-3-pyruvate decarboxylase; PDC, pyruvate decarboxylase; TPP, thiamin pyrophosphate; TK, transketolase.

The GenBank accession number for the sequence reported in this paper is AF354297.

(ADH, EC 1.1.1.1) during the regeneration of NAD⁺. PDC is widespread among plants, absent in animals, and rare in prokaryotes. This is in contrast to ADH, which is widespread among eukaryotes and prokaryotes. Prior to this study, the only bacterial *pdc* gene described was from the Gram-negative α -proteobacterium *Zymomonas mobilis* (Braü & Sahm, 1986; Conway *et al.*, 1987; Neale *et al.*, 1987a; Reynen & Sahm, 1988). *Z. mobilis* PDC has been purified to homogeneity, crystallized, and extensively characterized (Candy & Duggleby, 1998). PDC has also been purified from one other bacterium, *Sarcina ventriculi* (Lowe & Zeikus, 1992), an unusual Gram-positive organism.

S. ventriculi is an obligate anaerobe that grows from pH 2 to pH 10, fermenting hexose and pentose sugars to produce acetate, ethanol, formate, CO₂ and H₂ (Goodwin & Zeikus, 1987; Stephenson & Dawes, 1971). In this organism, the relative production of ethanol and acetate varies with environmental pH. Under acidic conditions, where acetic acid is toxic to cells, ethanol is the primary product (Goodwin & Zeikus, 1987). At neutral pH and above, a near-equimolar mixture of ethanol and acetate is produced, with low levels of formate (Lowe & Zeikus, 1991). These changes in fermentation profiles have been attributed to changes in the levels of two enzymes that metabolize pyruvate: PDC and pyruvate dehydrogenase (Goodwin & Zeikus, 1987; Lowe & Zeikus, 1991).

The properties of the *S. ventriculi* PDC are very different from those of the *Z. mobilis* enzyme. Unlike the Michaelis–Menten kinetics of *Z. mobilis* PDC (Hoppner & Doelle, 1983; Neale *et al.*, 1987a), the *S. ventriculi* enzyme was activated by pyruvate (Lowe & Zeikus, 1992), similar to PDCs from yeast and higher plants. *S. ventriculi* PDC was reported to have an unusually high K_m for pyruvate (13 mM), compared to K_m values of 0.3–4.4 mM for other PDCs (Hoppner & Doelle, 1983; Neale *et al.*, 1987a; Ullrich & Donner, 1970). The phenylalanine content of purified *S. ventriculi* PDC was four- to fivefold greater than that of other PDCs, suggesting significant differences in primary structure (Lowe & Zeikus, 1992).

To further examine the unusual nature of the *S. ventriculi* PDC, the *pdc* gene encoding it was cloned, sequenced, and expressed in recombinant *E. coli*. This approach provided a genetic determination of primary amino acid sequence and facilitated PDC purification for further kinetic and biophysical characterization.

METHODS

Materials. Biochemicals were purchased from Sigma. Other organic and inorganic analytical-grade chemicals were from Fisher Scientific. Restriction endonucleases and DNA-modifying enzymes were from New England BioLabs. Oligonucleotides were from Sigma-Genosys. 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 membranes for Southern hybridization were from Ambion.

Bacterial strains and media. Table 1 lists the *E. coli* strains used in this study, including strains TB-1 and DH5 α , which were used for routine recombinant DNA experiments. *E. coli* strain SE2309 was used to create a subgenomic DNA library in plasmid pBR322. *E. coli* strains ER1647, LE392 and BM25.8 were used in conjunction with λ BlueSTAR for a subgenomic DNA library. *E. coli* strains BL21(DE3), BL21-CodonPlus-RIL and BL21-CodonPlus-RIL/pSJS1240 were used to examine the expression of the *S. ventriculi pdc* gene from plasmid pJAM419. *E. coli* strains were grown in Luria–Bertani (LB) medium and supplemented with antibiotics as appropriate (30 mg chloramphenicol l⁻¹, 100 mg carbenicillin l⁻¹, 100 mg ampicillin l⁻¹ and/or 50 mg spectinomycin l⁻¹). *S. ventriculi* strain Goodsir was cultivated as described previously (Goodwin & Zeikus, 1987).

DNA isolation. Plasmid DNA was isolated and purified using a Quantum Prep Plasmid Miniprep Kit from Bio-Rad. DNA fragments were eluted from 0.8% SeaKem GTG agarose (FMC Bioproducts) using either Ultrafree-DA filters from Millipore or the QIAquick gel extraction kit from Qiagen. *S. ventriculi* genomic DNA was isolated and purified as described by Harwood & Cutting (1990).

Cloning of the *S. ventriculi pdc* gene. A degenerate oligonucleotide, 5'-AARGARGTNAAYGTNGARCAYATGTT-YGGNGT-3' (R is A or G; N is A, C, G or T; Y is C or T), was synthesized based on the N-terminal amino acid sequence of PDC purified from *S. ventriculi* (Lowe & Zeikus, 1992). This oligonucleotide was labelled at the 3'-end using terminal transferase with digoxigenin-11-dUTP and dATP as recommended by the supplier (Roche Molecular Biochemicals) and was used to screen genomic DNA from *S. ventriculi*.

For Southern analysis, genomic DNA was digested with *Bgl*I, *Eco*RI or *Hinc*II, separated by 0.8% agarose electrophoresis, and transferred to positively charged nylon membranes (Southern, 1975). Membranes were equilibrated at 58 °C for 2 h in 5 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate) containing 1% blocking reagent (Roche Molecular Biochemicals), 0.1% *N*-lauroylsarcosine, and 0.02% SDS. After the probe (0.2 pmol ml⁻¹) and poly(A) (0.01 mg ml⁻¹) were added, membranes were incubated at 58 °C for 18.5 h. Membranes were washed twice with 2 $\times$ SSC containing 0.1% SDS (5 min per wash) at 25 °C and twice with 0.5 $\times$ SSC containing 0.1% SDS (15 min per wash) at 58 °C. Signals were visualized using colorimetric detection according to the supplier (Roche Molecular Biochemicals).

For generation of a subgenomic library in plasmid pBR322, *S. ventriculi* chromosomal DNA was digested with *Hinc*II and fractionated by electrophoresis. The 2.5–3.5 kb *Hinc*II DNA fragments were ligated into the *Eco*RV site of pBR322 and transformed into *E. coli* SE2309. Colonies were screened with the degenerate oligonucleotide by colorimetric detection. By this method, plasmid pJAM400, which carries a *Hinc*II fragment containing 1350 bp of the *pdc* gene was isolated.

The λ BlueSTAR Vector System (Novagen) was used to create an additional subgenomic library to facilitate isolation of the full-length *pdc* gene from *S. ventriculi*. Genomic DNA was digested with *Bcl*I, separated by electrophoresis in 0.8% agarose, and the 6.5–8.5 kb fragments were ligated with the λ BlueSTAR *Bam*HI arms. *In vitro* packaging and plating of phage was performed according to the supplier (Novagen). A DNA probe was generated using an 800 bp *Eco*RI fragment of the *pdc* gene from pJAM400 that was labelled with digoxigenin-11-dUTP using the random-primed method as

Table 1. Strains and plasmids used in this study

Strain or plasmid	Phenotype, genotype, description, PCR primers	Source
<i>S. ventriculi</i> Goodsir	ATCC 55887	American Type Culture Collection
<i>E. coli</i> TB-1	F ⁻ <i>ara</i> Δ(<i>lac-proAB</i>) <i>rpsL</i> (Str ^r) [<i>φ80lac</i> Δ(<i>lacZ</i>)M15] <i>thi</i> <i>hsdR</i> (r _k ⁻ m _k ⁺)	New England Biolabs
DH5α	F ⁻ <i>recA1</i> <i>endA1</i> <i>hsdR17</i> (r _k ⁻ m _k ⁺) <i>supE44</i> <i>thi-1</i> <i>gyrA</i> <i>relA1</i>	Life Technologies
SE2309	F ⁻ <i>e14</i> ⁻ (McrA ⁻) <i>endA1</i> <i>supE44</i> <i>thi-1</i> <i>relA1</i> ? <i>rfdD1</i> ? <i>spoT1</i> ? Δ(<i>mcrC-mrr</i>)114::IS10 <i>pcnB80</i> <i>zad2084</i> ::Tn10	Provided by K. T. Shanmugam (Univ. of Florida)
ER1647	F ⁻ <i>fhuA2</i> Δ(<i>lacZ</i>) <i>r1</i> <i>supE44</i> <i>trp31</i> <i>mcrA1272</i> ::Tn10(Tet ^r) <i>his-1</i> <i>rpsL104</i> (Str ^r) <i>xyl-7</i> <i>mtl-2</i> <i>metB1</i> Δ(<i>mcrC-mrr</i>)102::Tn10(Tet ^r) <i>recD1040</i>	Novagen
LE392	F ⁻ <i>e14</i> ⁻ (McrA ⁻) <i>hsdR514</i> (r _k ⁻ m _k ⁺) <i>supE44</i> <i>supF58</i> <i>lacY1</i> or Δ(<i>lacIZY</i>) 6 <i>galK2</i> <i>galT22</i> <i>metB1</i> <i>trpR55</i>	Novagen
BM25.8	<i>supE</i> <i>thi</i> Δ(<i>lac-proAB</i>) [F' <i>traD36</i> <i>proA</i> ⁺ B ⁺ <i>lacI</i> ^q ZAM15] λimm ⁴³⁴ (Kan ^R) P1 (Cm ^R) <i>hsdR</i> (r _{k12} ⁻ m _{k12} ⁺)	Novagen
BL21-CodonPlus-RIL	F ⁻ <i>ompT</i> <i>hsdS</i> (r _B ⁻ m _B ⁻) <i>dcm</i> ⁺ Tet ^r <i>gal</i> λ (DE3) <i>endA</i> Hte [<i>argU</i> <i>ileY</i> <i>leuW</i> Cam ^r] (an <i>E. coli</i> B strain)	Stratagene
BL21(DE3)	F ⁻ <i>ompT</i> <i>gal</i> [<i>dcm</i>] [<i>lon</i>] <i>hsdS</i> _B (r _B ⁻ m _B ⁻ ; an <i>E. coli</i> B strain) with DE3, a λ prophage carrying the T7 RNA polymerase gene	Novagen
Plasmids		
pBR322	Ap ^r ; cloning vector	New England Biolabs
pUC19	Ap ^r ; cloning vector	New England Biolabs
pBlueSTAR-1	Ap ^r ; plasmid derived from λBlueSTAR-1	Novagen
pET21d	Ap ^r ; expression vector	Novagen
pJS1240	Sp ^r ; derivative of pACYC184 with <i>E. coli</i> <i>ileX</i> and <i>argU</i>	Kim <i>et al.</i> (1998)
pJAM400	Ap ^r ; two 3 kb <i>HincII</i> fragments of <i>S. ventriculi</i> genomic DNA ligated into the <i>EcoRV</i> site of pBR322; carries only 1350 bp of <i>pdC</i>	This study
pJAM410	Ap ^r ; 7 kb fragment of <i>S. ventriculi</i> genomic DNA with the complete <i>pdC</i> gene in pBlueSTAR-1	This study
pJAM411	Ap ^r ; 6 kb <i>SmaI</i> fragment from pJAM410 ligated into the <i>HincII</i> site of pUC19; carries 2103 bp of pBlueSTAR-1 vector and 4 kb of <i>S. ventriculi</i> genomic DNA with the complete <i>pdC</i> gene	This study
pJAM413	Ap ^r ; 3 kb <i>SacII</i> fragment of pJAM411 with the complete <i>pdC</i> gene in the <i>HincII</i> site of pUC19	This study
pJAM419	Ap ^r ; 1.7 kb <i>BspHI</i> -to- <i>XhoI</i> fragment generated by PCR amplification using pJAM410 as template, oligo 1 5'-ggcctcatgaaataacaattgcag-3', and oligo 2 5'-gcgggctcgagattagtagttatttg-3' (<i>BspHI</i> and <i>XhoI</i> sites in bold); ligated with an <i>NcoI</i> -to- <i>XhoI</i> fragment of pET21d; carries complete <i>pdC</i> gene with its start codon positioned 8 bp downstream of the Shine-Dalgarno sequence of pET21d	This study

recommended by the supplier (Roche Molecular Biochemicals). Plaques were screened using colorimetric detection. Cre-*loxP*-mediated subcloning was used to circularize the DNA of the positive plaques by plating λBlueSTAR phage with *E. coli* BM25.8, which expresses Cre recombinase (Novagen). The circularized plasmid pJAM410 was then purified and electroporated into *E. coli* DH5α.

For generation of a *pdC* expression vector, the promoterless *pdC* gene was subcloned into pET21d after amplification from pJAM413 (Table 1) by PCR. Primers were designed for directional insertion using *BspHI* (oligo 1) and *XhoI* (oligo 2) restriction sites. The resulting fragment was ligated into compatible *NcoI* and *XhoI* sites of pET21d (Novagen) to produce pJAM419 (Fig. 1). The fidelity of the *pdC* gene was confirmed by DNA sequencing.

Nucleotide and protein sequence analyses. DNA fragments of plasmids pJAM400 and pJAM410 (Fig. 1) were subcloned into plasmid vector pUC19 for determining the *pdC* sequence using the dideoxy termination method (Sanger *et al.*, 1977) and a LI-COR automated DNA sequencer (DNA Sequencing Facility, Department of Microbiology and Cell Science, University of Florida). The nucleotide sequence of the *S. ventriculi* *pdC* gene and surrounding DNA is deposited in the GenBank database (accession number AF354297).

Genepro 5.0 (Riverside Scientific), CLUSTAL W version 1.81 (Thompson *et al.*, 1994), Treeview version 1.5 (Page, 1996), and MultiAln (Corpet, 1988) were used for DNA and/or protein sequence alignments and comparisons. Deduced amino acid sequences were compared to protein sequences available in the GenBank, EMBL and SWISS-PROT databases

Fig. 1. Partial map of restriction endonuclease sites for a 7 kb *BclI* genomic DNA fragment from *S. ventriculi*. Plasmids used in this study include pJAM410, which carries the complete 7 kb *BclI* fragment, and pJAM400, pJAM411 and pJAM413, which were used for DNA sequence analysis. Plasmid pJAM419 was used for expression of the *S. ventriculi pdc* gene in recombinant *E. coli*. The location of the *pdc* gene and ORF1* (a partial open reading frame of 177 amino acids with no apparent start codon) are shown directly below the physical map, with large arrows indicating the direction of transcription. The dashed line below the physical map indicates the 3886 bp *HincII*-to-*HincII* region sequenced.

at the National Center for Biotechnology Information (Bethesda, MD) using the BLAST network server (Altschul *et al.*, 1990). The dense alignment surface (DAS) method was used for the prediction of transmembrane α -helices (Cserzo *et al.*, 1997).

Production of *S. ventriculi* PDC in recombinant *E. coli*. Plasmid pJAM419 was transformed into *E. coli* BL21-CodonPlus-RIL containing plasmid pSJS1240 (Table 1). Expression of the *pdc* gene in this plasmid is regulated by the bacteriophage T7 RNA polymerase promoter system (Novagen). Freshly transformed cells were inoculated into LB medium containing ampicillin, spectinomycin and chloramphenicol and grown at 37 °C (200 r.p.m.) until cells reached an OD₆₀₀ of 0.6–0.8. Transcription was then induced for 2–3 h by the addition of 1 mM IPTG. Cells were harvested by centrifugation at 5000 g (10 min, 4 °C) and stored at –70 °C or in liquid nitrogen.

Purification of the *S. ventriculi* PDC protein. All purification buffers contained 1 mM TPP and 1 mM MgSO₄ unless indicated otherwise. Recombinant *E. coli* cells (14.8 g wet wt) were thawed in 6 vols 50 mM sodium phosphate buffer at pH 6.5 (buffer A) and passed through a French pressure cell at 20000 p.s.i. (138 MPa). Cell debris was removed by centrifugation at 16000 g (20 min, 4 °C). Supernatant was removed and filtered through a 45 μ m filter membrane. Filtrate (692 mg protein) was applied to a Q Sepharose Fast Flow 26/10 column (Pharmacia) that was equilibrated at 2 ml min^{–1} with buffer A containing 300 mM NaCl. Fractions containing PDC activity (326 mg protein) were precipitated with 80% saturation (NH₄)₂SO₄. Protein was resuspended in buffer A, dialysed against buffer A (4 °C, 16 h), and filtered (45 μ m membrane). The filtrate (287 mg) was applied to a Q Sepharose column equilibrated with buffer A and developed

with a linear NaCl gradient (0–400 mM NaCl in 220 ml buffer A) at 4 ml min^{–1}. PDC active fractions eluted at 230–300 mM NaCl and were pooled. The pooled sample (23 mg) was applied to a 5 ml Bio-scale hydroxyapatite type I column (Bio-Rad) that was equilibrated at 1 ml min^{–1} with 5 mM sodium phosphate buffer at pH 6.5 (buffer B). The column was washed with 15 ml buffer B and developed with a linear sodium phosphate gradient (5–500 mM sodium phosphate at pH 6.5 in 75 ml). Protein fractions (11.4 mg) with PDC activity eluted at 200–300 mM sodium phosphate and were pooled. For further purification, portions of this material (0.25–0.5 mg protein per 0.25–0.5 ml) were applied to a Superdex 200 HR 10/30 column (Pharmacia) equilibrated at 0.25 ml min^{–1} in 50 mM sodium phosphate at pH 6.5 with 150 mM NaCl and 10% (v/v) glycerol in the presence and absence of 1 mM MgSO₄ and 1 mM TPP.

Activity assays and protein electrophoretic techniques. PDC activity was assayed by monitoring the pyruvic-acid-dependent reduction of NAD⁺ with baker's yeast alcohol dehydrogenase (ADH) (Sigma) as a coupling enzyme at pH 6.5 as previously described (Conway *et al.*, 1987). The assay also included the following modifications. Buffered enzyme (100 μ l) was added to a final volume of 1 ml containing 0.15 mM NADH, 0.1 mM TPP, 0–25 mM pyruvate and 10 U ADH in 50 mM potassium MES or 100 mM sodium hydrogen maleate buffer with 5 mM MgCl₂ at pH 6.5. Since this assay does not distinguish PDC from NADH-oxidizing enzymes such as lactate dehydrogenase, activity of cell lysate was estimated by correcting for control reactions performed in the absence of added ADH. One unit of enzyme activity is defined as amount of enzyme that generates 1 μ mol acetaldehyde min^{–1}. Thermostability was determined by incubating purified PDC in 50 mM

SvePDC	1	MTITTAAYLLRLKEVNVHMFVPGDYNLGLDYEDSKDIEVWGSCEINLNGYAADG
ScePDC1	1	MSEITFLGKYLIERLQVNVNVFGLPGDFNLILLDKIMEVEGMWAGNANELNARYAADG
ZmoPDC	1	MSYTVGVYLAERLVOIGLKHFAVAGDYNLVLLDNLILNKNMEQVCCNEINLNGFSAEQ
SvePDC	60	YARLRGTEVILLTYGVGSLSAINATIGSFAENVFVLHISGVPSALVQONRKLVHHSIARG
ScePDC1	61	YARIKGMSCILITTFGVGSLSALNGIAGSYAEHVGVLVHVGVPSSISOAKQLLLHHTLNG
ZmoPDC	60	YARAKGAAAVVTVSVGALSADDAIGGAYENLEPVITISGAPNNNDRAACHVLHHLGKT
SvePDC	120	EDTIERMFRITETESQSIISEYN-ADFEIDRVLESIYKVLGCIYELPVDIVSKETIDE
ScePDC1	121	DETVEHRSANISEETAMITDICTAPAEIDRCIRTTIVTORPVYGLLPANLVQINVPKIL
ZmoPDC	120	DYHYQLEMAKNITAAEAHYTFEPAPAKIDHVIKTALREKPVVLETAONIASMPCAAPG
SvePDC	179	MK-PENILMRSNEKFLKEFVNDVKEMVAS-SKGQHLADYEVLRAKAEKLEGEFNEAKI
ScePDC1	181	LOTPIIDMSLKENDAESEKEVIDTILVLAKDAKNEVILADACCSHDDVKAETKILIDLTF
ZmoPDC	180	PASALFNDEASDEASINAAVEETLKFIAN-RDKVAVLVGSKLRAGAEAAVKFADALGG
SvePDC	237	PVNTLSIGKTAVSESNFVAGLFSGETSSDLVKELCKASDIVLFGVKFIDTTAGRPVI
ScePDC1	241	PAFVTPMGKGSISSECHPRYGGVYVGTLSKPEVKEAVESADLLISVGCALLSDENIGSFSYS
ZmoPDC	239	AVATMAAKSFFPEONPHYICTSWGEVSYECVEKTKREADVATLAEVENDYSTTGWTDI
SvePDC	297	NADVMMIEGLTDCRIGETITVGLYIKDVTKALT-----DAKIKFENDVKEREAVEKF
ScePDC1	301	YETKILVGFHSDHKIRNATPPGVOMKEVLOKLLTNI-----ADAAGKYPVAVPARTPANAA
ZmoPDC	299	PDPKILVLAEPKSVVNGIRFESVHLKDYLRUAQRVSKKTGALDEKSKINAGELKKAAP
SvePDC	351	VEVDAKLQDRIYFKOMEAFLEPNDVLVGETGTSYSGACNMRFFPGSSFGSGSWMSIGVA
ScePDC1	359	VEASHPLEKOEVMNOLGNFLOEGDVVIAETGTSAFGINOTTFPNNITYGISQVWWSIGFT
ZmoPDC	359	ADPSAPLVNATIPARQVEALLPENTTVIAETGDSWFNAORMKLPNGARVEYEMCWGHGWS
SvePDC	411	TPAVLGTHLADKSR-----RNILSGDGSFQLTQVEVSTMIROKINTVLFVNNNGYTIER
ScePDC1	419	TCAPLGAFAFAEIDPKKRVILLEIGDGSFQLTQVEISTMIRWGLKPYLFVNNNGYTIER
ZmoPDC	419	VEAEGVAVGAFER-----RNILMVGDGSFQLTQVEAQMVRLLKLPVLEIFLINNGYTIET
SvePDC	467	LIHGEEYVNHQVQYAEVKTAT-----EDIQPTCFKVTTEKELAAMEEINKCTEGL
ScePDC1	479	LIHGKACQYNTIQVDHLSLIPTEG-----AKDYETHVATIGEDKLTQDKSFNDNSK-I
ZmoPDC	475	MIHDG--FYNNIKWWDYAGLEVEVNGNGGYDSGAGGLRAKTGGELAEATKVALANTDGP
SvePDC	524	AFVEVVMCKMDAPKSLRQEASLFSQNNY-----
ScePDC1	534	RMIEVMLPVVDAPONLVQAKLTATNTAKO-----
ZmoPDC	533	DLIECFIGREDCTEELVKWGRVAAANSRKPVNKL

Fig. 2. Multiple amino acid sequence alignment of *S. ventriculi* PDC with other PDC protein sequences. Abbreviations with GenBank or SWISS-PROT accession numbers: Sce, *Saccharomyces cerevisiae* P06169; Sve, *Sarcina ventriculi*; Zmo, *Zymomonas mobilis* P06672. Identical, functionally conserved and semi-conserved amino acid residues are shaded. Dashes indicate gaps introduced in protein sequence alignment. Indicated above the sequences are amino acid residues within 0.4 nm of the Mg²⁺- and TPP-binding site of yeast PDC1 (Dyda *et al.*, 1993) (▼), the Cys221 residue originally postulated to be required for pyruvate activation of yeast PDC1 (●), and the Tyr157 and Arg224 residues which form hydrogen bonds with allosteric activators such as pyruvamide (■). The underlined sequence is a conserved motif identified in TPP-dependent enzymes (Hawkins *et al.*, 1989).

sodium phosphate buffer at pH 6.5 with 1 mM TPP and 1 mM MgCl₂ for 90 min and then assaying for activity with 10 mM pyruvate. Protein concentration was determined using Bradford protein reagent with bovine serum albumin as the standard (Bio-Rad).

Molecular mass and N-terminal amino acid sequence analyses. Subunit molecular mass was estimated by reducing and denaturing SDS-PAGE using 12% polyacrylamide gels which were stained with Coomassie blue R-250. The molecular mass standards for SDS-PAGE were: phosphorylase *b* (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa) and lysozyme (14.4 kDa). For determination of native molecular mass, samples were applied to a Superdex 200 HR 10/30 column equilibrated in 50 mM sodium phosphate buffer at pH 6.5 with 150 mM NaCl, 10% glycerol, and no added cofactors. Molecular mass standards included: bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), α -amylase (200 kDa), apoferritin (443 kDa), and thyroglobulin (669 kDa).

The N-terminal sequence was determined for PDC protein purified from recombinant *E. coli*. The protein was separated by SDS-PAGE and electroblotted onto a PVDF membrane (Immobilon-P). The sequence was determined by automated Edman degradation at the protein chemistry core facility of the University of Florida Interdisciplinary Center for Biotechnology Research.

RESULTS AND DISCUSSION

S. ventriculi PDC operon

The N-terminal amino acid sequence of the PDC protein purified from *S. ventriculi* (Lowe & Zeikus, 1992) was used to generate a degenerate oligonucleotide for hybridization to genomic DNA. This approach facilitated the isolation of a 7.0 kb *Bcl*I genomic DNA fragment from *S. ventriculi*. The fragment was further subcloned in order to sequence both strands of a 3886 bp *Hinc*II-to-*Hinc*II region that hybridized to the oligonucleotide probe (Fig. 1). Analysis of the DNA sequence revealed an ORF of 1656 bp encoding a protein with an N-terminus identical to that of the previously purified *S. ventriculi* PDC. The ORF was therefore designated *pdc*. A canonical Shine-Dalgarno sequence is present 7 bp upstream of the *pdc* translation start codon. In addition, a region 82–110 bp upstream of *pdc* has limited identity to the eubacterial –35 and –10 promoter consensus sequence (aTaACA-N₁₆-TATtAa, where N is any nucleotide and upper-case nucleotides match the consensus). Downstream (43 bp) of the *pdc* translation stop codon is a 40 bp region predicted to form a stem-loop structure followed by an AT-rich region, consistent with a *p*-independent transcription terminator. Thus, the *S.*

Fig. 3. Dendrogram summarizing the relationships between selected PDCs and other TPP-dependent enzymes. Deduced protein sequences were aligned using CLUSTAL X. Amino acid extensions at the N- or C-terminus, and apparent insertion sequences, were removed. Remaining regions containing approximately 520–540 amino acids were compared. Treeview was used to display these results as an unrooted dendrogram. Protein abbreviations: PDC, pyruvate decarboxylase; IPD, indole-3-pyruvate decarboxylase; ALS, acetolactate synthase; PDH E1 or E1, the E1 component of pyruvate dehydrogenase; TK, transketolase. Organism abbreviations and GenBank or SWISS-PROT accession numbers: Abr, *Azospirillum brasilense* P51852; Aor, *Aspergillus oryzae* AAD16178; Apa, *Aspergillus parasiticus* P51844; Asy, *Ascidia sydneiensis samea* BAA74730; Ath, *Arabidopsis thaliana* BAB08775; Bfl, *Brevibacterium flavum* A56684; Bsu, *Bacillus subtilis* P45694; Cgl, *Corynebacterium glutamicum* P42463; Cpn, *Chlamydomonas pneumoniae* H72020; Dra, *Deinococcus radiodurans* A75387 (ALS), A75541 (E1); Ecl, *Enterobacter cloacae* P23234; Eco, *Escherichia coli* CAA24740; Ene, *Erwinia herbicola* AAB06571; Eni, *Aspergillus (Emerella) nidulans* P87208; Fan, *Fragaria × ananassa* AAG13131; Ghi, *Gossypium hirsutum* S60056; Gth, *Guillardia theta* NP_050806; Huv, *Hanseniaspora uvarum* P34734; Kla, *Kluyveromyces lactis* Q12629 (PDC), Q12630 (TK); Kma, *Kluyveromyces marxianus* P33149; Mav, *Mycobacterium avium* Q59498; Mja, *Methanococcus jannaschii* Q57725; Mie, *Mycobacterium leprae* CAC31122 (ORF), 033112 (ALS), CAC30602 (E1); Mth, *Methanobacterium thermoautotrophicum* A69081 (ORF), C69059 (ALS); Mtu, *Mycobacterium tuberculosis* E70814 (IPD), 053250 (ALS); Ncr, *Neurospora crassa* P33287; Nta, *Nicotiana tabacum* P51846 (PDC), P09342 (ALS); Osa, *Oryza sativa* P51847 (PDC1), P51848

ventriculi pdc appears to be transcribed as a monocistronic operon like the *Z. mobilis pdc* gene (Conway *et al.*, 1987).

A partial ORF was identified 722 bp upstream of *pdc* which encodes a 177 amino acid protein fragment (ORF1*) (Fig. 1). ORF1* has identity (28–29%) to several hypothetical membrane proteins (GenBank accession numbers CAC11620, CAC24018, CAA22902) and is predicted to form several membrane-spanning domains (data not shown).

S. ventriculi PDC protein sequence

The *S. ventriculi pdc* gene apparently encodes a protein of 552 amino acids (including the N-terminal methionine) with a calculated pI of 5.16 and anhydrous molecular mass of 61737 Da. Consistent with other *Z. mobilis* and fungal PDC proteins, the N-terminal extension of up to 47 amino acids that is common to plant PDC proteins is not conserved in the *S. ventriculi* PDC protein. Although the pI of the purified *S. ventriculi* PDC protein has not been experimentally determined, the calculated pI is consistent with the acidic pH optimum of 6.3–6.7 for stability and activity of this enzyme (Lowe & Zeikus, 1992). The amino acid composition of the protein deduced from the *S. ventriculi pdc* gene is similar to that determined for the *Saccharomyces cerevisiae* PDC1 (Kellermann *et al.*, 1986) and *Z. mobilis pdc* genes (Conway *et al.*, 1987; Neale *et al.*, 1987b). A notable exception is the alanine composition of *Z. mobilis* PDC, which is 1.8- to 2.2-fold higher than that of *S. cerevisiae* PDC1 and *S. ventriculi* PDC. Although the phenylalanine composition of the deduced *S. ventriculi* PDC protein is consistent with the other PDC proteins, it is almost 3.6-fold less than the composition previously reported for the purified *S. ventriculi* enzyme (Lowe & Zeikus, 1992). The reason for this discrepancy remains to be determined.

The amino acid sequence of *S. ventriculi* PDC was aligned with the sequences of the yeast (Sce) PDC1 and *Z. mobilis* (Zmo) PDC proteins, both of which have been analysed by X-ray crystallography (Arjunan *et al.*, 1996; Dobritzsch *et al.*, 1998; Dyda *et al.*, 1993; König, 1998) (Fig. 2). The conserved motif of TPP-dependent enzymes identified by Hawkins *et al.* (1989) and known to be involved in Mg^{2+} -TPP cofactor binding is highly conserved in all three PDC proteins. Amino acid residues located within 0.4 nm of the Mg^{2+} - and TPP-binding site of the PDC proteins which have been crystallized from yeast and *Z. mobilis* are conserved in primary sequence

with the *S. ventriculi* PDC protein. These include residues with similarity to the aspartate (SceD444, ZmoD440) and asparagine residues (SceN471, ZmoN467) that are involved in binding Mg^{2+} . The *S. ventriculi* PDC appears to be more similar to the yeast PDC than to that of *Z. mobilis* in binding the diphosphates of TPP, where serine and threonine side chains (SceS446 and T390) as well as the main chain nitrogen of isoleucine (SceI476) are conserved. This contrasts with the *Z. mobilis* enzyme, which utilizes a main-chain nitrogen of aspartate (ZmoD390) instead of the threonine hydroxyl group (SceT390) for binding the β -phosphate. *S. ventriculi* PDC residues are also similar to the aspartate, glutamate, threonine and histidine residues (SceD28, E477, T388, H114 and H115; ZmoD27, E473, T388, H113 and H114) which may potentially interact with intermediates during the decarboxylation reaction mediated by *Z. mobilis* and yeast PDCs. Furthermore, the isoleucine (Sce and Zmo I415) side chain, which appears to stabilize the V conformation of TPP through Van der Waals interactions, as well as the glutamate (SceE51, ZmoE50) which may donate a proton to the N1' atom of TPP, are conserved in the *S. ventriculi* PDC protein sequence.

A notable exception in conservation is the yeast C221 residue (Fig. 2), which is highly conserved among the majority of fungal PDCs but is not conserved in either bacterial or plant PDC proteins. Based on site-directed mutagenesis, chemical modification and kinetic studies, this C221 residue has been proposed to be a primary binding site of the regulatory substrate molecule and the starting point of a signal transfer pathway to the active-site TPP in the yeast enzyme (Baburina *et al.*, 1994, 1998; Hübner *et al.*, 1988). Consistent with these previous results the yeast C221 is positioned in a large cavity formed at the interface among all three PDC domains including the α or PYR (residues 1 to 189), β or R (residues 190 to 356) and γ or PP (residues 357 to 563) domains (Dyda *et al.*, 1993). However, recent high-resolution structural analysis of the brewer's yeast PDC crystallized in the presence of pyruvamide, a pyruvate analogue, enabled localization of the activator-binding site and revealed that cysteine does not play a direct role in this binding (Lu *et al.*, 2000). Additionally, kinetic studies using stopped-flow techniques revealed that the C221A variant of yeast PDC was still substrate activated, and the lag phase of product formation did not disappear with progressive thiol oxidation (Lu *et al.*, 2000). Instead, tyrosine (Y157) and arginine (R224) residues form hydrogen bonds with the amide group of pyruvamide. Both of these residues are conserved in the

(PDC2), P51849 (PDC3); Pae, *Pseudomonas aeruginosa* G83123; Pmu, *Pasteurella multocida* AAK03712; Pop, *Porphyra purpurea* NP_053940; Ppu, *Pseudomonas putida* AAG00523; Psa, *Pisum sativum* P51850; Pst, *Pichia stipitis* AAC03164 (PDC1), AAC03165 (PDC2); Rca, *Rhodobacter capsulatus* JC4637; Reu, *Ralstonia eutropha* Q59097; Sav, *Streptomyces avermitilis* AAA93098; Sce, *Saccharomyces cerevisiae* P06169 (PDC1), P16467 (PDC5), P26263 (PDC6), Q07471 (ORF), NP_011105 (E1 α), NP_009780 (E1 β); Sco, *Streptomyces coelicolor* T35828; Skl, *Saccharomyces kluyveri* AAF78895; Spl, *Spirulina platensis* P27868; Spo, *Schizosaccharomyces pombe* Q09737 (PDC1), Q92345 (PDC2); Sve, *Sarcina ventriculi* AF354297; Syn, *Synechocystis* sp. BAA17984; Vch, *Vibrio cholerae* A82375; Vvi, *Vitis vinifera* AAG22488; Zma, *Zea mays* P28516; Zbi, *Zygosaccharomyces bisporus* CAB65554; Zmo, *Zymomonas mobilis* P06672. Scale bar represents 0.1 nucleotide substitutions per site.

Table 2. Codon usage of *S. ventriculi* (Sve) and *Z. mobilis* (Zmo) *pdc* genes

Amino acid	Codon	Zmo*	Sve*	<i>E. coli</i> †
Ala	GCA	28.1	32.5	20.1
	GCC	35.1	0	25.5
	GCG	12.3	1.8	33.6
	GCU	73.8	34.4	15.3
Arg	AGA	1.8	39.8	2.1
	CGC	15.8	0	22.0
	CGG	1.8	0	5.4
	CGU	12.3	0	20.9
Asn	AAC	47.5	19.9	21.7
	AAU	10.5	27.1	17.7
Asp	GAC	28.1	3.6	19.1
	GAU	10.5	47.0	32.1
Cys	UGC	28.1	1.8	6.5
	UGU	10.5	7.2	5.2
Gln	CAA	0	28.9	15.3
	CAG	17.6	0	28.8
Glu	GAA	61.5	85.0	39.4
	GAG	3.5	5.4	17.8
Gly	GGA	1.8	50.6	8.0
	GGC	19.3	0	29.6
	GGG	0	0	11.1
	GGU	56.2	19.9	24.7
His	CAC	8.8	3.6	9.7
	CAU	14.1	12.7	12.9
Ile	AUA	0	39.8	4.4
	AUC	35.1	9.0	25.1
	AUU	14.1	21.7	30.3
Leu	CUA	0	7.2	3.9
	CUC	19.3	0	11.1
	CUG	33.4	0	52.6
	CUU	12.3	14.5	11.0
	UUA	1.8	59.7	13.9
	UUG	5.4	0	13.7
Lys	AAA	33.4	63.3	33.6
	AAG	31.6	5.4	10.3
Met	AUG	21.1	27.1	27.9
	UUC	31.6	18.1	16.6
	UUU	0	28.9	22.3
Pro	CCA	3.5	18.1	8.4
	CCC	3.5	0	5.5
	CCG	29.9	1.8	23.2
	CCU	10.5	7.2	7.0
Ser	AGC	14.1	12.7	16.1
	AGU	7.0	12.7	8.8
	UCA	1.8	32.5	7.2
	UCC	15.8	0	8.6
Thr	UCU	5.3	7.2	8.5
	ACA	0	34.4	7.1
	ACC	28.1	0	23.4
	ACG	10.5	0	14.4
Trp	ACU	8.8	30.7	9.0
	UGG	12.3	5.4	15.2
Tyr	UAC	12.3	7.2	12.2
	UAU	26.4	32.5	16.2
Val	GUA	0	36.2	10.9
	GUC	26.4	0	15.3
	GUG	7.0	0	26.4
	GUU	43.9	43.4	18.3

* Codon usage for amino acids is represented as frequency per thousand bases. Stop codons are not indicated. CGA and AGG codons for Arg, UCG for Ser, and GGG for Gly were not used for either of the *pdc* genes.

† Mean usage in frequency per thousand bases for genes in *E. coli* K-12. Codons for accessory tRNAs essential for high-level synthesis of *S. ventriculi* PDC in recombinant *E. coli* are highlighted in bold.

S. ventriculi PDC and not found in the *Z. mobilis* enzyme, which displays Michaelis–Menten kinetics. These results suggest that residues of the *S. ventriculi* PDC protein may allosterically bind the substrate activator with a mechanism common to the majority of fungal PDC proteins. Interestingly, these two residues that bind pyruvamide in the yeast enzyme are not universally distributed among the substrate-activated PDCs, most notably the plant PDCs.

Cluster analysis was performed to compare the PDC proteins and other TPP-dependent enzymes, including indole-3-pyruvate decarboxylase (IPD), the E1 component of pyruvate dehydrogenase (E1), acetolactate synthase (ALS), and transketolase (TK) (Fig. 3). The comparison reveals that all of these proteins are related in primary sequence and that there is a significant clustering of the sequences into families based on specific enzyme function. Of these, the *S. ventriculi* PDC appears most closely related to eubacterial IPD proteins as well as the majority of fungal PDC proteins. In contrast, the *Z. mobilis* PDC protein is most closely related to plant PDCs in addition to a couple of outgrouping fungal PDCs. Thus, it appears that the IPD protein family has close evolutionary roots with the PDC family and specifically the *S. ventriculi* PDC protein. In addition, the distant relationship of the *S. ventriculi* and *Z. mobilis* PDC proteins is consistent with the biochemical differences between these two enzymes (Hoppner & Doelle, 1983; Lowe & Zeikus, 1992; Neale *et al.*, 1987a; Ullrich & Donner, 1970). In contrast to *Z. mobilis* PDC, which may have originated by the horizontal transfer of a plant *pdc* gene, *S. ventriculi* PDC appears to have diverged quite early during evolution and last shared a common ancestor with most eubacterial IPD and fungal PDCs.

Production of *S. ventriculi* PDC protein

Significant differences in codon usage between the *S. ventriculi* and *Z. mobilis* *pdc* genes are evident and may impact translation in engineered host organisms (Table 2). In particular, the *pdc* gene of *S. ventriculi* encodes elevated use of tRNA_{AUA} and tRNA_{AGA}, both of which are relatively rare in *E. coli*. This is in contrast to the *Z. mobilis* *pdc* gene, which does not use the AUA codon and has only minimal use of the AGA codon. This suggests that production of the *S. ventriculi* PDC protein in recombinant *E. coli* may be limited by mRNA translation due to the requirement for elevated use of rare tRNAs.

To further investigate this, the levels of the tRNA genes that are rare in *E. coli* were modified during *pdc* expression by including multiple copies of these genes on the chromosome (*E. coli* strain BL21-CodonPlus-RIL) and/or on a complementary plasmid (pSJS1240) (Table 1). These modified *E. coli* strains were transformed with plasmid pJAM419, which carries the *S. ventriculi* *pdc* gene positioned 8 bp downstream of an optimized Shine–Dalgarno consensus sequence and controlled at the transcriptional level by T7 RNA

Fig. 4. *S. ventriculi* PDC protein synthesized in recombinant *E. coli*. Proteins were analysed by reducing SDS-PAGE using 12% polyacrylamide gels and stained with Coomassie blue R-250. Lanes 1 and 4, molecular mass standards (5 μ g). Lanes 2 and 3, cell lysate (20 μ g) of IPTG-induced *E. coli* BL21-CodonPlus-RIL/pSJS1240 transformed with pET21d or pJAM419, respectively. Lane 5, *S. ventriculi* PDC protein (2 μ g) purified from recombinant *E. coli*.

polymerase promoter and terminator sequences from plasmid vector pET21d. Detectable levels of PDC activity (0.16 U per mg protein at 5 mM pyruvate) were observed after induction of *pdc* transcription in *E. coli* host strains with additional chromosomal and/or plasmid copies of the *ileU/X*, *argU* and *leuW* genes encoding the rare tRNA_{AUA}, tRNA_{AGG/AGA} and tRNA_{CUA}. A 5- to 10-fold increase in the levels of a 58 kDa protein with a molecular mass comparable to the *S. ventriculi* PDC (Fig. 4) were produced in these strains compared to a similar *E. coli* strain without added tRNA genes [BL21(DE3)] (data not shown). These results suggest that the *S. ventriculi* PDC protein is synthesized in recombinant *E. coli* and that the high percentage of AUA and AGA codons of the *pdc* gene limits translation.

Interestingly, additional proteins of 43 and 27 kDa were also observed when the PDC protein was synthesized in *E. coli* compared to control strains (Fig. 4). The origin of these proteins remains to be determined. They may be fragments of PDC generated by proteolysis or truncated PDC produced from errors in translation/transcription. Alternatively, increased production of PDC may increase the levels of acetaldehyde, which can be toxic to the cell, and may subsequently induce the levels of proteins in response to this stress.

Biochemical and biophysical properties of the *S. ventriculi* PDC protein from recombinant *E. coli*

The *S. ventriculi* PDC protein was purified over 136-fold from recombinant *E. coli*. The N-terminal amino acid sequence of this protein (MKITIAEYLLXR, where X is an unidentified amino acid) was identical to the sequence

Fig. 5. Kinetics for pyruvate of the *S. ventriculi* PDC purified from recombinant *E. coli*. The data represent mean results from triplicate determinations of PDC activity by the ADH coupled assay using 1 μ g purified enzyme in 1 ml final assay volume with either potassium MES (●) or sodium hydrogen maleate (○) buffer as described in Methods.

of PDC purified from *S. ventriculi* (Lowe & Zeikus, 1992). Both PDC proteins have an N-terminal methionine residue, which suggests that this residue is not accessible for cleavage by either the *S. ventriculi* or *E. coli* aminopeptidases.

The thermostability of the purified *S. ventriculi* PDC was examined in the presence of 1 mM cofactors TPP and Mg²⁺ at pH 6.5. Enzyme activity was stable up to 42 °C but was abolished after incubation for 60–90 min at temperatures of 50 °C and above. This is consistent with the significant loss of PDC activity observed when a thermal treatment step (60 °C for 30 min) was included in the purification (data not shown). In contrast, methods used to purify *Z. mobilis* and other PDC proteins (König, 1998) typically incorporate thermal treatment to remove unwanted proteins. These results suggest that the recombinant *S. ventriculi* PDC protein is not as thermostable as other PDC proteins, including that of *Z. mobilis*.

PDC proteins have been shown to bind TPP and Mg²⁺ cofactors with high affinity at slightly acidic pH (König, 1998). Consistent with this, the recombinant *S. ventriculi* PDC retains full activity after incubation at 37 °C for 90 min in the presence of 25 mM EGTA or EDTA in pH 6.5 buffer without cofactors. This is similar to the PDC protein purified directly from *S. ventriculi*, which is fully active after similar treatment with metal chelators.

The recombinant *S. ventriculi* enzyme displays sigmoidal kinetics (Fig. 5), suggesting that the enzyme is substrate activated, similar to the fungal and plant PDC proteins (König, 1998). This contrasts with the *Z. mobilis* PDC, which is the only PDC known to display non-sigmoidal, Michaelis–Menten kinetics. Based on the kinetic data, it appears that the *S. ventriculi* PDC has two affinities for pyruvate (K_m values of 2.8 and 10 mM) (Fig. 5). The K_m value for the low-affinity site is comparable to the K_m value (13 mM) reported by Lowe & Zeikus (1992). The fivefold lower K_m observed in this

present study is apparently due to the difference in the buffer, since maleate buffer significantly inhibited the PDC activity at low pyruvate concentrations (Fig. 5). This difference in activity of *S. ventriculi* in maleate buffer is similar to that observed with *Z. mobilis* PDC, with a K_m of 4.4 mM for pyruvate in Tris/maleate buffer at pH 6 (Hoppner & Doelle, 1983) compared to K_m values of 0.3–0.4 mM in potassium MES buffer at pH 6 (Bringer-Meyer *et al.*, 1986) and sodium citrate buffer at pH 6.5 (Neale *et al.*, 1987b).

The high-affinity K_m value of 2.8 mM obtained for the *S. ventriculi* PDC is similar to the K_m values obtained with many fungal and plant PDCs, including those purified from *Saccharomyces cerevisiae* (1–3 mM) (Boiteux & Hess, 1970; Hübner *et al.*, 1978), *Zygosaccharomyces bisporus* (1.73 mM) (Neuser *et al.*, 2000), orange (0.8–3.2 mM) (Raymond *et al.*, 1979), and wheat germ (3 mM) (Zehender *et al.*, 1987). However, these K_m values are several-fold higher than those reported for the PDCs of *Z. mobilis* (0.3–0.4 mM) (Bringer-Meyer *et al.*, 1986; Neale *et al.*, 1987b) and rice (0.25 mM) (Rivoal *et al.*, 1990).

At pH 6.5, the recombinant *S. ventriculi* PDC forms a 235 kDa homotetramer consisting of a 58 kDa protein, as determined by Superdex 200 gel filtration chromatography and SDS-12% PAGE (see Methods). Exclusion of the cofactors from the buffer during gel filtration chromatography at pH 6.5 did not alter the tetramer configuration or enzyme activity, suggesting that the cofactors are tightly bound. The configuration of the PDC complex is consistent with that purified from *S. ventriculi* as well as the majority of those isolated from fungi, plants and *Z. mobilis*. There are, however, plant PDCs which have been reported to form larger complexes, including the PDC from *Neurospora crassa*, which forms aggregated filaments of 8–10 nm (Alvarez *et al.*, 1993) as well as the PDC from *Pisum sativum*, which forms up to 960 kDa complexes (Mücke *et al.*, 1995).

Conclusions

Based on this study, the *S. ventriculi* PDC protein appears to share similar primary sequence structure with TPP-dependent enzymes and is highly related to the fungal PDC and eubacterial IPD enzymes. The close relationship of the *S. ventriculi* and fungal PDC structures is consistent with the similar biochemical and biophysical properties of these enzymes. Both types of enzymes display substrate cooperativity with similar affinities for pyruvate. The structure and biochemistry of the *S. ventriculi* PDC, however, dramatically contrast with the only other bacterial PDC (*Z. mobilis*) which has been characterized. The *Z. mobilis* PDC is closely related to plants in primary structure and it is the only PDC known to display Michaelis–Menten kinetics.

This study also demonstrates the synthesis of active, soluble *S. ventriculi* PDC protein in recombinant *E. coli*. Only two other genes, the *Z. mobilis* *pdc* and *Sac-*

charomyces cerevisiae *PDC1* genes, have been reported to synthesize PDC protein in recombinant bacteria (Bräü & Sahm, 1986; Candy *et al.*, 1991; Conway *et al.*, 1987). Of these, at least 50% of the *S. cerevisiae* *PDC1* forms insoluble inclusions in *E. coli* and thus has not been useful in engineering bacteria for high-level ethanol production (Candy *et al.*, 1991). Due to codon bias, accessory tRNA is essential for efficient production of *S. ventriculi* PDC in recombinant *E. coli*. However, the low-G+C codon bias of the *S. ventriculi* *pdc* gene should broaden the spectrum of bacteria that can be engineered as hosts for high-level production of PDC protein and the engineering of homo-ethanol pathways (Ingram *et al.*, 1999).

ACKNOWLEDGEMENTS

We thank Francis Davis and Jack Shelton for DNA sequencing. Thanks also to K. T. Shanmugam for generously providing *E. coli* strain SE2309, and Sung-Hou and Rosalind Kim for providing plasmid pSJS1240. This work was supported in part by the Department of Energy (ZDH-9-29009-04 and DE-FG02-96ER20222) and the Florida Agricultural Experiment Station (Journal Series R-08274).

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Alvarez, M. E., Rosa, A. L., Temporini, E. D., Wolstenholme, A., Panzetta, G., Patrito, L. & Maccioni, H. J. (1993). The 59-kDa polypeptide constituent of 8–10-nm cytoplasmic filaments in *Neurospora crassa* is a pyruvate decarboxylase. *Gene* **130**, 253–258.
- Arjunan, P., Umland, T., Dyda, F. & 7 other authors (1996). Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 Å resolution. *J Mol Biol* **256**, 590–600.
- Baburina, I., Gao, Y., Hu, Z., Jordan, F., Hohmann, S. & Furey, W. (1994). Substrate activation of brewers' yeast pyruvate decarboxylase is abolished by mutation of cysteine 221 to serine. *Biochemistry* **33**, 5630–5635.
- Baburina, I., Li, H., Bennion, B., Furey, W. & Jordan, F. (1998). Interdomain information transfer during substrate activation of yeast pyruvate decarboxylase: the interaction between cysteine 221 and histidine 92. *Biochemistry* **37**, 1235–1244.
- Barbosa, M. F. S. & Ingram, L. O. (1994). Expression of the *Zymomonas mobilis* alcohol dehydrogenase II (*adhB*) and pyruvate decarboxylase (*pdc*) genes in *Bacillus*. *Curr Microbiol* **28**, 279–282.
- Boiteux, A. & Hess, B. (1970). Allosteric properties of yeast pyruvate decarboxylase. *FEBS Lett* **9**, 293–296.
- Bräü, B. & Sahm, H. (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., Schimz, K.-L. & Sahm, H. (1986). Pyruvate decarboxylase from *Zymomonas mobilis*. Isolation and partial characterization. *Arch Microbiol* **146**, 105–110.
- Candy, J. M. & Duggleby, R. G. (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., Duggleby, R. G. & Mattick, J. S. (1991). Expression of active yeast pyruvate decarboxylase in *Escherichia coli*. *J Gen Microbiol* **137**, 2811–2815.
- Conway, T., Osman, Y. A., Konnan, J. I., Hoffmann, E. M. & Ingram, L. O. (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.
- Cserzo, M., Wallin, E., Simon, I., von Heijne, G. & Elofsson, A. (1997). Prediction of transmembrane α -helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng* **10**, 673–676.
- Dobritzsch, D., König, S., Schneider, G. & Lu, G. (1998). High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases. *J Biol Chem* **273**, 20196–20204.
- Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B. & Jordan, F. (1993). Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-Å resolution. *Biochemistry* **32**, 6165–6170.
- Gold, R. S., Meagher, M. M., Tong, S., Hutkins, R. W. & Conway, T. (1996). Cloning and expression of the *Zymomonas mobilis* “production of ethanol” genes in *Lactobacillus casei*. *Curr Microbiol* **33**, 256–260.
- Goodwin, S. & Zeikus, J. G. (1987). Physiological adaptations of anaerobic bacteria to low pH: metabolic control of proton motive force in *Sarcina ventriculi*. *J Bacteriol* **169**, 2150–2157.
- Harwood, C. R. & Cutting, S. M. (1990). *Molecular Biological Methods for Bacillus*. New York: Wiley.
- Hawkins, C. F., Borges, A. & Perham, R. N. (1989). A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Lett* **255**, 77–82.
- Hoppner, T. C. & Doelle, H. W. (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.
- Hübner, G., Weidhase, R. & Schellenberger, A. (1978). The mechanism of substrate activation of pyruvate decarboxylase: a first approach. *Eur J Biochem* **92**, 175–181.
- Hübner, G., König, S. & Schellenberger, A. (1988). The functional role of thiol groups of pyruvate decarboxylase from brewer's yeast. *Biomed Biochim Acta* **47**, 9–18.
- Ingram, L. O., Aldrich, H. C., Borges, A. C. & 9 other authors (1999). Enteric bacterial catalysts for fuel ethanol production. *Biotechnol Prog* **15**, 855–866.
- Kellermann, E., Seeboth, P. G. & Hollenberg, C. P. (1986). Analysis of the primary structure and promoter function of a pyruvate decarboxylase gene (PDC1) from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **14**, 8963–8977.
- Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). Overexpression of archaeal proteins in *Escherichia coli*. *Biotechnol Lett* **20**, 207–210.
- König, S. (1998). Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. *Biochim Biophys Acta* **1385**, 271–286.
- Lowe, S. E. & Zeikus, J. G. (1991). Metabolic regulation of carbon and electron flow as a function of pH during growth of *Sarcina ventriculi*. *Arch Microbiol* **155**, 325–329.
- Lowe, S. E. & Zeikus, J. G. (1992). Purification and characterization of pyruvate decarboxylase from *Sarcina ventriculi*. *J Gen Microbiol* **138**, 803–807.
- Lu, G., Dobritzsch, D., Baumann, S., Schneider, G. & König, S. (2000). The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study. *Eur J Biochem* **267**, 861–868.
- Mücke, U., König, S. & Hübner, G. (1995). Purification and characterisation of pyruvate decarboxylase from pea seeds (*Pisum sativum* cv. Miko). *Biol Chem Hoppe-Seyler* **376**, 111–117.
- Neale, A. D., Scopes, R. K., Wettenhall, R. E. & Hoogenraad, N. J. (1987a). Nucleotide sequence of the pyruvate decarboxylase gene from *Zymomonas mobilis*. *Nucleic Acids Res* **15**, 1753–1761.
- Neale, A. D., Scopes, R. K., Wettenhall, R. E. & Hoogenraad, N. J. (1987b). Pyruvate decarboxylase of *Zymomonas mobilis*: isolation, properties, and genetic expression in *Escherichia coli*. *J Bacteriol* **169**, 1024–1028.
- Neuser, F., Zorn, H., Richter, U. & Berger, R. G. (2000). Purification, characterisation and cDNA sequencing of pyruvate decarboxylase from *Zygosaccharomyces bisporus*. *Biol Chem* **381**, 349–353.
- Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**, 357–358.
- Raymond, W. R., Hostettler, J. B., Assar, K. & Varsel, C. (1979). Orange pyruvate decarboxylase: isolation and mechanistic studies. *J Food Sci* **44**, 777–781.
- Reynen, M. & Sahm, H. (1988). Comparison of the structural genes for pyruvate decarboxylase in different *Zymomonas mobilis* strains. *J Bacteriol* **170**, 3310–3313.
- Rivoal, J., Ricard, B. & Pradet, A. (1990). Purification and partial characterization of pyruvate decarboxylase from *Oryza sativa* L. *Eur J Biochem* **194**, 791–797.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503–517.
- Stephenson, M. P. & Dawes, E. A. (1971). Pyruvic acid and formic acid metabolism in *Sarcina ventriculi* and the role of ferredoxin. *J Gen Microbiol* **69**, 331–343.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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.
- Ullrich, J. & Donner, I. (1970). Kinetic evidence for two active sites in cytoplasmic yeast pyruvate decarboxylase. *Hoppe Seylers Z Physiol Chem* **351**, 1026–1029.
- Zehender, H., Trescher, D. & Ullrich, J. (1987). Improved purification of pyruvate decarboxylase from wheat germ. Its partial characterisation and comparison with the yeast enzyme. *Eur J Biochem* **167**, 149–154.

Received 8 March 2001; revised 26 May 2001; accepted 4 June 2001.