Engineering pyruvate decarboxylase–mediated ethanol production in the thermophilic host *Geobacillus thermoglucosidasius*

Van Zyl L.[1], Taylor M. P.[1,2], Eley K.[2], Tuffin M.[1], Cowan D.A.[1,3].

[1] Institute for Microbial Biotechnology and Metagenomics (IMBM), University of the Western Cape, Modderdam Road, Bellville, Cape Town, South Africa.

[2] TMO Renewables Limited, 40 Alan Turing Road, The Surrey Research Park, Guildford, Surrey, GU2 7YF, UK

[3] Centre for Microbial Ecology and Genomics, Department of Genetics, University of Pretoria, Pretoria 0028, South Africa

Keywords: Pyruvate decarboxylase, bioethanol, *Gluconoacetobacter* spp., thermophilic expression

*Corresponding author: Professor Don Cowan, don.cowan@up.ac.za, Tel: +27 [0]12 4205873, Institutional Research Theme in Genomics, University of Pretoria, Pretoria 0028, South Africa.

Abstract

This study reports the expression, purification and kinetic characterization of a PDC from *Gluconobacter oxydans*. Kinetic analyses showed the enzyme to have high affinity for pyruvate (120µM at pH 5), high catalytic efficiency (4.75 x 10^5 $M^{-1}s^{-1}$ at pH 5), a pH_{opt} of approximately 4.5 and an *in vitro* temperature optimum at approximately 55°C (the highest yet reported for a bacterial PDC). Due to good *in vitro* thermostablity (approximately 40% enzyme activity retained after 30 minutes at 65°C) this PDC was considered to be a suitable candidate for heterologous expression in the thermophile *Geobacillus thermoglucosidasius*. Initial studies using a variety of methods failed to detect activity at any growth temperature. However, the application of codon harmonization (i.e., mimicry of the heterogeneous host's transcription and translational rhythm) yielded a protein that was fully functional in the thermophilic strain at 45°C (as determined by enzyme activity, Western blot, mRNA detection and ethanol

productivity). Here we describe the successful expression of PDC in a true thermophile. Yields as high as 0.35 g/g ±0.04 ethanol per gram of glucose consumed were detected, highly competitive to those reported in ethanologenic thermophilic mutants. Although activities could not be detected at temperatures approaching the growth optimum for the strain, this study highlights that the possibility that previously unsuccessful expression of *pdc*s in *Geobacillus* spp. may be the result of ineffective transcription / translation coupling.

Introduction

Pyruvate decarboxylase (PDC, EC 4.1.1.1) is the enzyme responsible for the non-oxidative decarboxylation of pyruvate to acetaldehyde and carbon dioxide. PDCs are common in the plant and fungal kingdoms and at least in the latter, together with alcohol dehydrogenase (ADH, EC 1.1.1.1) form part of an ethanol fermentation pathway (Konig, 1998). Several plant and yeast PDC's have been isolated and characterized, but as yet only four of bacterial origin have been described; from *Zymomonas mobilis*, *Zymobacter palmae, Acetobacter pasteurianus* and *Sarcina ventriculi* (Raj *et al.*, 2002). As in yeast, these bacterial PDCs participate in ethanol production, but are incorporated into the Entner Doudoroff pathway as opposed to glycolysis for pyruvate production.

There has been an increased interest in the use of thermophiles, such as *Geobacillus thermoglucosidasius,* for ethanol production, primarily because of their catabolic promiscuity, an important benefit for a second generation bioprocess design (Cripps *et al.*, 2009; Taylor *et al.*, 2009). Other advantages include improved product removal, reduced incidence of contamination and high ethanol yields in selectively mutated strains (Taylor *et al.*, 2009).

Ethanol production in *G. thermoglucosidasius* and mutants with enhanced ethanologenic phenotypes relies on endogenous metabolic pathways, generating acetyl CoA via pyruvate dehydrogenase and its subsequent conversion to acetaldehyde and ethanol by aldehyde dehydrogenase and alcohol dehydrogenase, respectively (Cripps *et al.*, 2009). An alternative to

further develop *G. thermoglucosidasius* as an ethanologenic strain is to engineer the expression of the PDC pathway. The *in vitro* high specificity and thermostability (half-life of 30min at 60°C) of the *Z. mobilis* PDC (ZmoPDC) has made it the main candidate for such engineering (Pohl *et al.*, 1995). However, both the ZmoPDC and *Zymobacter palmae* PDCs have been expressed in *G. thermoglucosidasius* but do not function at temperatures exceeding 55°C, despite good *in vitro* thermostability at these temperatures. The reasons for the low levels of activity are not fully understood (Taylor *et al.*, 2008; Thompson *et al.*, 2008), but protein misfolding resulting in inactive protein has been proposed (Thompson *et al.*, 2008). Attempts to express these proteins in mesophilic Gram positive hosts (notably lactic acid bacteria and *Bacillus megatarium*) have also had limited success (Gold *et al.*, 1996; Bongers *et al.*, 2005; Kaczowka *et al.*, 2005; Liu *et al.*, 2005; Talarico *et al.*, 2005; Liu *et al.*, 2006; Liu *et al.*, 2007; Orencio-Trejo *et al.*, 2008; Bi *et al.*, 2009).

The role that codon usage plays in heterologous protein expression has been recognized, but is not well understood (Gustafsson *et al.*, 2004). It has been demonstrated that the position and usage frequency of codons, together, play a role in correct protein folding and that "codon harmonization" could be used to overcome poor expression, at least in *E. coli* (Angov *et al.*, 2008; Rosano *et al.*, 2009). Incompatibilities in codon usage and their effect on expression of PDCs have been reported (Lowe *et al.*, 1992; Talarico *et al.*, 2001; Talarico *et al.*, 2005). Two examples of the effect of codon usage on PDC production include the 5-10 fold increase in soluble SvePDC when expressed in an *E. coli* strain with or without accessory tRNAs (specifically those which are rarely used in *E. coli*), as well as the superior production of this PDC relative to those from *A. pasteurianus* and *Z. mobilis* in *B. megatarium* (Talarico *et al.*, 2001; Raj *et al.*, 2002; Talarico *et al.*, 2005).

Despite the rarity of prokaryotic PDCs, we have identified a *pdc*-like gene sequence in the genome sequence of a Gram negative acetic acid bacterium, *Gluconobacter oxydans*. A PDC enzyme from *G. oxydans* has been previously characterized (King *et al.* 1954). *G. oxydans* is often associated with sugar rich environments such as ripe fruit, honeybees and cider as well as

in a variety of soil types and is used industrially to produce L-sorbose from D-sorbitol; D-gluconic acid, 5-keto- and 2-ketogluconic acids from D-glucose; and dihydroxyacetone from glycerol (Gupta *et al.*, 2001). This organism uses the PDC as part of the well characterized lactate oxidation and acetate excretion pathways (Chandra *et al.*, 2001).

The initial aims of this study were to clone, express and fully characterize the PDC from *Gluconobacter oxydans*. Due to the relatively high *in vitro* thermostability this PDC was chosen for expression studies in *G. thermoglucosidasius* NCIMB 11955 and its progeny. The principle of "codon harmonization" was applied to this gene to enhance expression in a Gram positive thermophilic host (*G. thermoglucosidasius* NCIMB 11955).

Materials and methods

Media, bacterial strains and plasmids

Bacterial strains and plasmids used in this study are shown in Table 1. *E. coli* strains were grown in Luria–Bertani (LB) broth (Sambrook *et al.*, 1989), with 200µg/ml ampicillin or 50 µg/ml kanamycin added as required. *G. thermoglucosidasius* strains were cultured either in LB, 2TY, TGP media or modified Urea Sulphates Medium (USM). In general, *E. coli* DH5α was used for plasmid construction

One liter of TGP broth contains 17 g tryptone, 3 g Soy peptone, 2.5 g K_2HPO_4 and 5 g NaCl. The pH was adjusted to 7.3 before autoclaving, after which 4 g Na-pyruvate and 4 mL glycerol were added in the form of filter-sterilized 10x concentrates. For solid media, 15 g/L agar was added before autoclaving. LB and TGP were used during genetic manipulation and general maintenance of cultures.

Per liter, 2TY medium contained 10 g yeast extract, 5 g NaCl, 20 g tryptone and 15 g agar (where applicable), with a final pH of 7.0. Urea Sulphates Medium (USM) supplemented with

yeast extract (USMYE) contained 10 g glucose, 0.42 g citric acid, 0.31 g $MgSO_4$, 3.1 g NaH_2PO_4, 3.5 g K_2SO_4, 3 g urea, 2.2 mg $CaCl_2$, 0.4 mg Na_2MoO_4, 1 g yeast extract, 1 g typtone, Bis-Tris 8.36 g, PIPES 12.08 g, HEPES 10.4 g, 1ml silicone antifoam and 5ml trace elements solution per liter. The trace element solution contained (per liter) 1.44 g $ZnSO_4.7H_2O$, 0.56 g $CoSO_4.6H_2O$, 0.25 g $CuSO_4.5H_2O$, 5.56 g $FeSO_4.6H_2O$, 0.89 g $NiSO_4.6H_2O$; 1.69 g $MnSO_4$ and 5.0 ml 12M H_2SO_4. The trace elements solution and glucose (50ml of a 20% w/v solution) were added aseptically after autoclaving. The pH of USM was adjusted to pH of 7 using 10 M NaOH. Cultures were incubated at 45°C, 52°C or 60°C as required with vigorous aeration.

G. oxydans was cultured in medium containing, per liter: 8g yeast extract, 15g peptone, 10 g glucose, 0.5% (w/v) ethanol and 0.3% (w/v) acetic acid. The final pH was between 3.5 and 4. Ethanol, acetic acid and glucose were added after autoclaving. Cultures were incubated at room temperature.

DNA manipulations and sequencing

Plasmid preparation, restriction endonuclease digestion, gel electrophoresis, ligation, and Southern/colony blot hybridization were performed using standard methods or following the manufacturers' recommendations (Sambrook *et al.*, 1989). Ultrapure plasmid DNA was obtained using the Wizard Plus SV miniprep DNA purification system (Promega™). Total DNA from all bacterial strains was prepared as described (Kotze *et al.*, 2006). The QIAGEN plasmid midi kit was used for Large-scale plasmid preparations. DNA was sequenced using an ABI Prism 377 automated DNA sequencer and sequences were analyzed with DNAMAN (version 4.1, Lynnon BioSoft). Codon usage in *G. thermoglucosidasius* NCIMB 11955 (with particular reference to the PDC genes from *G. oxydans*) were analyzed using the web-servers http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=1426&aa=1&style=N and http://gcua.schoedl.de/.

Polymerase chain reaction (PCR)

PCR was performed using Phusion DNA polymerase (New England Biolabs™). Generally, 50 ng DNA were used in a 50 µl reaction volume containing 2 mM $MgCl_2$, 0.125 µM of each primer, 0.2 mM of each deoxynucleoside triphosphate, and 1 U DNA polymerase. Reactions were carried out in a Hybaid Sprint thermocycler, with an initial denaturation at 94°C for 60 s, followed by 30 cycles of denaturation (30 s at 94°C), annealing (30 s) and variable elongation (72°C), where annealing temperatures and elongation times were adjusted as required. Primers are listed in Table 1.

Table 1: Bacterial strains, plasmids and primers used in this study. Underlined sections in primer sequences indicate restriction endonuclease sites

Strain or plasmid	Genotype or Description	Source or reference
Strains		
G. thermoglucosidasius TM89	*ldh*A$^-$ variant of *G. thermoglucosidasius* NCIMB 11955	TMO Renewables
G. oxydans	Wild type isolate	IWBT, University of Stellenbosch lab collection
E. coli DH5α	F´/*endA1 hsdR17* ($r_K^- m_K^+$) *supE44 thi-1 reacA1 gyrA* (Nal^r) *relA1* *Δ(lacZYA-argF)U169 (ϕ80dlacΔ(lacZ)M15)*	Promega Corp.
E. coli BL21-DE3	*E. coli* B F$^-$ *dcm ompT hsd*S(r_B^- m_B^-) *gal* λ(DE3)	Invitrogen Corp.
E. coli JM109	F´ *traD36 proA$^+$B$^+$ lacIq Δ(lacZ)M15/ Δ(lac-proAB) glnV44 e14$^-$ gyrA96 recA1 relA1 endA1 thi hsdR17*	New England Biolabs, Beverley, MA, USA
Plasmids		
pET28a	Kan^r; ColE1 replicon, HIS-tag expression vector	Novagen Corp.
pGO	Kan^r; ColE1 replicon; *G. oxydans pdc* gene cloned into pET28a	This study
pldhGO	Kan^r; ColE1 replicon; Lactate dehydrogenase (P*ldh*) gene promoter region (±170bp *Nco*I-*Nde*I) from *G. thermoglucosidasius* NCA1503 cloned upstream of the *G. oxydans pdc* gene in pET28a	This study
pTM049	Derivative of pUB190 containing the *ldh* promoter from *G. stearothermophilus* NCA1503.	TMO Renewables, Cripps 2009
pTMO111	Amp^r, Kan^r (in *G. thermoglucosidasius*) ColE1 replicon, pUB110 IncA replicon, *E. coli - G. thermoglucosidasius* shuttle-suicide (>55°C) vector	TMO Renewables, Cripps 2009

	containing a truncated *pflB* gene	
pGO111	Amp^r, Kan^r (in *G. thermoglucosidasius*), 3603bp *Dra*III-*Eco*RV fragment, blunted at the *Dra*III end, from pldhGO cloned into the unique *Swa*I site of pTMO111	This study
pGOF111	Amp^r, Kan^r (in *G. thermoglucosidasius*), 1887bp *Not*I-*Not*I fragment, containing the fully codon optimized *G. oxydans* PDC with the P*ldh* upstream, cloned into the unique *Not*I site of pTMO111	This study
Primers		
LDHF	5'-TATACCATGGGCGGGACGGGGAGCTGAGTGCTC-3'	Cripps 2009
LDHR	5'-GCCGCATATGATTCATCCTCCCTCAATAT-3'	Cripps 2009
GOPDCpETF	5'-GGAATTCCATATGACTTATACTGTCGG-3'	This study
GOPDCpETR	5'-CCGCTCGAGTCAGACGCTCTGCGG-3'	This study

Cloning of the *G. oxydans pdc* gene

The *pdc* gene was amplified using genomic DNA isolated from a wild type isolate of *G. oxydans* using primers GoxPDCpETF and GoxPDCpETR. The gene encoding the *G. oxydans* pyruvate decarboxylase was cloned into the pET28a expression vector in two parts. The 5' 913 bp fragment was cloned by digesting both PCR product and target vector, pET28a with *Nde*I and *Xho*I followed by sequencing using the T7 promoter and terminator primers. The 779bp 3'-fragment was cloned into pBluescriptSK by digesting the GoxPDC PCR product with *Xho*I (XhoI-XhoI) and sequenced using the primers M13F and M13R. The fragment was again excised by *Xho*I digestion and cloned into *Xho*I-digested pET28–GoxPDC-5'. Clones with the correct orientation of this 779bp fragment were identified by restriction enzyme digest of the final construct (pGO) using *Sph*I.

Purification of PDC protein

An overnight culture of pGO in *E. coli* BL21-DE3 with kanamycin was used to inoculate fresh LB broth (1 ml / 100 ml) and then incubated with aeration (120 rpm) overnight at room temperature to express the protein without IPTG induction. The cells were collected by centrifugation (5000 rpm for 10 min). BugBuster™ was used to lyse cells (3 ml/g of wet cells)

and the suspension incubated at room temperature for 20 min with shaking. After centrifugation to remove cell debris (10000 rpm for 20 min), DNaseI and RNaseA (Fermentas) were added (10 U/ml) to the lysate to reduce the viscosity and incubated at room temperature with shaking for thirty minutes. The HisBind™ resin and buffer kit (Novagen) was used to purify the protein. After elution with 9ml imidazole buffer (100mM), the protein was dialysed against 200 volumes of buffer (50 mM MES pH 6.4) containing 1 mM TPP and 1 mM $MgCl_2$. Purity was estimated by reducing SDS-PAGE gel (12%) and the protein concentrations determined using Bradford reagent (Bio-Rad) with bovine serum albumin as the standard (Laemmli, 1970).

Steady state kinetics and substrate specificity

PDC activity was measured using a coupled assay with baker's yeast ADH (Sigma-Aldrich) as described (Conway *et al.*, 1987). The reaction mixture (1ml final volume) contained 0.25 mM NADH, 5 mM $MgCl_2$, 0.1 mM TPP, 5 mM pyruvate (unless stated otherwise), and 10 U of ADH in 50 mM MES or 200mM Na citrate pH 6.4 or 6.0, respectively. For substrate range determination, ADH was replaced with 1 U/ml baker's yeast aldehyde dehydrogenase (ALDH, Sigma-Aldrich). β-mercaptoethanol was added to a final concentration of 3 mM and NADH replaced with NAD^+. Assays were performed in 100mM citric acid/K_2HPO_4 buffer pH7.0 (Vuralhan *et al.*, 2005). Activities were recorded at 25°C unless otherwise indicated using a Cary 50 temperature controlled spectrophotometer (Varian). In the case where aldehydes produced by the decarboxylation of certain substrates by PDC were not recognized as a substrate for ALDH, HPLC analysis was used to determine activity on that substrate. Reactions were run on a Rezex RHM monosaccharide column (Phenomonex), using 5mM H_2SO_4 as mobile phase under isocratic elution (0.6ml/min, 40°C) on a Dionex Ultimate 3000 machine. 20µl samples were injected by autosampler and the components detected using either a refractive index detector or a UV/Vis photodiode array at 215 nm. For kinetic data, initial rates were measured over the substrate range of 0.1mM to 30mM for pyruvate or 0.1mM to 50mM for other 2-keto acids. Kinetic parameters were determined by non-linear fitting of data to hyperbolic curves according to Michaelis-Menten (GraphPad Prism v. 4.00, GraphPad Software, San Diego, CA, USA). One unit of enzyme activity corresponds to the amount of enzyme that generates 1 µmol

of acetaldehyde per minute. k_{cat} values were calculated based on the MW of the monomer with one active site.

Construction of the *pdc* knock-in vectors and transformation

For expression of PDC in *G. thermoglucosidasius*, the 170-bp promoter region of the lactate dehydrogenase gene P*ldh* from *G. thermoglucosidasius* NCA1503 was cloned upstream of the Gox*pdc* gene. This promoter has been shown to be induced under microaerobic conditions in *G. thermoglucosidasius* (Cripps *et al.*, 2009). The promoter region was amplified from pTMO49 using the LDHF and LDHR primers and cloned into pGEM-T Easy. Sequencing confirmed that no DNA base changes had occurred. P*ldh* was cloned into pGO (Table 1) using the *Nco*I and *Nde*I sites such that the promoter was functional for *pdc* expression. This construct was digested with *Dra*III and the ends filled in using T4 DNA polymerase (Fermentas, ThermoFisher). Digestion with *Eco*RV yields a 3603 bp *Dra*III (blunt) - *Eco*RV fragment. The plasmid TMO111 was digested with *Swa*I and treated with rAPid™ alkaline phosphatase (Roche) to prevent self ligation. The *Dra*III (blunt) – *Eco*RV fragment and *Swa*I digested pTMO111, were ligated using T4 DNA ligase (Fermentas). Insertion at the *Swa*I site leaves 809 bp and 436 bp of the *pflB* gene on either side of the *ldh-pdc* for recombination with its chromosomal counterpart.

The constructs were passaged through *E. coli* JM109 for DNA methylation prior to transformation to prevent endonuclease degradation in *G. thermoglucosidasius*. *G. thermoglucosidasius* competent cells were prepared and transformed (Cripps *et al.*, 2009).

Fermentative product profile quantification

Cultures expressing *pdc* gene were grown overnight at 37°C for 16 hours and 200 rpm in LB media (*E. coli*) and 2TY media (*G. thermoglucosidasius*). A volume of 0.5ml of this culture was transferred to 10 ml of USMYE media, contained separately, in 50ml and 15ml screw cap universal tubes. This effectively generated 40ml and 5ml headspaces respectively, mimicking aerobic and microaerobic or fermentative culture conditions. These cultures were grown overnight at 37°C (*E. coli*) or 52°C (*G. thermoglucosidasius*) for 16 h and 200 rpm and the

supernatant removed by centrifugation (4000 rpm, 10 min). Metabolite concentrations in culture supernatants were determined by HPLC (see above) and products were compared to suitable standards of know concentration and against the media in which the cultures were grown. Experiments were carried out in triplicate for *E. coli* and in duplicate for *G. thermoglucosidasius*.

Western blotting

Rabbit anti-GoxPDC polyclonal antibodies were made by Antibodies Incorporated (Davis, CA, USA) using His-tag purified GoxPDC protein. Cells were harvested directly after fermentation by centrifugation at 5000rpm for 10min. The cell pellet was resuspended in MES buffer pH6.5 and sonicated using five pulses of thirty seconds each. Cell debris was removed by centrifugation (13000rpm for 20min), and the supernatant decanted. Protein concentrations were determined by Bradford assay. 40µg total protein was loaded for each sample and run on a 12% SDS-PAGE gel. Protein was transferred from the gel to Biotrace™PVDF membrane by semi-dry blotting. For signal detection, the anti-rabbit Super Signal West Femto Chemiluminescent Substrate (Pierce) kit was used and the signal visualized using a chemiluminescent camera.

Results

Amino acid sequence considerations in the *G. oxydans* PDC

The GoxPDC gene was amplified, cloned and sequenced as described in the materials and methods. The DNA sequence generated from our study differed from the published genome sequence (NC_006677.1) at 22 positions (indicated in Figure 1) which in turn gave rise to 5 amino acid changes (Y163F, S207N, A209T, I469M and D517E) but no frame shifts or deleterious events. The sequence alignment (Figure 1) indicates that none of the affected residues have been shown to be directly involved in catalysis, substrate or co-factor binding. Most changes are conservative (F163, N207 and E517) and/or are located far away from the active site (F163, N207, T209 and E517). Ile469 is, however, located in a region of the enzyme which may be sensitive to changes. It is positioned adjacent to Glu468, important in catalytic

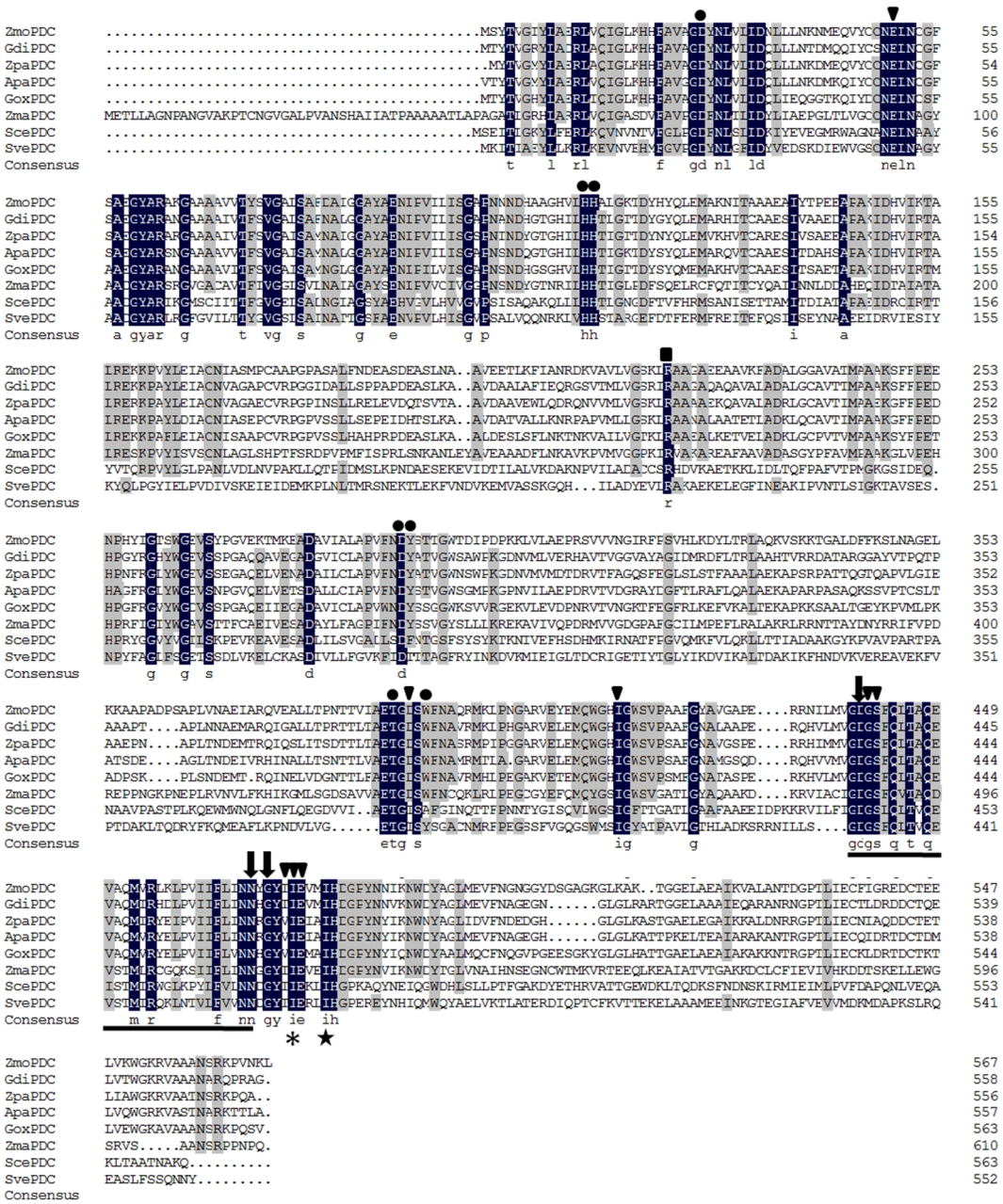


Figure 1: Multiple sequence alignment of the protein sequences from selected PDC proteins. The alignment was generated using the "full alignment" feature in DNAman with default setting. Residues shaded in black are 100% conserved while those in grey are 75% conserved The underlined region shows the conserved ThDP-binding motif and triangles indicate those residues which bind ThDP. Arrows indicate Mg^{2+} binding residues. Circles indicate residues which line the catalytic pocket and are thought to play a role in catalysis. Asterisk indicates the residue involved in substrate specificity. The star shows a residue thought to be needed for positioning of the substrate for catalysis. The square indicates the arginine residue involved in substrate activation of ScePDC and SvePDC. GdiPDC - *G. diazotrophicus*; GoxPDC - *G. oxydans*; ZpaPDC - *Z. palmae*; ZmoPDC - *Z. mobilis*; ZmaPDC - *Z. mays*; ScePDC - *S. cerevisiae*; SvePDC - *S. ventriculi*.

activity, Ile467, involved in substrate recognition, and Ile471, crucial for substrate positioning (Pohl *et al.*, 1998; Siegert *et al.*, 2005; Meyer *et al.*, 2010). Equivalent residues in other PDCs have generally not been associated with catalysis or substrate recognition (Figure 1). It is noted that the sequence alignment shows that in most PDC's used in the alignment, except SvePDC, position 163 is occupied by a tyrosine residue: in GoxPDC residue 163 is phenylalanine. GoxPDC, unlike ZpaPDC and ApaPDC, also contains an extra five amino acid loop from position 498-503 (GEESGKY), similar to positions 503-508 in ZmoPDC (DSGAGK). Unlike any of the other Gram negative PDC's, ZmoPDC has another region (359-362) containing four extra amino acids (ADPS). As the encoding gene was amplified using a polymerase with 3'-5' exonuclease activity and the differences were consistent in several independent clones, we infer that the changes represent natural variations in the GoxPDC and are not artifacts of the cloning procedure.

The protein demonstrates all the features typical of ThDP-binding enzymes including the conserved ThDP binding motif GDGS-XXX-NN as well as several conserved residues required for substrate binding and catalysis (indicated in Figure 1). It does, however, lack a cysteine residue equivalent to Cys221 in ScePDC, thought to be involved in allosteric substrate activation (Lu *et al.*, 2000; Konig *et al.*, 2009).

Kinetic characterization of the GoxPDC enzyme

GoxPDC was purified to homogeneity by affinity chromatography, as judged by reducing SDS-PAGE analysis (Figure 2). The estimated molecular weight of the protein, at ±60 kDa, corresponds well to the theoretical molecular mass of 60.8 kDa for GoxPDC. The predicted pI value is 6.0.

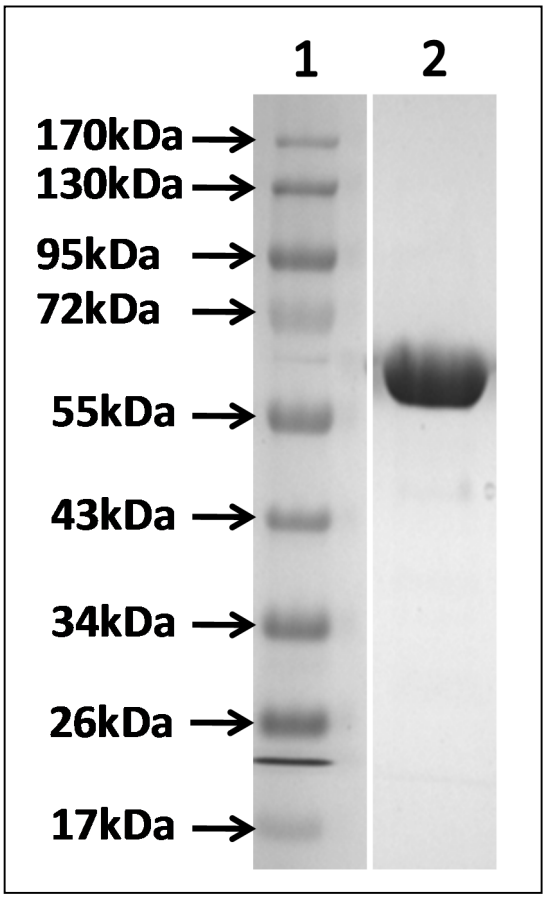


Figure 2: A denaturing SDS-PAGE gel showing purified GoxPDC protein. Lane 1, molecular weight marker (#SM0671), Lanes 2, HIS-tag purified GoxPDC protein. The GoxDC protein was approximately to 59kDa in size

Conventional enzyme characterization was performed using pyruvate as a substrate (kinetic data for GoxPDC are summarized in Table 2). The K_M value for pyruvate was found to be in the range of those determined for other PDCs from Gram negative bacteria assayed under similar conditions. The enzyme also displayed a ±20-fold decrease in the K_M for pyruvate with a decrease in pH from 7 to 5, without an equivalent change in the catalytic rate (k_{cat} showed an approximate 2 fold decrease). This is in line with previous observations in related enzymes (Raj *et al.*, 2002; Meyer *et al.*, 2010) and supports the interpretation that PDCs requires a protonated residue for efficient binding of the substrate. The ionizable group is thought to be the aminopyrimidine ring of the ThDP coenzyme (Meyer *et al.*, 2010). The GoxPDC enzyme displayed normal Michaelis Menten kinetics with pyruvate as the substrate and was not subject to allosteric substrate activation as has been seen in PDCs from plants, yeasts and the

Table 2: Steady state kinetic constants for GoxPDC compared with those from other Gram negative bacteia. Values given in brackets indicate the pH at which measurements were made.

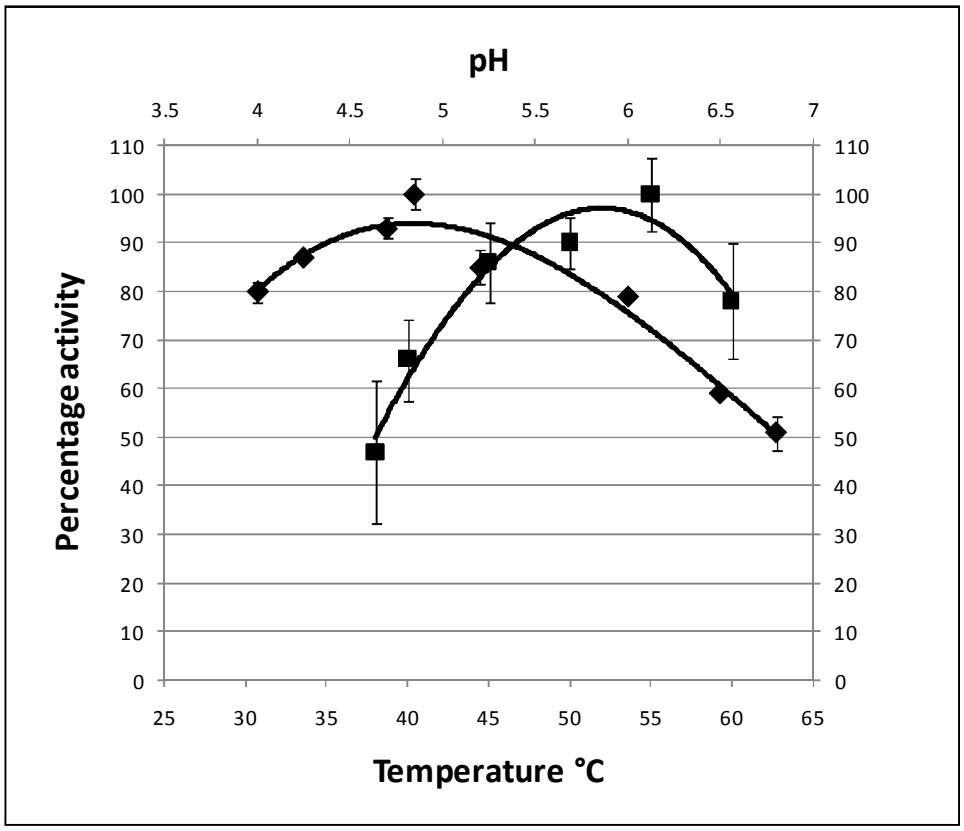
PDC	K_M (mM) assay pH shown in brackets	Specific activity in (U/mg)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}.s^{-1}$)	T_{opt} (°C)	pH_{opt}
GoxPDC	0.12 (5.0) 1.4 (6.5) 3.0 (7.0)	57 (5.0) 47 (6.5) 125 (7.0)	57 (5.0) 50 (6.5) 125 (7.0)	4.75×10^5 (5.0) 3.6×10^4 (6.5) 4.2×10^4 (7.0)	53	4.5-5.0
ApaPDC	2.8 (6.5)[a] / 0.39 (5.0)[e]	110 (6.5)[a] / 97 (5.0)[e]	341-508[e]	N/A	65[a]	3.5 - 6.5[a]
ZpaPDC	2.5 (6.5)[a] / 0.24 (6.0)[e]	116 (6.5)[a] / 130 (6.0)[e]	341-508[e]	N/A	55[a]	7.0[a]
ZmoPDC	1.3 (6.5)[a] / 0.31 (6.0)[b] / 1.1[c]	120 (6.5)[a] / 120[c]	150 (6.0)[b] / 486 (6.5)[c]	4.8×10^5 (6.0)[b] / 4.4×10^5 (6.5)[c]	60[a]	6.0-6.5[a]
SvePDC	13[d]	103[d]	412 [d]	3.2×10^4 [d] / 0.87×10^4 (7.0)	N/A	6.3 - 6.7

a – Gocke *et al*, 2009; *b* – Meyer *et al*, 2010 ; *c* - Siegert *et al*, 2005; *d* - Lowe *et al* 1992; *e* - Raj *et al* 2002

Values in brackets indicate pH

bacterium *S. ventriculi* (Konig *et al.*, 2009). The GoxPDC has a lower catalytic rate than the *A. pasteuranius* homologue but the catalytic efficiencies were similar to those reported for *Z. mobilis* and *S. ventriculi* PDCs.

Figure 3: Effect of pH on the activity of GoxPDC (♦) when using pyruvate as substrate. The assay buffer used was



100mM Na_2HPO_4/citrate buffer. The 100% activity was analogous to a specific activity of 72 U/mg for Topt and 160 U/mg for pHopt.

The pH optimum of GoxPDC was determined to lie between 4.5 and 5.0, similar to that of the PDC from *A. pasteurianus* and slightly lower than for other PDCs from Gram negative bacteria (Figure 3) (Gocke *et al.*, 2009). The temperature optimum of GoxPDC was between 50 to 55°C. Thermal inactivation studies demonstrated that the enzyme was fully stable at the experimental T_{opt}, with no loss of activity after an hour of incubation. However, at temperatures ≥ 60°C, moderate to rapid loss of activity was recorded, with only 30% of the initial activity retained at 65°C after an hour of incubation (Figure 4). These data demonstrate that GoxPDC is more thermostable than homologous enzymes from other Gram negative bacteria.

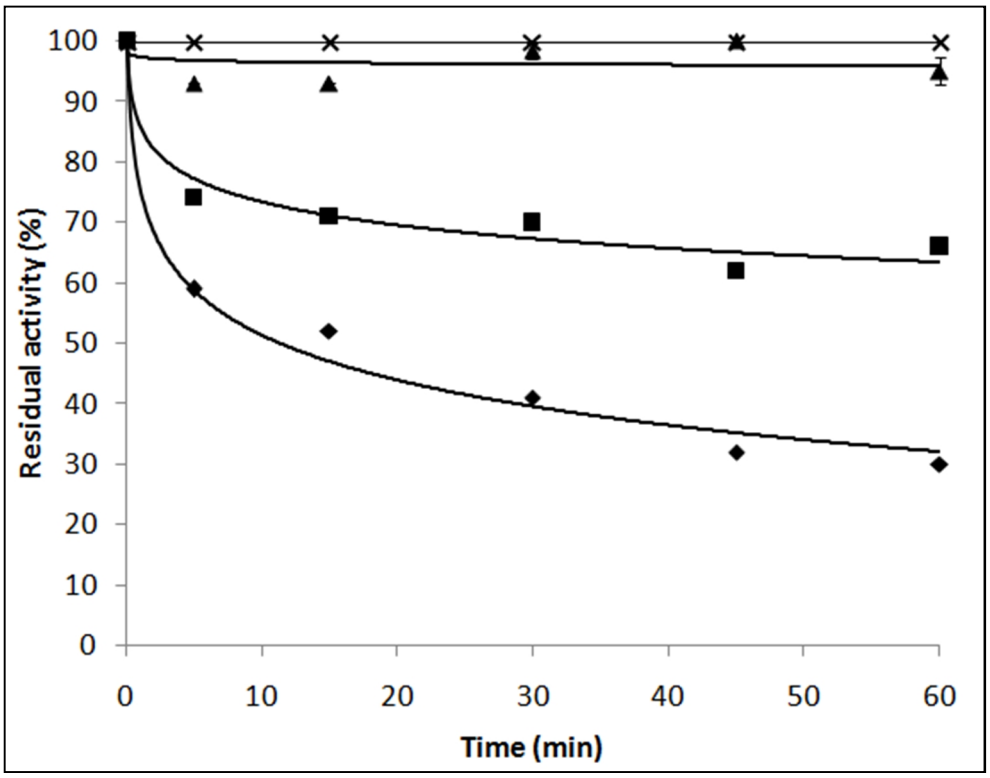


Figure 4: A thermal inactivation profile of GoxPDC at 25°C (x), 55°C (▲), 60°C (■) and 65°C (♦). Activity is expressed as a percentage of that at time zero in the standard assay at 25°C. The activity at 100% correlates to a specific activity of 26 U/mg. Assays were performed in 50mM MES buffer at pH6.5

The following substrates were shown by HPLC analysis to be decarboxylated by GoxPDC; 2-ketopropanoate (pyruvate), 2-ketobutanoate, 2-ketopentanoate, 2-keto-4-methylpentanoate and 4-hydroxy-phenyl-pyruvate. Calculated specific activities were 32 U/mg for 2-keto-butanoate, 1.2 U/mg for 2-keto-pentanoate and 0.2 U/mg for 2-keto-4-methylpentanoate. No activity was recorded on 3-phenyl-2-oxopropanoate, benzoyl formate, 4-hydroxy-phenyl pyruvate and indole-3-pyruvate.

Native *pdc* (GoxPDC$_{WT}$) expression in *G. thermoglucosidasius* and *E. coli*

The expression of the GoxPDC$_{WT}$ gene was initially attempted in *G. thermoglucosidasius* TM89 (a Δ*ldh* mutant of *G. thermoglucosidasius* NCIMB 11955). The gene was cloned into the *G. thermoglucosidasius* specific expression vector (pTMO111) under the control of the strong *ldh*

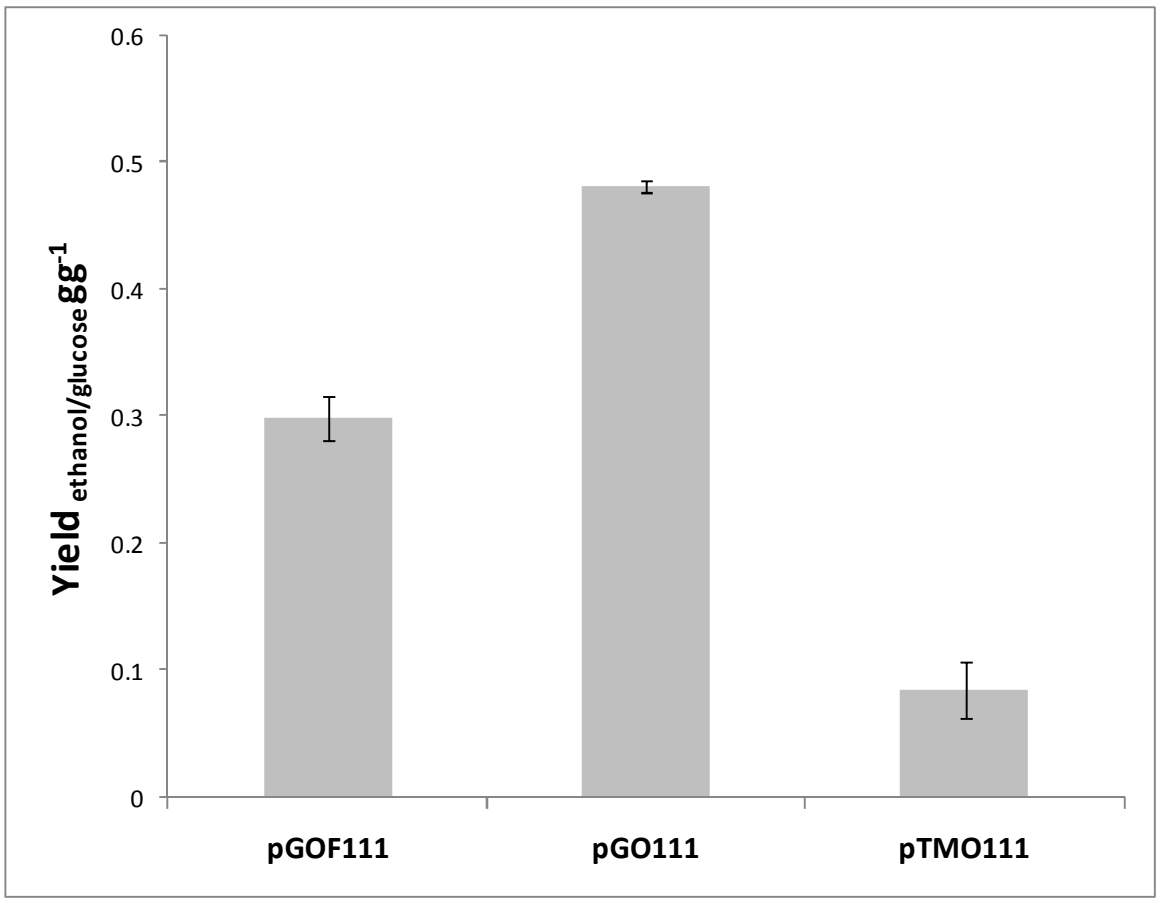
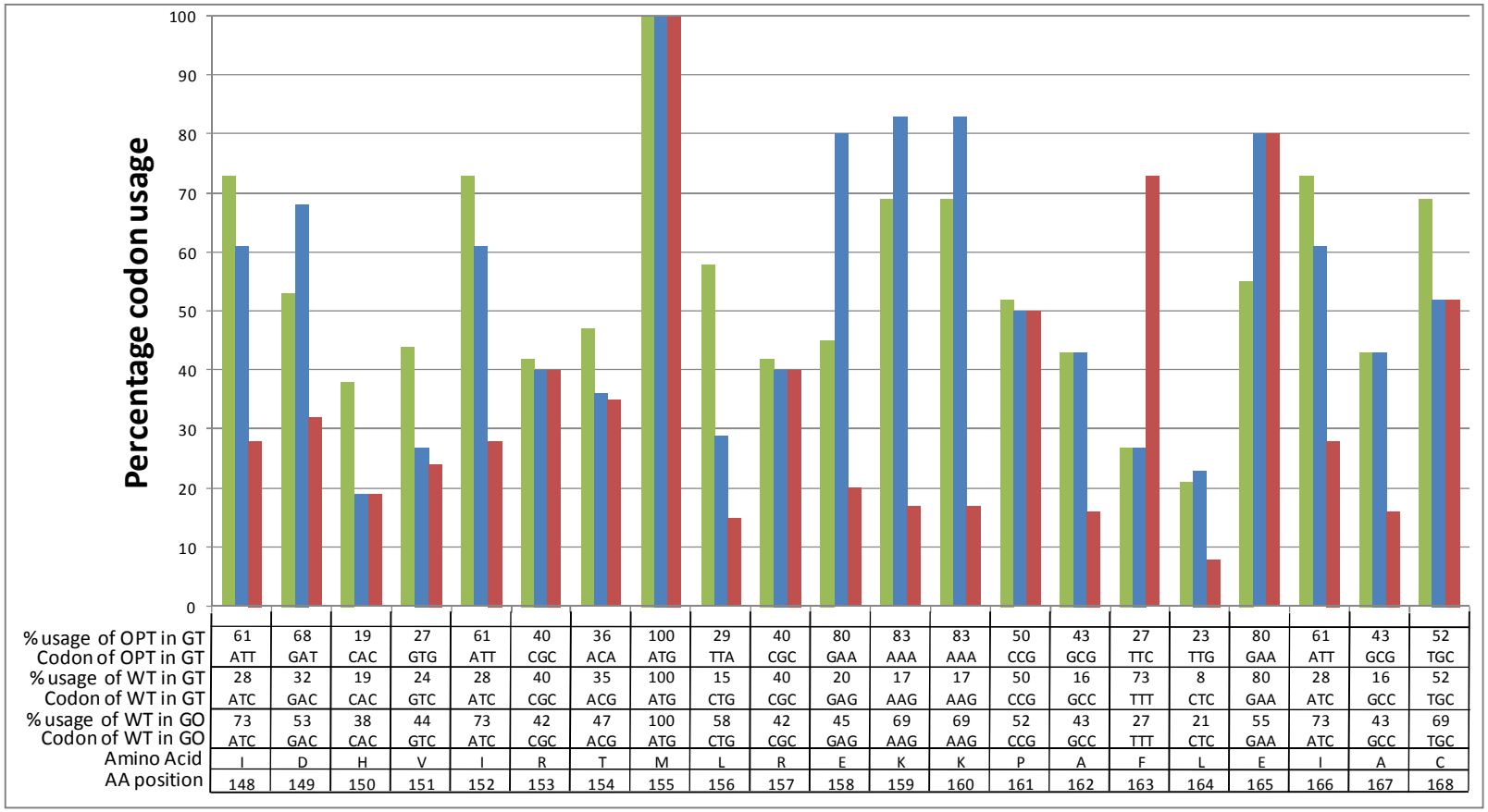


Figure 5. Comparison of the yield of ethanol produced per gram of glucose consumed during 10/15 model fermentations performed in *E. coli* DH5α. The data represents an average of three independent fermentations (batches of medium) with three repeats of each culture in each fermentation after 48H at the given temperature

promoter (Cripps *et al.*, 2009) and transformed into TM89 (pGOF111). Initial growth studies under microaerobic conditions (10ml of USM media, 1ml inoculum in a 15ml screw capped tube) showed no improvement in ethanol yield per gram glucose consumed compared to the GoxPDC-minus strain, irrespective of the growth temperature (Figure 8).

In further assessing the failure of functional GoxPDC expression, RT-PCR confirmed that the gene was transcribed (data not shown). However, no soluble protein could be detected by Western blotting for cultures grown at 45°C (Figure 7) and no PDC activity was detected using cell free extracts from the same cultures. Together these data suggest a failure at the level of



% usage of OPT in GT	61	68	19	27	61	40	36	100	29	40	80	83	83	50	43	27	23	80	61	43	52
Codon of OPT in GT	ATT	GAT	CAC	GTG	ATT	CGC	ACA	ATG	TTA	CGC	GAA	AAA	AAA	CCG	GCG	TTC	TTG	GAA	ATT	GCG	TGC
% usage of WT in GT	28	32	19	24	28	40	35	100	15	40	20	17	17	50	16	73	8	80	28	16	52
Codon of WT in GT	ATC	GAC	CAC	GTC	ATC	CGC	ACG	ATG	CTG	CGC	GAG	AAG	AAG	CCG	GCC	TTT	CTC	GAA	ATC	GCC	TGC
% usage of WT in GO	73	53	38	44	73	42	47	100	58	42	45	69	69	52	43	27	21	55	73	43	69
Codon of WT in GO	ATC	GAC	CAC	GTC	ATC	CGC	ACG	ATG	CTG	CGC	GAG	AAG	AAG	CCG	GCC	TTT	CTC	GAA	ATC	GCC	TGC
Amino Acid	I	D	H	V	I	R	T	M	L	R	E	K	K	P	A	F	L	E	I	A	C
AA position	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168

Figure 6: A comparison of codon usage frequencies of GoxPDC in the native host (*G. oxydans* - green bars), the heterologous host (*G. thermoglucosidasius* - red bars) and after harmonization (blue bars), for the amino acid residues 148-168. Disparages exisit between the % usage of codons in G. oxydans for the wild type PDC and those used in Geobacillus(large differences between the green and red bars). In particular the rarely used codons in Geobacillus (<10%), such as amino acid 165, would represent likely loci where a hiatus would be reached in translation in this strain. This would not correlate with a locus of low codon usage in G. oxydans and therefore would not be suitable stop point, for protein folding to occur. Post harmonization, it can be seen that the blue and green bars are more synergized and the harmonized gene mimics codon usage much closer to that

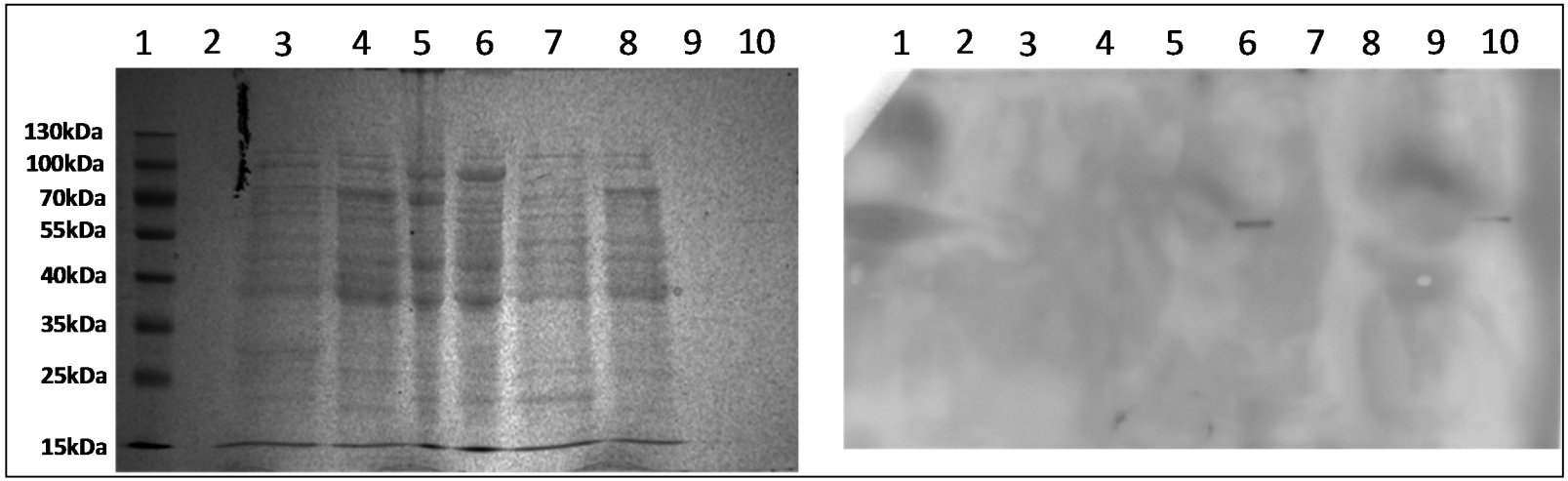


Figure 7 : SDS-PAGE and Western blots of cell extracts from *G. thermoglucosidasius* TM89 harboring either pGO111 (lane 4 and 8) or pGOF111 (3, 5, 6 and 7). Cultures were grown in tubes at either 45°C (lane 6, 7, 8) or 52°C (lane 3, 4, 5) respectively. Purified GoxPDC protein served as positive control (lane 10). Lanes 2 and 9 are empty

translation, possibly with the generation of misfolded protein which would be targeted for intracellular proteolysis.

E. coli containing the same expression vector construct (pGOF111) produced 0.5 g/g ± 0.005 ethanol per gram of glucose consumed under fermentative conditions, substantially higher than control constructs (DH5α-pTMO111) which produced only 0.1 g/g ± 0.01 (Figure 5). Cell densities were of the same order of magnitude, demonstrating that the higher ethanol concentrations were not simply the result of higher biomass levels in GoxPDC-positive strains. It is noted that these elevated ethanol yields were achieved in the absence of a recombinant *adh*II, which has previously been shown to be essential for enhanced ethanol production in *E. coli* strains expressing *pdc* from *Z. mobilis* (Lawford *et al.*, 1991). No previous studies reporting the functional expression of the *pdc* gene alone in *E. coli* have shown any evidence of elevated ethanol yields (Talarico *et al.*, 2001; Liu *et al.*, 2005; Talarico *et al.*, 2005).

The successful expression of functional GoxPDC in the gram-negative *E. coli* host and the absence of such expression in the gram-positive *Geobacillus* host led us to investigate strategies for enhancing protein folding in the latter. Codon harmonization (Angov *et al.*, 2008; Angov, 2011; Angov *et al.*, 2011) is thought to assist protein folding in heterologous host systems by mimicking the translation rates of the native host in a heterologous strain. The translation rate is determined by codon usage frequencies, where the presence of infrequently used codons forces a reduction of the translation rate, allowing evolving protein to fold in phases (Angov *et al.*, 2008; Angov, 2011; Angov *et al.*, 2011). Both the frequency and positioning of infrequently used codons is critical for protein folding.

Assessing codon usage and predicting a gene sequence for harmonization of GoxPDC expression in *G. thermoglucosidasius*

In order to establish a basis for codon optimization and harmonization, a comparison of codon usage frequencies for PDCs expressed in their native hosts, in *E. coli* and in *G. thermoglucosidasius* was performed (summarized in Table 3).

Table 3. Comparison of *E. coli* and *G. thermoglucosidasius* codon usage with respect to bacterial PDC sequences. The number in brackets represents the percentage rare codons compared to full length gene.

G. thermoglucosidasius		
	% codons in gene with <10% usage	**% codons in gene with 10-20% usage**
GoxPDC	21/564 (3.7%)	185/564 (32.8%)
ApaPDC	18/558 (3.2%)	165/558 (29.6%)
ZpaPDC	9/557 (1.6%)	171/557 (30.7%)
ZmoPDC	16/569 (2.8%)	156/569 (27.4%)
SvePDC	0/553 (0%)	25/553 (4.5%)
E. coli		
	% codons in gene with <10% usage	**% codons in gene with 10-20% usage**
GoxPDC	14/564 (2.5%)	76/564 (13.5%)
ApaPDC	9/558 (1.6%)	79/558 (14.2%)
ZpaPDC	7/557 (1.3%)	100/557 (18%)
ZmoPDC	13/569 (2.3%)	109/569 (19.2%)
SvePDC	48/553 (8.7%)	187/553 (33.8%)
Natural host		
GoxPDC	5/564 (0.9%)	36/564 (6.4%)
ApaPDC	9/558 (1.6%)	47/558 (8.4%)
ZpaPDC	27/557 (4.8%)	26/557 (4.66%)
ZmoPDC	11/569 (1.9%)	90/569 (15.8%)
SvePDC	14/553 (2.5%)	36/553 (6.5%)

It is well known that the presence of rarely used codons is detrimental to efficient heterologous protein expression (Kim et al., 2006; Rosano *et al.,* 2009). The data presented in Table 3 suggest that, based on the total number of codons in Gram negative PDC genes with less than 10% usage or 10 to 20% usage, they should be poorly expressed in the Gram positive host *G. thermoglucosidasius*. In contrast, the data suggest that SvePDC should be poorly expressed in *E. coli* and would be more efficiently expressed in a Gram positive host such as *B. megaterium*. Both these predictions are supported by expression data for SvePDC in *E. coli* and in *B. megaterium*, where expression in the *E. coli* host was only seen in BL21-Codon plus-RIL strains while the protein expressed to high levels in *B. megaterium*, without the need to engineer the gene sequence or host strain (Talarico *et al*., 2005). A similar prediction that Gram negative PDCs should express well in an *E. coli* background and poorly in a Gram positive host is supported by the wealth of expression data of these genes in various *E. coli* hosts (Ingram *et al.*, 1987; Chandra Raj *et al.*, 2001; Talarico *et al.*, 2001; Raj *et al.*, 2002; Talarico et al., 2005).

However, expression of three Gram negative PDC's (ZmoPDC, ZpaPDC and GoxPDC) in *G. thermoglucosidasius* showed substantial qualitative differences (Thompson et al, 2008). Production of soluble ZmoPDC in cell free extracts of *G. thermoglucosidasius* grown at 52°C, 54°C, 56°C and 58°C was observed to decrease with increased temperature and PDC activity was undetectable above 52°C (Thompson *et al.*, 2008). For ZpaPDC, activity was absent at growth temperatures above 45°C (Taylor *et al.*, 2008), while in this study, the $GoxPDC_{WT}$ showed no activity at 45°C (see above). These data collectively suggest that an analysis of the number of rare or infrequently used codons is insufficient to accurately predict the ability of genes to express in heterologous hosts.

Further analysis of rare codons only (<10% usage; shown in Table 4) demonstrated that ZmoPDC has a high coincidence of both frequency of usage and positioning of rare codons for expression in *G. thermoglucosidasius*. Of the eleven codons (CTC, 8% usage) which are rarely used by *Z. mobilis* in translation of the *pdc* gene, all are recognized as rare in *G. thermoglucosidasius* (i.e., the eleven CTC codons are in the same position and used at the same

Table 4: The correspondence of rare codons (<10% usage) for each PDC, between their native host and *G. thermoglucosidasius*

	All codons with <10% usage in the native host (number of codons in respective *pdc* gene)	**All codons with <10% usage in *G. thermoglucosidasius* (amino acid position)**	**% codon usage for selected codons in their native host**
ZmoPDC	CTC(11)	CTC (2,18,30,56,164,174,215,236,306,320,348,362,400,436,567,568), CCC (174,320), AGT (2,56,362)	CCC 10-20%, AGT 10-20%
ZpaPDC	CTA(4), TTG (3), ACA(3), CTC (3), GTG(2), CTT(2), GCG(2), TCA(2) , CCC (1), AGT (1), GGA (1), TCG (1), AAG (1), CCA (1)	CTC (35,187,517), CCC (356), CTA (12,155,231,509), AGT (193)	
ApaPDC	TCA(4), ACT(2), GGA(2), AGG (1)	CTC (30,37,38,84,352), CCC (101,174,212,251,268,285,299,302,309,336,415,454), AGG (515)	CTC 10-20%, CCC >20%
GoxPDC	ACT(2), CGA (1), TCT (1), TCA ()1	CTC (30,33,84,95,164,206,226,305,431,505,531,536,545), CCC (174,239,251,255,309,348,396), CGA (12)	CTC >20%, CCC >20%

N/A – Not applicable

frequency in both the native and heterologous hosts). Although CCC (2% usage) and AGT (8% usage) are also rarely used in *G. thermoglucosidasius*, both are used between 10 and 20% (CCC, 20% usage, AGT, 12% usage) in *Z. mobilis* which does not represent a large deviation from the frequency in the native host.

In GoxPDC there are five native rare codons, one of which (CGA) is also rarely used in *G. thermoglucosidasius*. However, for *G. thermoglucosidasius* twenty codons in the PDC gene (CCC and CTC) are recognized as rare. For expression in *G. thermoglucosidasius*, these codons constitute sites of translation limitation and might be detrimental to folding efficiency.

The principal of codon harmonization (Angov *et al.*, 2008) is to substitute synonymous codons from the heterologous host such that the codon usage frequency, positioning and therefore rhythm of translation follows that of the native host. In order to codon harmonize the GoxPDC gene for *G. thermoglucosidasius* expression, codon substitutions were made to the wild type gene sequence so as to match the usage frequencies found in the native host for each codon position while maintaining the amino acid composition of the wild-type protein. Figure 6 shows a twenty amino acid section of the protein to demonstrate how the harmonization was performed. This harmonization strategy should demonstrate whether the naturally evolved translation frequency alone would enable correct folding of the protein in *G. thermoglucosidasius*, without the need to calculate link / end segments (Thanaraj *et al.*, 1996; Angov *et al.*, 2008).

Expression of GoxPDC$_{WT}$ and GoxPDC$_{OPT}$ in *E. coli* and *G. thermoglucosidasius*

The fully codon harmonized gene was synthesized, cloned into an expression vector under the control of an *ldh* promoter (as described for the wild type gene; pGOF111) and the construct transformed into *G. thermoglucosidasius* TM89. Expression was evaluated at 45°C and 52°C (Figure 8) using expression of the wild-type *Gox pdc* gene in *G. thermoglucosidasius* as a control.

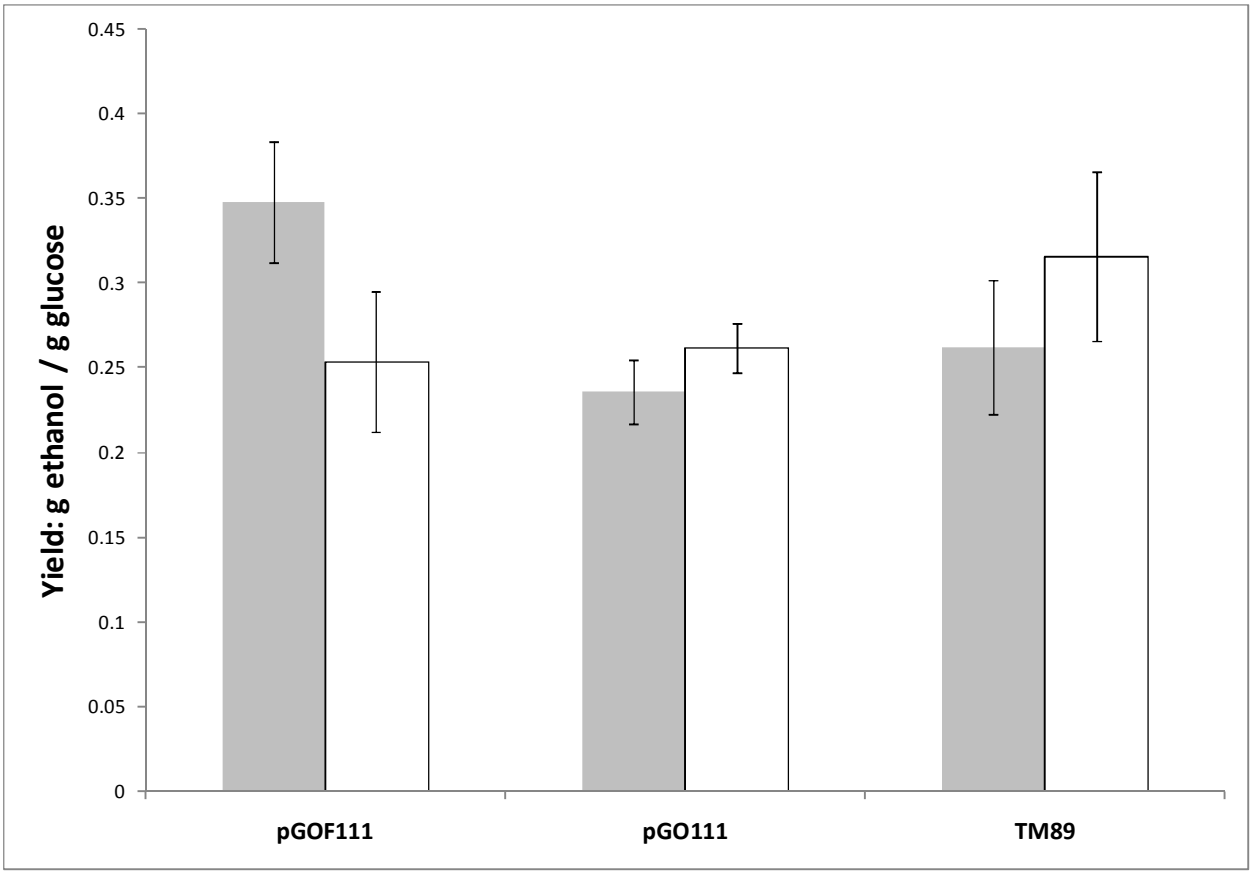


Figure 8: Comparison of the yield of ethanol produced per gram of glucose consumed during 10/15 model fermentations in *G. thermoglucosidasius*. Grey bars respresent fermentations performed at 45°C and white bars 52°C. The data represents an average of three independent fermentations (batches of medium) with three repeats of each culture in each fermentation after 48H at the given temperature.

G. thermoglucosidasius TM89 expressing the codon harmonized GoxPDC ($GoxPDC_{OPT}$) was shown to produce 0.35 ± 0.04 ethanol per gram of glucose consumed, compared to 0.26 g/g ± 0.04 for TM89 alone or 0.24 g/g ± 0.02 for TM89-pGO111 at 45°C (Figure 8). This result clearly demonstrates that the previous limitations in functional *pdc* expression in *Geobacillus* were, at least in part, due to a lack of codon harmonization. However, at 52°C strains expressing either the WT (0.26 g/g ± 0.01) or harmonized gene (0.25 g/g ±0.04) produced lower ethanol yields than wild-type TM89 (0.32 g/g ± 0.05) (Figure 8). We speculate that the reduced performance of TM89 when expressing either the unmodified or codon-harmonized GoxPDC at 52°C may be

due to the metabolic load imposed by the presence of the shuttle vector and/or the metabolic load imposed by misfolded PDC protein (proteotoxic stress).

The lack of improved ethanol yields at 52°C cannot be attributed to inadequate *in vitro* GoxPDC thermostability at that temperature (GoxPDC retains 100% enzymic activity at 55°C after 1 hour incubation). However, as *in vitro* thermostability is not necessarily a reliable guide of *in vivo* behavior, intracellular PDC activity was determined using the ADH coupled assay for $GoxPDC_{OPT}$-*G. thermoglucosidasius* cultures grown at 45°C and 52°C. A specific activity of 0.22 U/mg was determined for the codon harmonized gene product from cultures grown at 45°C, consistent with the observed phenotype. Neither the wild type gene product nor the pTMO111 control cultures demonstrated detectable activity under similar assay conditions. However, no PDC activity was detectable in $GoxPDC_{OPT}$-*G. thermoglucosidasius* cultures grown at 52°C. Furthermore, a positive Western blotting signal was only detectable in soluble protein extracts of recombinant *G. thermoglucosidasius* cultures grown at 45°C (Figure 7). Western blot analysis of extracts from cells grown at 52°C ($GoxPDC_{WT,}$ pTMO111 and $GoxPDC_{OPT}$) were reproducibly negative. These results, taken together, indicate that although codon harmonization enhanced the expression of GoxPDC in *G. thermoglucosidasius* TM89 cultures grown at 45°C, a second limitation, probably thermally related, prevents functional enzyme accumulation at temperatures closer to the growth optimum. The lack of detectable protein by Western blotting either indicates that a second limitation is present at the transcription – translation interface, or that the protein produced is extremely unstable and the resultant aggregates or proteolysis products are not suitable for antibody binding.

The high ethanol yield at 45°C is significant and consistent with yields reported in other multiply mutated thermophiles (Shaw *et al.*, 2008; Cripps *et al.*, 2009). It has been demonstrated that effective partitioning of carbon into biosynthesis and fermentation is critical in achieving optimal production of ethanol under fermentative conditions (Underwood *et al.*, 2002). Typically, *G. thermoglucosidasius* TM89 fermentations are characterized by the production of formate (average of 40.5 mM ± 7.2 after 48h at 45°C) and acetate, along with ethanol (Cripps *et*

al., 2009). However, it was noted that during TM89-pGOF111 fermentations at 45°C, no formate was produced but low level (± 185µM) fumarate accumulation was detected in culture supernatants. This may be due to a reduced metabolic flux through *pfl* where the active $GoxPDC_{OPT}$ may outcompete the *pfl* enzyme for pyruvate (Tolan *et al.*, 1987; Feldmann *et al.*, 1989; Underwood *et al.*, 2002; Orencio-Trejo *et al.*, 2008).

Conclusions

We have characterized a PDC from the acetic acid bacterium *G. oxydans*. The substrate recognition and decarboxylation range of the enzyme is similar to that of other Gram negative PDCs , showing a preference for short-chain aliphatic 2-keto acids (Gocke *et al.*, 2009). The value of k_{cat}/K_M for pyruvate compared to those for 2-ketobutanoate and 2-ketopentanoate, the nearest analogues, and the retention of Ile468, thought to be involved in substrate specificity, suggests that this enzyme favors pyruvate as its physiological substrate (Siegert *et al.*, 2005; Gocke *et al.*, 2009). In terms of its kinetic behavior, GoxPDC appears to behave in a manner similar to other Gram-negative bacterial PDCs, displaying the same pH dependent increase in k_{cat}/K_M while catalytic efficiency (k_{cat}) remains largely unchanged.

Our analysis of the codon usage pattern of all PDCs expressed in *G. thermoglucosidasius* to date has demonstrated that there is a correlation between the position and usage frequency of rare codons between the natural host and the heterologous host. This correlation may be responsible for the reported variations in PDC-expression efficiency: ZmoPDC > ZpaPDC > GoxPDC. It is well known that reducing the temperature of expression of a heterologous protein can improves the production of soluble protein (Correa *et al.*, 2011).

Although we could detect mRNA transcripts for $GoxPDC_{WT}$ when expressed in *G. thermoglucosidasius*, we could not detect any $GoxPDC_{WT}$ protein by Western blotting with cultures grown at 45°C and no PDC activity was found in a coupled assay with ADH. In contrast we could detect enzyme activity when the $GoxPDC_{OPT}$ was expressed in *G. thermoglucosidasius* and we could detect soluble protein by Western blot in cultures grown at 45°C. As a polyclonal

antibody was used for detection, the absence of signal for $GoxPDC_{WT}$ at 45°C suggests that the protein is either present but not exposing surface epitopes, or (more likely) rapidly turned over due to misfolding. As the cofactors ThDP and Mg^{2+} are bound in the cleft between two monomers, which requires an extensive network of interfacial contacts to form an active site and produce a functional enzyme, the detection of activity is a very good indicator of correct protein folding and subunit assembly into its quaternary structure.

When considered together with the increase in ethanol produced per gram of glucose consumed observed, the activity data suggest that codon harmonization improved translational folding and increasing the yield of correctly folded, active protein. Codon usage therefore appears to play a significant role in correct processing of GoxPDC protein, when expressed in *G. thermoglucosidasius,* and serves to reiterate the importance of codon usage in heterologous protein expression. This study also represents the first account of improved expression of a protein of mesophilic origin in a thermophilic host using this technique.

A second factor which plays a major role in determining whether or not the protein is active at any temperature is the innate ability of the protein monomers to fold correctly at that temperature. When correctly folded the GoxPDC protein displays relatively high thermostability when assayed *in vitro*. However, this does not necessarily translate to the ability to fold correctly at elevated temperatures, offering a possible explanation for the apparent failure of functional expression at 52°C, unlike ZmoPDC when expressed in the same host (Thompson *et al.*, 2008). We suggest that although codon optimization contributes to the correct folding of a nascent protein during translation of the mRNA, it cannot necessarily compensate for the kinetics involved in protein folding in temperature ranges outside those for which the protein had been selected for or evolved under. The further stabilization of the enzyme therefore represents an area of future improvement for the use of PDC in engineering superior homo-ethanolic pathways in *G. thermoglucosidasius*.

Acknowledgements

DAC wishes to thank the National Research Foundation of South Africa for financial contributions.

References

Adina-Zada, A., Hazra, R., Sereeruk, C., Jitrapakdee, S., Zeczycki, T. N., Maurice, M. S., Cleland, W. W., Wallace, J. C. and Attwood, P. V. 2011. Probing the allosteric activation of pyruvate carboxylase using 2',3'-O-(2,4,6-trinitrophenyl) adenosine 5'-triphosphate as a fluorescent mimic of the allosteric activator acetyl CoA. Arch Biochem Biophys 5092: 117-126.

Angov, E. 2011. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol J 66: 650-659.

Angov, E., Hillier, C. J., Kincaid, R. L. and Lyon, J. A. 2008. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. PLoS One 35: e2189.

Angov, E., Legler, P. M. and Mease, R. M. 2011. Adjustment of codon usage frequencies by codon harmonization improves protein expression and folding. Methods Mol Biol 705: 1-13.

Bi, C., Zhang, X., Ingram, L. O. and Preston, J. F. 2009. Genetic engineering of Enterobacter asburiae strain JDR-1 for efficient production of ethanol from hemicellulose hydrolysates. Appl Environ Microbiol 7518: 5743-5749.

Bongers, R. S., Hoefnagel, M. H. and Kleerebezem, M. 2005. High-level acetaldehyde production in Lactococcus lactis by metabolic engineering. Appl Environ Microbiol 712: 1109-1113.

Chandra Raj, K., Ingram, L. O. and Maupin-Furlow, J. A. 2001. Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactic acid by Acetobacter pasteurianus. Arch Microbiol 1766: 443-451.

Conway, T., Osman, Y. A., Konnan, J. I., Hoffmann, E. M. and Ingram, L. O. 1987. Promoter and nucleotide sequences of the Zymomonas mobilis pyruvate decarboxylase. J Bacteriol 1693: 949-954.

Correa, A. and Oppezzo, P. 2011. Tuning different expression parameters to achieve soluble recombinant proteins in E. coli: advantages of high-throughput screening. Biotechnol J 66: 715-730.

Cripps, R. E., Eley, K., Leak, D. J., Rudd, B., Taylor, M., Todd, M., Boakes, S., Martin, S. and Atkinson, T. 2009. Metabolic engineering of Geobacillus thermoglucosidasius for high yield ethanol production. Metab Eng 116: 398-408.

Feldmann, S., Sprenger, G. A. and Sham, H. 1989. Ethanol production from xylose with a pyruvate-formate-lyase mutant of Klebsiella planticola carrying a pyruvate-decarboxylase gene from Zymomonas mobilis. Appl Microbiol Biotechnol 31: 152-157.

Gocke, D., Graf, T., Brosi, H., Frindi-Wosch, I. and Walter, L. 2009. Comparitive characterisation of thiamin diphosphate-dependent decarboxylases. Journal of Molecular Catalysis B: Enzymatic 6130-35.

Gold, R. S., Meagher, M. M., Tong, S., Hutkins, R. W. and Conway, T. 1996. Cloning and expression of the Zymomonas mobilis "production of ethanol" genes in Lactobacillus casei. Curr Microbiol 334: 256-260.

Gupta, A., Singh, V. K., Qazi, G. N. and Kumar, A. 2001. Gluconobacter oxydans: its biotechnological applications. J Mol Microbiol Biotechnol 33: 445-456.

Gustafsson, C., Govindarajan, S. and Minshull, J. 2004. Codon bias and heterologous protein expression. Trends Biotechnol 227: 346-353.

Ingram, L. O., Conway, T., Clark, D. P., Sewell, G. W. and Preston, J. F. 1987. Genetic engineering of ethanol production in Escherichia coli. Appl Environ Microbiol 5310: 2420-2425.

Kaczowka, S. J., Reuter, C. J., Talarico, L. A. and Maupin-Furlow, J. A. 2005. Recombinant production of Zymomonas mobilis pyruvate decarboxylase in the haloarchaeon Haloferax volcanii. Archaea 15: 327-334.

King, T. E., and Cheldelin, V. H. 1954. Pyruvic carboxylase of Acetobacter suboxydans. J. Biol. Chem. 208:821-831

Kim, S., and Lee, S. B. 2006. Rare codon clusters at 5'-end influence heterologous expression of archaeal gene in Escherichia coli. Prot. Expr. Purif. 50: 49-57

Kane, J. F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr Opin Biotechnol 65: 494-500.

Konig, S. 1998. Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. Biochim Biophys Acta 13852: 271-286.

Konig, S., Spinka, M. and Kutter, S. 2009. Allosteric activation of pyruvate decarboxylases. A never-ending story? Journal of Molecular Catalysis B: Enzymatic 61: 100-110.

Kotze, A. A., Tuffin, I. M., Deane, S. M. and Rawlings, D. E. 2006. Cloning and characterization of the chromosomal arsenic resistance genes from Acidithiobacillus caldus and enhanced arsenic resistance on conjugal transfer of ars genes located on transposon TnAtcArs. Microbiology 152Pt 12: 3551-3560.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 2275259: 680-685.

Lawford, H. G. and Rousseau, J. D. 1991. Ethanol production by recombinant Escherichia coli carrying genes from Zymomonas mobilis. Appl Biochem Biotechnol 28-29: 221-236.

Libor, S. M., Sundaram, T. K. and Scrutton, M. C. 1978. Pyruvate carboxylase from a thermophilic Bacillus. Studies on the specificity of activation by acyl derivatives of coenzyme A and on the properties of catalysis in the absence of activator. Biochem J 1693: 543-558.

Liu, S., Dien, B. S. and Cotta, M. A. 2005. Functional expression of bacterial Zymobacter palmae pyruvate decarboxylase gene in Lactococcus lactis. Curr Microbiol 506: 324-328.

Liu, S., Dien, B. S., Nichols, N. N., Bischoff, K. M., Hughes, S. R. and Cotta, M. A. 2007. Coexpression of pyruvate decarboxylase and alcohol dehydrogenase genes in Lactobacillus brevis. FEMS Microbiol Lett 2742: 291-297.

Liu, S., Nichols, N. N., Dien, B. S. and Cotta, M. A. 2006. Metabolic engineering of a Lactobacillus plantarum double ldh knockout strain for enhanced ethanol production. J Ind Microbiol Biotechnol 331: 1-7.

Lowe, S. E. and Zeikus, J. G. 1992. Purification and characterization of pyruvate decarboxylase from Sarcina ventriculi. J Gen Microbiol 1384: 803-807.

Lu, G., Dobritzsch, D., Baumann, S., Schneider, G. and Konig, S. 2000. The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study. Eur J Biochem 2673: 861-868.

Meyer, D., Neumann, P., Parthier, C., Friedemann, R., Nemeria, N., Jordan, F. and Tittmann, K. 2010. Double duty for a conserved glutamate in pyruvate decarboxylase: evidence of the participation in stereoelectronically controlled decarboxylation and in protonation of the nascent carbanion/enamine intermediate. Biochemistry 4937: 8197-8212.

Orencio-Trejo, M., Flores, N., Escalante, A., Hernandez-Chavez, G., Bolivar, F., Gosset, G. and Martinez, A. 2008. Metabolic regulation analysis of an ethanologenic Escherichia coli strain based on RT-PCR and enzymatic activities. Biotechnol Biofuels 11: 8.

Pohl, M., Mesch, K., Rodenbrock, A. and Kula, M. R. 1995. Stability investigations on the pyruvate decarboxylase from Zymomonas mobilis. Biotechnology and Applied Biochemistry 2295-105.

Pohl, M., Siegert, P., Mesch, K., Bruhn, H. and Grotzinger, J. 1998. Active site mutants of pyruvate decarboxylase from Zymomonas mobilis--a site-directed mutagenesis study of L112, I472, I476, E473, and N482. Eur J Biochem 2573: 538-546.

Raj, K. C., Talarico, L. A., Ingram, L. O. and Maupin-Furlow, J. A. 2002. Cloning and characterization of the Zymobacter palmae pyruvate decarboxylase gene (pdc) and comparison to bacterial homologues. Appl Environ Microbiol 686: 2869-2876.

Rosano, G. L. and Ceccarelli, E. A. 2009. Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. Microb Cell Fact 8: 41.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular cloning: A laboratory manual. NY, Cold Spring Harbour Laboratory Press.

Shaw, A. J., Podkaminer, K. K., Desai, S. G., Bardsley, J. S., Rogers, S. R., Thorne, P. G., Hogsett, D. A. and Lynd, L. R. 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. Proc Natl Acad Sci U S A 10537: 13769-13774.

Siegert, P., McLeish, M. J., Baumann, M., Iding, H., Kneen, M. M., Kenyon, G. L. and Pohl, M. 2005. Exchanging the substrate specificities of pyruvate decarboxylase from Zymomonas mobilis and benzoylformate decarboxylase from Pseudomonas putida. Protein Eng Des Sel 187: 345-357.

Talarico, L. A., Gil, M. A., Yomano, L. P., Ingram, L. O. and Maupin-Furlow, J. A. 2005. Construction and expression of an ethanol production operon in Gram-positive bacteria. Microbiology 151Pt 12: 4023-4031.

Talarico, L. A., Ingram, L. O. and Maupin-Furlow, J. A. 2001. Production of the Gram-positive Sarcina ventriculi pyruvate decarboxylase in Escherichia coli. Microbiology 147Pt 9: 2425-2435.

Taylor, M. P., Eley, K. L., Martin, S., Tuffin, M. I., Burton, S. G. and Cowan, D. A. 2009. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. Trends Biotechnol 277: 398-405.

Taylor, M. P., Esteban, C. D. and Leak, D. J. 2008. Development of a versatile shuttle vector for gene expression in Geobacillus spp. Plasmid 601: 45-52.

Thanaraj, T. A. and Argos, P. 1996. Protein secondary structural types are differentially coded on messenger RNA. Protein Sci 510: 1973-1983.

Thompson, A. H., Studholme, D. J., Green, E. M. and Leak, D. J. 2008. Heterologous expression of pyruvate decarboxylase in Geobacillus thermoglucosidasius. Biotechnol Lett 308: 1359-1365.

Tolan, J. S. and Finn, R. K. 1987. Fermentation of d-Xylose to Ethanol by Genetically Modified Klebsiella planticola. Appl Environ Microbiol 539: 2039-2044.

Underwood, S. A., Buszko, M. L., Shanmugam, K. T. and Ingram, L. O. 2002. Flux through citrate synthase limits the growth of ethanologenic Escherichia coli KO11 during xylose fermentation. Appl Environ Microbiol 683: 1071-1081.

Underwood, S. A., Zhou, S., Causey, T. B., Yomano, L. P., Shanmugam, K. T. and Ingram, L. O. 2002. Genetic changes to optimize carbon partitioning between ethanol and biosynthesis in ethanologenic Escherichia coli. Appl Environ Microbiol 6812: 6263-6272.

Vuralhan, Z., Luttik, M. A., Tai, S. L., Boer, V. M., Morais, M. A., Schipper, D., Almering, M. J., Kotter, P., Dickinson, J. R., Daran, J. M. and Pronk, J. T. 2005. Physiological characterization of the ARO10-dependent, broad-substrate-specificity 2-oxo acid decarboxylase activity of Saccharomyces cerevisiae. Appl Environ Microbiol 716: 3276-3284.