Directed Evolution of Yeast Pyruvate Decarboxylase 1 for Attenuated Regulation and Increased Stability

Bradley J. Stevenson,* Jian-Wei Liu, and David L. Ollis

Research School of Chemistry, Australian National University, Canberra 0200, Australia

Received September 11, 2007; Revised Manuscript Received December 4, 2007

ABSTRACT: Five generations of directed evolution resulted in yeast pyruvate decarboxylase 1 (Pdc1) variants with improved activity for 1 mM pyruvate at pH 7.5 in the presence of phosphate. The best variant, named 5LS30, contained the following mutations: A143T, T156A, Q367H, N396I, and K478R. In comparison with native Pdc1, 5LS30 had the substrate concentration required for half-saturation reduced by almost 3-fold at pH 7.5 and the phosphate inhibition reduced by 4-fold at pH 6.0. The apparent cooperativity for pyruvate displayed by 5LS30 was also reduced since it appeared to be activated by pyruvate more easily than the native enzyme. The temperature at which half of the Pdc1 activity was irreversibly lost in 5 min increased from 52.6 °C, seen with the native form, to 61.8 °C for 5LS30. Curiously, the optimal temperature for Pdc1 activity was found to be dependent upon pyruvate concentration. In 1 mM pyruvate, native Pdc1 performed optimally at 30 °C and 5LS30 at 40 °C, whereas in 25 mM pyruvate native activity peaked at 45 °C and 5LS30 at 55 °C. Two screening processes were developed for directed evolution of Pdc1 expressed in *Escherichia coli*: colony screening and culture screening. The latter proved to be an ideal method for isolating PCR-generated variants of the *pdc1* gene with the desired phenotype. In this process, cultures were diluted and partitioned within 96-well plates such that each culture aliquot contained an average of two unique genotypes. This allowed rapid preparation of libraries for analysis of activity in crude lysates and can be applied to other directed evolution projects.

Pyruvate decarboxylase 1 (Pdc1)[1] from *Saccharomyces cerevisiae* catalyzes the essentially irreversible conversion of pyruvate to acetaldehyde according to the reaction:

$$\text{pyruvate} + H_2O \rightarrow \text{acetaldehyde} + CO_2 + OH^-$$

Thiamin diphosphate (ThDP) is essential for catalysis and can catalyze pyruvate decarboxylation without Pdc1, albeit 10^{12} times slower than the Pdc holoenzyme (*1*). Pdc1 increases the ionization rate of ThDP's thiazolium ring and provides a catalytic site that increases the chance that the resulting ThDP C2 anion will attack pyruvate to form α-lactyl-ThDP.

Yeast requires Pdc1 activity for anaerobic glycolysis, but this activity would be wasteful during aerobic metabolism. As a result, Pdc1 has evolved exquisite regulation mechanisms such as the requirement for activation by pyruvate and the resulting positive cooperativity for this substrate. These traits can be abolished *in vitro* by including a nonmetabolizable substrate analogue, such as pyruvamide (*2*). By way of contrast, pyruvate decarboxylase from *Zymomonas mobilis* (PdcZm) does not require substrate activation and displays Michaelis–Menten kinetics. It is likely that the unregulated PdcZm is closely related to an ancestral form and did not evolve regulation mechanisms since *Z. mobilis* has simple metabolic requirements for rapid fermentation in the sugar-rich environments of plant saps.

Cooperativity and substrate activation in Pdc1 create a threshold in pyruvate concentration, below which the rate of pyruvate decarboxylation is very low. This threshold increases with increasing concentration of phosphate, a competitive inhibitor that also increases the cooperativity of Pdc1 for pyruvate (*3*). Pdc1 activity can therefore be rapidly attenuated *in vivo* when the rate of phosphate uptake through glycolysis is reduced. In addition, all versions of Pdc analyzed to date show optimal activity in slightly acidic environments. In solutions above neutral pH, Pdc1 oligomers dissociate along the monomer interface that forms the catalytic site (*4*). This could be viewed as a necessary regulation mechanism, since Pdc1 activity would increase the cytosolic pH if it was not balanced by glycolytic activity.

We present research that aimed to reduce Pdc1's exquisite regulation, while maintaining catalytic activity, in order to find new insights into the structural mechanisms involved. In practice, we developed a directed evolution method to search for Pdc1 mutations that would improve (1) the expression in *Escherichia coli* at 37 °C, (2) the activity in crude lysates at pH 7.5 with only 1 mM pyruvate, and (3) the activity in the presence of phosphate. Mutations were introduced by error-prone PCR (EPCR) at either high (10.8 substitutions per gene) or low (4.4 substitutions per gene) rates, and the influence of mutation rate was examined. We

* Corresponding author. Phone: +44 151 7053107. Fax: +44 151 7053369. E-mail: B.J.Stevenson@liverpool.ac.uk.

[1] Abbreviations: EPCR, error-prone PCR; *h*, Hill coefficient; k_{cat}, maximum catalytic turnover; K_{iS} substrate inhibition constant; K_{iP} phosphate inhibition constant; PIPES, piperazine-1,4-bis(2-ethanesulfonic acid); Pdc1, pyruvate decarboxylase 1 from *Saccharomyces cerevisiae*; $S_{0.5}$, substrate concentration at half of the maximum activity; StEP, staggered extension process; $T_{1/2}$, temperature at which half of the enzyme activity is lost in 5 min.

10.1021/bi701858u CCC: $40.75
Published on Web 01/31/2008

aimed to recombine mutations in every other generation and found that EPCR was capable of recombining mutations itself.

Five rounds of directed evolution were carried out to produce five generations of progeny (optimal variants). Each round consisted of four levels of screening to find about ten progeny from initial libraries with tens of thousands of variants. The primary screen was designed to rapidly identify candidates that could be more accurately screened in 96-well plate assays in secondary screens. This was followed by tertiary and quaternary screens to isolate and validate the optimal variants. The screening conditions were made more stringent over the course of the project such that each generation was suitably challenged. In addition, the primary screening method was improved from colony screening in generations 1 and 2 to culture screening in generations 3−5.

With the directed evolution of Pdc1, we have identified mutations for both altered regulation and increased stability of Pdc1 that will be suitable for future analysis. We also present the location of mutations within the crystal structure of Pdc1 (*5, 6*) and suggest connections between structure and function.

EXPERIMENTAL PROCEDURES

Materials. Oligonucleotides were synthesized by Geneworks (SA, Australia). The *E. coli* strain BL21(DE3)$^{RecA-}$ and the plasmid pETMCSIII (*7*) were kindly provided by Nicholas Dixon at the Research School of Chemistry (Australian National University). Reagents were purchased from Sigma (St. Louis, MO) unless otherwise specified. Water used in molecular biology, protein preparation, and analysis was purified using a MilliQ system (Millipore, Billercia, MA). Tryptone, yeast extract, and agar for media preparation were supplied by Difco (Becton, Dickinson & Co., MD) while amplicillin was from Astral Pharmaceuticals (Mumbai, India).

Creating a Strain for Pdc1 Expression. The first requirement for the directed evolution of Pdc1 was a system for consistent expression of the protein. This was achieved with a method similar to that presented by Candy et al. (*8*), using the T7 RNA polymerase expression system (*9*) in *E. coli*. The *pdc1* gene was amplified from *S. cerevisiae* genomic DNA using the forward primer, 5′ TCAAATCAATCCAT*ATG*TCTG 3′, and the reverse primer, 5′ CTTTAACTAATAATTCTAGA*TTA*AATC 3′. The *Nde*I restriction site in the forward primer and the *Xba*I site in the reverse primer are underlined whereas the start and stop codons for the *pdc1* gene are italicized. These primers were designed using the sequence of chromosome VII available online (www.ncbi.nlm.nih.gov). The *pdc1* gene was cloned into pETMCSIII via the *Nde*I and *Xba*I sites, and the construct was transformed into BL21(DE3)$^{RecA-}$ to produce strain RSC1426 (local nomenclature). All transformations were performed using a MicroPulser (Bio-Rad, Hercules, CA) electroporation device according to the manufacturer's instructions. This vector allows a recombinant protein to be expressed with an N-terminal His tag (MHHHHHH). BL21(DE3)$^{RecA-}$ was also transformed with empty pETMCSIII for the negative control culture RSC1425. These cultures were grown at 37 °C in Luria–Bertani media with 50 mg/L ampicillin (LBA) and stored at −70 °C with 16% (v/v) glycerol.

DNA Sequencing. The *pdc1* gene could not be reliably sequenced from the pETMCSIII construct. Instead, the construct region containing the gene was amplified using primer pET-3, 5′ CGACTCACTATAGGGAGACCACAAC 3′, and primer pET-4, 5′ CCTTTCGGGCTTTGTTAGCAG 3′. The PCR product was sequenced using BigDye terminator (Applied Biosystems) with primer pET-3 sequencing from the 5′ end and pET-4 from the 3′ end of the gene. Analysis was performed by the Biomolecular Resource Facility (John Curtin School of Medical Research, Australian National University). DNA sequences were studied and aligned using GeneJockey II (BioSoft, Cambridge, U.K.).

Alcohol Dehydrogenase for Pdc1 Assays: Expression and Purification. The *E. coli* alcohol dehydrogenase P (AdhP) (*10*) was selected for Pdc1 assays in which acetaldehyde production is coupled to NADH oxidation. An efficient source for AdhP was prepared by amplifying the *adhP* gene from the *E. coli* K12 genome by PCR with 5′ CGCGGTACCCAT*ATG*CAGAACATCATCCGAAAAGGA 3′ as the forward primer (*Nde*I site underlined and *adhP*'s start codon in italics) and 5′ CCCAGGCCTGAATTC*TTA*GTGACGGAAATCAATCACCAT 3′ as the reverse primer (*Eco*RI site underlined and *adhP*'s stop codon in italics). Primers were designed using the *E. coli* genome sequence available online (www.ncbi.nlm.nih.gov). The *adhP* gene was cloned into pETMCSIII via the restriction sites *Nde*I and *Eco*RI, and the construct was used to transform BL21(DE3)$^{RecA-}$, resulting in culture RSC1427 for expression of His-tagged AdhP. This culture was grown overnight at 37 °C in LBA before the cells were harvested by centrifugation. The lactose present in the tryptone used in LBA preparation was ideal for induction (*11*) and resulted in high expression. During subsequent purification steps the preparations were kept between 0 and 6 °C. The cell pellet was resuspended in 20 mL of lysis buffer [50 mM piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES) at pH 7.5 with 0.5 M NaCl] per gram of wet cell pellet. The cells were lysed using a French pressure cell press (SLM Instruments, Urbana, IL) with the pressure maintained at 12000 psi. Insoluble cell debris was removed by centrifugation at 30000*g* for 30 min. The supernatant was then passed through a 5 mL HiTrap chelating HP column (Amersham Biosciences, Buckinghamshire, U.K.) that had been charged with Ni^{2+} ions according to the manufacturer's instructions. The loaded column was washed with 50 mL of 0.1 M imidazole (pH 7.5) in lysis buffer before the His-tagged AdhP was eluted with 0.5 M imidazole (pH 7.5) in lysis buffer. The AdhP was then dialyzed and stored in 10 mM PIPES at pH 7.5 with 5 mM 2-mercaptoethanol with a final protein concentration of 5 mg/mL. Typically, 80 mg of AdhP/L of culture was recovered. This preparation retained activity over several weeks during storage at 4 °C and for at least a year when stored at −70 °C.

Preparation of Purified Pdc1. His-tagged native or variant Pdc1 was produced and purified as described for AdhP with the following alterations. First, the column wash was performed with 20 mL of lysis buffer followed by 15 mL of 0.15 M imidazole in lysis buffer. Second, after elution with 0.5 M imidazole in lysis buffer, the purified Pdc1 samples were dialyzed into 10 mM PIPES at pH 6.5 with 1 mM dithiothreitol, 1 mM $MgCl_2$, and 0.1 mM ThDP. Third, the

preparations were diluted to 1 μM concentration, with concentrations determined by measuring A_{280} (the extinction coefficient was estimated by the ProtParam tool, ExPASy Web site, http://au.expasy.org/tools/protparam.html), and greater than 95% purity was verified by SDS−PAGE analysis (*12*).

Introduction of Genetic Variation. In generations 1 and 3, random point mutations were introduced by Mn^{2+}-induced EPCR (*13*) at two rates. The "low" error rate used PCR with final concentrations of 0.1 mM $MnCl_2$ and 6.5 mM $MgCl_2$ for 4.4 substitutions per *pdc1* gene; the "high" error rate used 0.4 mM $MnCl_2$ and 6.2 mM $MgCl_2$ for 10.8 substitutions per *pdc1* gene. *Taq* DNA polymerase (Roche, Basel, Switzerland) was used with the primers pET-3 and pET-4 in EPCR with 30 cycles of 94 °C for 10 s, 50 °C for 10 s, and 2 min at 72 °C. The average error rates were determined by randomly selected *pdc1* variants from generation 1. Estimates were based on nine variants from the low error rate library and seven from the high error rate library. Recombination of point mutations was achieved by DNA shuffling (*14*) in generation 2 according to the method by Lorimer and Pastan (*15*). High fidelity recombination was trialed in generation 4 using the staggered extension process [StEP (*16*)] modified for use with *Pfu* DNA polymerase (Stratagene, La Jolla, CA). In this case, thermocycling involved 100 cycles of 94 °C for 10 s, 50 °C for 10 s, and 72 °C for 10 s. The inclusion of 10 s for DNA elongation was to account for the slower extension for *Pfu* in comparison with *Taq* (based on the manufacturer's information). After the observation that EPCR resulted in recombination of existing mutations, generation 5 used a modified form of EPCR to provide both low error rate and StEP. In this protocol, EPCR was performed with the 72 °C incubation time for DNA elongation reduced to 1 min.

Preparing pdc1 Variant Libraries for Expression. Pools of *pdc1* gene variants were digested with *Nde*I, *Xba*I, and *Dpn*I (New England Biolabs, Ipswich, MA) and isolated by gel extraction. *Dpn*I digestion ensured that the parental genes, synthesized in pDNA *in vivo*, were removed from the library (ref *17*, p 13.19). Ligation was performed overnight at 14 °C using T4 DNA ligase (Roche, Basel, Switzerland) according to the manufacturer's directions. The ligation products were purified (PCR purification kit, Qiagen) and eluted in 30 μL of water, 2 μL of which was used in the electrotransformation of 50 μL of *E. coli* DH5α. This strain allowed high transformation efficiency and could be used to amplify the pDNA without bias since it was unable to transcribe the gene variants. Outgrowth was performed in 1 mL of autoclaved YENB media [7.5 g/L Bacto yeast extract (Difco) and 8 g/L Bacto nutrient broth (Difco)] at 37 °C for 1 h. This culture was diluted 10-fold by addition of 9 mL of LBA media, and 50 μL was plated out on LBA agar to determine the transformant density and variant library size. The 10 mL culture was grown overnight at 37 °C, and 2 mL was used in plasmid isolation (Miniprep, Qiagen). Subsequently, 1 μL of Miniprep was used for electrotransformation of BL21(DE3)$^{RecA-}$. After 1 h recovery in 1 mL of YENB media at 37 °C, 250 μL of 80% glycerol was added and mixed. The transformants were divided into aliquots of 50 μL and stored at −70 °C. The transformant density was estimated by thawing one aliquot and plating out on LBA agar.

Colony Screening. To screen for colonies of BL21-(DE3)$^{RecA-}$ transformants with Pdc1 activity, an aliquot of frozen transformants was thawed and grown overnight on LBA agar at 37 °C with a density less than 6 colonies per 1 cm^2. As with the liquid cultures of BL21(DE3)$^{RecA-}$, the lactose present in the media provided consistent and high expression of Pdc1. Indicator agar (adapted from ref *18*) was prepared immediately before use with 0.2 M sodium citrate at pH 4.5, 6 mM sodium pyruvate, 0.33 g/L sodium bisulfate, and 2% (v/v) ethanol saturated with *p*-rosaniline. These conditions were optimized to differentiate between colonies of RSC1425 (negative control) and RSC1426 (native Pdc1 expression). The agar plates were cooled before a volume of molten indicator agar (at 50 °C), equal to that volume of growth media, was carefully poured on top to cover all colonies evenly. Colonies with Pdc1 activity would develop a red halo and were screened 10 min after exposure to the indicator agar.

Culture Screening. Cultures of BL21(DE3)$^{RecA-}$ transformants were grown in LBA media at 37 °C overnight before being cooled to room temperature (room temperature, 22−25 °C). Cultures were grown either in stationary 96-well plates (rounded wells; Sarstedt, Nümbrecht, Germany) with petroleum jelly to seal the lids or as 1 mL cultures in closed 1.5 mL tubes (Sarstedt) with 200 rpm shaking. Cultures in 96-well plates were resuspended prior to assaying using either gentle vortexing, for 100 μL cultures, or 960 rpm translational shaking for 50 μL cultures. Three assays were used throughout the directed evolution of Pdc1, all of which began by mixing 5 μL of an overnight culture with 5 μL of bacterial protein extraction reagent (B-PER; Pierce, Rockford, IL) in the well of a 96-well plate (flat-based wells; Sarstedt).

In assay A, crude lysate was incubated at room temperature for 15 min before 290 μL of assay mix was added to give 50 mM HEPES·Na^+ at pH 7.5, 5 mM $MgCl_2$, 0.1 mM ThDP, 1 mM sodium pyruvate, 0.15 mM NADH, and 2 μg/mL AdhP. The maximum rate of A_{340} decline over 5 min during the next 30 min was measured with a 96-well plate reader (Multiskan Ascent; Labsystems, Helsinki, Finland) and used as an indicator of recoverable activity at room temperature.

Assay B was developed to increase the throughput of culture analysis. After mixing a culture with B-PER, as in assay A, the crude lysate was left for 2 min at room temperature before addition of 290 μL of assay mix (staggered to maintain constant incubation time between samples). The final assay conditions were 100 mM HEPES·Na^+ at pH 7.5, 10 mM phosphate, 1 mM sodium pyruvate, 0.15 mM NADH, and 2 μg/mL AdhP. The rate of A_{340} decline over 5 min at room temperature, 1 min after addition of assay mix, was used to indicate the Pdc1 activity.

Assay C was identical to assay B, except that the reaction mix provided 50 mM sodium phosphate at pH 7.5 instead of HEPES and phosphate.

Screening Strategies. There were two strategies used to screen for desirable Pdc1 variants. Both strategies used four levels of screening to isolate and validate optimal variants. Strategy 1 used colony screening to identify the two colonies with the most vivid red halos from each culture plate (ca. 300 colonies). Each colony from this primary screen was used to inoculate 100 μL of LBA in 96-well plate format and grown overnight at 37 °C. The cultures were then used

in a secondary screen with assay A. Cultures with greater activity than controls expressing the best parent Pdc1 were streaked out on LBA agar for overnight growth at 37 °C to obtain isolated colonies. These were then used to inoculate individual 1 mL LBA cultures, incubated overnight at 37 °C with shaking, and used in a tertiary screen with assay A. Cultures that had significantly better activity than the optimal parent were mixed with glycerol (final concentration of 16%, v/v) and stored at −70 °C. After several batches of screens had resulted in a collection of frozen cultures, each culture was streaked out on LBA agar, and single colonies were used in quaternary screening (identical to tertiary screening). The best variants from this final screen became the progeny.

Strategy 2 used culture screening throughout, with either assay B or assay C. To manually prepare many cultures with different Pdc1 variants for analysis, a diluted suspension of transformed cells in ice-cold LBA was made with approximately two viable cells per 50 μL. This suspension was distributed as 50 μL aliquots into 96-well plates and incubated as described for culture screening. The culture with the highest activity from each plate was then streaked out on LBA agar and grown overnight at 37 °C. Four isolated colonies from each primary culture were used to inoculate 50 μL of LBA in 96-well plate format alongside control cultures inoculated with a colony of the best parent. The secondary screen was performed as for the primary screen, and cultures with greater activity than the best parent were streaked out on LBA and grown overnight at 37 °C. Tertiary and quaternary screens were performed as in strategy 1, except using either assay B or assay C. The progeny from each generation were stored in 16% glycerol at −70 °C, and pDNA was isolated (Qiagen) for DNA sequencing.

Kinetic Analysis of Pdc1. Assays for Pdc1 activity were performed in a triple buffer consisting of 100 mM Tris, 50 mM MES, and 50 mM acetic acid (*19*). This allowed the pH to be varied by addition of NaOH or HCl while keeping the ionic strength constant. Reactions were performed at 25 °C with 0.02 μM Pdc1, 1 mM $MgCl_2$, 0.1 mM ThDP, 0.1 mM NADH, and 10 μg/mL AdhP. The integrity of NADH and activity of AdhP were not rate-limiting within the pH range used (4.9−7.5), as determined by addition of acetaldehyde in place of the Pdc1 sample. Experiments included sodium pyruvate and sodium phosphate in a range of concentrations as required. Absorption was monitored with a Cary 1E spectrophotometer (Varian, Palo Alto, CA) with a temperature-controlled cuvette mount. Each reaction was started by injecting and mixing 50 μL of the Pdc1 sample into 450 μL of reaction solution in the cuvette. A_{340} was monitored for 20 s, 1 min after the injection of enzyme, to ensure that the Pdc1 was assayed when activated and before the NADH supply was exhausted. The change in A_{340} was assumed to be entirely due to the oxidation of NADH (extinction coefficient of 6220 M^{-1} cm^{-1}) for determining the acetaldehyde production rate. The catalytic turnover was related to the substrate concentration using the Hill equation with substrate inhibition (*19*) and nonlinear regression (Kaleidagraph, version 3.6.2; Synergy Software):

$$\text{turnover} = k_{cat}[S]^h/(S_{0.5}{}^h + [S]^h(1 + [S]/K_{iS}))$$

Analysis of Pdc1 Activity with Respect to Temperature. Reactions were performed as described for Kinetic Analysis of Pdc1 with the following differences. The reaction buffer (containing PIPES·Na^+, pyruvate, $MgCl_2$, and ThDP) was equilibrated at a defined temperature for 20 min. PIPES was used since its pK_a has the lowest temperature sensitivity ($\Delta pK_a/°C = -0.0085$) of the buffers developed by Good et al. (*20*). The reactions were started by mixing in 12 μL of purified Pdc1 sample with AdhP and NADH to produce a 0.5 mL reaction with 0.1 M PIPES (pH 6.2 at 30 °C), 1 or 25 mM pyruvate, 1 mM $MgCl_2$, 0.1 mM ThDP, 0.1 mM NADH, 0.53 μM AdhP, and 0.01 μM Pdc1. The reaction rates were determined by measuring the slope from 20 to 60 s after the start. The activity of AdhP, or integrity of NADH, was not rate-limiting at any of the temperatures used.

Analysis of Thermal Stability. Samples of 1 μM His-tagged Pdc1 in 10 mM PIPES at pH 6.5 with 1 mM dithiothreitol, 1 mM $MgCl_2$, and 0.1 mM ThDP were incubated at defined temperatures (T) for 5 min. All incubated samples were stored on ice for at least 2 h prior to assaying with conditions described for the kinetic analysis of Pdc1. Assays for Pdc1 activity were performed at 25 °C as described for Assays of Pdc1 Activity with Respect to Temperature. The activity (a) of heat-treated samples was compared to the activity (a_0) of unheated control samples that were prepared identically. The temperature at which half of the Pdc1 was irreversibly inactivated over 5 min ($T_{1/2}$) was calculated by fitting a modified form of the Hill equation by nonlinear regression (Kaleidagraph):

$$a = a_0 - (a_0T^h)/(T_{1/2}{}^h + T^h)$$

Mutation Location Analysis. Structures of Pdc1 (*5, 6*) were analyzed using the Swiss PDB Viewer (Macintosh version 3.7). The images were rendered using POV-Ray (Macintosh version 3.6).

Table 1: Methods Used in Directed Evolution of Pdc1 for Each Generation

generation	diversity method	screening strategy	primary screening		secondary screening	
			method	no.[a]	method	no.[a]
1	EPCR	1	colonies	30000	assay A	500
2	shuffling	1	colonies	16000	assay A	500
3	EPCR	2	assay B	21000	assay B	100
4	StEP	2	assay B	19000	assay B	100
5	EPCR	2	assay C	55000	assay C	100

[a] An estimate of the number of genotypes examined during screening.

RESULTS

Five generations of mutagenesis and screening were performed, as outlined in Table 1. This table summarizes how the screening process developed during the course of the directed evolution. The progeny (10 and 14 isolated clones) from each generation were sequenced, and then the unique variant genes became the parents of the next generation.

Mutagenesis. The intent of these experiments was to alternate between nucleotide point mutagenesis in one generation and recombination in the next. However, all methods for introducing genetic mutations provided both point mutations and recombinations. The progeny derived from DNA shuffling in generation 2 had 1.2 point mutations per gene whereas the StEP with *Pfu* DNA polymerase in generation 4 had only 0.5 on average. Generations 1, 3, and 5 used EPCR and resulted in a mean of 2.4, 2.3, and 0.8

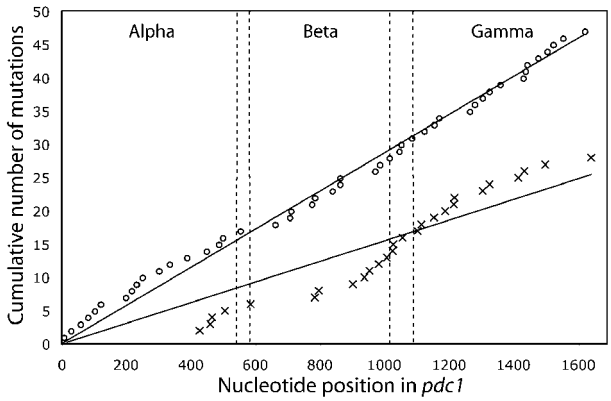

FIGURE 1: The positions of silent (○) or expressed (×) mutations are plotted with respect to the cumulative number of mutations from the start codon. The least-squares regression was calculated with a fixed intercept at the origin and plotted for both sets of data (the r^2 value is 0.991 for silent mutations and 0.848 for expressed mutations). The horizontal axis is divided into regions (dashed lines) to indicate each of the three domains of Pdc1 (alpha, beta, and gamma).

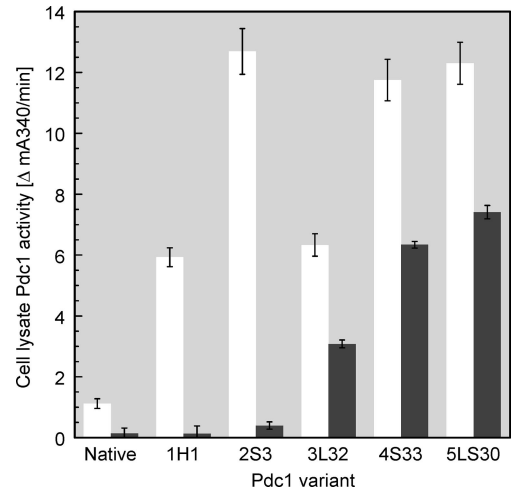

FIGURE 2: The Pdc1 activity in whole cell lysate was determined for the best variant from each generation. Activity was assayed as described for the final screening conditions used in generation 1 (assay A, white bars) or generation 5 (assay C, dark bars). Measurements were corrected for the background activity from BL21(DE3) with pETMCSIII but without a *pdc1* gene inserted. The error bars represent the standard deviation ($n = 3$) for each culture.

substitutions per progeny gene, respectively. Generation 5 stood out since its progeny had fewer substitutions and no expressed mutations. The progeny had fewer mutations than expected by the EPCR mutation rates, and there was no significant difference between progeny from the high or low error rate libraries.

Recombination was estimated by analyzing the DNA sequences of the progeny from one generation and comparing it with their parents' sequences. Mutations present in both parent and progeny were assumed to be inherited rather than generated by another PCR error. New combinations could therefore arise by a minimum number of recombination events and could be observed in all generations except the first, since the parent gene had no mutations to observe. On average, there were two of these events per progeny gene in each generation, showing that DNA shuffling, StEP, and EPCR were all capable of recombining parental genes. StEP produced the greatest recombination observed in an individual, with a minimum of five recombination events in one of the generation 4 progeny (data not presented).

The distribution of silent mutations in the progeny from all generations did not appear to be biased (Figure 1), and it is assumed that the mutagenesis methods were capable of influencing any region within *pdc1*. In contrast, the progeny had an uneven distribution of expressed mutations. None occurred in the first 75% of the alpha domain but were more frequent at the end of the beta domain and the start of the gamma domain.

The type of point mutations observed were biased as expected for Mn^{2+}-influenced *Taq* DNA polymerase (*21*) with transitions accounting for 78% of the point mutations and 74% of these substituting AT with GC. The frequency of mutations in the progeny of generations 1 and 3 was much lower than what was engineered for the library. Silent mutations were more common in the progeny, accounting for 50% of substitutions. Given the codon use in *pdc1* and mutation bias of *Taq* (*21*), only 32% of point mutations would be silent if mutations occurred, and were selected, randomly.

Directed Evolution. The progeny from each generation were named according to their generation (1−5), mutagenesis method (L, low error rate; H, high error rate; S, DNA shuffled or StEP; or LS, low error rate StEP), and an identification number. These Pdc1 variants were isolated, first, for stable expression in *E. coli* grown at 37 °C and, second, for activity in unfavorable conditions. The screening conditions progress during the directed evolution (Table 1) leading up to assays of crude cell lysate in 1 mM pyruvate at pH 7.5 with 50 mM phosphate (assay C). However, native Pdc1 has insignificant activity with assay C due to inhibition by phosphate, high pH, and low substrate concentration. This is confounded by the background activity from the *E. coli* lactate dehydrogenase (*8*), an enzyme that oxidizes NADH while reducing pyrutate in the assay conditions. Therefore, to begin directed evolution, the secondary screening for activity in *E. coli* crude cell lysates was performed with phosphate absent and cofactors supplied (assay A).

Performance in assay A increased rapidly in the best variant from the first two generations: 1H1 and 2S3 (Figure 2). These generations both used assay A in the final screen. More stringent screening was applied in generations 3 and 4 with assay B. This was similar to assay A but with 10 mM phosphate and no added ThDP or Mg^{2+}. Generation 3 progeny had reduced performance in assay A, with its best variant (3L32) reverting to the activity of its grandparent, 1H1. However, the optimal variants from the following generations, 4S33 and 5LS30, regained the high assay A activity seen with 2S3.

Pdc1 activity only became significantly higher than the background activity of crude *E. coli* lysate in assay C after the isolation of variant 2S3. In contrast to assay A performance from generation 2 to 3, assay C performance increased by approximately 6-fold. Generation 4's 4S33 improved by another 2-fold. Assay C was finally introduced for screening in generation 5; however, the best variant, 5LS30, had only a small increase of 10% over 4S33.

Optimal Mutations for Pdc1. The best variant for activity in assay C, 5LS30, incorporated five expressed mutations.

Table 2: Expressed Mutations in Unique Pdc1 Variants Isolated during Directed Evolution[a]

	Pdc1 residue no.																											reps	
variant	143	154	156	169	196	262	266	301	313	318	328	336	342	343	352	367	372	385	396	405	406	435	442	472	478	500	547	G	P
native	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	G	K	I	D	K	P	Q		
1H1	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	**V**	N	Y	**C**	**R**	I	D	K	P	Q		
1H2	A	R	T	A	S	V	T	Y	H	N	F	**N**	K	G	R	Q	N	I	N	Y	G	K	I	D	K	P	Q	2	
1L1	A	R	T	**T**	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	G	K	I	D	K	P	Q		
1L2	**T**	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	G	K	I	D	K	P	Q		
1L3	A	R	T	A	**P**	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	G	K	I	**N**	K	P	Q	2	
1L4	A	R	T	A	S	V	T	Y	H	N	F	T	**M**	G	R	Q	N	I	N	Y	G	K	I	D	K	P	**R**	2	
1L6	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	**H**	I	N	Y	G	K	I	D	K	P	Q		
1L8	A	R	T	A	S	V	T	Y	H	N	**L**	T	K	**C**	R	Q	N	I	N	Y	G	K	I	D	K	P	Q		
2S1	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	**V**	N	Y	**C**	**R**	I	D	K	P	**R**		
2S2	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	**V**	N	Y	**C**	**R**	I	D	K	P	Q		
2S3	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	**C**	**R**	I	**N**	K	P	**R**		
2S4	A	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	N	Y	G	K	I	D	K	P	**R**		
2S5	A	R	T	A	**P**	V	T	Y	**Y**	N	F	T	K	G	R	Q	N	I	N	Y	**C**	K	I	D	K	P	Q		
2S6	**T**	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	**V**	N	Y	**C**	**R**	I	D	K	P	Q		2
2S8	**T**	R	T	A	S	V	T	Y	H	N	F	**N**	K	G	R	Q	N	I	N	Y	G	K	I	D	K	P	Q		2
2S10	**T**	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	G	K	**T**	D	K	P	**R**		
3H13	**T**	R	T	A	S	V	**A**	Y	H	N	F	**N**	K	G	R	Q	N	I	N	Y	G	K	I	D	K	**S**	Q		
3L26	**T**	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	N	Y	**C**	**R**	I	D	**R**	P	**R**		
3L32	A	R	T	A	S	V	?	**C**	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**		
3L34	**T**	R	T	A	S	V	?	Y	H	N	F	T	K	G	R	**H**	N	I	N	Y	**C**	K	I	D	K	P	Q		
3L41	A	**K**	T	A	S	**I**	T	Y	H	N	F	T	K	G	R	Q	N	I	N	**F**	**C**	**R**	I	**N**	K	P	Q		
3L46	**T**	R	T	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	Y	**C**	**R**	I	**N**	K	P	**R**		
3L51	A	R	T	A	**P**	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	N	Y	**C**	**R**	I	D	K	P	**R**		
3L53	**T**	R	T	A	?	V	T	Y	H	**S**	F	T	K	G	R	Q	N	I	N	Y	**C**	K	I	D	K	P	Q		
3L56	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	Q	N	I	N	Y	**C**	**R**	I	**N**	K	P	Q		
4S17	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	Q		
4S110	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	Q		
4S113	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	Q	N	I	N	**F**	**C**	**R**	I	D	**R**	P	Q		
4S25	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	N	Y	**C**	**R**	I	D	**R**	P	Q		
4S29	**T**	R	**A**	A	S	V	T	**C**	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	Q		
4S33	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**		3
4S37	**T**	R	**A**	A	S	V	T	**C**	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**	2	
4S39	**T**	R	**A**	A	S	V	T	**C**	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	**R**	I	D	K	P	**R**		
4S42	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**		
4S48	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	**K**	Q	N	I	N	Y	**C**	**R**	I	D	**R**	P	**R**		
4S411	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	Q	N	I	N	Y	**C**	**R**	I	D	**R**	P	**R**		
5L39	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**		
5L48	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	Q		2
5LS17	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	**R**	P	**R**		2
5LS26	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	K	P	**R**		
5LS30	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	**R**	P	Q		4
5LS40	**T**	R	**A**	A	S	V	T	Y	H	N	F	T	K	G	R	**H**	N	I	**I**	Y	G	**R**	I	D	**R**	P	**R**		3
5LS50	**T**	R	**A**	A	S	V	T	Y	H	N	F	**N**	K	G	R	**H**	N	I	**I**	Y	G	K	I	D	**R**	P	**R**		

[a] Mutated residues are in bold type, and an underline on a mutation indicates its first occurrence. Only the unique protein sequences discovered in each generation are presented. Some Pdc1 variants were isolated repeatedly during the final screen in each generation. There were several examples of parental duplication: five native *pdc1* genes (two of which had silent mutations) were isolated in generation 1, and one generation 2 gene (2S3) was isolated in generation 3. The number of variant replicates (reps) seen within a generation is presented for identical genes (G) and unique genes, but identical proteins due to silent mutations (P). Amino acid residues are represented by single letter codes, and a question mark indicates that the sequencing data was of insufficient quality in a particular region.

Of the 14 progeny from generation 5, four shared the same collection of expressed mutations, differing only in silent mutations. These five mutations originated throughout the first four generations and were optimally recombined in generation 5 (Table 2). A143T first occurred in generation 1 and was present in all of the progeny of generation 4; at this point it was fixed in the population of variants. Q367H originated in generation 2 and became fixed in generation 5. In contrast, T156A was highly favorable after originating in generation 3, becoming fixed in generation 4. N396I was also favored after originating in generation 3; however, during generation 4 it appeared to compete with a mutation from generation 1, G406C. By generation 5 all progeny included N396I but not G406C. Generation 3 was the most successful period for discovering beneficial point mutations, being the source for K478R as well. This fifth and final expressed point mutation for 5LS30 was favored but was only present in 71% of generation 5 progeny.

Contrasting Native and 5LS30 Pdc1 Kinetics. The reaction progress curves show a striking difference between native and 5LS30 Pdc1: the variant reaches a stable catalytic rate more quickly (Figure 3A). With about 1 mM pyruvate, native Pdc1 only reaches a stable reaction rate after exposure to pyruvate for 1 min, whereas 5LS30 requires only 10 s. The difference was less pronounced with 10 mM pyruvate (Figure 3B). Furthermore, unlike the native enzyme, the activation of 5LS30 did not decrease with increasing phosphate concentration (Figure 3B).

As expected from the screening conditions used, 5LS30 shows the greatest improvement in activity over native Pdc1 at low pyruvate concentration, in the presence of phosphate and at higher pH (Figure 4). The data presented in Figure 4

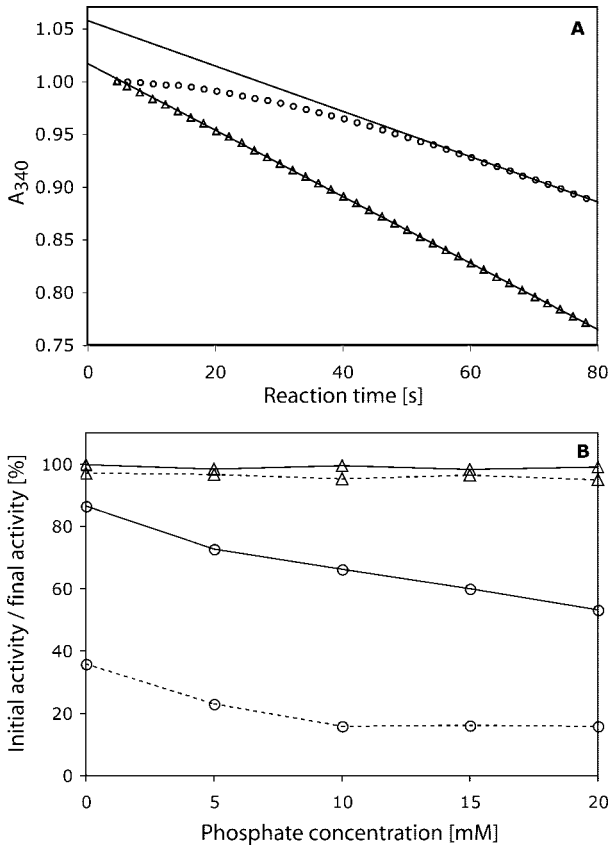

FIGURE 3: The reaction progress for native (○) and 5LS30 (△) Pdc1 was analyzed at pH 6 and 30 °C. Part A has the progress curves for reactions with 1.25 mM pyruvate. Regression lines were calculated using the points from 60 to 80 s after the reaction began. The earliest observation, set to 1 unit, was made 5 s after the reactions began. Part B shows the initial activity (6−26 s after addition of enzyme) as a percentage of the final activity (70−80 s after addition of enzyme). Observations are presented for 1.25 mM (- - -) and 10 mM pyruvate (—) with respect to phosphate concentration.

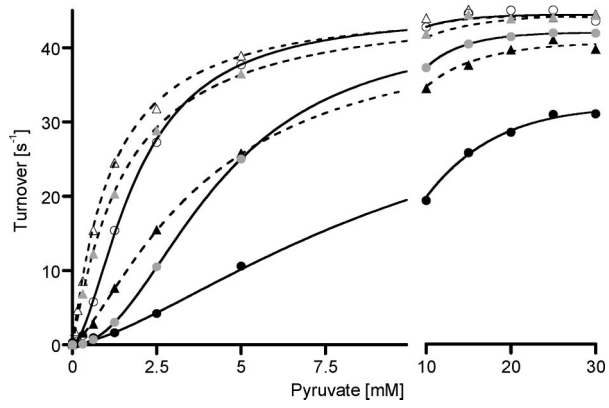

FIGURE 4: The initial rate of native or 5LS30 Pdc1 at 25 °C as a function of pyruvate concentration. Reactions were performed as described in Kinetic Analysis of Pdc1 at pH 6.0 without phosphate (native, ○; 5LS30, △), pH 6.0 with 20 mM phosphate (native, shaded circle; 5LS30, shaded triangle), and pH 7.5 in the absence of phosphate (native, ●; 5LS30, ▲). The data points represent single measurements, and the results from nonlinear regression (Hill equation with substrate inhibition) are shown for results with native (—) or 5LS30 Pdc1 (- - -).

are examples of the observations used to estimate the values presented in Figure 5. However, the observed turnover never reaches the calculated k_{cat} value on account of the substrate

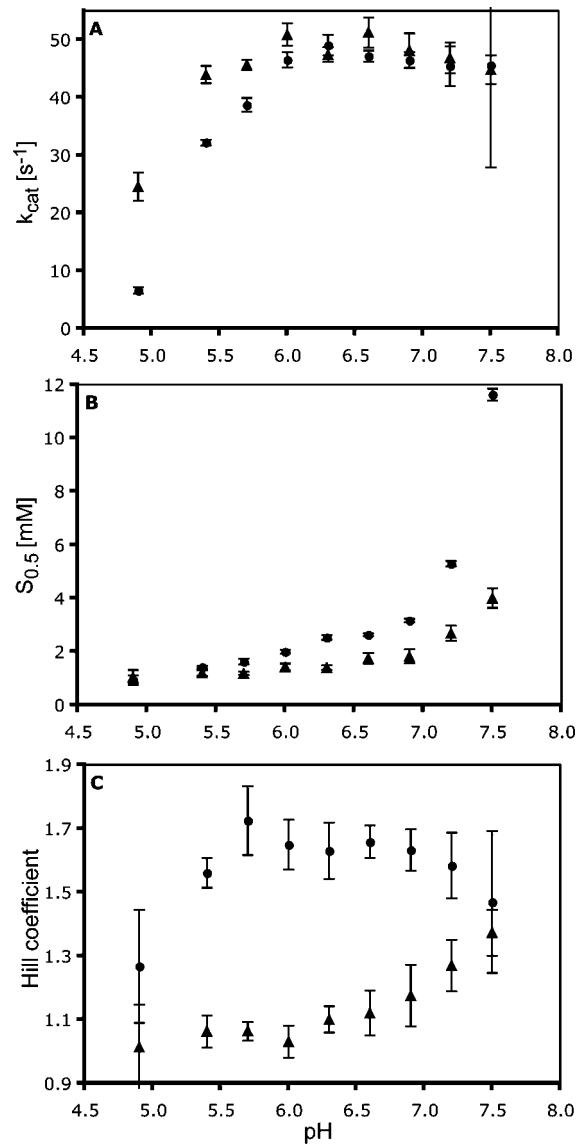

FIGURE 5: Pdc1 kinetic parameters for pyruvate decarboxylation with respect to pH at 25 °C. The parameters for native Pdc1 (●) and 5LS30 Pdc1 (▲) are presented for the maximal activity (A), the pyruvate concentration required for half of the maximal activity (B), and cooperativity (C). These data and their standard error (error bars) were calculated by nonlinear regression, as illustrated in Figure 4, from a single experiment for each Pdc1 version and pH as described in Kinetic Analysis of Pdc1. The substrate inhibition constant (K_{iS}) could not be accurately estimated, typically between 100 and 1000 mM, and did not show any significant differences (data not presented).

inhibition. In some cases, such as at pH 7.5, the observations could be easily explained by a standard Hill equation. The inclusion of the substrate inhibition term in nonlinear regression of such observations resulted in greater uncertainty in k_{cat}. Despite this, the Hill equation with substrate inhibition was used because all observations needed to be analyzed with the same equation, and this equation explained the majority of substrate-saturation observations.

The k_{cat} for native and 5LS30 Pdc1 was maximal between pH 6 and pH 6.6, dropping gradually as pH increased to 7.5 and falling rapidly as pH decreased to 4.9. 5LS30 shows the most significant improvement in k_{cat} at low pH rather than high pH, where screening occurred (Figure 5A). The screens were designed to find improved activity at low

Table 3: Kinetic Constants (±Standard Error) of Native and 5LS30 Pdc1 in the Presence of Phosphate at 25 °C and pH 6.0, Determined as Described in Kinetic Analysis of Pdc1[a]

[phosphate], mM	$S_{0.5}$		h	
	native	5LS30	native	5LS30
0	1.89 ± 0.05	1.39 ± 0.07	1.76 ± 0.05	1.03 ± 0.03
5	2.39 ± 0.06	1.37 ± 0.02	1.96 ± 0.06	1.06 ± 0.01
10	2.99 ± 0.05	1.45 ± 0.05	1.93 ± 0.04	1.03 ± 0.02
15	3.54 ± 0.04	1.65 ± 0.04	2.04 ± 0.03	1.01 ± 0.02
20	4.39 ± 0.05	1.89 ± 0.09	2.08 ± 0.03	1.02 ± 0.03
slope	0.123	0.026	0.015	− 0.001
r^2	0.990	0.845	0.852	0.327

[a] The k_{cat} showed no significant trend with an average (±standard deviation) of 52 ± 1 s^{-1} for native and 54 ± 1 s^{-1} for 5LS30 Pdc1. These values are 10% higher than the k_{cat} values determined in the analysis of pH (Figure 5). The higher k_{cat} is probably due to the Pdc1 preparations having a longer storage period in the presence of cofactors after purification.

pyruvate concentrations at high pH, and this is reflected in the comparison of $S_{0.5}$. Both native and 5LS30 Pdc1 have $S_{0.5}$ increase with increasing pH, most rapidly above pH 7, but the rate of this increase is reduced for 5LS30 (Figure 5B). At pH 7.5, 5LS30 has almost one-third of the native $S_{0.5}$.

The apparent cooperativity of 5LS30 Pdc1 is significantly lower than native Pdc1 from pH 5.4 to pH 7.2. At pH 6.0 the Hill coefficient of 5LS30 Pdc1 was not significantly different from unity (Figure 5C). The Hill coefficient of native Pdc1 increased in the presence of phosphate for native Pdc1, as observed previously (*3*), but phosphate had no such influence on 5LS30 Pdc1 at pH 6.0 (Table 3). In addition, the competitive inhibition by phosphate was reduced in 5LS30 (Table 3). The phosphate inhibition constant (K_{iP}) for native and 5LS30 Pdc1 at pH 6.0 and 30 °C was 14.7 ± 1.6 mM and 51 ± 15 mM, respectively. The substrate inhibition (K_{iS}) constant for native and 5LS30 Pdc1 was the same within experimental error in each of the conditions tested.

Native Pdc1 Activity Was Not Altered by the His Tag. The similarity between the kinetic constants determined for native Pdc1 and those reported by Liu et al. (*19*) suggests that the N-terminal His tag had no influence on activity. The only difference occurred in estimates of k_{cat}. This value varied between Pdc1 preparations and appeared to be inversely proportional to the preparation time (data not presented). Typically, the k_{cat} of native Pdc1 was 20% higher than that reported by Liu et al. but showed the same response to pH.

5LS30 Had Improved Thermal Stability. As well as screening for increased activity in 1 mM pyruvate at a higher than optimal pH, there was also screening for improved stability. This was because variants were expressed at 37 °C, 7 °C higher than the optimal temperature for *S. cerevisiae*. Stability was quantified by determining the residual activity after high-temperature treatment. The temperature at which half of native Pdc1's activity was irreversibly lost during a 5 min incubation ($T_{1/2}$) was 52.6 ± 0.4 °C, whereas for 5LS30 it was 61.8 ± 0.1 °C (Figure 6). It was hypothesized that these variants could also have a greater optimal temperature for catalysis. Surprisingly, the optimal temperature was dependent on the pyruvate concentration (Figure 7). Native Pdc1 could decarboxylate 1 mM pyruvate optimally at 30 °C and 25 mM pyruvate at 46 °C. Under the same conditions 5LS30 was optimal at 38 °C with 1 mM and 55 °C with 25 mM pyruvate.

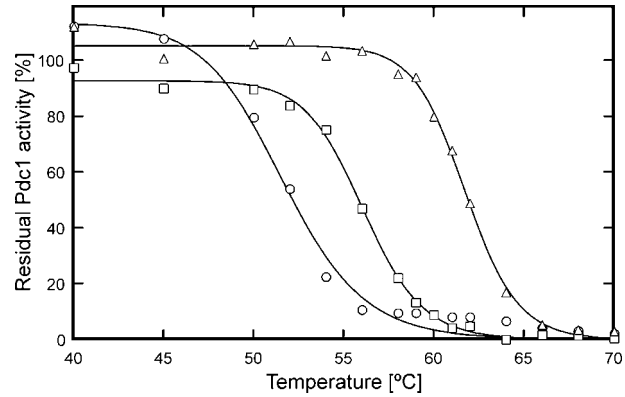

FIGURE 6: The irreversible thermal inactivation of purified native (○), 4S17 (□), and 5LS30 (△) Pdc1 was investigated by incubating samples at specific temperatures for 5 min. The residual activities were determined and analyzed as described in Analysis of Thermal Stability. The $T_{1/2}$ (±standard error) for native, 4S17, and 5LS30 Pdc1 was 52.6 ± 0.4, 56.1 ± 0.1, and 61.8 ± 0.1 °C, respectively, and their Hill coefficient for denaturation, h (±standard error), was 22 ± 3, 34 ± 2, and 43 ± 3, respectively. The turnover observed at 100% was 12 s^{-1} for native, 7 s^{-1} for 4S17, and 13 s^{-1} for 5LS30 Pdc1.

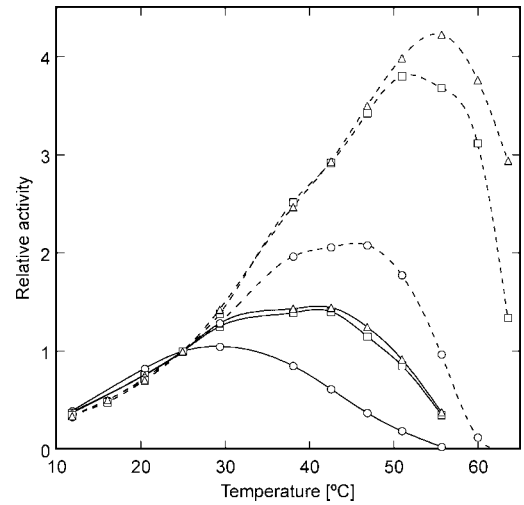

FIGURE 7: The catalytic activity of native (○), 4S17 (□), and 5LS30 (△) Pdc1 is presented with respect to temperature. The activity was determined as described in Kinetic Analysis of Pdc1 with Respect to Temperature with 100 mM PIPES to maintain the pH within an optimal range for Pdc1, ranging from pH 6.35 at 12 °C to pH 5.90 at 64 °C. Single measurements with 1 mM (solid line) or 25 mM (dashed line) pyruvate are plotted relative to their sample's activity at 25 °C. The observed turnover at 25 °C with 1 mM pyruvate was 8 s^{-1} for native, 7 s^{-1} for 4S17, and 12 s^{-1} for 5LS30; at 25 mM pyruvate the turnover was 15 s^{-1} for native, 8 s^{-1} for 4S17, and 15 s^{-1} for 5LS30 Pdc1.

The generation 4 variant, 4S17, differed from 5LS30 only in that it lacked the mutation K478R. It showed similar kinetic traits but was less stable than 5LS30 with a $T_{1/2}$ of 56.1 ± 0.1 °C (Figure 6) and had poorer catalytic performance with 25 mM pyruvate at temperatures above 50 °C (Figure 7), although, in 1 mM pyruvate, the catalytic response to temperature was not significantly different to 5LS30 (Figure 7).

The Pdc1 samples used in these thermal stability experiments (Figures 6 and 7) were prepared with a delay in the purification: all samples were stored for 2 weeks before cofactors were included. By comparison, the samples used to determine kinetic constants (presented in Figure 5 and

Table 3) were dialyzed into the storage buffer with cofactors immediately after purification. The specific activity of both native and 5LS30 Pdc1 was reduced to 25% due to a period without cofactors, while the 4S17 Pdc1 sample was reduced to 15%. Further investigation into why 4S17 had lower activity after storage without cofactors has not yet been conducted.

DISCUSSION

Directed Evolution Techniques. The colony screening in strategy 1 was set up to screen many thousands of Pdc1 variants for greater activity in *E. coli* before screening for improved activity in defined conditions. However, by the end of generation 2 it was discovered that this method could only distinguish between presence and absence of Pdc1 activity, not the amount of activity. The progeny were therefore screened from a pool of about 500 active Pdc1 variants in each of the first two generations (Table 1). Remarkably, the secondary screen succeeded in finding variants with greater activity in assay A. These variants provided a foothold for evolving activity in low pyruvate with higher pH and the presence of phosphate. Improvements in Pdc1's performance might have been possible from such a small sample since many different mutations could potentially alter the regulation of the enzyme.

Primary screening with cultures in strategy 2 was made practicable by circumventing the labor-intensive tasks of colony picking and culture inoculation by using dilute culture arrays. The library of *E. coli* transformants was diluted to what seemed a reasonable genotype density of two transformants per culture aliquot. The genotype density could be accurately estimated by assuming that the number of initial cells per culture would follow a Poisson distribution and determining the proportion of sterile cultures: a genotype density of 2 gave 13 sterile wells per 96-well plate. Generations 3−5 demonstrated that variants with significant improvements could be found while in a mixture with other variants.

Genotype mixtures in the primary screen made it more difficult to find mutations that offered minor improvements toward activity. This could be one reason for the improvements tapering off in generation 5 with no new expressed mutations discovered. Another possibility is that the deregulation of Pdc1 activity had a negative influence upon the *E. coli* host. The evolution process was producing an enzyme that would irreversibly decarboxylate pyruvate to a very low concentration *in vivo*. Perhaps transformants expressing exceptional versions of Pdc1 would be outcompeted by other transformants in the genotype mixture. Alternatively, the activity reached by variant 5LS30 could represent the limit for performance in assay C obtainable with point mutations and medium-throughput analysis.

Mutagenesis. Two mutation rates were used to produce libraries of *pdc1* variants in generations 1 and 3. The low error rate was intended to be at the recommended level of between 1 and 3 nucleotide substitutions per gene on average (*22*) but proved to be 4.4 per *pdc1* gene. The high error rate, at 10.8 substitutions per *pdc1* gene, was used to test whether the Pdc1 protein could withstand a higher mutation rate. The benefit of a higher mutation rate for directed evolution would be the ability to search through more mutations and to increase the chance of finding codependent mutations.

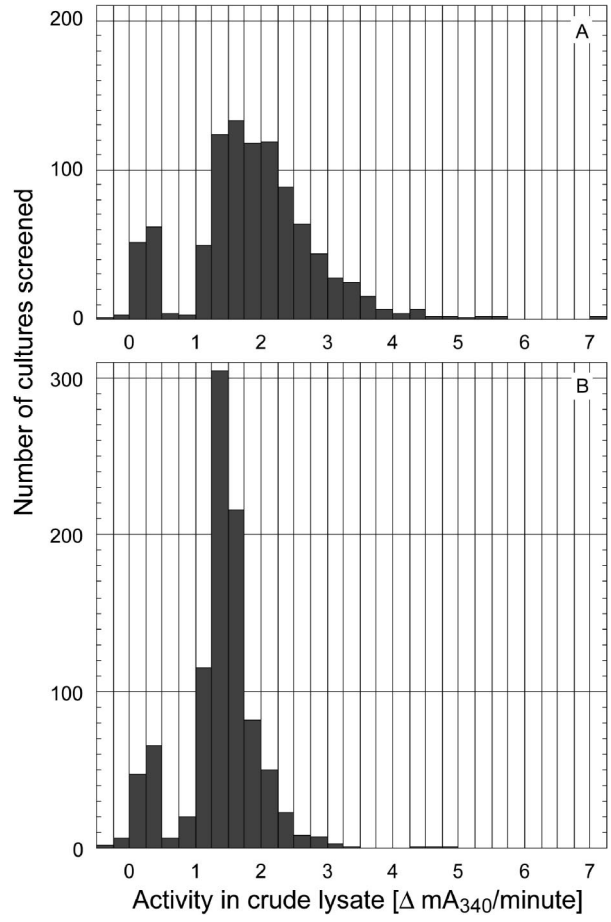

FIGURE 8: Distribution of crude lysate activity in 960 cultures expressing a *pdc1* library with either 4 (A) or 11 (B) nucleotide substitutions per gene on average in generation 3. Both libraries had a genotype density of 2.1 (transformed cells starting each culture on average) and were analyzed with assay B as described in Culture Screening. Wells with sterile media had less than 0.5 Δ mA_{340}/min, whereas the best parental form (expressing 2S3 Pdc1) gave almost 5 Δ mA_{340}/min.

Analysis of the culture activities during screening in generation 3 showed that the high error rate was too high because most variants were inactive, whereas the low error rate provided a distribution in culture activity that extended beyond the capabilities of the parents (Figure 8). This was also evident in the progeny of generations 1 and 3: although the number of variants analyzed from the high and low error libraries was equal, 70% of the progeny originated from the low error rate gene pool (variants with an "L" in their name, Table 2). These two generations also showed that the progeny were biased toward fewer mutations overall and more silent mutations. Beneficial mutations were only being found when they occurred with no more than one or two other new expressed mutations. In the next round of EPCR, generation 5, the high error rate was abandoned since Pdc1 was clearly unable to tolerate the high mutation rate.

The recombination observed in generations 3 and 5 was probably due to template switching during elongation of nascent genes, as in StEP (*16*), since the Mn^{2+} ions added for EPCR reduce the productivity of *Taq* DNA polymerase (*13*). The potential for recombination from EPCR was increased in generation 5 with half of the variants being

derived from EPCR for low error rate with the extension time reduced to 1 min. The majority of the progeny for generation 5 were derived from this library (indicated by "LS" in a variant's name) rather than the low error rate library (Table 2). Silent mutations were present in generation 5, but no new expressed mutations were found; the improvement in this generation was entirely due to recombination. The simplicity of the EPCR with abbreviated extension and its ability to introduce point mutations while facilitating recombination make it ideal for future directed evolution projects.

Mutation Bias. The mutagenesis method was, however, limited to the mutation bias imposed by *Taq* DNA polymerase. Given the codon use in *pdc1*, residues only had the potential to be mutated to one of two to four other amino acids. This is about half of the expressed variability available if point mutations could be generated randomly (*23*). Mutations found in this study have not necessarily found the optimal amino acid for a particular residue, and site-saturation mutagenesis, as used by Miyazaki and Arnold (*23*), would be one option to explore further improvements.

The location bias observed for expressed mutations (Figure 1) indicates not only where beneficial changes were made but also the regions where Pdc1 is resilient enough to accept changes. Of the three domains, the gamma domain tolerated the most mutations, possibly because this domain provides ligands for the Mg^{2+} ion and most of the ThDP-binding pocket (*24*). Although specific residues need to be conserved for cofactor binding, the cofactors could act as a scaffold for this domain and allow variability in other residues. This is exemplified by the mutation D472N, discovered in generation 1, which lies between two of the residues involved in binding Mg^{2+}. Other regions prone to mutation were toward the end of the alpha and beta domains (Figure 1). In spite of the number of expressed mutations in the beta domain, none persisted into 5LS30.

Locations of Mutations in the Alpha Domain of 5LS30. Two of the five expressed mutations in 5LS30 were located in the alpha domain. A143T alters a residue that, in the crystal structure by Arjunan et al. (*5*), is close to the same residue on opposite dimers within the Pdc1 tetramer. This mutation could therefore have some influence on oligomerization. Residue 143 occurs near the start of an α helix, and toward the end of this same helix is the site of 5LS30's other mutation in the alpha domain, T156A. Residue 156 is intriguing since it is adjacent to a residue, Y157 (Figure 9), that appears to bind pyruvamide in the crystal structure by Luet al. (*6*). The relevance of this binding site to Pdc1's substrate activation has been disputed because pyruvamide activates Pdc1 in a different fashion to pyruvate and the site between Y157's hydroxyl group and the backbone oxygen of R224 would be unfavorable for pyruvate (*25*). Previous research (*26, 27*) has implicated the involvement of C221 and nearby H92 (*28*). However, the dominance of T156A in generation 4 suggests that the potential binding site at Y157 requires further examination.

Mutations Responsible for Greater Thermal Stability. K478R was present in 5LS30 but not in 4S17 and must be responsible for an increase of 5.5–5.9 °C in $T_{1/2}$. Compared to a lysine at residue 478, the extra length and hydrogen bonding potential from an arginine residue could increase the hydrogen bond network and possibly replace the bound water molecule within Pdc1's structure (Figure 10). The mutation from lysine to arginine is often observed to increase the thermal stability of proteins (*29*); however, another mutation seen in this project, K435R, had no such effect on the stability of 5LS40 (data not shown). This is probably because residue 435 is on the surface of Pdc1, pointing into the solution, rather than interacting with other residues.

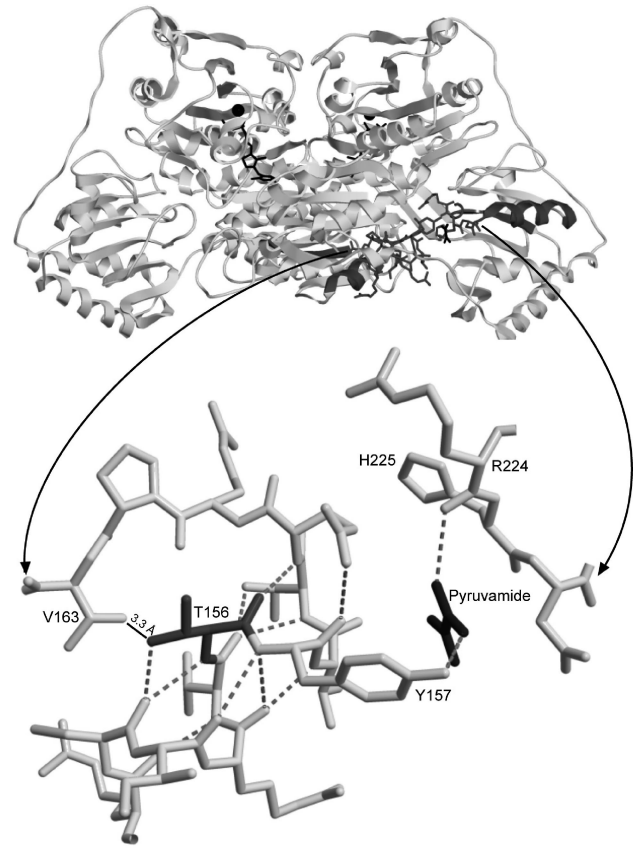

FIGURE 9: The Pdc1 dimer (*6*) is illustrated in light gray ribbons except for the cofactors and the protein strands involved in binding pyruvamide away from the catalytic site (dark gray). A section of the selected protein strands (stick form) has been enlarged and pitched 90° about its horizontal axis to be viewed above (relative to the image of the dimer). This enlarged view shows threonine 156 (T156) and pyruvamide in dark gray with potential hydrogen bonds shown as dotted lines.

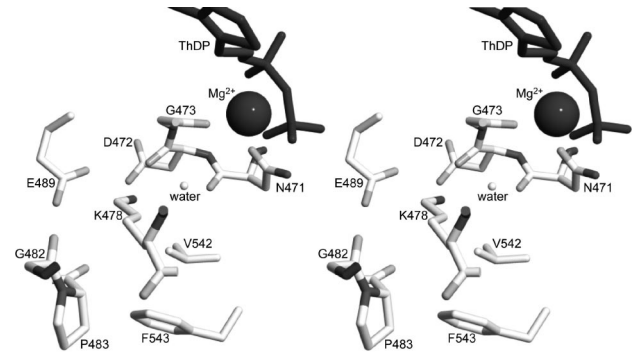

FIGURE 10: The environment of residue 478 (K478) in a crystal structure of Pdc1 (*5*) is presented as a stereo diagram with residues and cofactors labeled. An arginine residue at position 478 could usurp the crystallographic water molecule (water) to interact with the backbone oxygens of residues 471 and 472 (N471 and D472). Alternatively, the arginine residue could interact with the acid group of residue 489 (E489) and the backbone oxygen of residue 482 (G482).

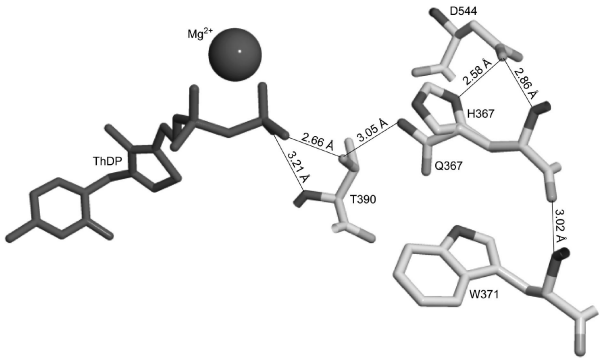

FIGURE 11: Glutamine 367 (Q367) is adjacent to a threonine involved in forming a pocket for ThDP's diphosphate group (T390). The best fitting rotamer of a histidine at position 367 (H367) has been superimposed onto the original structure of Pdc1 (*5*). The putative hydrogen bond network between residues T390, H367, and D544 appears to be possible after adjusting their torsion angles.

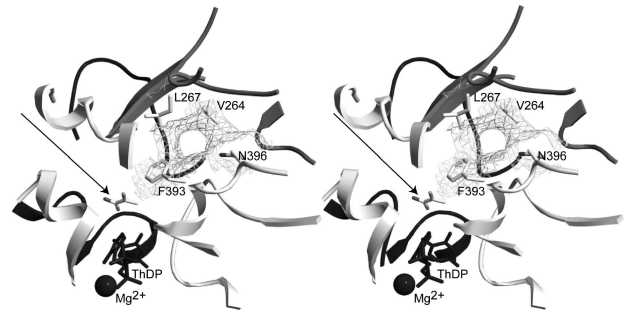

FIGURE 12: The crystallographic structure of Pdc1's catalytic site with pyruvamide bound (*6*) is displayed in this stereo diagram with the regions that define it. These regions of Pdc1 are displayed as ribbons: white for those in the gamma domain, gray for the beta domain, and dark gray for those in the opposite monomer's alpha domain. The cofactors, in dark gray, are labeled, and an arrow points to the pyruvamide molecule through the largest opening to the catalytic site. Another opening exists between the beta and gamma domains and is defined by the residues V264, L267, F393, and N396 (their side chains are illustrated as rods and labeled). The surface defined by these four residues (rendered in light gray mesh) was calculated with the Swiss PDB Viewer (quality level 3).

Of the four mutations common to 5LS30 and 4S17 Pdc1, Q367H is probably responsible for the 3.0–4.0 °C increase in $T_{1/2}$ for 4S17 compared to native Pdc1. This is the only change that could increase the network of interactions within Pdc1 and therefore increase the protein stability. Unlike the glutamine in native Pdc1, a histidine at residue 367 could form hydrogen bonds with the side chains of D544 and T390 (Figure 11). The positions of both T390 and D544 would need to be altered slightly to accommodate this putative hydrogen bond network.

Overall, the $T_{1/2}$ of 5LS30 was 8.7−9.7 °C higher than that of native Pdc1. This improvement in Pdc1 was serendipitously selected for: greater thermal stability would allow more of the Pdc1 expressed overnight at 37 °C to remain active for the screening assays.

N396I Is Located near an Open Catalytic Cavity. N396I has highlighted an overlooked feature of Pdc1's catalytic site: multiple openings to the bulk solvent. The largest opening could be sealed off by a lid-like C-terminus (*30*), but another two openings are present in the Pdc1 crystal structure (*5*) due to what appears to be a loose fit between the beta domain and the complex of alpha and gamma domains. If complete closure of the catalytic site is necessary for activity, then N396I is ideal, since it provides a hydrophobic residue at an auxiliary opening to the bulk solvent (Figure 12). This opening is defined by the side chains of K269, Q397, and N548 on the surface of Pdc1 and roughly 5 Å in diameter (not including hydrogen atoms). It leads to an opening to the catalytic site defined by V264, F297, F393, and N396 restricting the internal opening to about 7 Å in diameter (not including hydrogen atoms). With the mutation N396I, the aforementioned opening is surrounded by only hydrophobic residues that could more easily interact to seal off this region of the catalytic cavity from the bulk solution. Another mutation, G406C, occurred in the region of residue 396 but was never isolated along with N396I during directed evolution. Whether this was due to linkage between these two mutations or some deleterious effect of both mutations occurring has not yet been investigated. It is possible that N396I was discovered because it enables the beta domain to associate with the other domains such that activation occurs more easily.

5LS30 Is Activated More Easily Than Native Pdc1. The directed evolution of Pdc1 has led to a version that can be activated more easily than the native form (Figure 3). Rather than arising from changes to the catalytic site, it is likely that the kinetic characteristics of 5LS30 Pdc1 variants stem from this increased ease of activation. By decreasing the activation threshold, 5LS30 Pdc1 was able to have a greater proportion of its population active in low pyruvate concentration, giving rise to a reduced $S_{0.5}$ value. The change in activation by pyruvate could also be responsible for the reduction of $S_{0.5}$ sensitivity to pH. Since high pH causes the Pdc1 holoenzyme to split apart at the interface that forms its catalytic sites (*4*), an active Pdc1 with ligands in its catalytic site is more likely to remain active.

If cooperativity is a reflection of the proportion of Pdc1 holoenzymes in an active confirmation, then the reduction in the Hill coefficient also stems from the increased ease of activation. The cooperative nature of 5LS30 activation occurs when the catalytic activity is very low so that its influence on the catalytic rate is reduced. As a result, the Hill coefficient was close to 1 at low pH for 5LS30 (Figure 5C), even though cooperativity was not completely abolished.

Pdc1 Sensitivity to Phosphate Was Reduced. The influence of phosphate on native Pdc1 suggests that it competes with pyruvate in the catalytic site. Furthermore, since it increases cooperativity, it must prevent a Pdc1 oligomer from shifting to an active conformation while present in at least one of the sites. For this reason, phosphate could provide further insight into the mechanism of Pdc1 activation. 5LS30 Pdc1 still has phosphate competing at the catalytic site, albeit with a higher K_{iP}, but its presence no longer influences the cooperativity (Table 3) or activation (Figure 3). This suggests that 5LS30 Pdc1 can be activated by pyruvate even when phosphate is bound in the catalytic site.

Pdc1 Optimal Temperature Was Increased. Pdc1's catalytic rate in 1 mM pyruvate has a counterintuitive response: it decreases with increasing temperature above 30 °C. This previously unobserved sensitivity to temperature could be physiologically relevant; perhaps yeast growing anaerobically benefit from limiting exergonic processes such as pyruvate decarboxylation and fermentation as the temperature increases. The decrease in activity occurs well below the temperature required for irreversible denaturation of the

protein and is probably related to the ability for activation to occur. Only when the pyruvate is saturating at 25 mM does irreversible inactivation influence the optimal catalytic temperature. In this situation, the presence of K478R in 5LS30 increases the optimal temperature compared to variant 4S17 (Figure 7). The increased optimal temperature of 5LS30 at low pyruvate concentration, 8 °C higher than native Pdc1, probably stems from easier activation, analogous to the increased activity in high pH at low pyruvate concentration.

The dependence of Pdc1's optimal catalytic temperature on the pyruvate concentration suggests that pyruvate confers additional stability, possibly through the conformational change involved in activation. Pyruvate's influence on stability has also been observed with a similar Pdc from *Candida utilis* (Torula yeast). Rosche et al. (*31*) found that this version of Pdc is inactivated after 30 min of exposure to benzaldehyde but retains activity if pyruvate is present during exposure. Alternatively, increases in temperature above 30 °C could reduce the affinity for pyruvate at the activation site, possibly through increased motion of the beta domain. In either case, it is likely that the structural changes responsible for easier activation are also responsible for the increased optimal temperature for decarboxylation of pyruvate.

CONCLUSIONS

Pdc1 has evolved with sensitivity to pH, temperature, phosphate, and pyruvate that is perfectly suited to the metabolism in yeast. Substrate activation is probably the key mechanism that underpins this sensitivity. This research has succeeded in developing a directed evolution method for Pdc1 and has allowed new mutations to be discovered. These mutations are not detrimental to catalytic activity and can alter substrate activation such that Pdc1 is activated more easily. Rather than finding Pdc1 variants with eliminated regulation, we have found a variant with attenuated sensitivity for pyruvate concentration, phosphate concentration, pH, and temperature. The mutations within 5LS30 Pdc1 will be ideal targets for site-directed mutagenesis and further research into how Pdc1's structure facilitates intricate regulation.

The directed evolution method developed for pyruvate decarboxylase has consequences for other biotechnology projects. First of all, it has been possible to increase the thermal stability of Pdc1 by almost 10 °C. Directed evolution with more stringent screening for thermal stability (i.e., heating cultures prior to assays) could be applied to Pdc1 or the naturally more stable Pdc from *Z. mobilis*. A thermostable Pdc would be useful for thermophilic fermentation and more efficient ethanol production. In addition, this project introduces the concept of genomic density: analyzing mixtures of genotypes and isolating the optimal variants after a primary screening. Compared to colony picking, this approach allows bacterially expressed enzymes to be analyzed in a 96-well plate format with higher throughput.

ACKNOWLEDGMENT

Analysis for DNA sequencing was performed by the Biomolecular Resource Facility at the Australian National University.

REFERENCES

1. Alvarez, F. J., Ermer, J., Hübner, G., Schellenberger, A., and Schowen, R. L. (1991) Catalytic power of pyruvate decarboxylase. Rate-limiting events and microscopic rate constants from primary carbon and secondary hydrogen isotope effects. *J. Am. Chem. Soc. 113*, 8402–8409.
2. Hübner, G., Weidhase, R., and Schellenberger, A. (1978) The mechanism of substrate activation of pyruvate decarboxylase: a first approach. *Eur. J. Biochem. 92*, 175–181.
3. Boiteux, A., and Hess, B. (1970) Allosteric properties of yeast pyruvate decarboxylase. *FEBS Lett. 9*, 293–296.
4. Sergienko, E. A., and Jordan, F. (2002) Yeast pyruvate decarboxylase tetramers can dissociate into dimers along two interfaces. Hybrids of low-activity D28A (or D28N) and E477Q variants, with substitution of adjacent active center acidic groups from different subunits, display restored activity. *Biochemistry 41*, 6164–6169.
5. Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D., and Jordan, F. (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.
6. Lu, G., Dobritzsch, D., König, S., and Schneider, G. (1997) Novel tetramer assembly of pyruvate decarboxylase from brewer's yeast observed in a new crystal form. *FEBS Lett. 403*, 249–253.
7. Neylon, C., Brown, S. E., Kralicek, A. V., Miles, C. S., Love, C. A., and Dixon, N. E. (2000) Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry 39*, 11989–11999.
8. Candy, J. M., Duggleby, R. G., and Mattick, J. S. (1991) Expression of active yeast pyruvate decarboxylase in *Escherichia coli*. *J. Gen. Microbiol. 137*, 2811–2815.
9. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol. 185*, 60–89.
10. Shafqat, J., Hoog, J. O., Hjelmqvist, L., Oppermann, U. C., Ibanez, C., and Jornvall, H. (1999) An ethanol-inducible MDR ethanol dehydrogenase/acetaldehyde reductase in *Escherichia coli*: structural and enzymatic relationships to the eukaryotic protein forms. *Eur. J. Biochem. 263*, 305–311.
11. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif. 41*, 207–234.
12. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage. *Nature 227*, 680–685.
13. Leung, D. W., Chen, E., and Goeddel, D. V. (1989) A method for random mutagenesis of a defined DNA fragment using a modified polymerase chain reaction. *Technique 1*, 11–15.
14. Stemmer, W. P. C. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U.S.A. 22*, 10747–10751.
15. Lorimer, I. A. J., and Pastan, I. (1995) Random recombination of antibody single chain Fv sequences after fragmentation with DNase1 in the presence of Mn^{2+}. *Nucleic Acids Res. 23*, 3067–3068.
16. Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol. 16*, 258–261.
17. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning. A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York.
18. Conway, T., Sewell, G. W., Osman, Y. A., and Ingram, L. O. (1987) Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. *J. Bacteriol. 169*, 2591–2597.
19. Liu, M., Sergienko, E. A., Guo, F., Wang, J., Tittmann, K., Hubner, G., Furey, W., and Jordan, F. (2001) Catalytic acid-base groups in yeast pyruvate decarboxylase. 1. Site-directed mutagenesis and steady-state kinetic studies on the enzyme with the D28A, H114F, H115F, and E477Q substitutions. *Biochemistry 40*, 7355–7368.
20. Good, N. E., Winget, D., Winter, W., Connolly, T. N., Izawa, S., and Singh, R. M. M. (1966) Hydrogen ion buffers for biological research. *Biochemistry 5*, 467–477.
21. Vanhercke, T., Ampe, C., Tirry, L., and Denolf, P. (2005) Reducing mutational bias in random protein libraries. *Anal. Biochem. 339*, 9–14.
22. Moore, J. C., and Arnold, F. H. (1996) Dircted evolution of a *para*-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol. 14*, 458–467.
23. Miyazaki, K., and Arnold, F. H. (1999) Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J. Mol. Evol. 49*, 716–720.

24. Candy, J. M., and Duggleby, R. G. (1998) Structure and properties of pyruvate decarboxylase and site-directed mutagenesis of the *Zymomonas mobilis* enzyme. *Biochim. Biophys. Acta 1385*, 323–238.
25. Sergienko, E. A., and Jordan, F. (2002) New model for activation of yeast pyruvate decarboxylase by substrate consistent with the alternating sites mechanism: demonstration of the existence of two active forms of the enzyme. *Biochemistry 41*, 3952–3967.
26. Baburina, I., Gao, Y., Zhixiang, H., Jordan, F., Hohmann, S., and Furey, W. (1994) Substrate activation of brewer's yeast pyruvate decarboxylase is abolished by mutation of cysteine 221 to serine. *Biochemistry 33*, 5630–5635.
27. Baburina, I., Moore, D. J., Volkov, A., Kahyaoglu, A., Jordan, F., and Mendelsohn, R. (1996) Three of four cysteines, including that responsible for substrate activation, are ionized at pH 6.0 in yeast pyruvate decarboxylase: evidence from fourier transform infrared and isoelectric focusing studies. *Biochemistry 35*, 10249–10255.
28. Baburina, I., Li, H., Bennion, B., Furey, W., and 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.
29. Mrabet, N. T., Van den Broeck, A., Van den Brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M., and Wodak, S. J. (1992) Arginine residues as stabilizing elements in proteins. *Biochemistry 31*, 2239–2253.
30. Lobell, M., and Crout, D. H. G. (1996) Pyruvate decarboxylase: a molecular modelling study of pyruvate decarboxylase and acyloin formation. *J. Am. Chem. Soc. 118*, 1867–1873.
31. Rosche, B., Breuer, M., Hauer, B., and Rogers, P. L. (2005) Role of pyruvate in enhancing pyruvate decarboxylase stability towards benzaldehyde. *J. Biotechnol. 115*, 91–99.

BI701858U