**RESEARCH**

# An iterative computational design approach to increase the thermal endurance of a mesophilic enzyme

Deanne W. Sammond[1*†], Noah Kastelowitz[2†], Bryon S. Donohoe[1], Markus Alahuhta[1], Vladimir V. Lunin[1], Daehwan Chung[1], Nicholas S. Sarai[1], Hang Yin[2], Ashutosh Mittal[1], Michael E. Himmel[1], Adam M. Guss[3] and Yannick J. Bomble[1*]

## Abstract

**Background:** Strategies for maximizing the microbial production of bio-based chemicals and fuels include eliminating branched points to streamline metabolic pathways. While this is often achieved by removing key enzymes, the introduction of nonnative enzymes can provide metabolic shortcuts, bypassing branched points to decrease the production of undesired side-products. Pyruvate decarboxylase (PDC) can provide such a shortcut in industrially promising thermophilic organisms; yet to date, this enzyme has not been found in any thermophilic organism. Incorporating nonnative enzymes into host organisms can be challenging in cases such as this, where the enzyme has evolved in a very different environment from that of the host.

**Results:** In this study, we use computational protein design to engineer the *Zymomonas mobilis* PDC to resist thermal denaturation at the growth temperature of a thermophilic host. We generate thirteen PDC variants using the Rosetta protein design software. We measure thermal stability of the wild-type PDC and PDC variants using circular dichroism. We then measure and compare enzyme endurance for wild-type PDC with the PDC variants at an elevated temperature of 60 °C (thermal endurance) using differential interference contrast imaging.

**Conclusions:** We find that increases in melting temperature ($T_m$) do not directly correlate with increases in thermal endurance at 60 °C. We also do not find evidence that any individual mutation or design approach is the major contributor to the most thermostable PDC variant. Rather, remarkable cooperativity among sixteen thermostabilizing mutations is key to rationally designing a PDC with significantly enhanced thermal endurance. These results suggest a generalizable iterative computational protein design approach to improve thermal stability and endurance of target enzymes.

**Keywords:** Computational protein design, Biofuels, Thermal stability, Pyruvate decarboxylase

## Background

Advanced biofuels can meet up to three billion barrels of liquid fuel demand on a yearly basis, but face high costs and process complexity as obstacles to commercial implementation [1, 2]. Consolidated bio-processing (CBP) offers a way to address both cost and process complexity by combining cellulase production, enzymatic digestion, and fermentation into a one-pot reaction [3]. The challenge of CBP is to engineer a single microbe that has the capability of deconstructing cellulosic biomass and fermenting the released sugars to biofuels or bio-based chemicals at sufficiently low cost. Promising CBP microbes are able to effectively deconstruct plant biomass to elemental sugars [4–6],

*Correspondence: Deanne.Sammond@nrel.gov;
Yannick.Bomble@nrel.gov
†Deanne W. Sammond and Noah Kastelowitz contributed equally to this work
[1] Biosciences Center, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA
Full list of author information is available at the end of the article

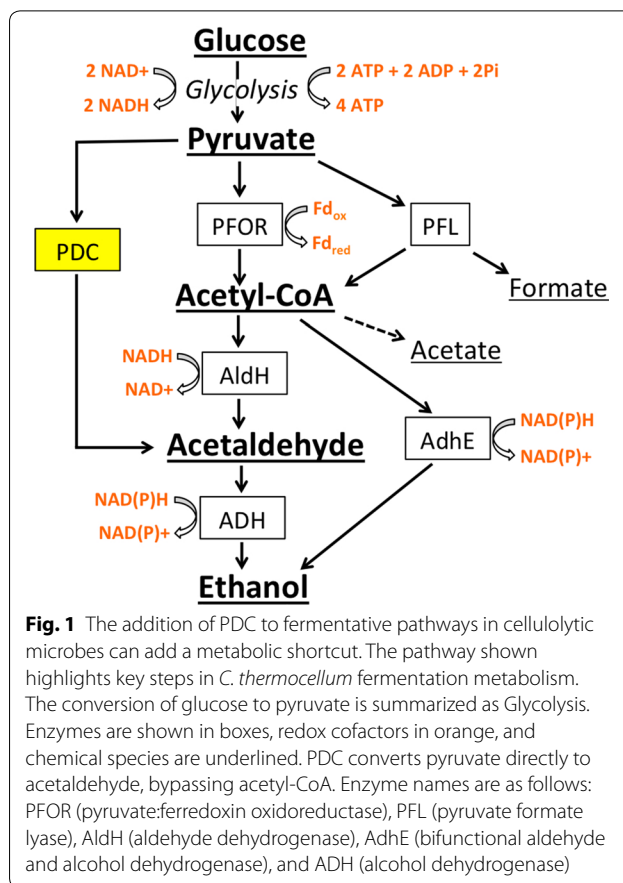Sammond *et al. Biotechnol Biofuels (2018) 11:189*

Page 2 of 13

but still require optimization to increase the fermentation of those sugars to products such as bioethanol [7–9].

Metabolic pathway engineering of CBP microbes has focused on achieving chemical outputs near theoretical yields. Streamlining metabolic pathways by removing enzymes at branched points is a common approach to increase chemical yields. Alternatively, branch points can be bypassed entirely through the introduction of nonnative enzymes. Introducing a nonnative enzyme can be difficult, however, if the enzyme has not evolved to function in an environment similar to that of the host organism. In these cases, the hope is that protein engineering can be used to rationally engineer enzymes, increasing the range of tools available for metabolic pathway engineering.

Pyruvate decarboxylase (PDC) converts pyruvate to acetaldehyde through nonoxidative decarboxylation, a reaction that typically occurs in anaerobic fermentations. An alternative metabolic route for converting pyruvate to ethanol found in fermentative cellulolytic organisms is branched and produces several additional products including acetate, formate, and lactate [10]. Expression of a PDC in such CBP microbes can bypass these branched points, channeling more pyruvate directly to acetaldehyde for conversion to ethanol (Fig. 1).

Incorporation of the *Z. mobilis* PDC and alcohol dehydrogenase (ADH) into the cellulolytic mesophile, *Clostridium cellulolyticum*, resulted in a 53% increase in the production of ethanol and a 48% decrease in the production of lactate [11]. *Clostridium thermocellum* is a highly cellulolytic thermophile that has emerged as one of the most promising CBP microbes [12]. To date, however, PDC has not been identified in any thermophilic organism, challenging the incorporation of PDC into microbes such as *C. thermocellum*. The potential for using enzymes like PDC to increase yields of renewable fuels and chemicals demonstrates the value of combining computational protein design and metabolic engineering [13].
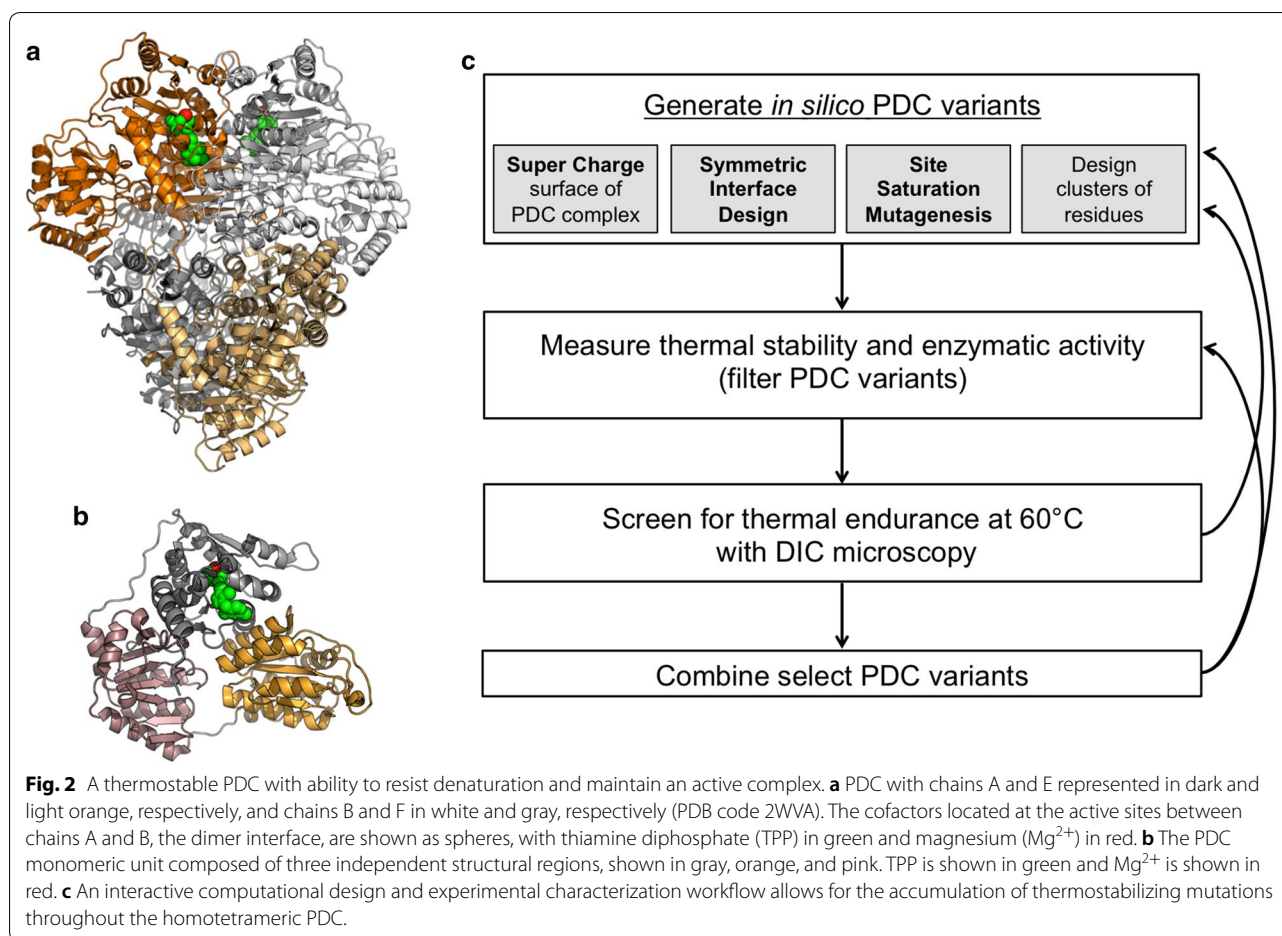
Protein design tools enhance our ability to manipulate the physicochemical properties of target enzymes. While a majority of randomly selected mutations are deleterious, substitutions identified using computational approaches have a much higher rate of success. In fact, previous computational protein design results demonstrate that multiple independent mutations can be combined to efficiently enhance the thermostability of a target protein without having to screen each substitution individually [14, 15]. Although there are examples of proteins that have been made more thermostable by computational protein design methods, many previous examples have been smaller monomeric proteins [16, 17].



**Fig. 1** The addition of PDC to fermentative pathways in cellulolytic microbes can add a metabolic shortcut. The pathway shown highlights key steps in *C. thermocellum* fermentation metabolism. The conversion of glucose to pyruvate is summarized as Glycolysis. Enzymes are shown in boxes, redox cofactors in orange, and chemical species are underlined. PDC converts pyruvate directly to acetaldehyde, bypassing acetyl-CoA. Enzyme names are as follows: PFOR (pyruvate:ferredoxin oxidoreductase), PFL (pyruvate formate lyase), AldH (aldehyde dehydrogenase), AdhE (bifunctional aldehyde and alcohol dehydrogenase), and ADH (alcohol dehydrogenase)

Engineering a larger tetrameric protein like PDC poses additional design challenges. For example, large, multidomain proteins likely have complex folding trajectories and therefor may not have energy minima corresponding to single stable conformations [18]. Thus, larger proteins may require stabilization in more than one structural region in order to observe significant increases in thermal stability.

Engineering an enzyme to function in a thermophilic organism requires increasing the thermal endurance of the enzyme at the growth temperature of the host and may not directly correlate with increases in melting temperature. *C. thermocellum*, for example, grows at 60 °C [19], which is well above the growth temperature of *Z. mobilis* (maximum growth temperature 42 °C) [20] but just below the melting temperature ($T_m$) of the *Z. mobilis* PDC (62 °C) (Additional file 1: Figure S1A). Changes in ellipticity at 60 °C indicate PDC is within a structurally unstable temperature. When seeking to engineer a mesophilic enzyme for use in a thermophilic host organism it is unknown what the target melting temperature should be to achieve an extended enzyme lifetime, or thermal endurance, at the target temperature.

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 3 of 13



**Fig. 2** A thermostable PDC with ability to resist denaturation and maintain an active complex. **a** PDC with chains A and E represented in dark and light orange, respectively, and chains B and F in white and gray, respectively (PDB code 2WVA). The cofactors located at the active sites between chains A and B, the dimer interface, are shown as spheres, with thiamine diphosphate (TPP) in green and magnesium ($Mg^{2+}$) in red. **b** The PDC monomeric unit composed of three independent structural regions, shown in gray, orange, and pink. TPP is shown in green and $Mg^{2+}$ is shown in red. **c** An interactive computational design and experimental characterization workflow allows for the accumulation of thermostabilizing mutations throughout the homotetrameric PDC.

Here we engineer PDC enzymes with increased thermal stability using the Rosetta protein modeling software [21]. We screen the PDC designs using circular dichroism to identify designs with increased resistance to thermal denaturation. We then monitor the lifespan of the enzymes at 60 °C using phase-contrast microscopy and find that while we are able to engineer enzymes with significantly longer life-spans, increases in melting temperature do not correlate to increases in endurance at elevated temperature.

## Methods

### In silico design

Substitutions were selected based on Rosetta energy units as well as identifying substitutions displaying features associated with increased thermal stability, including increasing buried hydrophobic surface area [22], increasing quality of side-chain packing [23], increase surface charge of the tetrameric complex [24], decreasing conformational flexibility and improved packing at the symmetric dimer interface [25]. Computational approaches can be clustered into four topics:

1. Single substitutions were identified with in silico site saturation mutagenesis, using the Rosetta protein modeling software [21]. The in silico experiments used either a fixed-backbone approximation, where the backbone coordinates from the X-ray crystal structure were held fixed in the simulation, or a minimization of backbone and side-chain torsion angles prior to evaluating amino acid identities at each position.

2. Previous work comparing thermophilic and mesophilic enzymes showed that the backbone of structurally equivalent clusters can accommodate alternate sequence combinations by adjusting an average of 3.5 Å for each Cα atom [23]. Here, clusters of interacting residues were determined using a distance cutoff, where residues were considered interacting when any side-chain heavy-atoms were within 4 Å. Two clusters were identified for design targets based on poor side-chain packing, G6, I97, I166, A171 and H21, M42, Q44, Y46. These two clusters were designed by iterating between sequence design

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 4 of 13

and minimization of backbone and side-chain torsion angles.

3. The PDC dimer displays symmetry at the interface resulting in a design constraint where any mutation made to one monomer must be mirrored in the bound partner. The symmetric design protocol simultaneously models the effects of an amino acid substitution at both positions at the interface [25].

4. Engineering a protein surface to have a high net charge can increase solubility and enhance thermal stability [24, 26]. Designing the PDC complex on the hypothesis that supercharging the surface can stabilize the interactions between domains, we identified substitutions to impart a high net charge while modeling the energetic consequences of the substitutions.

We selected $PDC_{1.01}$ and $PDC_{1.10}$ as the foundations for the second round of protein design based on the observed enhancements in thermal stability. We identified four additional mutations by selecting the next set of substitutions that would remove glycine residues (G491A, G515A, G516A, and G540A; Additional file 1: Table S1). We targeted the identification of additional glycine-removing mutations due to the successful thermal stabilization of $PDC_{1.01}$, which contained two such mutations. These cases where a substitution removed a glycine residue underwent additional evaluation. Glycine, which has only a hydrogen atom at its side-chain position, can accommodate phi–psi backbone angles not accessible to any other natural amino acid. For this reason, any in silico mutations replacing a glycine residue underwent a filter based on Rosetta energy scores. The Rosetta energy term, p_aa_pp, gives the probability of finding an amino acid at given phi and psi dihedral angles. In silico mutations that removed a glycine residue were only considered if the p_aa_pp energy score was below 0.2. Additionally, glycine is the smallest amino acid. Mutations that passed the above filter were then sorted by the Lennard–Jones repulsive energy term (LJrep) to ensure significant atomic clashes were not present that might indicate the protein region could not accommodate the substitution (Additional file 1: Table S1). Glycine residues within 6 Å of the dimer active sites were not considered for redesign. We additionally avoided mutations introducing cysteine residues, as cysteine residues can participate in redox reactions [27].

A solvent exposed mutation in $PDC_{1.01}$, A189K, was identified as a potentially beneficial single mutation in an in silico point mutation scan, but places a positively charged residue at the PDC surface and thus does not fit the goal of the negatively supercharged surface for $PDC_{1.10}$. This mutation was subsequently removed in $PDC_{2.02}$. A version of $PDC_{2.02}$ that included the A189K

mutation was characterized and evaluated using CD to confirm the mutation does not alter thermal stability based on observed changes in molar ellipticity (Additional file 1: Figure S3).

## Construction, cloning, and expression of PDC variants

DNA sequences for each PDC variant were codon optimized and cloned into a pET22b(+) vector (GenScript, Piscataway, NJ). The sequence for a hexahistidine tag was placed at the C terminus of the constructs. PDC protein variants were expressed either for 4 h at 37 °C or overnight at 25 °C with 0.1 mM IPTG in the BL21 (DE3) strain of *Escherichia coli*. Proteins were concentrated using Vivaspin spin columns with a molecular weight cutoff of 10,000 Da (GE Healthcare Life Sciences, Pittsburgh).

Protein purification. The frozen cell pellets were thawed at room temperature with equal volume of buffer A (50 mM Tris pH 7.5, 100 mM NaCl, 10 mM imidazole, 0.1 mM thiamine pyrophosphate (TPP), 0.5 mM dithiothreitol (DTT) and 1 mM $MgCl_2$) and lysed with lysozyme and sonication. One mg/mL lysozyme (Hampton Research, Aliso Viejo, CA), 1.0 U/mL Pierce Universal Nuclease (Thermo Scientific, Rockford, IL) and EDTA-free protease inhibitor (Thermo Scientific, Rockford, IL) according to manufacturer instructions were added in the lysis mixture and incubated for 30 min at room temperature with occasional vortexing. Sonication was done at room temperature for 2 min using a Branson 5510 water bath sonicator (Branson Ultrasonics Corporation, Danbury, CT). Cell debris was removed by centrifugation at $15,000 \times g$ for 15 min. The supernatant was loaded onto an eight mL HisPur Cobalt column (Thermo Scientific, Rockford, IL) using an Akta FPLC system (GE Life Sciences, Piscataway, NJ) with buffer A (50 mM Tris pH 7.5, 100 mM NaCl, 20 mM imidazole, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$). After loading and washing the unbound proteins from the column, PDC samples were eluted using 100% of Buffer B (50 mM Tris pH 7.5, 100 mM NaCl, 250 mM imidazole, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$). Final purification was performed by size-exclusion chromatography using a HiLoad Superdex 200 (26/60) column (GE Healthcare, Piscataway, New Jersey, USA) in buffer C (20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$).

## Differential scanning calorimetry

Protein samples were analyzed by differential scanning calorimetry (DSC) using a Microcal VP-DSC instrument (Malvern, Worcestershire, UK) to measure the excess heat capacity of protein unfolding as a function of temperature. These measurements were used to directly

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 5 of 13

calculate the enthalpy of unfolding ($\Delta H_{cal}$) for each protein sample according to the equation:

$$\Delta H = \int\limits_{T_f}^{T_u} C_p \mathrm{d}T$$

Protein samples were measured over a temperature range of 10–90 °C and at a scan rate of 60 °C/h. No feedback mode was used for each DSC experiment. Buffer baseline scans were established before each protein run by loading both the sample and reference cells with buffer and performing 3 or more heating and cooling cycles until the deviation in scans was less than 0.01 mcal/min. Buffer was then removed from the sample cell and replaced with the protein sample during the temperature range of 25–15 °C during the cool down cycle. All protein samples were diluted to 0.2 mg/mL in 20 mM TRIS pH 7.5, 100 mM MnCl, 1 mM $MgCl_2$, 0.5 mM DTT, 0.1 mM TPP buffer and run at that concentration in order to minimize post-unfolding aggregation.

DSC data on each sample was analyzed by the Origin 7.0 software [28] coupled to a DSC data analysis module provided with the Microcal VP-DSC. Sample scans were buffer subtracted, normalized to molar concentration and corrected by baseline correction options in the DSC data analysis module prior to least squares analysis using a non-two-state model option. The fitted sample curves produced the melting temperature ($T_m$), enthalpy of unfolding ($\Delta H_{cal}$), and van't Hoff enthalpy ($\Delta H_{vH}$) for each transition.

### Circular dichroism spectroscopy and thermal melts
Circular dichroism (CD) spectra and thermal melt measurements were performed on a Chirascan-plus spectrometer (Leatherhead, Surrey, UK) using a 0.5-mm path-length quartz cuvette. WT and mutant PDC proteins were buffer exchanged into a potassium phosphate buffer (10 mM potassium phosphate, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM TPP, 0.5 mM DTT, pH 7.5) and prepared to a final concentration of 0.2 mg/mL. CD spectra were measured at the stated temperatures with a step size of 0.25 nm. The thermal melts were performed in continuous ramp mode at a rate of 2 °C/min while measuring CD at 222 nm. The $T_m$ of each protein was determined from the first derivative of the thermal melt curves using Prism 6 (GraphPad, La Jolla, CA). Each thermal melt experiment was performed in triplicate.

### ThermoFluor high-throughput protein stability assay
The high-throughput ThermoFluor assay was used to evaluate the effects of the glycine-to-alanine substitutions G491A, G515A, G516A, and G540A. For the ThermoFluor assay, a hydrophobic dye, SYPRO Orange, binds to exposed hydrophobic regions of the protein. As proteins begin denaturation upon heating, increasing amounts of hydrophobic regions are exposed resulting in an increased signal. The assay is performed in an RT-PCR machine using 96-well plates, allowing the simultaneous characterization of many protein variants. This assay has been used to rapidly screen protein mutants generated from computational design approaches [29, 30], and thus allowed us to evaluate the threshold for selecting these mutations for combinatorial libraries.

SYPRO® Orange Protein Gel Stain, supplied at 5000× concentrate in dimethyl sulfoxide (Thermo Scientific, Waltham, MA), was diluted to 200× in buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$). Wild-type PDC and variants were diluted to 5 μM in buffer, with protein concentrations determined by measuring absorbance at 280 nm. Extinction coefficients were calculated using the method described by Gill and von Hipple [31]. 50 μL samples were made by combining 45 μL of protein with 5 μL of 200× SYPRO® Orange stain. Protein variants were measured in triplicate, with 50 μL samples placed in Hard-Shell® 96-well PCR plates with clear wells (BioRad, Hercules, CA). Plates were covered with MicroAmp® optical adhesive film (Thermo Scientific, Rockford, IL) to prevent sample evaporation. Spectra were obtained on a BioRad Real-Time C1000 Touch Thermal Cycler (BioRad, Hercules, CA). Thermal denaturations were done by increasing temperature from 25 to 95 °C at a rate of 1 °C/min, taking a plate read every minute using the FRET scan mode. Circular dichroism and ThermoFluor spectra were generated using IGOR Pro (WaveMetrics Inc., Lake Oswego, OR).

### PDC activity assay
PDC activity was measured using an assay where decarboxylation is coupled with alcohol dehydrogenase [32], and the conversion of NADH to NAD+ by the alcohol dehydrogenase was monitored for 5 min with a Varian Cary 400 (Agilent Technologies, Santa Clara, CA) temperature-controlled spectrophotometer at 25 °C. Four cuvettes, a blank and a triplicate of one sample of interest, were prepared and measured at a time. Each cuvette contained one mL of reaction mix and 20 μL of PDC with a suitable dilution in buffer D (20 mM BIS–Tris pH 6.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM TPP, 0.5 mM DTT), mixed using a 1 mL pipette (all samples in triplicate). Reaction mix contained 20 mM Na pyruvate (Fisher Scientific, Fair Lawn, New Jersey), 0.288 mM NADH (Sigma Chemical CO, St. Louis, MO) and 40 U yeast alcohol dehydrogenase (MP Biomedicals LLC, Solon, OH) in buffer D. Blank contained only the reaction

Sammond *et al. Biotechnol Biofuels  (2018) 11:189*

Page 6 of 13

mix. The Protein concentration was determined using the Bradford protein reagent with bovine serum albumin as the standard (BioRad, Hercules, CA).

### Crystallization

$PDC_{2.03}$ crystals were initially obtained with sitting drop vapor diffusion using a 96-well plate with Grid Screen Salt HT from Hampton Research (Aliso Viejo, CA). Fifty microliter of well solution was added to the reservoir, and drops were made with 0.2 μL of well solution and 0.2 μL of protein solution using a Phoenix crystallization robot (Art Robbins Instruments, Sunnyvale, CA). The crystals were grown in 0.1 M MES monohydrate pH 6.0 and 2.4 M ammonium sulfate at 20 °C. The protein solutions contained 6 mg/mL of protein in 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 0.5 mM DTT, and 0.1 mM TPP.

### Data collection and processing

The $PDC_{2.03}$ crystals were flash frozen in a nitrogen gas stream at 100 K before home source data collection using an in-house Bruker X8 MicroStar X-ray generator with Helios mirrors and Bruker Platinum 135 CCD detector. Data were indexed and processed with the Bruker Suite of programs version 2014.9 (Bruker AXS, Madison, WI).

### Structure solution and refinement

Intensities were converted into structure factors, and 5% of the reflections were flagged for Rfree calculations using programs F2MTZ, Truncate, CAD, and Unique from the CCP4 package of programs [33]. The program MOLREP [34] version 11.4.06 was used for molecular replacement using wild-type PDC (PDB code 2WVG [35]) as the search model. Refinement and manual correction was performed using REFMAC5 [36] version 5.8.0155, and Coot [37] version 0.8.6. The MOLPRO-BITY method [38] was used to analyze the Ramachandran plot, and root-mean-square deviations (RMSD) of bond lengths and angles were calculated from ideal values of Engh and Huber stereo chemical parameters 47 [39]. Wilson B-factor was calculated using CTRUNCATE version 1.15.10 [33]. The data collection and refinement statistics are shown in Additional file 1: Table S2.

### Structure analysis

Programs Coot 45, PyMOL (http://www.pymol.org) and ICM (http://www.molsoft.com) were used for comparing and analyzing structures. Figures 2a, b, 5 were created using PyMOL. The root-mean-square deviation (RMSD) between the monomeric unit of PDC and $PDC_{2.03}$ was computed using PyMol (The PyMOL MolecularGraphics System, Version 1.5.0.4 Schrödinger, LLC.)

### DIC microscopy

Pyruvate decarboxylase (PDC) protein preps were diluted in buffer to 1 mg/mL concentration. 7.5 μL of protein solution was placed between two glass coverslips separated by a 0.15 mm deep SecureSeal imaging spacer (Grace Bio-labs, Bend, OR). Heating was controlled using a Linkam FTIR600 temperature-controlled microscope stage (Linkam Scientific Instruments, UK) and heated from 24 °C to either 60 or 120 °C at a ramp rate of 1 °C/min. The optics were set up for bright field differential interference contrast imaging on a NikonE800 microscope (Nikon, Tokyo, Japan), using a 20× 0.75 NA PlanApo ELWD objective. Images were captured every 30 s over the 3 h using a SPOT RTKE CCD camera (Diagnostic Instruments, Sterling Heights, MI) as TIFF stacks. TIFF stacks were analyzed using FIJI (ImageJ).

## Results

### Computational design of a thermostable PDC

PDC self-associates to form a 260 kDa dimer of dimers, with two active sites buried at each dimer interface (Fig. 2a) [35]. A thermostable and active PDC will likely require stabilization of the monomer to resist denaturation at high temperatures and also stabilization of the interactions between the monomers to maintain the active complex. In addition, the monomeric unit of PDC is composed of three distinct structural regions (Fig. 2b). Our protein-engineering goal was to identify thermostabilizing mutations located in each structural region to address the distinct components and interactions of the PDC complex.

We used multiple design strategies to identify amino acid substitutions distributed throughout the PDC structure (Table 1). First, single substitutions were identified with in silico site saturation mutagenesis, combining mutations to decrease the number of variants selected for experimental characterization ($PDC_{1.01}$ through $PDC_{1.04}$). Next, clusters of spatially interacting residues were redesigned simultaneously to improve regions with poor atomic packing, iterating between backbone minimization and amino acid design ($PDC_{1.05}$ and $PDC_{1.06}$). Additionally, the PDC dimer displays symmetry at the interface resulting in a design constraint where any mutation made to one monomer must be mirrored in the bound partner. A symmetric interface design protocol simultaneously models the effects of an amino acid substitution at both positions at the interface ($PDC_{1.07}$ and $PDC_{1.08}$) [25]. Lastly, engineering a protein surface to have a high net charge can increase solubility and enhance thermal stability [24, 26]. Designing the PDC complex on the hypothesis that supercharging the surface can stabilize the interactions between domains, we identified substitutions to impart a high net charge while

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 7 of 13

**Table 1 Designing thermostabilizing mutations for the PDC homotetrameric complex**

| PDC variant | Mutations | Activity at 25 °C |
|---|---|---|
| 1.01 | G224A, V374I, G540A | Yes |
| 1.02 | F55W, G224A, V374I | Yes |
| 1.03 | F55W, G224A, V374I, G540A | Yes |
| 1.04 | G224A, G540A | Yes |
| 1.05 | G6A, I97V, I166F, A171F | ND[a] |
| 1.06 | H21I, M42F, Q44V, Y46F | Yes |
| 1.07 | A77V, L78F, S79M, A80M, A83V, I84V, G85V, G86M, M127V, A128M, Y163F | ND[a] |
| 1.08 | A77V, L78F, S79M, A80M, D82V, A83V, I84V, G85V, G86M, M127V, A128M | ND[a] |
| 1.09 | S2R, N41K, A207K, Q333K, A357K, T379R, N402K, V524K, A527K, A557K | ND[a] |
| 1.10 | S2D, L38D, A189K, A207E, Q333E, A357D, A376D, A519E, A527D, K553E | Yes |
| 2.01 | G109A, G224A, V374I, G491A, G515A, G516A, G540A | Yes |
| 2.02 | S2D, L38D, A207E, G224A, Q333E, A357D, V374I, A376D, A519E, A527D, G540A, K553E | Yes |
| 2.03 | S2D, L38D, G109A, A207E, G224A, Q333E, A357D, V374I, A376D, G491A, G515A, G516A, A519E, A527D, K553E, G540A | Yes |

PDC variants were expressed in *E. coli* and evaluated for activity

[a] ND. The activity for these designs was not determined as they did not express on the first attempt or expressed in such a small amount that not enough protein was obtained to evaluate the activity or thermal stability

modeling the energetic consequences of the substitutions (net positive charge for $PDC_{1.09}$ and net negative charge for $PDC_{1.10}$) [40].

The mutations display features associated with increased thermal stability, including increased buried hydrophobic surface area [22], enhanced quality of side-chain packing [23], increased surface charge of the tetrameric complex [24], decreased conformational flexibility and improved packing at the symmetric dimer interface [25]. A summary of the designs is shown in Table 1. Our approach utilizes an iterative workflow, where the most promising designs are selected after experimental characterization, then combined, and/or further enhanced with additional mutations (Fig. 2c). The iterative computational design approach is key to identifying thermostabilizing mutations that work together to stabilize any structurally weak regions.
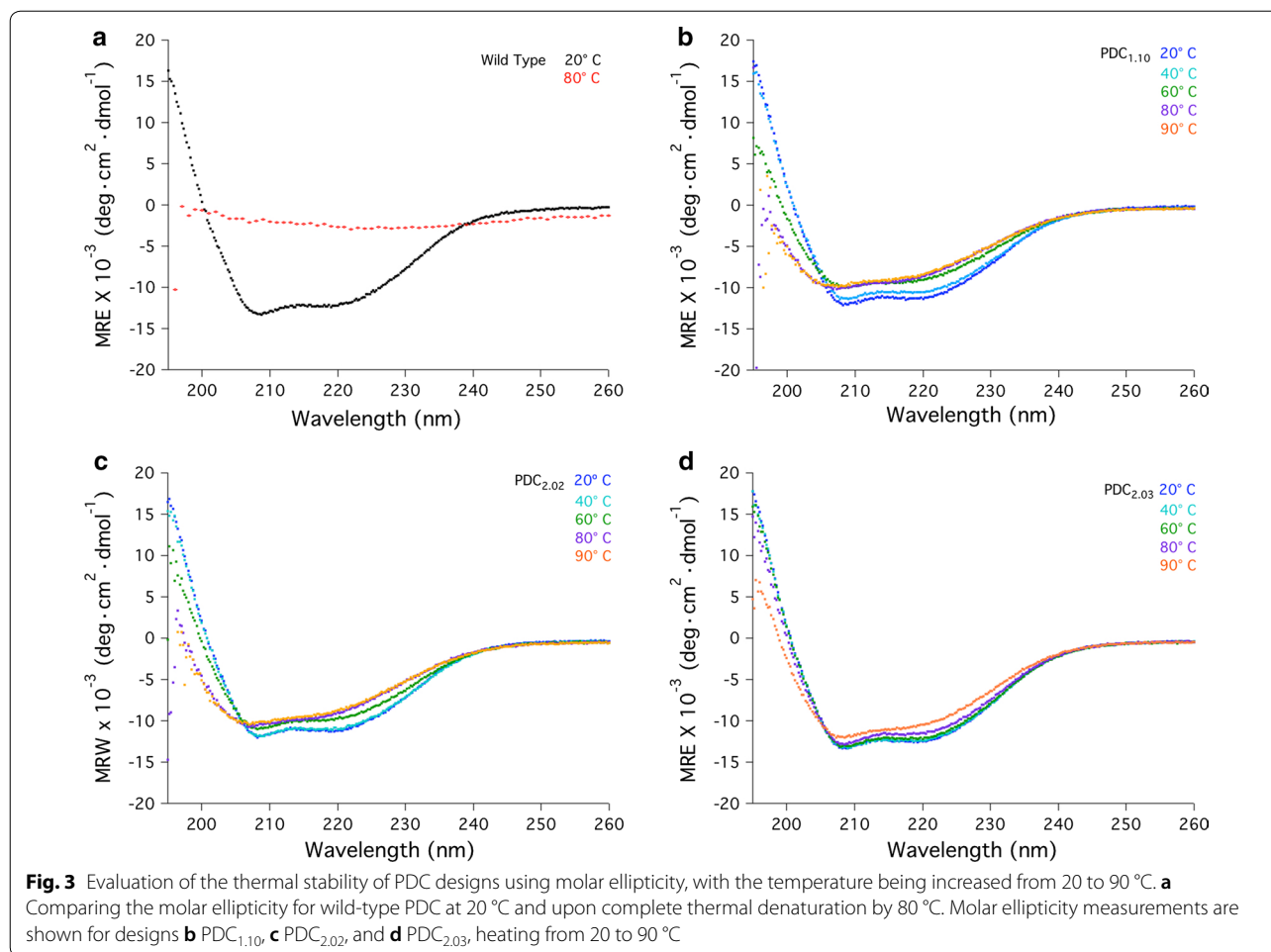
### Identification of PDC designs with enhanced thermostability

To evaluate whether the PDC designs resulted in an enhancement in thermal stability, genes encoding wild-type PDC and variants listed in Table 1 were expressed in *Escherichia coli*. Enzymatic activity for all PDC enzymes was determined, with all successfully expressed PDC enzymes showing similar levels of activity. Pyruvate decarboxylation is measured by coupling PDC to an NADH-dependent alcohol dehydrogenase (ADH). The oxidation of NADH by ADH upon the conversion of acetaldehyde to ethanol results in a colorimetric change. NADH is not stable at elevated temperatures,

thus activity was determined at 25 °C [41]. Any designs that did not produce a detectable level of soluble protein in the first attempt were discarded, and activity for these designs was labeled "ND", or not determined, in Table 1.

Differential scanning calorimetry (DSC) was initially used to evaluate the thermal stability of the PDC variants. DSC scans displayed multiple peaks, possibly due to the multimeric attribute of PDC, challenging the comparison of some PDC variants (Additional file 1: Figure S2). DSC results served to eliminate a few underperforming designs ($PDC_{1.02}$, $PDC_{1.04}$, $PDC_{1.03}$, $PDC_{1.07}$), but we instead relied on circular dichroism (CD) to characterize the remaining variants. PDC was shown to have irreversible unfolding, thus the melting temperatures reported here for PDC variants indicate relative thermal stabilities.

Two PDC variants were selected for additional characterization from the first round of design. Evaluating ellipticity at 222 nm $PDC_{1.01}$ displayed an increased $T_m$ of 4 °C relative to wild type, with melting temperatures of 66 and 62 °C, respectively (Additional file 1: Figure S1). $PDC_{1.10}$, interestingly, did not undergo a large change in ellipticity at 222 nm, indicating the protein did not fully denature or precipitate by 90 °C. We therefore assessed the thermal stability of $PDC_{1.10}$ by measuring molar ellipticity from 195 to 260 nm to evaluate changes in a broader range of secondary structural elements (Fig. 3b). For comparison, wild-type PDC was evaluated from 195 to 260 nm at 20 °C and also at 80 °C, the temperature by which PDC is fully denatured (Fig. 3a). Unlike the wild-type PDC, $PDC_{1.10}$ resists thermal denaturation up to

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 8 of 13



**Fig. 3** Evaluation of the thermal stability of PDC designs using molar ellipticity, with the temperature being increased from 20 to 90 °C. **a** Comparing the molar ellipticity for wild-type PDC at 20 °C and upon complete thermal denaturation by 80 °C. Molar ellipticity measurements are shown for designs **b** $PDC_{1.10}$, **c** $PDC_{2.02}$, and **d** $PDC_{2.03}$, heating from 20 to 90 °C

90 °C. However, loss of some secondary structure indicated by a loss of molar ellipticity is observed at 60 °C.

Three tiers of design were attempted for the second round of PDC engineering. First, $PDC_{1.01}$ was further improved by the addition of four substitutions, generating $PDC_{2.01}$. Second, $PDC_{1.01}$ and $PDC_{1.10}$ were combined to evaluate the additivity of these two successful designs, generating $PDC_{2.02}$. In the third design, the same four substitutions were combined with $PDC_{1.01}$ and $PDC_{1.10}$, generating $PDC_{2.03}$. The third approach was the most thorough test of both the accuracy of our computational predictions and the combinability of the designs.
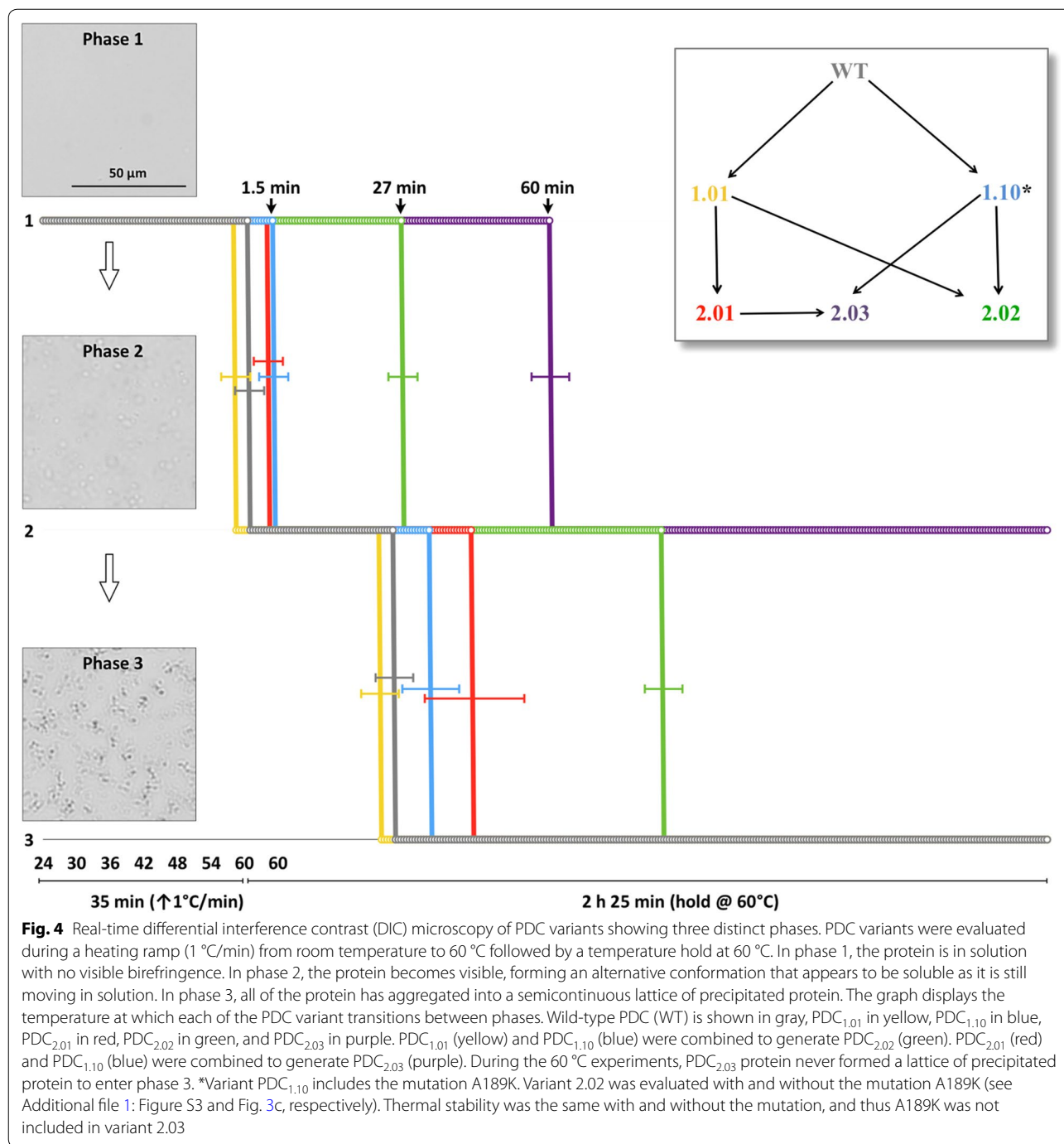
The thermal stability of $PDC_{2.01}$, containing the four substitutions and the parent design $PDC_{1.01}$, was evaluated using the high-throughput ThermoFluor assay (Additional file 1: Figure S1B, A, respectively). $PDC_{2.01}$ exhibited a melting temperature nearly identical to $PDC_{1.01}$, displaying no improvement in thermal stability. Nearly all of the pairwise combinations displayed a $T_m$ within one or two degrees of the parent $PDC_{1.01}$. The most significant increase in $T_m$ was seen with the

addition of G491A and G516A, resulting in an increase of 4 °C.

$PDC_{2.02}$ and $PDC_{2.03}$ were evaluated by measuring molar ellipticity from 195 to 260 nm, as done with the parent design, $PDC_{1.10}$ (Fig. 3c). $PDC_{2.02}$ resists thermal denaturation, as seen with $PDC_{1.01}$. $PDC_{2.02}$ also loses some secondary structure at 60 °C, as seen with $PDC_{1.01}$. Thus, while $PDC_{2.02}$ does not fully denature up to 90 °C, it also does not display an apparent increase in thermal stability compared to $PDC_{1.01}$ based on changes in molar ellipticity.

A significant increase in thermal stability is observed with $PDC_{2.03}$, which combines $PDC_{1.01}$, $PDC_{1.10}$ and the additional four substitutions. $PDC_{2.03}$ does not show loss of secondary structure until 90 °C, well above the target growth temperature of 60 °C, at which point small changes in molar ellipticity are observed (Fig. 3d). Importantly, $PDC_{2.03}$ differs from $PDC_{2.02}$ by only four substitutions. The combined ThermoFluor and CD results indicate, therefore, the same four mutations have no apparent thermostabilizing effect in the context of
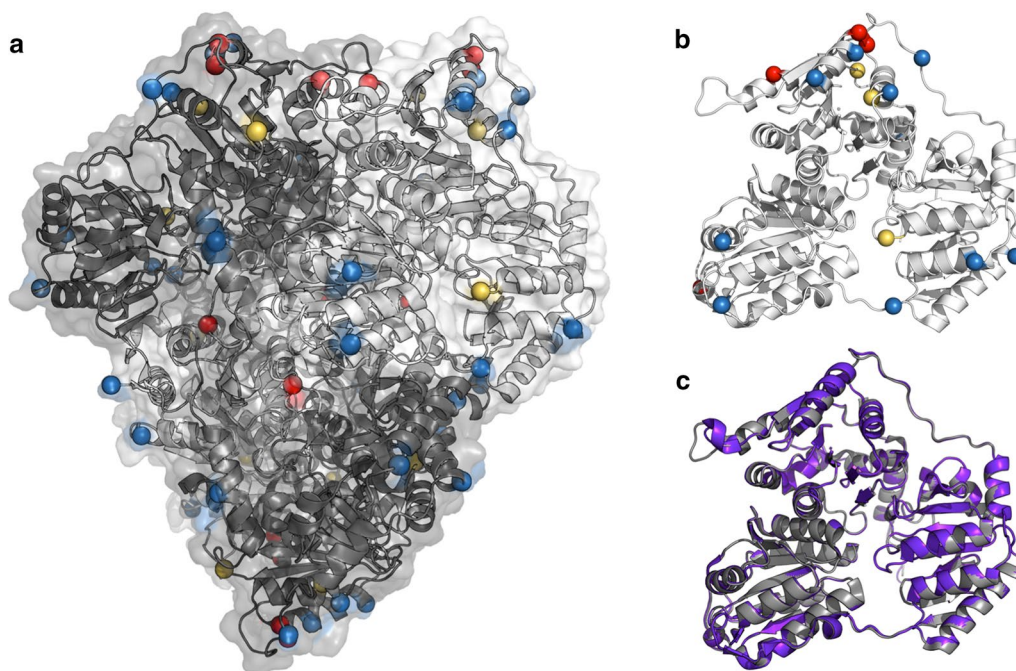
Sammond *et al. Biotechnol Biofuels*  (2018) 11:189

Page 9 of 13



**Fig. 4** Real-time differential interference contrast (DIC) microscopy of PDC variants showing three distinct phases. PDC variants were evaluated during a heating ramp (1 °C/min) from room temperature to 60 °C followed by a temperature hold at 60 °C. In phase 1, the protein is in solution with no visible birefringence. In phase 2, the protein becomes visible, forming an alternative conformation that appears to be soluble as it is still moving in solution. In phase 3, all of the protein has aggregated into a semicontinuous lattice of precipitated protein. The graph displays the temperature at which each of the PDC variant transitions between phases. Wild-type PDC (WT) is shown in gray, $PDC_{1.01}$ in yellow, $PDC_{1.10}$ in blue, $PDC_{2.01}$ in red, $PDC_{2.02}$ in green, and $PDC_{2.03}$ in purple. $PDC_{1.01}$ (yellow) and $PDC_{1.10}$ (blue) were combined to generate $PDC_{2.02}$ (green). $PDC_{2.01}$ (red) and $PDC_{1.10}$ (blue) were combined to generate $PDC_{2.03}$ (purple). During the 60 °C experiments, $PDC_{2.03}$ protein never formed a lattice of precipitated protein to enter phase 3. *Variant $PDC_{1.10}$ includes the mutation A189K. Variant 2.02 was evaluated with and without the mutation A189K (see Additional file 1: Figure S3 and Fig. 3c, respectively). Thermal stability was the same with and without the mutation, and thus A189K was not included in variant 2.03

$PDC_{2.01}$ while significantly increasing thermal stability in the context of $PDC_{2.03}$.

### Measuring the structural endurance of PDC variants

The aim of this work was to develop an iterative computational protein design protocol to rationally engineer a mesophilic enzyme, PDC from *Z. mobilis*, to function like a thermophilic enzyme. The maximum growth temperature for *Z. mobilis* is 42 °C [20], well below the optimum growth temperature of 60 °C for *C. thermocellum* [19]. While our target temperature of 60 °C is below the $T_m$ of wild-type PDC (Additional file 1: Figure S1), where half the protein has been thermally inactivated, changes in ellipticity indicate PDC is at or approaching a structurally unstable temperature.

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 10 of 13



**Fig. 5** Mutations in PDC$_{2.03}$ imparting an increase in thermal stability and endurance without changing the backbone. **a** The Cα atoms at mutated positions are highlighted as spheres. The mutations that comprise PDC$_{2.03}$ are located throughout the PDC enzyme. Mutations from PDC$_{1.01}$ are shown in yellow, PDC$_{1.10}$ in blue, and the mutations added to generate PDC$_{2.03}$ are shown in red. Chains A and E are shown in dark gray, and chains B and F are shown in light gray, respectively. **b** The mutations in PDC$_{2.03}$ are found in each of the three structural regions of PDC monomer. **c** The X-ray crystal structure of PDC$_{2.03}$ (PDB code 5TMA) is shown in purple aligned with the wild-type PDC in gray (PDB code 2WVA). The Cα backbone root-mean-square deviation (RMSD) for the PDC monomeric unit is 0.19 Å

Seeking an approach capable of making time-resolved measurements of PDC variants at the target temperature of 60 °C, we evaluated variants using differential interference contrast (DIC) microscopy during a heat and hold experiment. This approach is similar to optical density and static light scattering experiments, where blockage or scattering of light indicates changes in the size of protein species, generally due to aggregation [42, 43]. In the approach used here to evaluate the structural endurance of PDC variants, we combined measured changes in birefringence with precise temperature control to evaluate protein aggregation upon thermal denaturation.

The proteins were heated from 25 °C at a ramp rate of 1 °C/min to a final temperature of 60 °C and then held at 60 °C for a total time of 3 h. Using DIC analysis we observe three distinct phases of protein behavior. In the starting condition (phase 1), the protein is in solution and invisible to DIC imaging, producing no birefringent margins. As the protein is heated, it undergoes a structural change, forming an alternate but stable conformation, (phase 2) and becomes visible. This alternate conformation appears to be soluble, as the phase 2 structures are observed moving in solution. Finally, all of the protein appears to aggregate into a semicontinuous lattice

of precipitated protein (phase 3). Representative micrographs of these three phases are shown at the left side of Fig. 4. These observed structural phases fit the model proposed by Tomazic and Klibanov, where soluble proteins (phase 1) can undergo a partial unfolding into kinetically stable structures (phase 2) followed by the irreversible formation of aggregates (phase 3) [44, 45].

Wild-type PDC entered phase 2 immediately upon reaching 60 °C, remaining in phase 2 for approximately 25 min before reaching phase 3. PDC$_{1.01}$, which exhibited a four degree increase in $T_m$ based on ellipticity measurements, showed no improvement over wild-type PDC when held at 60 °C. PDC$_{1.10}$, which resisted denaturation and precipitation in CD experiments to at least 90 °C, similarly showed little to no improvement over wild-type PDC when held at 60 °C. PDC$_{1.10}$ results suggest that the loss of molar ellipticity observed at 60 °C in CD experiments indicates a structural instability and does not represent a stable conformation.

All second round designs (PDC$_{2.01}$, PDC$_{2.02}$ and PDC$_{2.03}$) outperform the first round designs, (PDC$_{1.01}$ and PDC$_{1.10}$). PDC$_{2.01}$ enters phase 2 within minutes after reaching 60 °C but remains in phase 2 for approximately 35 additional minutes before entering phase 3. PDC$_{2.02}$

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 11 of 13

and $PDC_{1.10}$, which displayed nearly identical changes in molar ellipticity, have markedly different behavior when monitored at 60 °C. $PDC_{2.02}$ remains in phase 1 for 27 min at 60 °C before entering phase 2. $PDC_{2.02}$ then lasts another 90 min before entering phase 3, outperforming the parent designs $PDC_{1.01}$ and $PDC_{1.10}$.

$PDC_{2.03}$, which displayed no loss of molar ellipticity until 90 °C when monitored by CD, lasted 60 min at 60 °C before entering phase 2, and never entered phase 3 for the duration of the experiment. Thus, the four additional mutations (G491A, G515A, G516A, and G540A) displayed manifestly different stabilizing effects depending on whether they were evaluated in the context of design $PDC_{2.01}$ or $PDC_{2.03}$.

### Structural evaluation of $PDC_{2.03}$

$PDC_{2.03}$ contains sixteen mutations in the monomeric unit and sixty-four mutations in the homotetrameric complex (Fig. 5a). The mutations are found throughout the PDC complex as well as in each of the three structural regions of the monomeric unit (Fig. 5b). The wild-type PDC has a $T_m$ of 62 °C, while $PDC_{2.03}$ resists thermal denaturation to at least 90 °C. We used several approaches to identify potential thermostabilizing mutations, but the mutations included in $PDC_{2.03}$ were all generated using a fixed-backbone approach, where the backbone coordinates were held fixed to the dihedral angles observed in the crystal structure (PDB code 2WVA).

To determine whether $PDC_{2.03}$ retained the backbone conformation from the design runs, we elucidated the X-ray crystal structure of the PDC variant (PDB code 5TMA). The structure of $PDC_{2.03}$ was refined to a resolution of 1.67 Å with R and $R_{free}$ of 0.170 and 0.203, respectively. There are two molecules in the asymmetric unit in complex with TPP and magnesium. Each monomer has three domains with an open α/β topology [46] with several ethylene glycol and sulfate molecules on the surface.

The backbone root-mean-square deviation (RMSD) between chain A of the wild-type PDB and $PDC_{2.03}$ is 0.19 Å (Fig. 5c). The largest conformational difference, located in the upper left of Fig. 5c, is located at a crystal contact point and distant from both the active sites and any of the 16 mutations. Thus, the mutations impart significant enhancement in thermal stability without unintended structural alterations.

### Conclusions

In this study, we used iterative computational protein design to rationally engineer the mesophilic enzyme, pyruvate decarboxylase, to exhibit enhanced thermal endurance at the growth temperature of a thermophilic organism. The target temperature of 60 °C is below the

$T_m$ of the wild-type PDC, at 62 °C, but is in a thermally unstable region as determined by changes in ellipticity (Additional file 1: Figure S1A). We identified a series of thermostabilizing mutations, increasing the $T_m$ of PDC by 4 °C in one case ($PDC_{1.01}$) and generating a variant that resists thermal denaturation to at least 90 °C ($PDC_{1.10}$) in another. While increases in structural endurance at elevated temperature are not achieved by any single design, the combination of designs resulted in a PDC variant with a large increase in thermal endurance at 60 °C.

We find, however, that increases in $T_m$ do not directly correlate with enhanced thermal endurance. Designs $PDC_{1.10}$ and $PDC_{2.02}$, for example, exhibit no observable difference in thermal stability based on measured changes in molar ellipticity, but display strikingly different thermal endurance; $PDC_{2.02}$ survives approximately twice as long at 60 °C compared to $PDC_{1.10}$. Similarly, $PDC_{1.10}$ shows significant thermal stabilization compared to wild-type PDC, resisting thermal denaturation to at least 90 °C, yet survives only minutes longer at 60 °C.

The explanation for these observed incongruities may have to do with cooperativity between the mutations. Fersht introduced the concept of a double-mutant cycle to determine whether two residues are energetically coupled, either through a direct or indirect interaction [47]. Based on this analysis, two residues are coupled if the change in free energy for the double mutant is not equal to the sum of the changes in free energy for each mutation individually. This concept has been expanded to include changes in free energy associated with structural or functional properties upon a double mutation [48]. Here we have generated PDC variants which all included multiple mutations. However, the energetic nonadditivity expected between mutations of coupled residues is clearly observed when comparing the thermal endurance of PDC variants (Fig. 4). $PDC_{2.02}$, which resulted from the combination of $PDC_{1.01}$ and $PDC_{1.10}$, displays an increase in structural endurance that far exceeds what would be achieved if the mutations were additive. Similarly, $PDC_{2.03}$ survives much longer at 60 °C than would be expected if the residues in the parent designs, $PDC_{1.10}$ and $PDC_{2.01}$, were not energetically coupled. None of the sixteen mutations that comprise $PDC_{2.03}$ are directly interacting and are generally structurally removed from one another. Thus, the observed energetic coupling is likely due to structural stabilizations in interacting regions of the PDC.

Further, the glycine-to-alanine mutations (G109A, G491A, G515A, and G516A) included in $PDC_{2.01}$ and $PDC_{2.03}$ display context-dependent behavior. These four mutations do not increase thermal stability when added to $PDC_{1.01}$ to generate $PDC_{2.01}$, and similarly do not result in a significant increase in thermal endurance.

Sammond *et al. Biotechnol Biofuels* (2018) 11:189

Page 12 of 13

However, these same four mutations increase both thermal stability and thermal endurance when added to $PDC_{2.02}$ to generate $PDC_{2.03}$. On the one hand, these results suggest that the high-throughput thermal stability screens may fail to identify thermal stabilizing mutations if they are not evaluated in the favorable context of a more thermostable protein. However, the context dependence of these mutations also demonstrates the importance of cooperativity, such that even mutations that by themselves yield only the smallest gains in stability can ultimately make a significant contribution when part of a larger design.

These results suggest stabilizing all regions of a protein at or above the target temperature is critical to achieve the desired enhancement in thermal endurance. $PDC_{1.10}$ and $PDC_{2.02}$ both exhibit some loss of molar ellipticity at 60 °C, the temperature at which we evaluate thermal endurance by DIC. The improvement in endurance for $PDC_{2.02}$ compared to $PDC_{1.10}$, which is not observed as an improvement in stability based on the thermal melt, may therefore be the result of structural stabilization near 60 °C. In addition, $PDC_{2.03}$ exhibits the highest degree of thermal endurance at 60 °C and also shows no loss of molar ellipticity until 90 °C. Both the context dependence of the four glycine-to-alanine mutations as well as the pronounced cooperativity observed in the combined mutations of $PDC_{2.03}$ may be a reflection of the thermal stabilization of structurally cooperative regions of the multimeric PDC.

Much of protein design and structure prediction is predicated on the hypothesis that the native structure for a given protein is the conformation with the lowest Gibbs free energy. However, there is significant evidence that for many proteins the active and native conformation may not be the lowest energy conformation, but instead kinetic energy barriers drive the folding trajectories to achieve the observed native conformations [18]. For proteins that fall into such kinetically driven folding trajectories, perturbing physiological conditions such as temperature or pH can alter protein folding and subsequently alter the sampled conformations. Thus, a critical next step and test of our ability to computationally design thermostable enzymes is to evaluate protein folding under elevated temperatures.

## Additional file

**Additional file 1.** Additional figures and tables.

## Authors' contributions
DWS performed in silico modeling, protein expression, and ThermoFluor stability assays, assisted in protein purification, wrote the manuscript. NK performed circular dichroism experiments. BSD performed differential interference contrast microscopy. MA performed protein purification, activity assays, and X-ray crystallography. VVL performed X-ray crystallography. DC assisted in protein expression. NSS assisted in protein purification and ThermoFluor stability assays. HY assisted in circular dichroism experiments, and analysis. AM assisted in activity assays. MEH assisted in protein purification and activity assays. AMG assisted in experimental design and development of the project idea and protein expression. YJB assisted in development of the project idea and protein purification. All authors read and approved the final manuscript.

## Author details
[1] Biosciences Center, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA. [2] Department of Chemistry & Biochemistry and the BioFrontiers Institute, University of Colorado, Boulder, CO 80309, USA. [3] Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

## Competing interests
The authors declare that they have no competing interests.

## Availability of data and materials
The X-ray crystal structure of $PDC_{2.03}$ generated during the current study is available in the Protein Data Bank [PDB] repository with the PDB ID 5TMA, [https://www.rcsb.org/structure/5tma].

## Consent for publication
Not applicable.

## Ethics approval and consent to participate
Not applicable.

## Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References
1. Armstrong RC, Wolfram C, de Jong KP, Gross R, Lewis NS, Boardman B, Ragauskas AJ, Ehrhardt-Martinez K, Crabtree G, Ramana MV. The frontiers of energy. Nat Energy. 2016. https://doi.org/10.1038/nenergy.2015.20.
2. McCann MC, Penning BW, Dugard CK, Carpita NC. Tailoring plant cell wall composition and architecture for conversion to liquid hydrocarbon Biofuels. New York: Elsevier; 2015.
3. Lynd LR, van Zyl WH, McBride JE, Laser M. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol. 2005;16(5):577–83.
4. Basen M, Rhaesa AM, Kataeva I, Prybol CJ, Scott IM, Poole FL, Adams MW. Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. Bioresour Technol. 2014;152:384–92.
5. Brunecky R, Alahuhta M, Xu Q, Donohoe BS, Crowley MF, Kataeva IA, Yang SJ, Resch MG, Adams MW, Lunin VV, et al. Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA. Science. 2013;342(6165):1513–6.
6. Xu Q, Resch MG, Podkaminer K, Yang S, Baker JO, Donohoe BS, Wilson C, Klingeman DM, Olson DG, Decker SR, et al. Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. Sci Adv. 2016;2(2):e1501254.

Sammond *et al. Biotechnol Biofuels  (2018) 11:189*

Page 13 of 13

7. Biswas R, Prabhu S, Lynd LR, Guss AM. Increase in ethanol yield via elimination of lactate production in an ethanol-tolerant mutant of *Clostridium thermocellum*. PLoS ONE. 2014;9(2):e86389.

8. Chung D, Cha M, Guss AM, Westpheling J. Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. Proc Natl Acad Sci USA. 2014;111(24):8931–6.

9. Chung D, Cha M, Snyder EN, Elkins JG, Guss AM, Westpheling J. Cellulosic ethanol production via consolidated bioprocessing at 75 degrees C by engineered *Caldicellulosiruptor bescii*. Biotechnol Biofuels. 2015;8:163.

10. Olson DG, Sparling R, Lynd LR. Ethanol production by engineered thermophiles. Curr Opin Biotechnol. 2015;33:130–41.

11. Guedon E, Desvaux M, Petitdemange H. Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. Appl Environ Microbiol. 2002;68(1):53–8.

12. Olson DG, McBride JE, Shaw AJ, Lynd LR. Recent progress in consolidated bioprocessing. Curr Opin Biotechnol. 2012;23(3):396–405.

13. Eram MS, Ma K. Decarboxylation of pyruvate to acetaldehyde for ethanol production by hyperthermophiles. Biomolecules. 2013;3(3):578–96.

14. Goldenzweig A, Goldsmith M, Hill SE, Gertman O, Laurino P, Ashani Y, Dym O, Unger T, Albeck S, Prilusky J, et al. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. Mol Cell. 2016;63(2):337–46.

15. Murphy GS, Mills JL, Miley MJ, Machius M, Szyperski T, Kuhlman B. Increasing sequence diversity with flexible backbone protein design: the complete redesign of a protein hydrophobic core. Structure. 2012;20(6):1086–96.

16. de Souza AR, de Araujo GC, Zanphorlin LM, Ruller R, Franco FC, Torres FA, Mertens JA, Bowman MJ, Gomes E, Da Silva R. Engineering increased thermostability in the GH-10 endo-1,4-beta-xylanase from *Thermoascus aurantiacus* CBMAI 756. Int J Biol Macromol. 2016;93(Pt A):20–6.

17. Lehmann M, Wyss M. Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. Curr Opin Biotechnol. 2001;12(4):371–5.

18. Baker D, Agard DA. Kinetics versus thermodynamics in protein folding. Biochemistry. 1994;33(24):7505–9.

19. Freier D, Mothershed CP, Wiegel J. Characterization of *Clostridium thermocellum* JW20. Appl Environ Microbiol. 1988;54(1):204–11.

20. Benschoter AS, Ingram LO. Thermal tolerance of *Zymomonas mobilis*: temperature-induced changes in membrane composition. Appl Environ Microbiol. 1986;51(6):1278–84.

21. Kaufmann KW, Lemmon GH, Deluca SL, Sheehan JH, Meiler J. Practically useful: what the Rosetta protein modeling suite can do for you. Biochemistry. 2010;49(14):2987–98.

22. Sammond DW, Eletr ZM, Purbeck C, Kimple RJ, Siderovski DP, Kuhlman B. Structure-based protocol for identifying mutations that enhance protein–protein binding affinities. J Mol Biol. 2007;371(5):1392–404.

23. Sammond DW, Kastelowitz N, Himmel ME, Yin H, Crowley MF, Bomble YJ. Comparing residue clusters from thermophilic and mesophilic enzymes reveals adaptive mechanisms. PLoS ONE. 2016;11(1):e0145848.

24. Lawrence MS, Phillips KJ, Liu DR. Supercharging proteins can impart unusual resilience. J Am Chem Soc. 2007;129(33):10110–2.

25. DiMaio F, Leaver-Fay A, Bradley P, Baker D, Andre I. Modeling symmetric macromolecular structures in Rosetta3. PLoS ONE. 2011;6(6):e20450.

26. Miklos AE, Kluwe C, Der BS, Pai S, Sircar A, Hughes RA, Berrondo M, Xu J, Codrea V, Buckley PE, et al. Structure-based design of supercharged, highly thermoresistant antibodies. Chem Biol. 2012;19(4):449–55.

27. Marino SM, Gladyshev VN. Analysis and functional prediction of reactive cysteine residues. J Biol Chem. 2012;287(7):4419–25.

28. Edwards PM. Origin 7.0: scientific graphing and data analysis software. J Chem Inf Model. 2002;42(5):1270–1.

29. Lavinder JJ, Hari SB, Sullivan BJ, Magliery TJ. High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. J Am Chem Soc. 2009;131(11):3794–5.

30. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E, Springer BA, Lane P, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen. 2001;6(6):429–40.

31. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 1989;182(2):319–26.

32. Gounaris AD, Turkenkopf I, Buckwald S, Young A. Pyruvate decarboxylase. I. Protein dissociation into subunits under conditions in which thiamine pyrophosphate is released. J Biol Chem. 1971;246(5):1302–9.

33. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):235–42.

34. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 1):22–5.

35. Pei XY, Erixon KM, Luisi BF, Leeper FJ. Structural insights into the prereaction state of pyruvate decarboxylase from *Zymomonas mobilis*. Biochemistry. 2010;49(8):1727–36.

36. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):355–67.

37. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486–501.

38. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 1):12–21.

39. Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein-structure refinement. Acta Crystallographica Sect A. 1991;47:392–400.

40. Der BS, Kluwe C, Miklos AE, Jacak R, Lyskov S, Gray JJ, Georgiou G, Ellington AD, Kuhlman B. Alternative computational protocols for supercharging protein surfaces for reversible unfolding and retention of stability. PLoS ONE. 2013;8(5):e64363.

41. Wu JT, Wu LH, Knight JA. Stability of NADPH: effect of various factors on the kinetics of degradation. Clin Chem. 1986;32(2):314–9.

42. den Engelsman J, Garidel P, Smulders R, Koll H, Smith B, Bassarab S, Seidl A, Hainzl O, Jiskoot W. Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. Pharm Res. 2011;28(4):920–33.

43. Samra HS, He F. Advancements in high throughput biophysical technologies: applications for characterization and screening during early formulation development of monoclonal antibodies. Mol Pharm. 2012;9(4):696–707.

44. Tomazic SJ, Klibanov AM. Mechanisms of irreversible thermal inactivation of Bacillus alpha-amylases. J Biol Chem. 1988;263(7):3086–91.

45. Vieille C, Zeikus GJ. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev. 2001;65(1):1–43.

46. Dobritzsch D, Konig S, Schneider G, Lu G. High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases. J Biol Chem. 1998;273(32):20196–204.

47. Carter PJ, Winter G, Wilkinson AJ, Fersht AR. The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*). Cell. 1984;38(3):835–40.

48. Horovitz A. Double-mutant cycles: a powerful tool for analyzing protein structure and function. Fold Des. 1996;1(6):R121–6.