

Modeling of pyruvate decarboxylases from ethanol producing bacteria

Anjala Shrestha¹, Srisuda Dhamwichukorn^{1,2}, Ekachai Jenwitheesuk^{3,4*}

¹Joint Graduate School of Energy and Environment, King Mongkut's University of Technology Thonburi, Prachautid Road, Toongkru, Bangkok 10140, Thailand; ²Pacific Time Co., Ltd. 77/123 Sinn Sathorn Tower, 29th Floor. Krungthoburi Road. Kongton sai, Klongsan, Bangkok 10600, Thailand; ³Bioinformatics program, King Mongkut's University of Technology Thonburi, Prachautid Road, Toongkru, Bangkok 10140, Thailand; ⁴National Center for Genetic Engineering and Biotechnology, 113 Thailand Science Park, Phahonyothin Road, Klongluang, Pathumthani 12120, Thailand. Phone: +662 654-6700 Fax: +662 564-6701 Email: ekachai@biotec.or.th; * Corresponding author

Received May 25, 2009; Revised December 19, 2009; Accepted February 10, 2010; Published February 28, 2010

Abstract:

Pyruvate decarboxylase (PDC) is a key enzyme in homoethanol fermentation process, which decarboxylates 2-keto acid pyruvate into acetaldehyde and carbon dioxide. PDC enzymes from potential ethanol-producing bacteria such as *Zymomonas mobilis*, *Zymobacter palmae* and *Sarcina ventriculi* have different K_m and K_{cat} values for the substrate pyruvate at their respective optimum pH. In this study, the putative three-dimensional structures of PDC dimer of *Z. palmae* PDC and *S. ventriculi* PDC were generated based on the X-ray crystal structures of *Z. mobilis* PDC, *Saccharomyces cerevisiae* PDC form-A and *Enterobacter cloacae* indolepyruvate decarboxylase in order to compare the quaternary structures of these bacterial PDCs with respect to enzyme-substrate interactions, and subunit-subunit interfaces that might be related to the different biochemical characteristics. The PROCHECK scores for both models were within recommended intervals. The generated models are similar to the X-ray crystal structure of *Z. mobilis* PDC in terms of binding modes of the cofactor, the position of Mg^{2+} , and the amino acids that form the active sites. However, subunit-subunit interface analysis showed lower H-bonding in both models compared with X-ray crystal structure of *Z. mobilis* PDC, suggesting a smaller interface area and the possibility of conformational change upon substrate binding in both models. Both models have predicted lower affinity towards branched and aromatic 2-keto acids, which correlated with the molecular volumes of the ligands. The models shed valuable information necessary for further improvement of PDC enzymes for industrial production of ethanol and other products.

Keywords: bacterial pyruvate decarboxylase; homology modelling; molecular docking, 2- keto acids.

Background:

The pyruvate decarboxylase enzyme (PDC, EC 4.1.1.1) can be found in plants and fungi, but is very rare in prokaryotes and absent in animals [1]. This enzyme requires thiamine diphosphate (ThDP) and Mg^{2+} cofactors for its function in catalyzing the non-oxidative decarboxylation of 2-keto-acids, of which its preferred substrate is pyruvate. Catalytically active ThDP, called ylide-ThDP is formed by deprotonation of the C2 atom of the thiazolium ring. This ylide attacks the pyruvate C2 atom forming lactyl-ThDP, followed by release of carbon dioxide and formation of a resonance-stabilized enamine. Proton addition to the enamine gives hydroxyethyl-ThDP, which releases acetaldehyde and regenerates ylide-ThDP [2]. The acetaldehyde produced by PDC pyruvate decarboxylation is then converted to ethanol by alcohol dehydrogenase (ADH, EC 1.1.1.1). In addition, PDC decarboxylates 2-keto acid [3], catalyzes acyloin (benzoin)-type condensation reaction of the aldehyde to form chiral 2-hydroxyketones [4], and has pharmaceutical application in synthesis of the L-ephedrine intermediate 2-hydroxyketone (R)-phenylacetylcarbinol ((R)-PAC).

Given the importance of the PDC enzyme in ethanol and chiral 2-hydroxyketone production, homoethanol-producing microorganisms are the focus of research in different aspects. Among them, Gram-negative *Zymomonas mobilis*, *Zymobacter palmae* and Gram-positive *Sarcina ventriculi* possess PDC enzymes, ZmPDC, ZpPDC and SvPDC respectively. The only available PDC crystal structure of a homoethanol-producing bacterial species is from *Z. mobilis* ATCC29291 (resolved at 1.9 Å, PDB ID: 1zpd [5]). The structure shows extensive interactions in the subunit-subunit interface, which are significantly different from the other PDCs, and might be responsible for the variation in biochemical properties between these species [6, 7]. Therefore, in this study, the three-dimensional structures of the ZpPDC and SvPDC were generated using the homology modelling technique in order to compare the quaternary structures of these bacterial PDCs with respect to the enzyme-substrate interactions, and subunit-subunit interfaces that might be related to the different biochemical characteristics.

Methodology:

Modelling of PDC of *Z. palmae* and *S. ventriculi*
The amino acid sequences of ZpPDC and SvPDC (GenBank accession numbers: AAM49566 and AAL18557, respectively)

along with the X-ray crystal structures of ZmPDC (PDB ID: 1zpd), indolepyruvate decarboxylase of *Enterobacter cloacae* (EclPDC; PDB ID: 1ovm) [8], and *Saccharomyces cerevisiae* (ScPDC form-A; PDB ID: 1pvd) [9] were used as inputs for homology modelling by Modeller version 9v3 [10]. ZpPDC shares 62%, 31% and 32% sequence identity with ZmPDC, ScPDC and EclPDC, respectively; while SvPDC shares 32% 38% and 40% sequence identity with ZmPDC, ScPDC and EclPDC, respectively (Figure 1).

Different models were generated based on the sequence alignment data and the X-ray crystallography templates. Two scoring functions, DOPE (discrete optimized protein energy) [11] score and RAPDF (residue-specific all-atoms conditional probability discriminatory function) [12] score were used to select the final models. The best model, with the lowest DOPE and RAPDF scores, was subjected to energy minimization by the NAMD program without water molecules using conjugate gradient [13]. The quality of the final model was checked using PROCHECK version 3.5.4 [14]. Docking of the ThDP and Mg^{2+} cofactors into the homology models was performed by superimposition with the template by using the protein structure visualization program DeepView (SWISS-PdbViewer, version 3.7) [15]. Since template structures do not contain cocrystallized substrate, the substrate analog pyruvamide was taken from the X-ray crystal structure of ScPDC form-B (PDB ID 1qpb) [16] to elucidate substrate interaction in the active site. Protein interface analysis was carried out using the Protein-Protein Interaction server [17].

Calculation of ligand volume

The volume of the ligand was calculated using Mol_Volume version 1 [18]. The van der Waals radii for each atom type was derived from the CHARMM 22 force field. The radius of the spherical probe (R_PROBE) was set to 2.0 Å, and the GRID_STEP was set to 0.5 Å.

Protein-ligand docking

The AutoDockTools version 1.4.6 [19] was used to prepare the ligand and the protein structures, and to generate the grid parameter and docking parameter files. The 2-keto acid structures were generated using ChemDraw Ultra 7.0 [20]. The ThDP and Mg^{2+} were included in docking calculations. The whole protein was treated as a rigid body while the ligand was allowed to move freely

in the grid box during the docking process. The grid box was generated using AutoGrid version 4 [21]. The size of the grid box was set to 40 x 40 x 40 points on each side with the distance of 0.375 Å between each point.

Docking simulations were carried out using the AutoDock version 4 [21] with the Lamarckian genetic algorithm (LGA). Each docking calculation was derived from 100 different runs that were set to terminate after a maximum of 2 500 000 energy evaluations. The population size was set to 150. The other parameters were set to the default value provided in the AutoDock4. AutoDock4 calculates Inhibition constant, K_i , as “ $\exp [(\Delta G \times 1000) / (Rcal \times TK)]$ ” where ΔG is docking energy which is a sum of intermolecular energy and internal energy of ligand, $Rcal$ is $1.98719 \text{ cal K}^{-1} \text{ mol}^{-1}$ and TK is 298.15 K. The equilibrium dissociation binding constant, K_d is calculated as “ $1/K_i$,” that is commonly used to describe the affinity between a ligand and a protein.

Discussion:

Homology models of ZpPDC and SvPDC

In general, the homology models of ZpPDC and SvPDC are similar to the X-ray crystal structure templates. The root mean square deviation (RMSD) between the α of ZpPDC and ZmPDC, and between the α of SvPDC and EcIPDC were 2.27 and 2.93 Å respectively (Figure 2). For ZpPDC, 90.4% of the residues were in most favorable regions; only two residues (Ser73 from each subunit) were in disallowed regions giving a PROCHECK score of -0.32. For SvPDC, 88.7% of the residues are in most favorable regions with only four residues (Arg187, Arg494, and Thr520 of subunit A and Lys333 of subunit B) in disallowed regions, with a PROCHECK score of -0.33. However, these residues were not involved in cofactor or substrate binding according to the equivalent residues in other species. After a short energy minimization, these residues were in the generously allowed regions.

Each subunit of both models consists of three domains, namely PYR, R and PP domains. All share the same α/β topology common to all thiamine dependent enzymes. The PYR and PP domains have six parallel beta sheets flanked by six alpha helices, whereas the R domain contains four parallel beta sheets (R β 4-7) with a pair of adjacent anti-parallel beta sheets (R β 2-3) comparable to ZmPDC [5] (Figure 1).

In all PDCs, two PDC monomers interact to form a dimer which plays a vital role in the catalytic mechanism of the enzyme since it comprises the active site. In both models, most of the interface residues come from the PYR and PP domains, while only 6 residues are from R domain. The contribution of interface residues by each domain reflects the average mobility of the domain in the resolved PDC crystal structure: e.g. for EcIPDC the PYR domain has the lowest average B-factor, 23 \AA^2 whereas the R domain has the highest overall B factor, 37 \AA^2 [8]. The ZpPDC interface accessible surface area is similar to ZmPDC (19%), but different from SvPDC (16%). Fewer H-bonds are present in the SvPDC model (30) compared with the ZpPDC model (44) and ZmPDC (66); in contrast the EcIPDC has the fewest (26). The number of hydrogen bonds at the interface between subunits was found to be proportional to the subunit interface area for experimentally determined structures with correlation coefficient of 0.75 (Figure 3). Extrapolation of this relationship was used to determine interface areas for ZpPDC and SvPDC, which were 3587 \AA^2 and 3284 \AA^2 respectively. These interfaces are much smaller than the ZmPDC interface area, and are comparable with ScPDC form-B and EcIPDC, suggesting some conformational changes in ZpPDC and SvPDC upon substrate binding.

Binding modes of the cofactors

In our homology models, the pyrimidine ring of ThDP interacted with the PYR domain of one subunit, and the rest of the molecule interacted with the PP domain of the other subunit giving “V” conformation of ThDP [22] that is in agreement with resolved X-ray crystal structures of PDCs [5, 16]. Only few acidic and basic residues that were in close contact with the cofactors among the X-

ray crystal structures and the models were similar i.e. Glu50, Asp27, His113, Glu473 (Table 1 in supplementary material).

Substrate binding sites of the homology models

In the ZmPDC X-ray crystal structure, the substrate atoms are surrounded by the side chains of Asp26, His113, His114, Tyr290, Thr388 and Glu473 (all conserved except Tyr290) which might be involved in substrate binding and catalysis [5]. Superimposition of three X-ray crystal structures (ScPDC form-B [16], ZmPDC and SvPDC) and the two homology models of ZpPDC and SvPDC showed equivalent conserved residues were also present in homology models with similar orientation, while Tyr290 of ZmPDC and ZpPDC was replaced by smaller residue, Thr288 in SvPDC and EcIPDC (Figure 4). Conversion of Thr288 to its ZmPDC analogue would decrease active site volume of SvPDC and subsequently decrease K_m value for pyruvate. The two conserved histidine residues (His113, His114 for ZmPDC and SvPDC; His112, His113 for ZpPDC) that have crucial roles in catalysis [23] and are in very close proximity to the pyrimidine ring of ThDP are orientated similar to X-ray crystal structure of ZmPDC (Figure 4).

Pyruvamide interacts with ThDP and side chains of Asp26, His113, Glu468 of ZpPDC, and Asp27, His114 of SvPDC. The distances between pyruvamide and these structures are within 4 Å (Figure 5A and 5B). The orientation of bulky residues in the active site is similar in ZmPDC and ZpPDC (Trp392 and Trp387, respectively) (Figure 4) while Tyr384 is found in SvPDC model at equivalent position with different orientation. This residue is replaced by alanine in ScPDC (Ala392) and EcIPDC (Ala387). The type and position of these residues indicate that ZpPDC and SvPDC may have similar carbonylation activity as of ZmPDC, since the second aldehyde molecule in the carbonylation reaction is accommodated in this region [5].

Substrate activation of homology models

SvPDC has been shown to display allosteric activation by the substrate with a K_m value of 4 to 6 mM [7], similar to the fungal and plant PDCs, while ZmPDC, ZpPDC [6] and EcIPDC [8] follow Michaelis-Menten kinetics. Superimposition of the ScPDC form-B with the homology model of SvPDC revealed Tyr155, Tyr321, Gln158, Ala221, Arg220, Lys225, within 5 Å distance of pyruvamide. Hydrogen bonds are formed between the amide group of pyruvamide and the side chain hydroxyl oxygen atom of Tyr155 and the main chain oxygen atom of Arg220 of same subunit (Figure 5C and 5D), which are similar as in ScPDC form-B. Comparison of X-ray crystal structures of non-activated ScPDC form-A and activated ScPDC form-B showed significant changes in three interfaces: between the R and the PYR domain, between monomer subunits, and between dimer subunits, suggesting the interaction of pyruvamide with ScPDC leads to large scale conformational changes during catalysis [14]. The SvPDC model suggests that similar conformational changes upon substrate binding at regulatory site could occur in this model. Another theory of activation by covalent binding of substrate to Cys221 which then interacts with His92 of ScPDC form-B [16] seems implausible for SvPDC since these residues are replaced by Glu216 and Asn91 at the equivalent positions respectively.

Affinity of ZpPDC and SvPDC towards other 2-keto acids

PDC homologues across species have conserved cofactor binding residues but different substrate preferences. The experimentally-determined K_m values for four different 2-keto acids binding to ZmPDC [3] suggest a relationship between molecular volume of substrate and K_m (Figure 6); therefore, dissociation values (K_d) for various 2-keto acids of varying molecular volumes docked with crystal structure of ZmPDC, ZpPDC and SvPDC models were determined. The larger 2-keto acids bind with higher K_d than those of the smaller 2-keto acids to both homology models (Table 2 in supplementary material), suggesting that the substrate binding preferences of the ZpPDC and SvPDC enzymes are similar to ZmPDC. Ile472 and Thr382 residues proposed to be responsible for making substrate binding site volume small in ZmPDC [5] are conserved in SvPDC (equivalent residues Ile464, Thr382) and ZpPDC (equivalent residues Ile467, Thr380). While in EcIPDC and

benzoylformate dehydrogenase whose preferred substrates are aromatic 2-keto acids, the equivalent residues at same positions are Val467, Gln383 and Ala460, Leu403, respectively. Therefore, we speculate that mutation of Ile467 of ZpPDC and Ile464 of SvPDC to

small residues like alanine or valine might produce variants capable to use aromatic 2-keto acids as substrate.

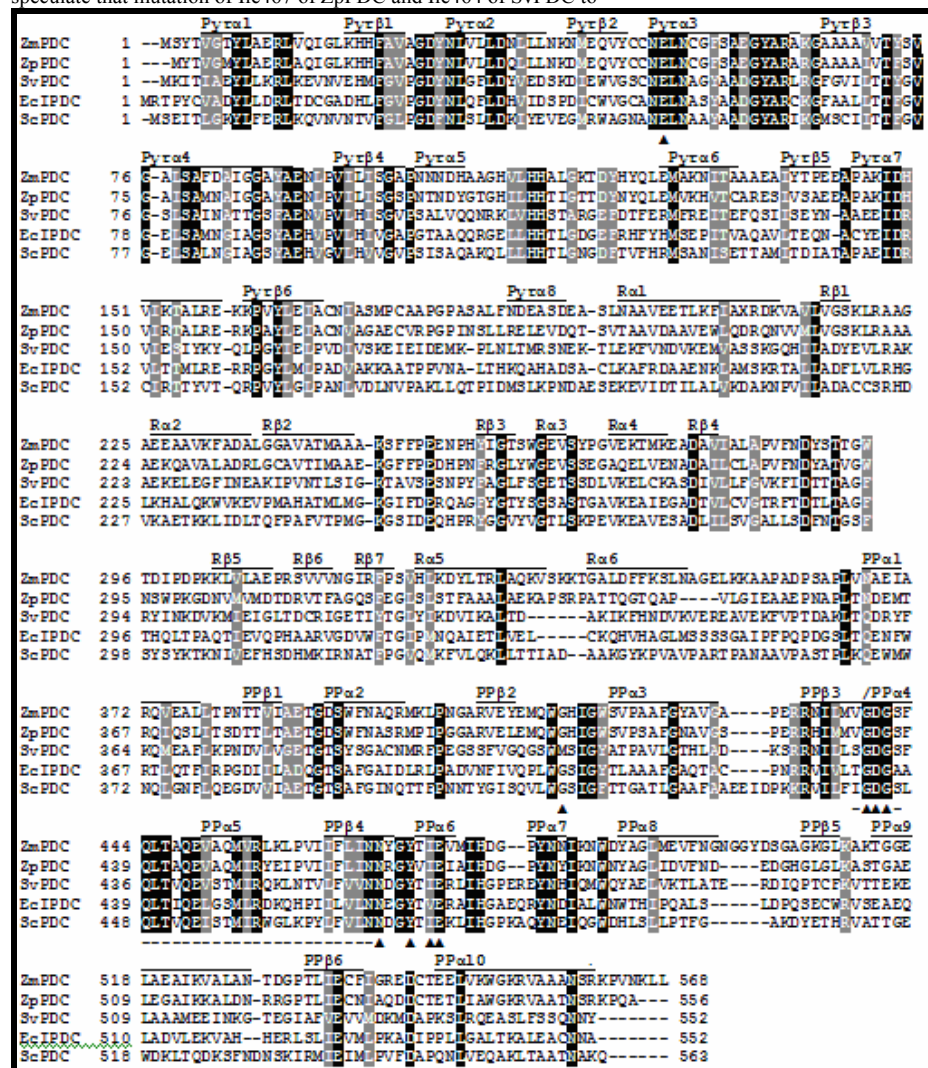


Figure 1: Multiple amino acid sequence alignment of ZmPDC, ZpPDC, SvPDC, EcIPDC and ScPDC (1pvm). Identical, functionally conserved and semi-conserved amino acid residues are shaded. Dashes indicate gaps introduced in alignment. The ThDP-binding motif is underlined with broken lines. Secondary structural motifs are indicated above the alignment according to the X-ray structure of ZmPDC [5]. (▲) are residues within 3 Å of ThDP in ZmPDC.

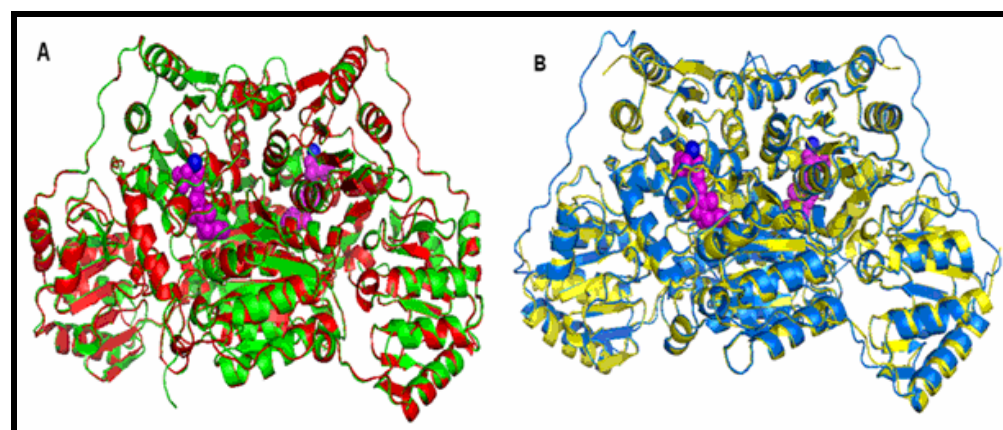


Figure 2: Superimposition of the C α of the X-ray crystal structure of ZmPDC (Green) and the homology model of ZpPDC (Red) (A), and the X-ray crystal structure of EcIPDC (Yellow) and the homology model of SvPDC (Blue) (B). The ThDP and Mg²⁺ are shown in space-filled model in magenta and dark blue, respectively.

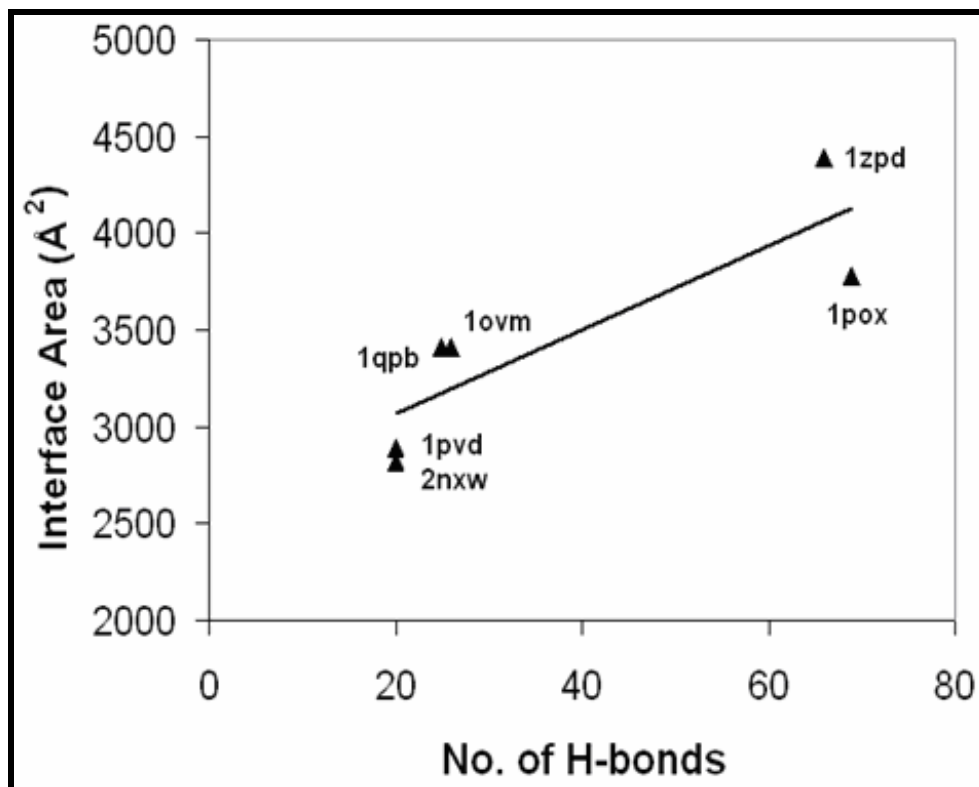


Figure 3: Correlation of interface area and number of hydrogen bonds of crystal structures of ZmPDC (1zpd) [5]; Pyruvate oxidase from *Lactobacillus plantarum* (1pox) [24]; EcIPDC (1ovm) [8]; ScPDC form-B (1qpb) [16]; ScPDC form-A (1pvd) [9] and phenylpyruvate decarboxylase of *Azospirillum brasilense* (2nxw) [25].

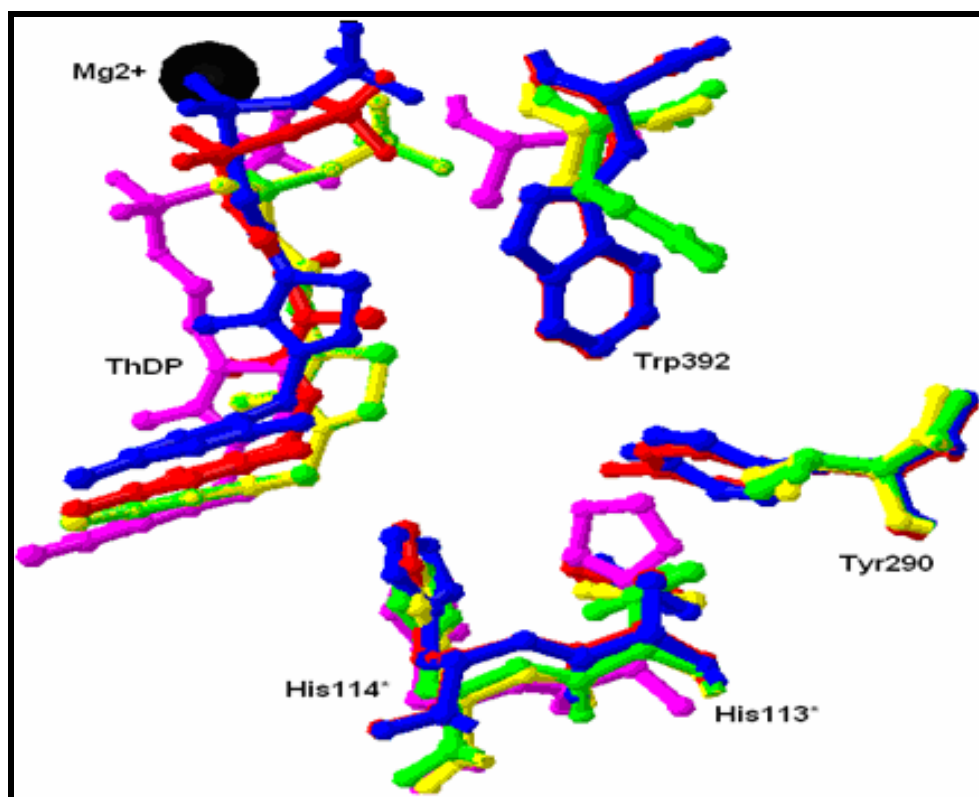


Figure 4: Superimposition of crystal structures of ZmPDC (Red), EcIPDC (Yellow), ScPDC form-A (Pink), and homology models of ZpPDC (Blue) and SvPDC (Green) showing different orientation of first Histidine residue of two consecutive Histidine residues. Residues are numbered according to ZmPDC. Residues labeled with (*) are from the other monomer. Figure generated using DeepView / Swiss-PdbViewer 3.7.

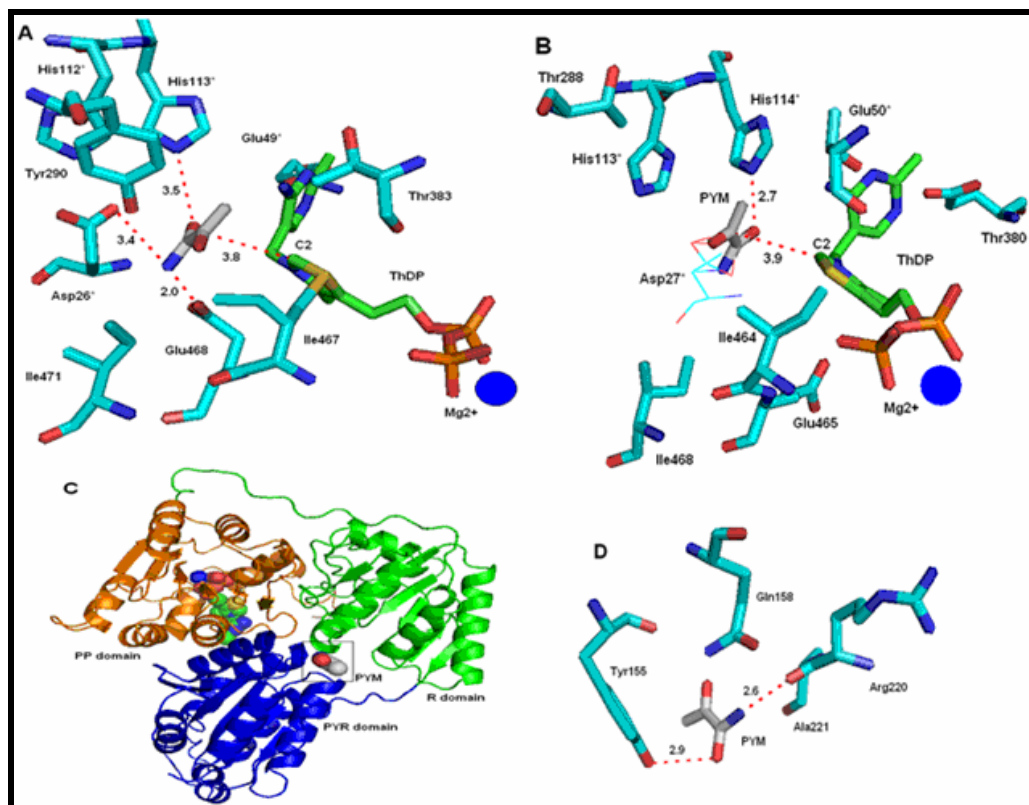


Figure 5: Putative active sites of homology models of ZpPDC (A) and SvPDC (B). Asp27* in figure B is shown in line structure for clarity. The regulatory site of SvPDC is shown in (C) and in close detail in (D). Cofactors, ThDP (Green) and Mg^{2+} (Blue), pyruvamide (PYM) (Red) at the regulatory site of SvPDC are shown in space-filled model in figure C. Residues labeled with (*) are from the other monomer. Potential H-bonds are indicated as red dotted lines.

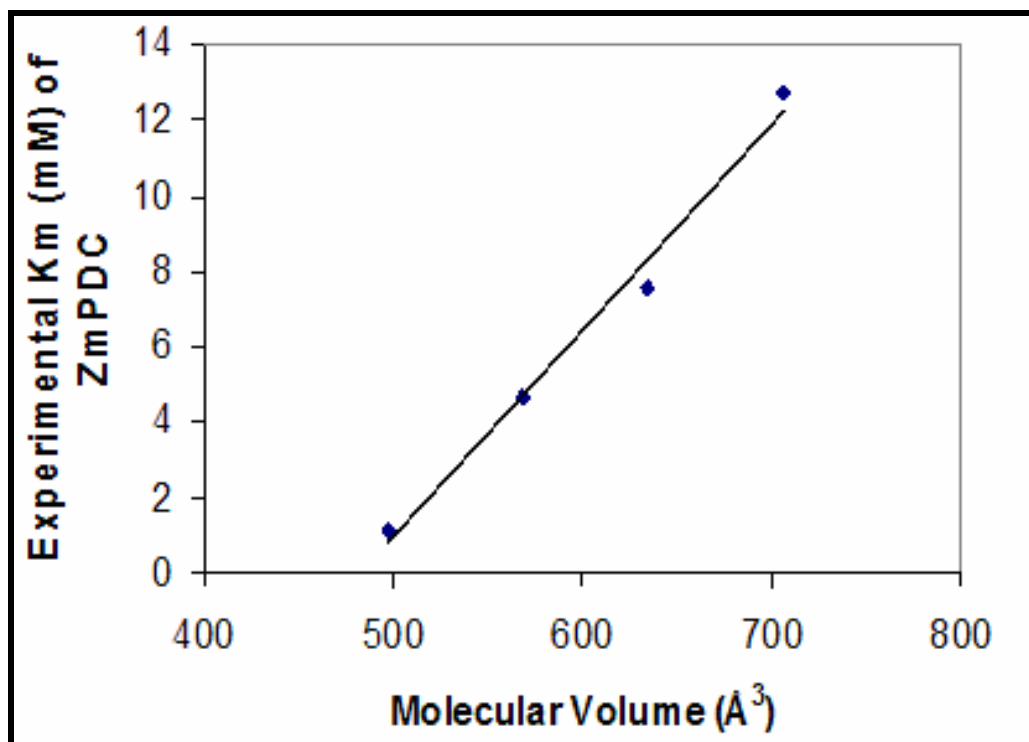


Figure 6: Correlation of ZmPDC experimental K_m values for four 2-keto acids; pyruvate, 2-ketobutanoic acid, 2-ketopentanoic acid, and 2-ketohexanoic acid, [23] and their calculated molecular volume.

Conclusion:

In this study, the homology models of the pyruvate decarboxylase of *Z. palmarum* and *S. ventriculi* were generated to explore the

structural similarities and differences among bacterial PDCs at the atomic level. The homology models and the X-ray crystal structure of ZmPDC were similar in terms of cofactor binding modes,

substrate binding residues, and active site volume. The mechanism of allosteric activation shown by SvPDC could be similar as in ScPDC form-B which needs further experimental verification. A relationship was found between the molecular volume of 2-ketoacid substrates and calculated dissociation constant K_d values, demonstrating preference of ZpPDC and SvPDC enzymes for aliphatic 2-ketoacids.

Authors' contributions:

AS performed modelling, docking and wrote the first draft of the manuscript. SD helped with data analysis and provided suggestion on manuscript. EJ supervised the project, provided mentorship and helped with drafting and revising of the manuscript. All authors read and approved the final manuscript.

Acknowledgment:

This investigation received financial support from the Joint Graduate School of Energy and Environment (JGSEE), King Mongkut's University of Technology Thonburi, Bangkok, Thailand. The authors are grateful to Philip Shaw, Sissades Tongsim, Pavita Tipsombatboon, Alisa Wilantho and Wanwimon Mokmak for their technical support and valuable comments.

References:

- [1] S Konig *et al.*, *Biochemistry* (1998) **37**: 5329 [PMID: 9655918]
- [2] A Schutz *et al.*, *Biochemistry* (2005) **44**:6164 [PMID: 15835904]
- [3] P Siegert *et al.*, *Protein Eng Des Sel.* (2005) **18**: 345 [PMID: 15930043]
- [4] M Pohl *et al.*, *Curr Opin Biotechnol.* (2004) **15**: 335 [PMID: 15296931]
- [5] D Dobritzsch *et al.*, *J Biol Chem.* (1998) **273**: 20196 [PMID: 9685367]
- [6] LA Talarico *et al.*, *Microbiology* (2001) **147**: 2425 [PMID: 11535783]
- [7] KC Raj *et al.*, *Appl Environ Microbiol.* (2002) **68**: 2869 [PMID: 12039744]
- [8] A Schutz *et al.*, *Eur J Biochem.* (2003) **270**: 2312 [PMID: 12752451]
- [9] P Arjunan *et al.*, *Mol Biol.* (1996) **256**: 590 [PMID: 8604141]
- [10] A Sali and TL Blundell, *J Mol Biol.* (1993) **234**: 779 [PMID: 8254673]
- [11] MY Shen and A Sali, *Protein Sci.* (2006) **15**: 2507 [PMID: 17075131]
- [12] R Samudrala and J Moult, *J Mol Biol.* (1998) **275**: 895 [PMID: 9480776]
- [13] JC Phillips *et al.*, *J Comput Chem.* (2005) **26**: 1781 [PMID: 16222654]
- [14] AL Morris *et al.*, *Proteins* (1992) **12**: 345 [PMID: 1579569]
- [15] <http://www.expasy.org/spdbv>.
- [16] G Lu *et al.*, *Eur J Biochem.* (2000) **267**: 861 [PMID: 10651824]
- [17] CD Reynolds *et al.*, *Bioinformatics* (2009) **25**: 413 [PMID: 19001476]
- [18] L Kalé *et al.*, *Journal of Computational Physics* (1999) **151**: 283
- [19] <http://autodock.scripps.edu>.
- [20] <http://www.CambridgeSoft.com>.
- [21] GM Morris *et al.*, *J Comput Chem.* (1998) **19**: 24
- [22] G Hubner *et al.*, *Biochim Biophys Acta* (1998) **1385**: 221 [PMID: 9655909]
- [23] G Schenk *et al.*, *Eur J Biochem.* (1997) **248**: 63 [PMID: 9310361]
- [24] YA Muller *et al.*, *J Mol Biol.* (1993) **237**: 315 [PMID: 8145244]
- [25] W Versees *et al.*, *FEBS J.* (2007) **274**: 2363 [PMID: 17403037].

Edited by Tan Tin Wee

Citation: Shrestha *et al.* Bioinformation 4(8): 378-384 (2010)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

Supplementary Material:

Table 1: Residues within 3 Å of ThDP in different X-ray crystallography structures and homology models.

Protein (PDB ID)	Contacted residue
ScPDC (1pvd)	Glu51*, Thr390, Gly413, Asp444, Gly445, Ser446, Thr475, Ile476
EcIPDC (1ovm)	Glu52*, Thr385, Gly408, Ala437, Gly464, Thr466
ZmPDC (1zpd)	Glu50*, Gly413, Asp440, Gly441, Ser442, Asn467, Tyr470, Thr471, Ile472
ZpPDC (model)	Ala24*, Glu49*, Gly408, His409, Asp435, Gly436, Ser437, Asn462, Gly464, Tyr465, Val466, Ile467
SvPDC (model)	Glu50*, Ser434, Met405, Ile407, Gly461, Try462, Thr463, Ile464

Residues labeled with (*) are from the other monomer.

Table 2: Calculated K_i and K_d of 2-keto acid derivatives against the X-ray crystal structure of ZmPDC and homology models of ZpPDC and SvPDC.

Ligands	Mol (Å ³)	Vol.	ZmPDC			ZpPDC			SvPDC		
			K_i (mM)	K_d (mM)	(1/ K_i)	K_i (mM)	K_d (mM)	(1/ K_i)	K_i (mM)	K_d (mM)	(1/ K_i)
Pyruvate	497		371	0.003		692	0.001		920	0.001	
2-ketobutanoic acid	568		719	0.001		534	0.002		583	0.002	
2-ketopentanoic acid	634		276	0.004		330	0.003		309	0.003	
2-ketohexanoic acid	706		NA	NA		166	0.006		185	0.005	
2-keto-3-methylbutanoic acid	597		662	0.002		267	0.004		310	0.003	
2-keto-3-methylpentanoic acid	845		717	0.001		148	0.007		147	0.007	
2-keto-3-cyclohexylpropanoic acid	825		NA	NA		NA	NA		NA	NA	
2-keto-3-3-dimethylpentanoic acid	735		NA	NA		80	0.012		75	0.013	
2-keto-4-methylhexanoic acid	756		786	0.001		73	0.014		98	0.010	
2-keto-4-methylpentanoic acid	694		580	0.002		130	0.008		140	0.007	
2-keto-4,4-dimethylpentanoic acid	700		NA	NA		111	0.009		99	0.010	
Acetaldehyde	373		41	0.024		44	0.022		27	0.038	

NA: Autodock calculated free energy of binding was positive.