Theor Chem Acc (2012) 131:1280
DOI 10.1007/s00214-012-1280-1

REGULAR ARTICLE

A QM/MM study on the catalytic mechanism of pyruvate decarboxylase

Qianqian Hou · Jun Gao · Yongjun Liu · Chengbu Liu

Received: 4 June 2012 / Accepted: 12 September 2012 / Published online: 29 September 2012

Abstract Pyruvate decarboxylase (PDC) is a typical thiamin diphosphate (ThDP)-dependent enzyme with widespread applications in industry. Though studies regarding the reaction mechanism of PDC have been reported, they are mainly focused on the formation of ThDP ylide and some elementary steps in the catalytic cycle, studies about the whole catalytic cycle of PDC are still not completed. In these previous studies, a major controversy is whether the key active residues (Glu473, Glu50′, Asp27′, His113′, His114′) are protonated or ionized during the reaction. To explore the catalytic mechanism and the role of key residues in the active site, three whole-enzyme models were considered, and the combined QM/MM calculations on the nonoxidative decarboxylation of pyruvate to acetaldehyde catalyzed by PDC were performed. According to our computational results, the fundamental reaction pathways, the complete energy profiles of the whole catalytic cycle, and the specific role of key residues in the common steps were obtained. It is also found that the same residue with different protonation states will lead to different reaction pathways and energy profiles. The mechanism derived from the model in which the residues (Glu473, Glu50′, Asp27′, His113′, His114′) are in their protonated states is most consistent with experimental observations. Therefore, extreme care must be taken when assigning the protonation states in the mechanism study. Because the experimental determination of protonation state is currently difficult, the combined QM/MM method provides an indirect means for determining the active-site protonation state.

Keywords Pyruvate decarboxylase · Decarboxylation · Protonation state · Combined QM/MM · Reaction mechanism

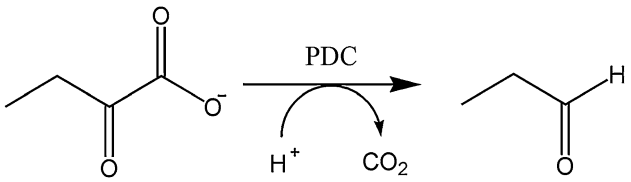
1 Introduction

Thiamin diphosphate (ThDP), the biological active form of vitamin B1, is a key organic cofactor required by a number of enzymes which participate in many biosynthetic pathways and catalyze a broad range of reactions, and most remarkably, the cleavage and formation of carbon–carbon bonds [1, 2]. Pyruvate decarboxylase (PDC; EC 4.1.1.1) is the simplest thiamin diphosphate (ThDP)-dependent enzyme that catalyzes the nonoxidative decarboxylation of pyruvate to acetaldehyde (see Scheme 1) [3, 4]. The active acetaldehyde intermediate could also be condensed with a second acetaldehyde to form acetoin through an aldol-type condensation reaction [5]. PDC has recently attracted much attention in the enantioselective synthesis of pharmaceutical intermediates, production of ethanol and chiral compounds in industry [6–8]. Therefore, it is very important to understand the catalytic mechanism of PDC. According to amounts of kinetic and mechanistic experimental studies, the catalytic cycle of PDC can be subdivided into several elementary steps and involves a series of substrate–

Electronic supplementary material The online version of this article (doi:10.1007/s00214-012-1280-1) contains supplementary material, which is available to authorized users.

Q. Hou · J. Gao · Y. Liu (✉) · C. Liu
Key Lab of Colloid and Interface Chemistry, Ministry of Education, School of Chemistry and Chemical Engineering, Shandong University, Jinan 250100, Shandong, China
e-mail: yongjunliu_1@sdu.edu.cn

Y. Liu
Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, Qinghai, China

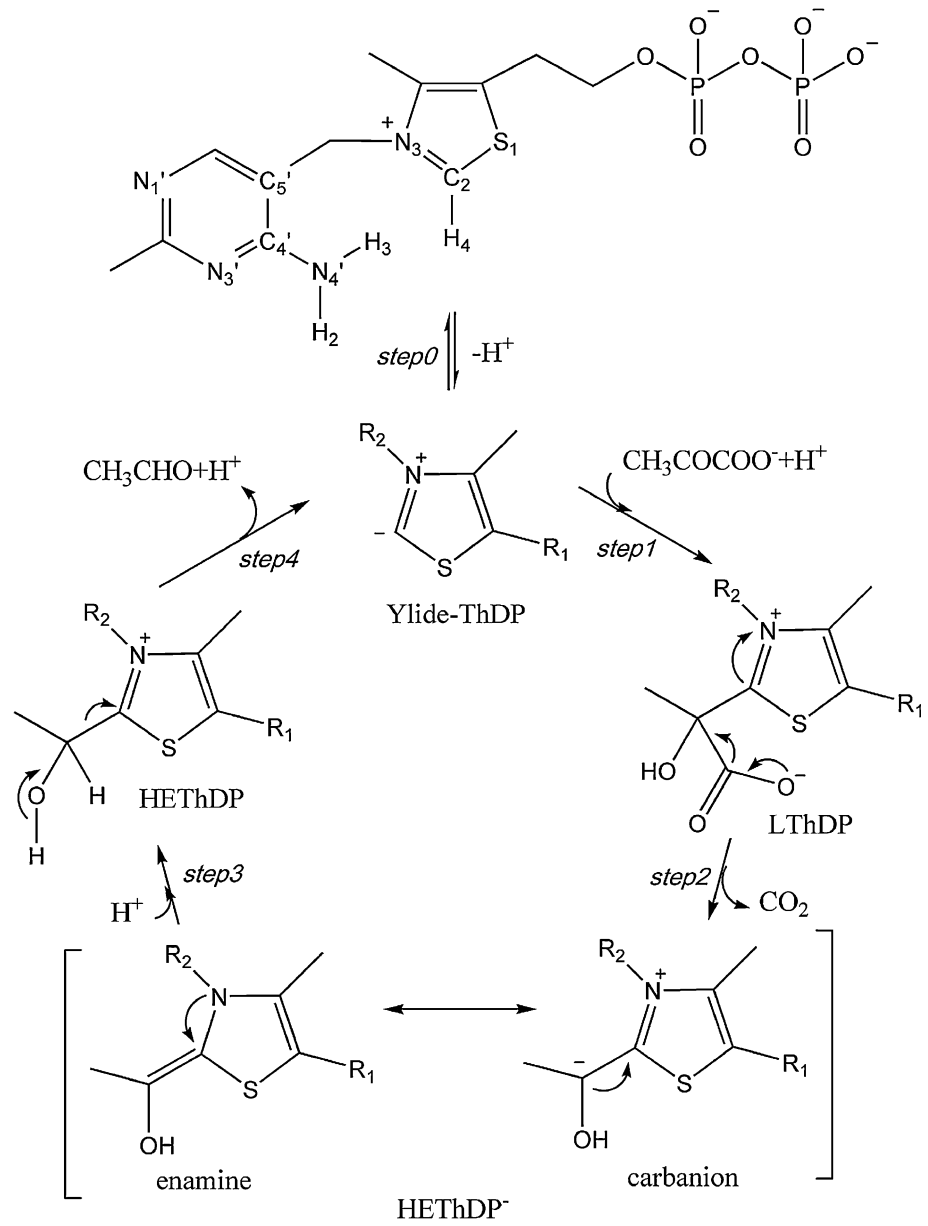


Scheme 1 Pyruvate decarboxylase (PDC) catalyzes pyruvate to acetaldehyde

cofactor conjugates as reaction intermediates, as described in Scheme 2 [9–12]. Initially, the reactive ylide is formed via the deprotonation of thiazolium ring (step 0). Subsequently, the ylide attacks pyruvate to generate the predecarboxylation intermediate 2-lactyl-thiamin diphosphate (LThDP) (step 1). Decarboxylation of LThDP yields the resonating α-carbanion and enamine forms of 2-hydroxyethyl-ThDP (HEThDP$^-$) (step 2), which is then protonated to give its conjugated acid (HEThDP) (step 3). Finally, the product of acetaldehyde is liberated, and the ylide is regenerated (step 4). In experiments, the resulting rate constant suggested that both of decarboxylation (step 2) and liberation of acetaldehyde (step 4) were probably to be rate-limiting, and the formation of enzyme-bound LThDP (step 1) was very fast [13].

Several high-resolution X-ray data provided insight into the structure of PDC [14–17]. The crystal structure of PDC variant Glu473Asp from *Zymomonas mobilis* in complex with the reaction intermediate 2-lactyl-ThDP (PDB code: 3OE1) [13] is shown in Fig. 1. Like other known ThDP-dependent enzymes, *Z. mobilis* PDC is tetrameric and the

Scheme 2 Catalytic cycle of *Zm*PDC with identified intermediates and elementary catalytic steps



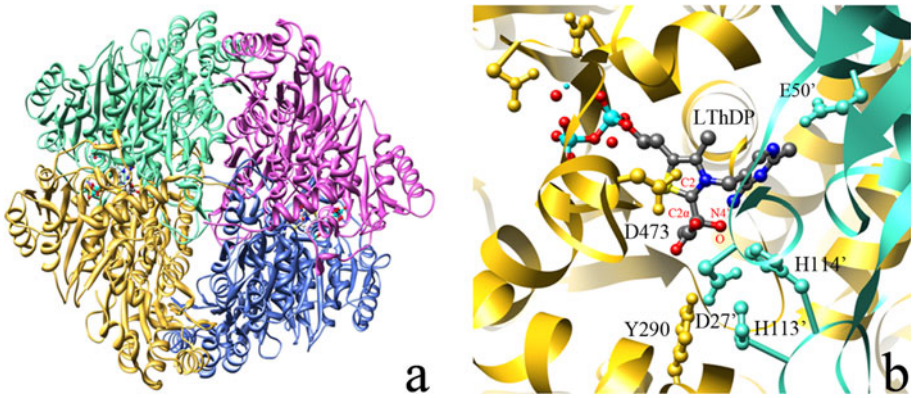


Fig. 1 **a** Crystal structure of PDC variant E473D in complex with LThDP (PDB code: 3OE1). **b** The active site structure of PDC

minimum quaternary structure for catalysis is a dimer. The active site is located between the interface of two monomers and anchored by the diphosphate group, Mg^{2+}, and several key surrounding residues. The coenzyme ThDP adopts a typical V-conformation that N4′ of the aminopyrimidine group is in close proximity to the C2 of the thiazolium ring. In addition, there are many charged and polar residues in the active sites, including Glu473, Tyr290, Asp27′ His113′, and His114′ as shown in Fig. 1b, involved in coenzyme binding and catalysis. For ThDP-dependent enzymes, many studies have focused on the tautomeric equilibrium between 4′-aminopyrimidine (amino-ThDP) and 1′,4′-iminopyrimidine ThDP (imino-ThDP), and the widely accepted mechanism is that an intramolecular proton transfers from C2 to N4′ assisted by the interaction between Glu50′ and the N1′ atom of the pyrimidine ring, generating the catalytic active ylide [18]. The active pocket also contains several other crucial residues (Glu473, Glu50′, Asp27′, His113′, and His114′), which have been judged by site-directed mutagenesis that mutation diminishes or abolishes catalytic activity [19–22]. For example, the enzyme only retained 0.46 % activity upon replacement of Glu50′Gln [19]. The mutant Glu473Gln had a very low, but measurable activity of 0.025 % of the wild type [19]. Determination of the kinetic properties showed that the affinity of Asp27′Ala was decreased 30-fold [20]. Experiments even detected that His113′Glu rendered the enzyme completely inactive [22]. His114′ played a comparatively minor role during the catalytic process because mutation to Gln resulted in an enzyme that retained 37 % of the wild-type activity [19].

For the decarboxylation process catalyzed by PDC, several different catalytic mechanisms have been proposed in the literature. In several experiments, 1H NMR spectroscopy was used to detect the key intermediates, and site-directed mutagenesis enabled the assignment of individual side chains to single steps in catalysis [23]. According to the experiment results, Tittmann proposed two independent proton relay systems catalyzing the activation of ThDP (Glu50′–N1′–N4′–NH2, step 0), substrate binding (Glu50′–N1′–N4′–NH2, step 1), acetaldehyde release (Glu50′–N1′–N4′–NH2 and His113–Asp27, step4) as well as the stereochemical control of decarboxylation by Glu473 inducing a perpendicular orientation of the substrate carboxylate to the thiazolium ring of the enzyme-bound ThDP [23]. A subsequent computational study focused on the details of the elementary reaction steps totally based on the above mechanism was carried out by using density functional theory (DFT) [24]. During the calculations, small models with different residues were constructed to estimate the contribution of individual residue to a certain step, but the bottom atoms of the residues were frozen at their crystallographic positions. The geometry parameters and activation energies of the formations of ylide and HEThDP, and the liberation of acetaldehyde were obtained. But there were still some limitations in their study. The covalent addition of pyruvate was not designed. The transition state for the decarboxylation step was not located, and only a rough estimate of the barrier was obtained by using a series of constrained optimizations. The rate-limiting step was the protonation of α-carbanion with an energy barrier of 33 kcal/mol, and this result was not consistent well with experiments in which the decarboxylation process was rate-limiting step [13]. Small models in which the protein framework was frozen may neglect the heterogeneous enzyme environment, and its applicability in modeling enzyme reactions has been well recognized to be very limited. Similarly, the decarboxylation reaction of LThDP had been studied with AM1 and ab initio HF or MP2 methods based on different models with or without the active-site residues, which led to different activation barriers and exothermicities [25]. Furthermore, long-term molecular dynamics simulations of PDC were performed by Lie, and different proton relays upon nucleophilic attack were proposed [26]. When the substrate attacked the ylide-ThDP(C2), Asp27′ transferred its proton to Tyr290(OH)

which then protonated the substrate forming LThDP. More recently, several crystal structures of PDC were reported by Luisi and Leeper [27], in which a complex hydrogen bonding network was present in the active site. Two water molecules were involved in the network, and one was present in the pyruvate-bound structure and the other in the protonation of $HEThDP^{-}$ after decarboxylation. Thus, a different mechanism involving two waters was proposed.

As outlined above, the experimental and theoretical studies have successfully described the PDC catalytic mechanism, but there are still some discrepancies between the published results [24–27]. A common problem is that different protonation states of key residues are not taken into account in their computational models [24–27]. The models involved inadequate treatment of protein environment is the other issue. To eliminate most of the aforementioned uncertainties of the computational studies, three hypotheses have been proposed on the protonation state of key residues, and different mechanisms are calculated and compared, which have not been done before. In recent years, QM/MM method has become increasingly popular and successfully applied in the fields of extended systems, especially enzymes. And it has been provided a powerful and effective tool for describing the enzymatic mechanism at the atomistic level and identifying additional properties that may lead to evidence in future experiments [28–31]. The extensive combined quantum mechanical/molecular mechanical (QM/MM) calculations on the whole catalytic cycle of PDC are performed for the first time, with aims to unveil how the well-structured microenvironment around the reactive site accelerates reactions. Emphasis is given on individual amino acid residues and their contribution to the catalysis.

2 Computational methods

The crystal structure of PDC variant Glu473Asp from *Z. mobilis* in complex with the reaction intermediate 2-lactyl-ThDP (PDB ID: 3OE1) [13] was used to prepare the computational model. The protein and coenzyme from one dimer were retained, and the mutated residue 473 was changed back to glutamic acid with VMD program [32]. After optimized at B3LYP/6-31G(d) level with the Gaussian 03 package [33], the substrate was placed into the active site using the Autodock program [34]. Then all the pretreatments and MM calculations were done with the CHARMM22/CMAP force field as implemented in the CHARMM program [35, 36]. The standard CHARMM topology and parameter files do not contain the information of ThDP; therefore, we describe ThDP by using the existing atom types of pyrimidines and the histidine acid, and the corresponding CHARMM parameters are given in the Supporting Information. The missing hydrogen atoms were added using the HBUILD facility after all of the charged residues were set to their protonated or deprotonated states in the toppar file. But, for the ionic residues in the active site, we will discuss their different protonation states in the following paragraph. Each system was neutralized by adding Na^{+} ion at random position and solvated into a water sphere with a 38 Å radius. The TIP3P potential [37] was used for the water molecules. Preliminary structural optimization was carried out to remove bad contacts and relax the complex. Then stochastic boundary MD simulations were carried out with CHARMM package. During the simulations, the reaction zone was defined as a sphere of 30 Å radius in which the atoms were propagated according to Newtonian mechanics. The buffer zone involved the atoms within 30–35 Å, which were retained by a harmonic restoring force and propagated by Langevin dynamic. The rest atoms beyond 35 Å were defined as the reservoir zone, which were fixed during the MD simulations. The effect imposed by the atoms in the reservoir zone was simulated by stochastic boundary potentials [35, 38]. About 800 ps simulations were performed for each prepared system to equilibrate the systems. The root-mean-squared deviations (RMSD) of the protein were derived and shown in Fig. S1. It can be seen that the backbone of the protein only changes slightly. We also calculated the root-mean-squared deviations (RMSD) between the crystal structure and the selected snapshots for QM/MM calculations, and the values are 1.25, 1.38, and 1.37 Å for model A, B, and C, respectively. Thus, the final snapshot was selected as the starting point for the QM/MM calculations.

Following equilibration, the ChemShell software package [39] was used to perform the QM/MM calculations by integrating the TURBOMOLE package [38] for QM subsystem and DL-POLY program [40] for MM subsystem. The resulting QM/MM system had a total of ~33,000 atoms. For the QM subsystem, the substrate, part of the coenzyme (including 4′-aminopyrimidine and the thiazole ring), Glu473 (monomer A), Asp27′, Glu50′, His113′, and His114′ (monomer B) were included and treated with B3LYP functional and 6-31G(d) basis set. All the other atoms were assigned to MM subsystem, which was described at CHARMM22 force field with DL-POLY program [41]. During the calculations, the QM region and 2,287 MM atoms (defined by including all the atoms around ThDP within a distance of 12 Å) were completely relaxed, whereas the remaining MM atoms were fixed. Hybrid delocalized internal coordinates (HDLC) optimizer [42] was employed for geometry optimizations, and a quasi-Newton limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [42] was used to search for minima. For the transition states, they were taken from the highest point on the potential energy profile along the

reaction coordinate, which were scanned over bond lengths, then optimized with partitioned rational function optimization (P-RFO) method [42] and characterized by a single negative eigenvalue. The electronic embedding scheme and the hydrogen link atoms with charge shift model [43] for QM/MM boundary were applied in the QM/MM treatment. No cutoffs were introduced for the nonbonding MM and QM/MM interactions. Finally, a high-level single point electronic energy calculation was performed at a larger basis set 6-311++G(d,p) to obtain accurate energies.

3 Results and discussion

3.1 The protonation states of key residues in active site

X-ray crystallographic studies of PDC have identified a number of titratable residues in the active site, such as Glu473, Asp27′, Glu50′, His113′, and His114′, which are all involved in our QM region, but the precise role of these residues is unclear. As discussed previously, the protonation states of these residues in active site are important for the catalysis. Glu50′ may be the best studied residue [44–46] which is conserved in the sequence of all ThDP-dependent enzymes and located at a similar position in those determined enzymes. This residue has been proved to initiate the ionization of C2, so it is protonated in our three models. Another residue, Glu473, plays an important role in the orientation of the substrate and avoids electrostatic repulsion, which is presumably protonated and behaves as a hydrogen bond donor [13]. However, there are not sufficient evidences for the protonation state of Asp27′, His113′, and His114′. The pKa value of Asp in its free state is 3.9, and that of His is 6.0. Under the standard assay conditions of pH 6.0, Asp prefers to be deprotonated, and both of the protonated and deprotonated states are possible for His. But when these residues exist within proteins, the situation will be different. We have calculated the pKa values of the crystal structure using PROPKA method [47, 48]. The pKa values of Asp27′, His113′, and His114′ are 3.9, 7.9, and 2.0, respectively. At pH 6.0, His113′ is expected to be protonated, and both of Asp27′ and His114′ are expected to be deprotonated. It should be noted that the calculated pKa values depend largely on the crystal structure which may be inaccurate to some extent. In order to explore the protonation state of these residues, we have examined three models (model A, B, and C) in this paper, and the different protonation states of the residues for these three models are described in Fig. S2. Model A defines all of the three residues as their protonation states. For model B, only His113′ is protonated, and both of Asp27′ and His114′ exist as their deprotonated states. Model C assumes all of the three residues as their deprotonated states. Thus, three different situations will be discussed.

3.2 Reaction mechanisms

3.2.1 Model A

The mechanism proposed from model A is firstly investigated with QM/MM strategy. The MD simulations and QM/MM optimization yield the structure of the reactant, as shown in Fig. 2. The relative positions of the coenzyme and residues show good agreement with the experimental structure [13]. The substrate pyruvate is stabilized by three hydrogen bonds from Asp27′, His113′, and the amino of ylide-ThDP. The reactive C2 atom of thiazolium ring positions a distance of 3.26 Å from the carbonyl carbon atom (C2α) of pyruvate and defines the angle with the C2α–O atoms at 71.25°. This optimized structure is quite compatible for the covalent addition of pyruvate. Structure of transition state ($TS1_A$) indicates that when the ylide attacks on the carbonyl carbon atom, the hydrogen atom of the amino group simultaneously transfers to the carbonyl oxygen atom and forms the tetrahedral substrate-LThDP ($IM1_A$), in which the bond order of C2–C2α has changed from double to single. The total net charge of Glu50′ decreases from −0.077 in R_A to −0.132 in $TS1_A$, and it is expected to disperse the negative charges of the aminopyrimidine ring to stabilize the $TS1_A$ structure. It can be seen from $IM1_A$ that two oxygen atoms of carboxyl group form two hydrogen bonds with Glu473 and Asp27′. Thus, these two residues may play important roles in the decarboxylation of LThDP (step 2). LThDP is decarboxylated though $TS2_A$ to give the resonating $HEThDP^-$ intermediate ($IM2_A$). Accompanying the decarboxylation, the hydrogen bond between CO_2 and Asp27′ is broke, and then, Asp27′ forms a hydrogen bond with the C2α atom, as evidenced by the C2α–H(Asp27′) distance of 1.98 Å in $IM2_A$. As a result, Asp27′ donates one proton to $HEThDP^-$ in the third step. At $TS3_A$, the distances of C2–H and O(Asp27′)–H are 1.38 and 1.53 Å, respectively. After the conjugated acid (HEThDP, $IM3_A$) is formed, the C2–C2α bond has increased 0.14 Å compared with $IM2_A$. During this step, the hydrogen bond between His113′ and the hydroxyl group is well kept. The imidazole group of His114′ is only weakly interacted with the substrate, and the temporary proton transferring from His114′ to Asp27′ is not observed, which was present in Li' calculations [24]. The release of acetaldehyde is the last step. It can be seen from IM3A that there exists a hydrogen bond between the hydroxyl and the 4′-amino group, which is expected to be the only reasonable group to accept the proton. In $TS4_A$, the breaking of C2–C2α bond and deprotonation of α-hydroxyl group are a concerted reaction, and this step is

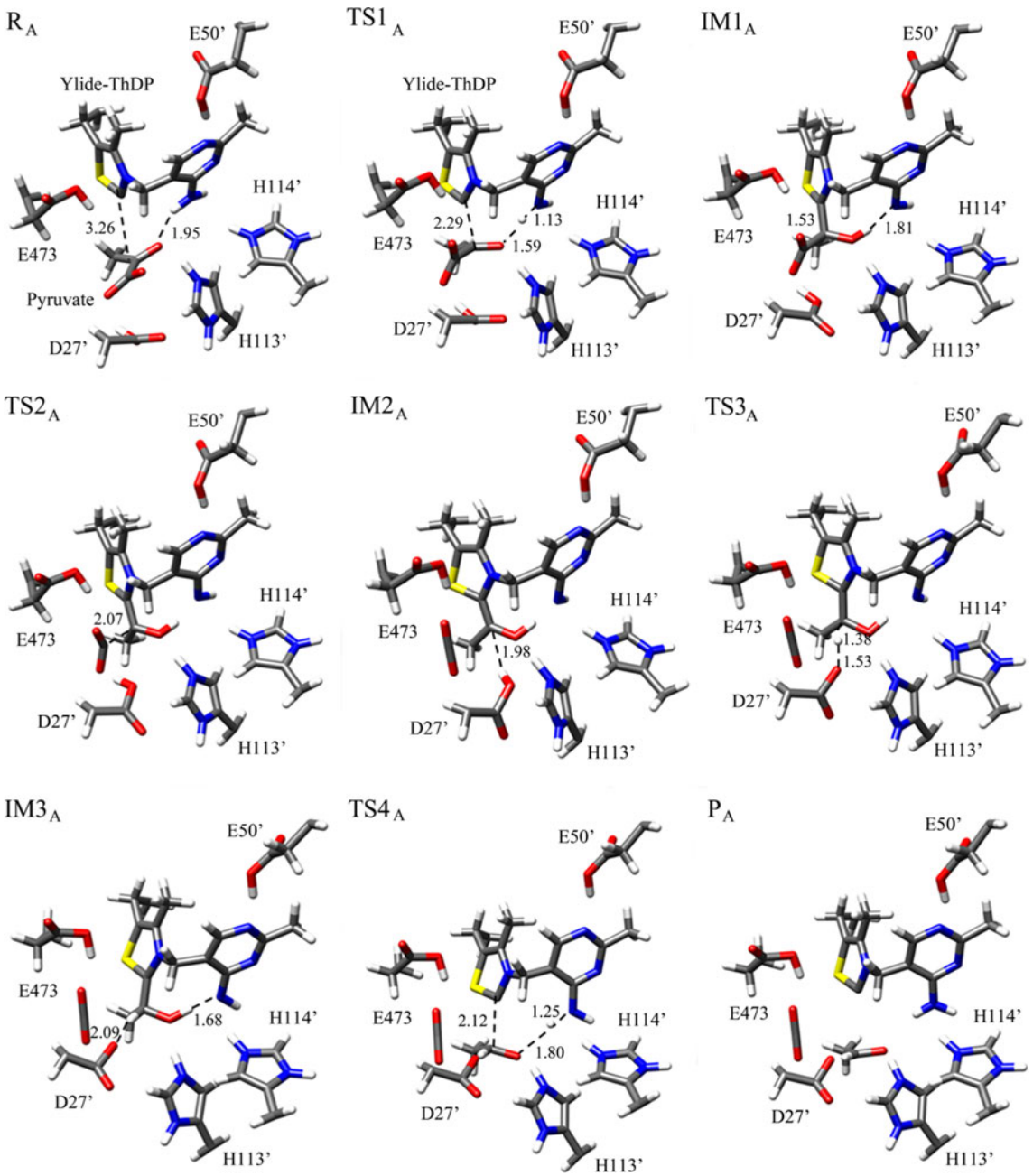


Fig. 2 Optimized structures of reactant (R_A), transition states ($TS1_A$, $TS2_A$, $TS3_A$, $TS4_A$), intermediate ($IM1_A$, $IM2_A$, $IM3_A$), and product (P_A) for model A. Distances are given in Å

calculated to be exothermic by 9.8 kcal/mol with an energy barrier of 8.9 kcal/mol. The relative energies for the whole catalysis are depicted in Fig. 3. The small barriers suggest that the reaction will probably proceed via this pathway.

3.2.2 Model B

As illustrated in our aforementioned discussion, model B focuses on the reaction from step 2, as depicted in Scheme 2. The QM/MM optimized geometries of $IM1_B$ are shown in Fig. 4, and the key internuclear distances in the intermediate complexes and transition states are listed in Table 1. In the QM region, Glu473, Glu50′, and His113′ are set to their protonation states, while Asp27′ and His14′ are in their deprotonation states. At $IM1_B$, the substrate forms four hydrogen bonds with Glu473, Asp27′, and the amino group of coenzyme. Similar to mechanism A, LThDP is first decarboxylated through an energy barrier of 3.8 kcal/mol. At $TS2_B$, the C2α–C2 distance is shortened from 1.48 to 1.38 Å, which is presumably arose from the

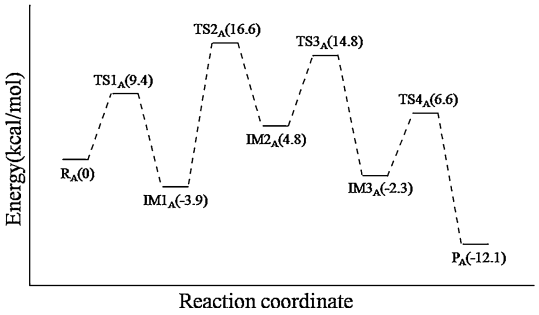


Fig. 3 The QM/MM energy profile for model A

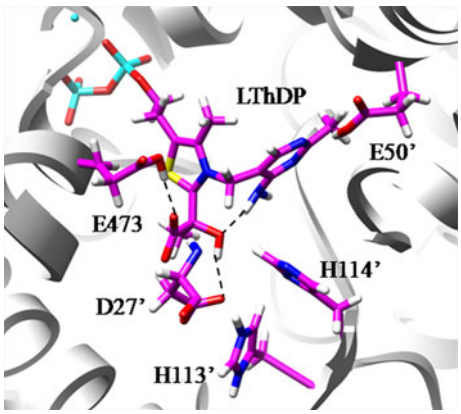


Fig. 4 The optimized structures of model B at its initial state ($IM1_B$)

Table 1 Key geometric parameters (Å) for model B

Label	C2–C2α	C2α–CO2	HN4′–N4′	C2α–HN4′	O–H	H–O (D27′)
$IM1_B$	1.48	1.64	1.01	2.66	0.99	2.50
$TS2_B$	1.38	2.23	1.01	2.50	0.98	2.46
$IM2_B$	1.35	2.92	1.01	2.42	0.98	2.71
$TS3_B$	1.48	3.08	1.72	1.14	0.98	2.09
$IM3_B$	1.51	3.14	1.98	1.10	0.98	2.03

loss of the conjugation in CO_2. When $HEThDP^-$ is formed, a proton transferring step occurs, in which the amino group of coenzyme is expected to be the proton donor. The transition state for the proton transferring step ($TS3_B$) is located as the distances of HN4′–N4′ and C2α–HN4′ are 1.72 and 1.14 Å, respectively, with the energy barrier about 30.5 kcal/mol. In $IM3_B$, there is a hydrogen bond of 2.03 Å between the hydroxyl and Asp27′, raising the possibility that Asp27′ is responsible for deprotonation of the hydroxyl group in the last step. For the next step, the reaction coordinate driving (RCD) method is used to map out the minimum energy path. Two pathways are defined. One of the reaction coordinates is defined as a combination of two bond lengths: $d = R_{C2-C2\alpha}-R_{H-O(Asp27')}$, and the other is $d = R_{H-O(Asp27')}$. That is, both of the concerted and stepwise reaction mechanisms for the proton transferring step are calculated. Unfortunately, both of the energies are increased steadily, and no transition state is found. The already obtained energy profile is presented in Fig. 5. The energy barrier of the proton transferring step is 30.5 kcal/mol, which is much higher than the decarboxylation step, and it is conflict with the experimental data [13].

3.2.3 Model C

Model C is also prepared from the intermediate lactylthiamin diphosphate (LThDP, $IM1_C$), and the QM/MM optimized geometries of $IM1_C$ are presented in Fig. 6. Compared with the crystal structure, the carboxyl group of residue Asp27′ has reversed due to several negative charged residues in the active site. The first step of decarboxylation of LThDP yields a barrier/endothermicity of 15.7/0.3 kcal/mol. When $IM2_C$ is formed, the distance between C2α and the hydrogen atom of amino is 2.31 Å. In the $TS2_C$, the atom C2α accepts a proton from the amino with an energy barrier of 34.8 kcal/mol. At $IM3_C$, His114′ forms a hydrogen bond with the hydroxyl group, as evidenced by the H–N1(His114′) distance of 1.92 Å. Furthermore, because there is an increase in the distance of C2–C2α bond, the proton atom transfers to the N1 atom of His114′ spontaneously, which is in agreement with a concerted mechanism. And then, the final product acetaldehyde is liberated. As revealed in Fig. 7, the activation energy for the rate-determining step is 34.8 kcal/mol, which is remarkably higher than the barrier of 20.5 kcal/mol in model A. Furthermore, the overall process is endothermic by 14.0 kcal/mol. Thus, the large barrier and instability of the $IM3_C$ and P_C implies this pathway unlikely to occur in practice (Table 2).

4 Discussion

The mechanistic details described above provide some supports for the mechanism of model A in which the calculated energy barriers are low. In experiments [13], microscopic rate constants for elementary steps in the reaction of wild-type *Zm*PDC were estimated on the basis of the relative concentrations of the intermediates. There was significant accumulation of both LThDP and HEThDP at the steady state, but HEThDP carbanion/enamine was not easy to be observed. The net rate constant of decarboxylation of LThDP (step 2) was about 397 s^{-1}. The constant of acetaldehyde liberation was about 265 s^{-1}, which was actually a composite rate constant and reflected both protonation of the carbanion/enamine intermediate (step 3) and liberation of acetaldehyde (step 4). We calculated the free energies from these experimental data.

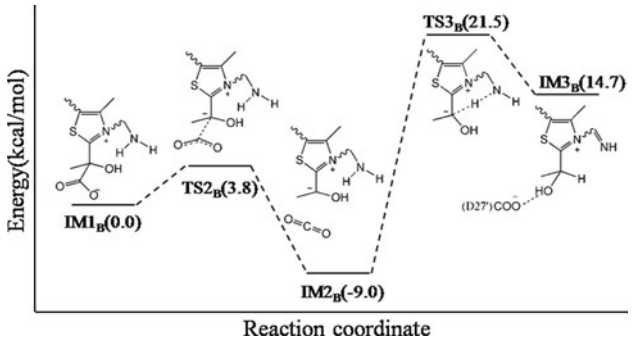


Fig. 5 The QM/MM energy profile for model B. The QM regions of model B and C contain the same residues as those of model A. For the sake of clarity, only part of the residues is shown

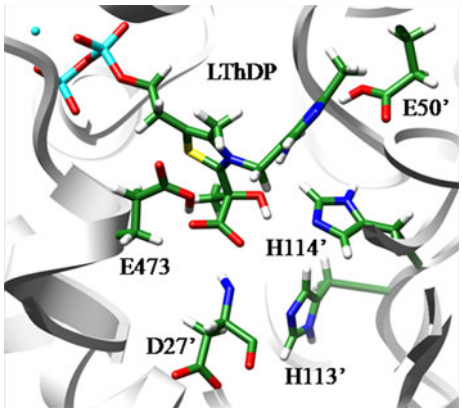


Fig. 6 The optimized structures of model C at its initial state ($IM1_C$)

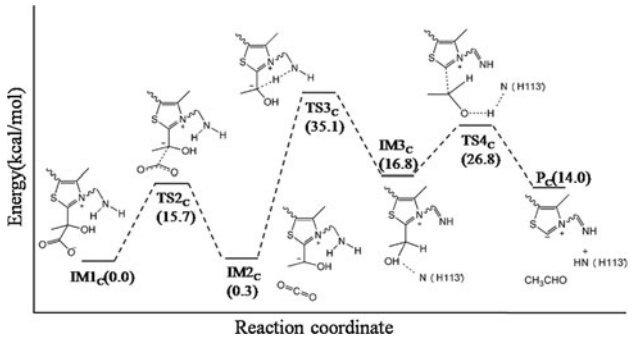


Fig. 7 The QM/MM energy profile for model C. For the sake of clarity, only part of the residues is shown

And it is 14.2 kcal/mol for decarboxylation of LThDP and 14.4 kcal/mol for acetaldehyde liberation, indicating that both of the steps are possibly rate-limiting for the overall catalytic reaction. According to our QM/MM calculation results shown in Fig. 3, the decarboxylation of LThDP turns out to be the most energy demanding and is the rate-limiting step. The free energy for this step is calculated to be 16.2 kcal/mol, which agrees well with the value of 14.2 kcal/mol obtained from experiments. It also can be seen that the intermediates LThDP and HEThDP lie in low energies (−3.9 and −2.3 kcal/mol, respectively) relative to the reactant, but the energy of the intermediate HEThDP carbanion/enamine (4.8 kcal/mol) is relatively high and will be instantaneously protonated in the next step. This will explain that why there was significant accumulation of both LThDP and HEThDP, but little of HEThDP carbanion/enamine in experiments [13].

Table 2 Key geometric parameters (Å) for model C

Label	C2–C2α	C2α–CO2	HN4′–N4′	C2α–HN4′	O–H	H–N1 (H114′)
$IM1_C$	1.52	1.61	1.02	2.80	0.99	2.29
$TS2_C$	1.43	2.01	1.01	2.53	1.00	1.74
$IM2_C$	1.36	3.15	1.01	2.31	1.01	1.75
$TS3_C$	1.46	3.56	1.51	1.24	1.00	1.86
$IM3_C$	1.50	3.56	1.87	1.11	1.00	1.92
$TS4_C$	2.47	3.66	1.91	1.11	1.10	1.47
P_C	2.52	3.51	2.00	1.10	1.92	1.03

Furthermore, the individual roles of the key residues are also illustrated clearly in model A. As mentioned, Glu473, Asp27′, Glu50′, His113′, and His114′ are included in the QM region. Glu50′ always forms a hydrogen bond to the 4′-aminopyrimidine ring of the coenzyme and disperse its negative charge, which would assist the tautomerism of amino and imino. This residue is very important because the enzyme only retained 0.46 % activity upon replacement of Glu50′Gln [19]. The presence of protonated Glu473 and Asp27′ is important in the optimal orientation of the substrate and the stereochemical control of the decarboxylation of LThDP via two hydrogen bonds to the carboxyl group. The mutant Glu473Gln had a very low, but measurable activity of 0.025 % of the wild type [19]. What is more, Asp27′ provides one proton to the enamine intermediate in the third step. Determination of the kinetic properties showed that the affinity of Asp27Ala was decreased 30-fold [20]. The side chain of His113′ is very close to the active center and interacts with the carbonyl group of the substrate. Experiments even detected that His113Glu rendered the enzyme completely inactive [22]. His114′ plays a comparatively minor role during the catalytic process because mutation to Gln resulted in an enzyme that retained 37 % of the wild-type activity [19]. According to our calculation, the most possible protonation state and the specific role of the key residues in catalysis have been illustrated. It should be noted that the previous calculation on small models has been proved valuable, but one of the disadvantages of such calculations is the neglect of the protein environment as well as the different protonation states of the key residues. In our work, we build on the previous knowledge and ameliorate it by using QM/MM methods. And investigation on the effect of the mutations on the barriers is our current ongoing work.

5 Conclusions

In this paper, a theoretical study of the nonoxidative decarboxylation of pyruvate to acetaldehyde catalyzed by PDC was presented, a systematic and detailed mechanism for PDC was determined, and the dependence of reaction mechanism on protonation state of titratable residues in the active site was elucidated. Three models with key residues in their different protonation states result in dissimilar reaction pathways. Compared with the unfavorable energies of two possible mechanisms proposed for model B and C, model A is qualitatively consistent with available experimental observations. In model A, Asp27′, His113′, and His114′ are all defined as their protonated states, and the rate-determining step of decarboxylation undergoes a free energy barrier of 16.2 kcal/mol. Our results will shed light on the development of potent PDC inhibitors and redesign of enzyme activities for biocatalytic applications.

Acknowledgments This work was supported by the National Natural Science Foundation of China (21173129).

References

1. Jordan F (2003) Nat Prod Rep 20:184–201
2. Jordan F, Baburina I, Gao Y, Guo F, Kahyaoglu A, Nemeria N, Volkov A, Yi JZ, Zhang D, Machado R, Guest JR, Furey W, Hohmann S (1996) New insights to the regulation of thiamin diphosphate dependent decarboxylases by substrate and THDP.Mg (II). In: Bisswanger H, Schellenberger A (eds) Biochemistry and physiology of thiamin diphosphate enzymes. A u. C. Intemann Wissenschaftlicher Verlag, Prien, pp 53–69
3. Candy JM, Duggleby RG (1998) Biochim Biophys Acta 1385:323–338
4. Raj KC, Ingram LO, Maupin-Furlow JA (2001) Arch Microbiol 176:443–451
5. Chen GC, Jordan F (1984) Biochemistry 23:3576–3582
6. Hyon SS, Peter LR (1996) Biotechnol Bioeng 49:52–62
7. Rogers PL, Shin HS, Wang B (1997) Adv Biochem Eng Biotechnol 56:33–59
8. Pohl M, Lingen B, Müller M (2002) Chem Eur J 8:5289–5295
9. Mizuhara S, Handler P (1954) J Am Chem Soc 76:571–573
10. Breslow R (1958) J Am Chem Soc 80:3719–3726
11. Schowen RL (1998) Thiamin-dependent enzymes. In: Sinnot M (ed) Comprehensive biological catalysis-A mechanistic reference. Academic Press, London, vol 2, pp 217–266
12. Karimian K, Mohtarami F, Askari M (1981) J Chem Soc Perkin Trans II 2:1538–1543
13. Meyer D, Neumann P, Parthier C, Friedemann R, Nemeria N, Jordan F, Tittmann K (2010) Biochemistry 49:8197–8212
14. Dyda F, Furey W, Swaminathan S, Sax M, Farrenkopf B, Jordan F (1993) Biochemistry 32:6165–6170
15. Dobritzsch D, König S, Schneider G, Lu GG (1998) J Biol Chem 273:20196–20204
16. Arjunan P, Umland T, Dyda F, Swaminathan S, Furey W, Sax M, Farrenkopf B, Gao Y, Zhang D, Jordan F (1996) J Mol Biol 256:590–600
17. Versees W, Spaepen S, Vanderleyden J, Steyaert J (2007) FEBS J 274:2363–2375
18. Franka RAW, Leeper FJ, Luisi BF (2007) Cell Mol Life Sci 64:892–905
19. Huang CY, Chang AK, Nixon PF, Duggleby RG (2001) Eur J Biochem 268:3558–3565
20. Wu YG, Chang AK, Nixon PF, Li W, Duggleby RG (2000) Eur J Biochem 267:6493–6500
21. Chang AK, Nixon PF, Duggleby RG (1999) Biochem J 339:255–260
22. Schenk G, Leeperz FJ, England R, Nixon PF, Duggleby RG (1997) Eur J Biochem 248:63–71
23. Tittmann K, Golbik R, Uhlemann K, Khailova L, Schneider G, Patel M, Jordan F, Chipman DM, Duggleby RG, Hübner G (2003) Biochemistry 42:7885–7891
24. Wang JY, Dong H, Li SH, He HW (2005) J Phys Chem B 109:18664–18672
25. Sakaki S, Musashi Y, Ohkubo K (1993) J Am Chem Soc 115:1515–1519
26. Lie MA, Celik L, Jørgensen KA, Schiøtt B (2005) Biochemistry 44:14792–14806
27. Pei XY, Erixon KM, Luisi BF, Leeper FJ (2010) Biochemistry 49:1727–1736
28. Zhang CH, Guo Y, Xue Y (2011) Theor Chem Acc 129:781–791
29. Smith CR, Smith GK, Yang ZX, Xu DG, Guo H (2011) Theor Chem Acc 128:83–90
30. Wu RB, Wang SL, Zhou NJ, Cao ZX, Zhang YK (2010) J Am Chem Soc 132:9471–9479
31. Gómez H, Polyak I, Thiel W, Lluch JM, Masgrau L (2012) J Am Chem Soc 134:4743–4752
32. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38
33. Frisch MJ et al. (2004) Gaussian 03, Revision C.02. Gaussian, Inc, Wallingford, CT
34. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
35. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187–217
36. Chen JH, Im W, Brooks CL (2006) J Am Chem Soc 128:3728–3736
37. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
38. Berkowitz M, McCammon JA (1982) Chem Phys Lett 90:215–217
39. Sherwood P, de Vries AH, Guest MF, Schreckenbach G, Catlow CRA, French SA, Sokol AA, Bromley ST, Thiel W, Turner AJ, Billeter S, Terstegen F, Thiel S, Kendrick J, Rogers SC, Casci J, Watson M, King F, Karlsen E, Sjovoll M, Fahmi A, Schäfer A, Lennartz C (2003) J Mol Struct (THEOCHEM) 632:1–28
40. Ahlrichs R, Bar M, Haser M, Horn H, Kolmel C (1989) Chem Phys Lett 162:165–169
41. Smith W, Forester TR (1996) J Mol Graph 14:136–141
42. Billeter SR, Turner AJ, Thiel W (2000) Phys Chem Chem Phys 2:2177–2186
43. Bakowies D, Thiel W (1996) J Phys Chem 100:10580–10594
44. Kern D, Kern G, Neef H, Tittmann K, Killenberg-Jabs M, Schneider CW, Hübner G (1997) Science 275:67–70
45. Jordan F, Nemeria NS, Zhang S, Yan Y, Arjunan P, Furey W (2003) J Am Chem Soc 127:12732–12738
46. Friedemanna R, Tittmannb K, Golbikb R, Hübner G (2009) J Mol Catal B Enzym 61:36–38
47. Li H, Robertson AD, Jensen JH (2005) Proteins Struct Funct Bioinf 61:704–721
48. Bas DC, Rogers DM, Jensen JH (2008) Proteins Struct Funct Bioinf 73:765–738