Adaptive Steered Molecular Dynamics Combined With Protein Structure Networks Revealing the Mechanism of Y68I/G109P Mutations That Enhance the Catalytic Activity of D-psicose 3-Epimerase From *Clostridium Bolteae*

Jingxuan Zhu[1], Yi Li[2], Jinzhi Wang[2], Zhengfei Yu[1], Ye Liu[1], Yi Tong[2]* and Weiwei Han[1]*

[1] *Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, School of Life Science, Jilin University, Changchun, China,* [2] *COFCO (Jilin) Bio-Chemical Technology Co., Ltd, Changchun, China*

The scarcity, richness, and other important physiological functions of D-psicose make it crucial to increase the yield of D-psicose. The production of D-psicose can be accomplished by D-psicose 3-epimerase (DPEase) from *Clostridium bolteae* (*Cb*DPEase) catalyzing the substrate D-fructose. Although the catalytic efficiency of the *Cb*DPEase has been raised via using the site-directed mutagenesis (Y68I/G109P) technique, structure-activity relationship in the wild-type *Cb*DPEase and Y68I/G109P mutant is currently poorly understood. In our study, a battery of molecular modeling methods [homology modeling, adaptive steered molecular dynamics (ASMD) simulations, and Molecular Mechanics/Generalized Born Surface Area (MM-GB/SA)], combined with protein structure networks, were employed to theoretically characterize the reasons for the differences in the abilities of the D-fructose catalyzed by the wild-type *Cb*DPEase and Y68I/G109P mutant. Protein structure networks demonstrated that site-directed mutagenesis enhanced the connectivity between D-fructose and *Cb*DPEase, leading to the increased catalytic efficiency mediated by the functional residues with high betweenness. During the dissociation of the D-fructose from the Y68I/G109P mutant, planes of benzene rings of F248 and W114 could be continuously parallel to the stretching direction of D-fructose. It made the tunnel have an open state and resulted in the stable donor-π interactions between D-fructose and the benzene rings around 18Å. The stronger substrate-protein interactions were detected in the Y68I/G109P mutant, instead of in the wild-type *Cb*DPEase, which were consistent with the binding free energy and Potential Mean of Force (PMF) results. The theoretical results illustrated the reasons that Y68I/G109P mutations increased the catalytic efficiency of *Cb*DPEase and could be provided the new clue for further DPEase engineering.

Keywords: D-fructose, D-psicose 3-epimerase (DPEase), catalytic efficiency, molecular modeling, adaptive steered molecular dynamics, protein structure networks

INTRODUCTION

D-psicose, as a C-3 epimer of D-fructose, has plentiful significant physiological functions, covering anti-hyperglycemia and anti-obesity effects (Granström et al., 2004; Fukada et al., 2010; Iida et al., 2010; Zhang et al., 2016; Shintani et al., 2017). It has a sucrose sweetness of 70%, and contributes less calories for its poor absorption in the digestive tract (Granström et al., 2004; Fukada et al., 2010). D-psicose is therefore commonly applied as a functional sugar in the treatment of diabetes and obesity (Ochiai et al., 2013). Furthermore, D-psicose has various physiological characteristics beneficial to health. For example, it can efficaciously reduce the accumulation of abdominal fat (Matsuo et al., 2001; Ochiai et al., 2014), remove reactive oxygen species (ROS) (Murata et al., 2003), and inhibit hepatic lipase activities (Matsuo et al., 2015). Besides, it can reduce the postprandial glycemic excursion (Matsuo and Izumori, 2006; Hayashi et al., 2010), improve the metabolism of serum lipids (Afach et al., 2008), and protect the nervous system (Takata et al., 2005). Also, it can strengthen insulin resistance (Hossain et al., 2011, 2012), treat atherosclerotic diseases and so on (Murao et al., 2007).

D-psicose is rare in nature, and its chemical synthesis is difficult to get (Zhang et al., 2016). Therefore, the technology of D-psicose production by the enzymatic reaction of ketose 3-epimerase has been widely concerned in commercial production (Kim et al., 2006; Yoshida et al., 2007; Mu et al., 2011; Chan et al., 2012). In general, D-fructose can serve as the substrate to produce the D-psicose and characterize the enzymic activity of the ketose 3-epimerase. D-tagatose 3-epimerase (DTEase) isolated from *Pseudomonas cichorii* ST-24 (*P. cichorii* DTEase) (Itoh et al., 1994; Izumori et al., 2008) is the original member of the ketose 3-epimerases. DTEase family enzymes are derived from a variety of microorganisms and their relative sequences are quite similar. It indicates that enzymes, with the significantly enhanced efficiency for D-psicose production, may have a common catalytic mechanism (Zhang et al., 2013) which catalyzes the bioconversion from D-fructose to D-psicose by the epimerization reaction at the C-3 position. Particularly, DTEase from *Agrobacterium tumefaciens* (*A. tumefaciens* DPEase), characterized by Kim et al. was defined as D-psicose 3-epimerase (DPEase) for it was mainly designed for D-psicose. Before long, puzzles concerning crystal structures of *P. cichorii* DTEase and *A. tumefaciens* DPEase were successively resolved. DPEase is a tetramer composed of four TIM-barrel shaped monomers, and the active site in each monomer differs from other TIM-barrel enzymes. The metal ion (Mn^{2+} or Co^{2+}) in the monomer of DPEase is coordinated as an octahedron by two water molecules and four residues (E150, D183, H209, and E244) that are entirely conserved in the DPEase family. In the case of D-fructose binding, the substrate replaces water molecules in the active site to form a conformation of the intermediate cis-enediolate. O1, O2, and O3 of the substrate are interacted with four key residues of the active site.

Recently, Jia et al. reported the characterization study of the wild-type *Clostridium bolteae* DPEase (*Cb*DPEase) (Jia et al., 2014). Also, the catalytic activity of the wild-type *Cb*DPEase

was promoted by mutations in the Y68 and G109 sites (Zhang et al., 2016). Compared with the wild-type, the Y68I/G109P mutant would be more appropriate for the industrial D-psicose production. To explore the structure–function relationships between the substrate and the wild/mutant type *Cb*DPEase, D-fructose was used as the substrate to investigate the mechanism of Y68I/G109P mutations that enhanced the catalytic activity of *Cb*DPEase by using adaptive steered molecular dynamics and the protein structure networks method.

It was very known that molecular recognition and specific ligand-receptor interactions are critical steps in numerous physiological processes, such as enzyme catalysis, the organelle assembly, and energy conversion. It requires the incorporation of non-covalent bonds between the protein and the ligand to induce the realization of its functional response (Deuflhard et al., 1999). And steered molecular dynamics (SMD) simulation can reveal the details of molecular interactions during dissociation, thus providing important information for the binding mechanism. In our study, adaptive steered molecular dynamics (ASMD) simulations were performed to quantitatively calculated the binding energy of the protein and the ligand, and qualitatively analyzed the protein-ligand interactions and conformational changes of the protein induced by ligand. The dissociation of the substrate is the reverse process of molecular binding and molecular recognition. Hence, the results of ASMD simulations could intuitively reflect the affinity of substrate binding. Besides, by analyzing the protein structure networks, we could figure out characteristics of the dual mutation-induced changes of the connectivity in the networks of the protein-ligand complex. The results of this study may be conductive to the exploration of efficient point mutations in the future, and the results also provided the theoretical basis for the discovery of the novel mechanism of protein design for the DPEase family.

MATERIALS AND METHODS

Modeling and Validation of Dpease, and Structure-Based Sequence Alignments

The *Cb*DPEase protein sequence (Genbank ID: EDP19602.1) was obtained from the NCBI server. A similarity search for *Cb*DPEase was performed by the Moddeller 9.20 (Fiser and Sali, 2003) to find potentially related sequences of known structures as suitable templates. Because of the high sequence similarity between *Cb*DPEase and templates (higher than 53%), crystal structures of D-psicose 3-epimerases from *Clostridium cellulolyticum* (*Cc*DPEase) (PDB ID: 3VNK) (Chan et al., 2012) and *Agrobacterium fabrum* (*Af*DPEase) (PDB ID: 2HK1) (Kim et al., 2006) were chosen as the multiple templates for homology modeling. Homology modeling was performed via using Modeler 9.20 (Fiser and Sali, 2003). It generated five models, among which the best one was selected based upon the least DOPE (Discrete Optimized Protein Energy) score (Shen and Sali, 2006). In this rating of the score, a statistical program named the atomic distance-dependent was applied to discern the similarity between templates and models. The lower is the DOPE score, the more reliable is the model (Shen and Sali, 2006). After

predicting the three-dimensional structure of *Cb*DPEase, we finished the validation of the modeled structure via using Structure Analysis and Verification Server (SAVS) to assess the overall stereo chemical quality of the modeled protein (www. mbi.ucla.edu/SAVS). Ramachandran plot was analyzed by the program PROCHECK (Laskowski et al., 1993). We further evaluated the modeled structure of PhMTNA by the use of VERIFY3D and ERRAT, a verification algorithm well-suited for the evaluation of the progress of crystallographic model building and refinement (Bowie et al., 1991).

To compare related structures (templates and models) and perform the interactive communication between sequences and structures, the MatchMaker extension of Chimera (Meng et al., 2006) was used for the structure-based sequence alignments and the structures superimposition. The standard Needleman-Wunsch (Needleman and Wunsch, 1970) algorithms BLOSUM-62 and 30% weighting of the secondary structure term (i.e., 70% weighting of the residue similarity term) were employed as the default settings in global and local sequence alignments production, respectively. Based on the structures and their corresponding sequence alignment suitable for Chimera, the AL2CO histogram (Pei and Grishin, 2001) of conservation style was applied to map sequence conservation onto structures with Chimera. And the secondary structure was presented in Pro-origami web service (Stivala et al., 2011) as a cartoon.

Integrative Visual Analysis for Residue Interaction Networks of Protein-Ligand Complexes

The Gaussian 09 software (Robb et al., 2009) with B3LYP 6-31+G* level was devoted to optimizing the structure of D-fructose. The optimized structures were illustrated in **Figure S6**. Autodock vina (Trott and Olson, 2010) was used to dock the D-fructose into the *Cb*DPEase binding pockets, and the lowest and the most stable model was selected for the further simulations.

The latest study of individual residues and their structure-based functions has demonstrated that residue interaction networks (RINs) obtained from three-dimensional structures are promising to investigate the relationship between structures and functions. In RINs, amino acid residues in the protein are referred to as nodes that are connected by edges over the noncovalent interactions identified by Probe (Word et al., 1999). In our study, the RINs for 3D structures of *Cc*DPEase, *Af*DPEase (templates) and *Cb*DPEase were generated from RINerator software (Doncheva et al., 2011). Using the RINalyzer (Doncheva et al., 2011), the resulting networks and accompanying analysis were visualized in Cytoscape (Shannon et al., 2003) and structureViz apps (Morris et al., 2007). The corresponding structures were displayed by UCSF Chimera (Pettersen et al., 2004; Meng et al., 2006). Specifically speaking, the bridge between biological networks visualized in Cytoscape and corresponding 3D structures was acted by the Cytoscape plugin structureViz.

Molecular Dynamics Simulation

The D-fructose bound to *Cb*DPEase was performed on 300 ns time scales' molecular dynamics simulation using the Amber 14 software package (Case et al., 2005) with TIP3P water molecules (Harrach and Drossel, 2014). For the ligand (D-fructose), generalized AMBER force field (GAFF) parameters and RESP partial charges (Wang et al., 2004; Maier et al., 2015) were assigned using the ANTECH AMBER program, which shipped with Amber Tools (Sousa da Silva and Vranken, 2012). The production MD was accomplished at 333 K and 1 atm for 60 ps with the desired temperature and pressure (NPT) coupled by Berendsen methods (Darden et al., 1993). The MD trajectory was saved and written to the output file every 2 ps. The RMSD plot in **Figure S7** reflected the stabilization of MD simulation for the modeled structure. After MD simulations, the averaged snapshots (within 60–85 ns and 210–265 ns) were chosen to generate the mutant type structures. Y68 was mutated into I68, and G109 was mutated into P109 using Discovery Studio 4.0 client software. The wild-type *Cb*DPEase from 60 to 85 ns and its Y68I/G109P mutant were equilibrated for 10 ns to generate the first two equilibrium trajectories. Besides, the wild-type *Cb*DPEase from 210 to 265 ns and its Y68I/G109P mutant were equilibrated for 10 ns to generate the latter two equilibrium trajectories. They correspond to steps of minimization, heating, density, slowly releasing restraints, and finally an unconstrained molecular dynamics simulation. All the simulations were performed by the same force field and parameter settings as the MD simulations. The trajectories generated from the equilibration were applied in the next Principle component analyses (PCA), network analysis, MM-GB/SA calculations, and CAVER calculations.

Principle Component Analyses and Free Energy Landscape Analysis

The AMBER14 plugin CPPTRAJ module (Lu and Luo, 2003) was used to perform Principle Component Analyses (PCA) for understanding the most significant conformational changes in the equilibration stage of the wild-type *Cb*DPEase and the Y68I/G109P mutant. All representative structures based on disparate conformational motions were extracted by using OpenGUI interface of VMD (Humphrey et al., 1996), and visualized through PyMOL (Delano, 2002). The energies of macromolecule conformations such as the first two principal components (PC1, PC2) were characterized in the Free Energy Landscape (FEL). By using converting dot distribution to probability distribution (ddtpd) program, the trajectories to the PC1 and PC2 of motion were mapped (Nicolaï et al., 2013; Iida et al., 2016; Tripathi et al., 2016).

Protein Structure Network Analysis

The network parameters (clusters, hubs, cliques and communities) were utilized to analyze the residue interaction networks. The clusters, and groups of interacting residues (nodes) (the number of interactions is at least two) in the network were calculated by MCODE algorithm (O'Driscoll et al., 2014) visualized in Cytoscape. The hubs are highly connected nodes in the network with at least 4 associated edges on the node. A $k = n$ clique is a group of "n" nodes, in which each one is connected to every other node in the group. A community of $k = n$ cliques is a collection of cliques that share n-1 nodes

among them. The network parameters for hubs, cliques and communities were computed by using the Clique Percolation Method (Palla et al., 2005) when implemented in the CFinder software (Kang et al., 2014). Edge attribute files were generated from RINerator. The protein structures used in the protein network analysis were the representative structures from 2500 snapshots of equilibration trajectories in the wild-type *Cb*DPEase and the Y68I/G109P mutant.

Network Centrality Analysis

Network Centrality analysis, one of key features of RINalyzer, can compute weighted centrality measures. It incorporates the residue degree, shortest path closeness and betweenness. Betweenness centrality depicts the amount of times that a residue (node) acts as a bridge over the shortest paths between two nodes that are not directly connected. It measures the degree of significance of the bridge node and the weighted of connectivity between nodes in the global network. A node that continually appears in the shortest paths among other nodes (the shortest paths often contain that node) is more capable of facilitating communication among other nodes. The formula is as follows:

$$C_B(i) = \sum_{j<k} \frac{P_{jk}(i)}{P_{jk}} \tag{1}$$

where P_{jk}, determined by the Floyd–Warshall algorithm (Pallottino, 2010), is the number of shortest paths between nodes j and k. $P_{jk}(i)$ is the amount of shortest paths through node i between nodes j and k (Daly and Haahr, 2009).

Here, the Cytoscape plugin RINalyzer (Doncheva et al., 2011) was utilized to compute weighted centrality measures. For shortest paths centralities, the NrInteractions was chosen for attribution as the edge weight. The average weight option was performed to handle multiple edges, and negative weights were ignored in the network. Besides, the cutoff for weighted degree was set to 1 Å. The edge weights in RINs representing distances indicate that if the weight is smaller, the interaction is stronger. Therefore, it is imperative to convert them to distance scores by the max-value option.

Free Energy Calculations

Molecular mechanics Generalized Born surface area (MM-GB/SA) methods (Hou et al., 2011, 2015; Sun et al., 2014a,b) implemented in Amber 14 package were applied to compute the

binding free energy for the protein-ligand complex. For each system, the key coordinates were extracted from equilibrium trajectories in 10 ns after the MD simulations. In addition, the energy contribution of each residue to the protein and ligand interaction was analyzed through the decomposition of the binding free energy. In those methods, the binding free energy (ΔG_{bind}) is deemed as the discrepancy of binding free energies among protein-ligand complex ($G_{complex}$), the receptor ($G_{protein}$) and ligand (G_{ligand}), which is summarized as

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) = E_{MM} + G_{sol} - T\Delta S \tag{2}$$

$$E_{MM} = E_{ele} + E_{vdw} \tag{3}$$

$$G_{sol} = G_{GB} + G_{SA} \tag{4}$$

In the Equations (2) and (3), E_{MM} denotes the gas phase interaction that contains electrostatic (E_{ele}) and van der Waals (E_{vdW}). In the Equation (4), G_{sol} represents solvation energy which separates into an electrostatic/polar solvation free energy (G_{GB}) and a nonpolar solvation free energy (G_{SA}). G_{GB} is obtained from the Generalized Born (GB) method, and G_{SA} is acquired from the molecular solvent accessible surface area (SASA). Moreover, the $-T\Delta S$ value in the Equation (2) signifies the variation of the conformational entropy, which was evaluated from the 200 frames randomly collected from the equilibrium trajectories using the *nmode* program in AMBER 14.0 (Case et al., 2005). In our calculations, the GB model was calculated by GB-OBC1. The atomic radii were both set to the default bondi. The values of the polar solvation free energy (G_{GB}) were set to default ($\varepsilon = 1$) and the nonpolar part (G_{SA}) was calculated by the LCPO method: $G_{SA} = 0.0072 \times \Delta SASA$.

Pathways Identified With CAVER Analyst 2.0

CAVER Analyst 2.0 is normally used to quantify pathways in proteins that approximate buried active sites, and to visualize static and dynamic structures in tunnels and channels in real time (Chovancova et al., 2012). At this point, the dynamic structures (1,000 snapshots) were originated from 10 ns equilibrium trajectores. In the tunnel computation setting, the starting point was set at the position of the ligand (D-fructose). The probe radius and the clustering threshold were set to 1 and 4.5 Å, correspondingly. The shell depth and radius were set to 4 and 3 Å, accordingly. And other parameters used default settings throughout the calculations.

Adaptive Steered Molecular Dynamics (ASMD) Simulations

ASMD simulations (Ozer et al., 2012; Bureau et al., 2016) were performed for wild-type and Y68I/G109P mutant complexes using Amber 14 package (Case et al., 2005). The predetermined

TABLE 1 | The results of Sequence alignment for *Clostridium cellulolyticum* DPEase (PDB ID 3VNK) and Model, *Agrobacterium fabrum* DPEase (PDB ID 2HK1) and Model generated with MatchMaker and Match-Align modules are visualized in UCSF Chimera.

Reference structure	Match structure	N/RMSD	Sequence alignment score
3VNK	Model	153/0.278	1030.5
2HK1	Model	136/0.599	967.1

reaction coordinate was divided into numerous smaller stages, with each one having multiple simulations performed parallelly. In those stages, a single trajectory was selected from simulations where the energy value was closest to the Jarzynski average (JA) (Jarzynski, 1997; Ozer et al., 2012). The swarm of trajectories were contracted into one single JA structure, which can help remove trajectories that may contribute least to the overall Potential Mean of Force (PMF). The PMF calculation can inspect the change of the energy of the system as a function of the geometrical coordinate (the distance between the ligand and the

protein). In the ASMD simulations, the pulling direction was determined with the vector of two atoms in accordance with the top frequently observed pathway identified by CAVER Analyst 2.0. The stretching velocity was 10 Å/ns, and the spring constant k was 40 kcal/(mol × Å^2). The ligand was pulled roughly along the reaction coordinate from 4 to 30 Å. The reaction coordinate could be split into 14 stages, with each one carrying out 24 simulations for each protein-ligand system. When the reaction coordinate reached 30 Å, the ligand was completely dissociated from binding sites of the protein.

FIGURE 1 | (A) The map of sequence conservation onto structures with Chimera. The secondary structure elements of (B) CbDPEase (model), (C) CcDPEase (template), (D) AfDPEase (template) are provided by Pro-origami.
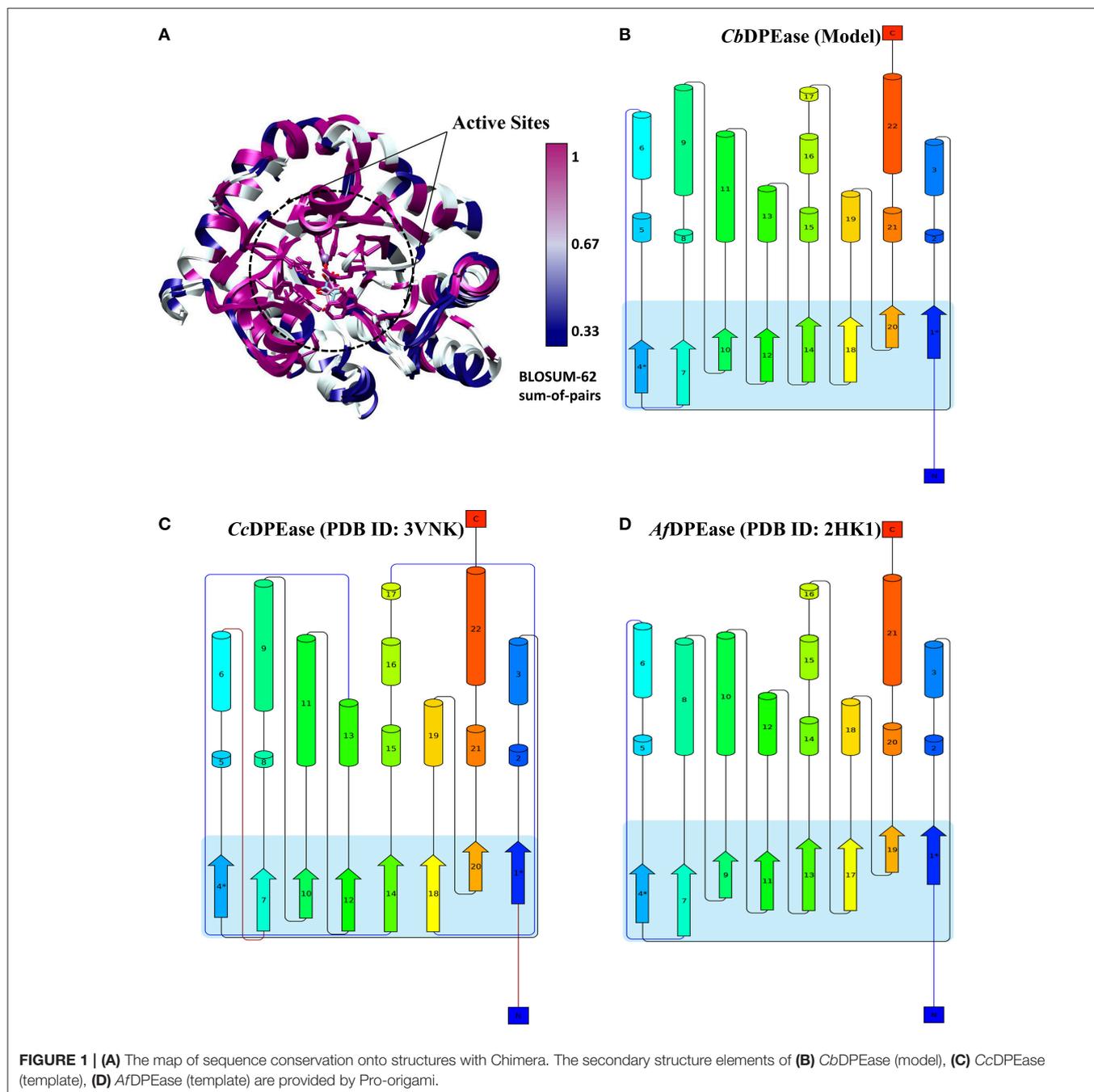
RESULTS AND DISCUSSIONS

Structural Validation of the Modeled DPEase Protein And Structure-Based Sequence Alignments

In our study, crystal structures of *Cc*DPEase (PDB ID: 3VNK) (Chan et al., 2012) and *Af*DPEase (PDB ID: 2HK1) (Kim et al., 2006) bound to D-fructose were chosen as multiple templates to the model. The modeled three-dimensional structure of *Cb*DPEase was selected on the grounds of the least DOPE

score −36561.49. The plotted DOPE score profile for per residue was demonstrated in **Figure S1**, which mirrored the relatively similar energy between models and templates. The psi and phi distribution of the Ramachandran plot for the modeled *Cb*DPEase protein showed that 93.0% of residues were present in the most favorable regions, and 6.2% of residues were in the additional allowed regions. There were 0.4% of the residues in generously allowed regions and 0.4% of them in the disallowed region (**Figure S2**). The average overall quality factor evaluated by ERRAT2 program was 87.633%, which was within

FIGURE 2 | Simultaneous visualization of biological information for templates and model structures using different complementary views. Three-dimensional structures of **(A)** *Cb*DPEase (model), **(B)** *Cc*DPEase (template), and **(C)** *Af*DPEase (template) are shown with UCSF Chimera. Two-dimensional view of the subnetwork of the substrate (Fructose) for **(D)** *Cb*DPEase (model), **(E)** *Cc*DPEase (template), and **(F)** *Af*DPEase (template). The interaction types: *cnt (interatomic contact):* sc-ligand colored by light blue, *combi (combined):* all-all colored by purple, *hbond:* mc-ligand colored by red, *hbond:* sc-ligand colored by orange, *ovl (overlap):* sc-ligand colored by yellow.

the expected range and ensures the model quality (**Figure S3**). The result of Verify3D indicated that 98.97% of residues had averaged 3D-1D score $\geq$ 0.2 (**Figure S4**). These validation studies concluded that bonded, non-bonded interactions and backbone conformations were within the limits of the reliable structure.

To explore the structural similarity between crystal structures of DPEase and modeled structures, a multiple sequence alignment and two superimpositions of the targeted protein and templates were displayed in **Figure S5** and **Table 1**. It showed that (1) the Root Mean Square Deviation (RMSD) of 0.278 Å was over 153 amino acid pairs for CcDPEase and RMSD of 0.599 Å was over 136 amino acid pairs for AfDPEase. It also reflected that the similar sequence alignment scores (1030.5 and 967.1 respectively) were related to the two templates. Moreover, the sequence conservation was mapped onto the structure to compare the conserved regions of templates and the modeled protein. The map in **Figure 1A** showed that structures at the active sites were mostly conserved and associated with the most similar sequence (shown in purple). Then, the secondary structure elements visualized in **Figures 1B–D** gave us a clear depiction of high similarity of chains and the secondary structure elements for templates and the modeled protein. All the above results provided an ample evidence to demonstrate that the modeled protein structurally homologous to the CcDPEase (PDB ID: 3VNK) and the

AfDPEase (PDB ID: 2HK1), showing the reliability for further simulations.

Integrative Visual Analyses for Residue Interactions Networks of Protein Structures

The optimized structure of D-fructose was docked in the model of CbDPEase. Superimposition of molecular docking conformation (blue) and crystal structure (red) was demonstrated in **Figure S6**, which indicated that the substrate was properly positioned in the active site. To survey the rationality of the docking from the perspective of the interaction between the substrate and the protein, we extracted subnetworks of the substrates of CbDPEase, CcDPEase, and AfDPEase (**Figure 2**). It was evident that the active residues around the D-fructose for the three structures were literally alike and conservative, such as W112 in CcDPEase and W112 in AfDPEase corresponding to W124 in CbDPEase ($\text{W112}_{Cc\text{DPEase}}/\text{W112}_{Af\text{DPEase}}/\text{W124}_{Cb\text{DPEase}}$), $\text{E150}_{Cc\text{DPEase}}/\text{E150}_{Af\text{DPEase}}/\text{E162}_{Cb\text{DPEase}}$, $\text{E156}_{Cc\text{DPEase}}/\text{E156}_{Af\text{DPEase}}/\text{E168}_{Cb\text{DPEase}}$, $\text{H186}_{Cc\text{DPEase}}/\text{E186}_{Af\text{DPEase}}/\text{H198}_{Cb\text{DPEase}}$, $\text{H209}_{Cc\text{DPEase}}/\text{E209}_{Af\text{DPEase}}/\text{H221}_{Cb\text{DPEase}}$, $\text{R215}_{Cc\text{DPEase}}/\text{E215}_{Af\text{DPEase}}/\text{R227}_{Cb\text{DPEase}}$, and $\text{E244}_{Cc\text{DPEase}}/\text{E244}_{Af\text{DPEase}}/\text{E256}_{Cb\text{DPEase}}$. Hence, the docking complex for CbDPEase could be used to the next simulations.

FIGURE 3 | The Network centrality analysis of the representative structures from the first two equilibrium trajectories. The frequency distributions of the betweenness values of the **(A)** wild-type structure (green bars), **(B)** Y68I/G109P mutant type structure (pink bars). Residue-based betweenness profiles of the **(C)**, wild-type structure (green) and **(D)** Y68I/G109P mutant type structure (pink). The residue-based betweenness distributions point to small-world organization of the interaction networks in the CbDPEase structures. The distributions of protein structure network parameters are obtained by the representative structures in equilibration stage.

Representative Structures of the Wild-Type and Y68I/G109P Mutant

The Free Energy Landscape (FEL) map shown in **Figures S8A–D** was devoted to characterizing the first two principal components of motion in equilibration stage simulations. From those analyses, the lowest-energy structures of the wild-type *Cb*DPEase and Y68I/G109P mutant from the first two equilibrium trajectories were achieved after 9.3 and 9.9 ns respectively (**Figures S8A,B**). Similarly, based on the latter two equilibrium trajectories of the wild-type *Cb*DPEase and Y68I/G109P mutant (see Materials and Methods), the structures located in the lowest-energy region of FEL, 7.6 and 6.9 ns, were considered as representative structures (**Figures S8C,D**). Consequently, all these structures were selected as representative structures and used for the next calculations (network analysis and ASMD simulations).

Structure-Activity Relationship of the *Cb*DPEase and Y68I/G109P Mutant: Network Centrality Analysis

To recognize crucial mediating nodes (residues) involved in network graphs for the representative structures from the first two equilibrium trajectories, frequency distributions of the

betweenness values were identified as shown in **Figures 3A,B**. It indicated the frequency of nodes lying on paths of other nodes in the networks graphs. The betweenness profile in the Y68I/G109P mutant manifested a greater distribution in low betweenness value (0.01∼0.02), which was a sharper decay and a longer tail of the distribution in comparison with the wild-type *Cb*DPEase. Thus, mutations at Y68 and G109 may trigger a few functional residues to act as significant bridges on the communication between other nodes to promote the catalytic activity of the *Cb*DPEase. To confirm this conjecture, the network centrality for every mediating residue was examined to track the relationship between centrality and functional activities of key residues in the wild-type *Cb*DPEase and Y68I/G109P mutant. Unlike the wild-type *Cb*DPEase, the betweenness in the Y68I/G109P mutant revealed the noticeable peaks in keeping with the functional residues in the active pocket (Y6, E152, L154, R217, and E246), mutated residues (Y68I, G109P) and certain neighboring residues nearby the active sites (F91, L94, L135, L138) (**Figures 3C,D**). Furthermore, highly similar results were also occurred in the betweenness centrality of the representative structures from the latter two equilibrium trajectories (**Figure S9**). Based on the proposed catalytic mechanism and substrates binding studies of *Cc*DPEase bound to D-fructose, the residues mentioned as follows are affordable for

FIGURE 4 | The small-world interaction networks analysis of the *Cb*DPEase (model) structures. **(A)** The distribution of cliques in the wild-type structure (green) and the Y68I/G109P mutant type structure (pink). **(B)** The distribution of communities in the wild-type structure (green) and the Y68I/G109P mutant type structure (pink). The number of community nodes for **(C)**, $k = 3$, **(D)** $k = 4$, **(E)** $k = 5$ in the wild-type structure (green) and the Y68I/G109P mutant type structure (pink). **(F)** The degree distribution of residue hubs in the wild-type structure (green) and the Y68I/G109P mutant type structure (pink). The distributions of network parameters are obtained by the representative structures in equilibration stage.

the catalytic activity of CbDPEase, E244$_{CcDPEase}$/E246$_{CbDPEase}$, which are coordinated with the Co^{2+}. Firstly, it removes a proton from C3, and then E150$_{CcDPEase}$/E152$_{CbDPEase}$ protonates C3 on the opposite side (Chan et al., 2012). And Y6$_{CcDPEase}$/Y6$_{CbDPEase}$, E150$_{CcDPEase}$/E152$_{CbDPEase}$, R215$_{CcDPEase}$/R217$_{CbDPEase}$, and E244$_{CcDPEase}$/E246$_{CbDPEase}$ can make hydrogen bonds with D-fructose (Chan et al., 2012). Moreover, the Y68I/G109P mutant enzyme is more thermostable and it demonstrates outstanding catalytic ability, which is 1.7-fold for t$_{1/2}$ value and 1.2-fold for k$_{cat}$/K$_m$ of the wild-type CbDPEase correspondently (Zhang et al., 2016).

Overall, these results of network centrality analysis suggested that catalytic activities of the wild-type CbDPEase and Y68I/G109P mutant could be mediated by residues in the active pockets. Interestingly, these functional residues could be surrounded and shielded by neighboring residues with high betweenness values, which could enhance the communications of nodes in networks. The betweenness analysis of the Y68I/G109P mutant detected noticeable peaks in several functional residues, which showed functional residues to be more capable of communicating with neighboring nodes in comparison with the wild-type CbDPEase. Thus,

site-directed mutagenesis of Y68 and G109 could enhance the substrate-binding affinity and catalytic efficiency of CbDPEase, which was consistent with experiments (Zhang et al., 2016).

Small-World Interaction Networks Analysis

To detect the close interaction cooperativity from the representative structures of the first two equilibrium trajectories, we gradually measured the small-world interaction networks mediated by functional residues in the following passages.

First, the number of cliques and communities were computed to characterize the stable module of interconnected residues and mutagenesis-induced variations in small-world interaction networks. The mutant cliques graph displayed a greater number of local cliques (Cliques number = 415) than the wild-type (Cliques number = 386) (**Figure 4A**). The more densely interconnected cliques could establish much steadier communities in the Y68I/G109P mutant. The analysis of community (**Figure 4B**) indicated that the total amounts of stable communities in the Y68I/G109P mutant were larger than the wild-type CbDPEase. Interestingly enough, the numbers of interconnecting residues (nodes) in certain

FIGURE 5 | The cluster network analysis. **(A)** General overview and **(B)** a detailed close-up of the network of cluster in the wild-type structure. **(C)** General overview and **(D)** a detailed close-up of the network of cluster in the Y68I/G109P mutant type structure.
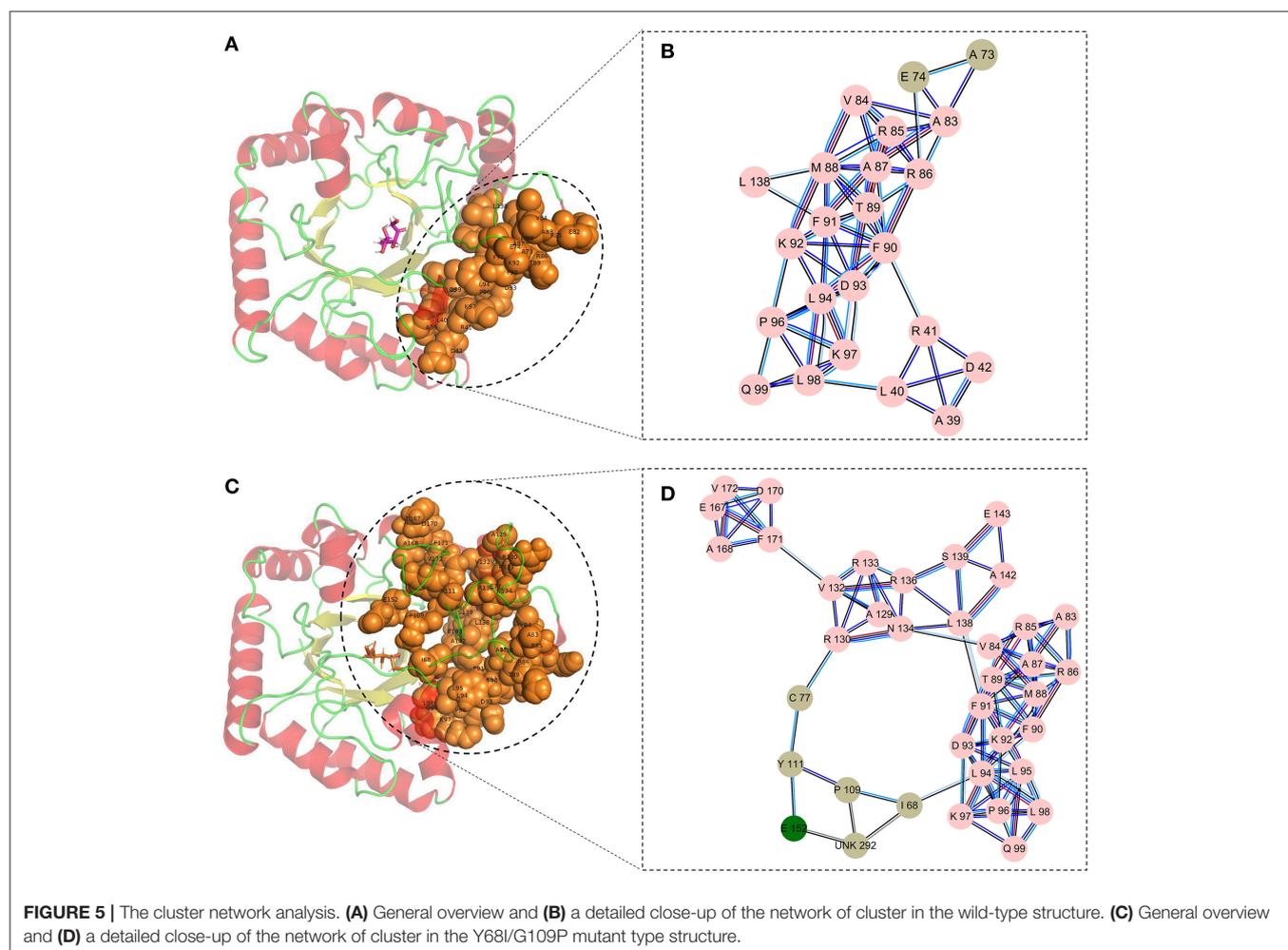
TABLE 2 | Energy terms obtained using MM-GB/SA for the D-fructose bound to the wild-type *Cb*DPEase and Y68I/G109P mutant (kcal/mol).

System	Wild-type *Cb*DPEase	Y68I/G109P mutant
ΔE_{vdW}	−16.67 ± 3.95	−15.61 ± 3.37
ΔE_{ele}	−59.39 ± 10.31	−72.82 ± 7.31
ΔG_{GB}	57.42 ± 4.19	59.62 ± 4.81
ΔG_{SA}	−3.81 ± 0.14	−3.88 ± 0.15
ΔG_{polar}[a]	−1.97	−13.20
$\Delta G_{nonpolar}$[b]	−20.48	−19.49
$\Delta G_{MM-GB/SA}$[c]	−22.44 ± 5.52	−32.69 ± 3.31
$-T\Delta S$	19.95 ± 5.37	19.46 ± 4.95
ΔG_{bind}[d]	−2.49	−13.23
k_{cat}/K_m (min^{-1} mM^{-1})[e]	59.2 ± 2.9	73.7 ± 6.2

[a]$\Delta G_{polar} = \Delta E_{ele} + \Delta G_{GB}$
[b]$\Delta G_{nonpolar} = \Delta E_{vdW} + \Delta G_{SA}$
[c]$\Delta G_{MM-GB/SA} = \Delta E_{ele} + \Delta G_{GB} + \Delta E_{vdW} + \Delta G_{SA}$
[d]$\Delta G_{bind} = \Delta G_{MM-GB/SA} - T\Delta S$
[e]k_{cat}/K_m values obtained from the experimental data (min^{-1} mM^{-1}) (Zhang et al., 2016).

communities consisting of different cliques ($k = 3$, $k = 4$, $k = 5$), compared to wild-type, were distinctly high in Y68I/G109P mutant (portrayed as pink peaks) (**Figures 4C–E**), which indicated the higher cooperativity in the mutant type.

Second, the distribution of local residue hubs in the conformational networks was also characterized and shown in **Figure 4F**. The hubs were highly connected nodes (at least 4) in the network. The result showed that the number of hub nodes increased markedly when the degree of the residue hub was low. Then, the exponentially decayed, as the degree of the residue hub increased in the wild-type *Cb*DPEase and Y68I/G109P mutant, which was in agreement with previous network-based studies of proteins (Atilgan et al., 2003; del Sol et al., 2005). More interestingly, the larger distribution of residue hubs detected in the Y68I/G109P mutant could be consistent with the tighter cooperativity of communities in the mutant protein, which may be the crucial evidence of increased activities of *Cb*DPEase.

Finally, the cluster network analysis was displayed as a detailed map of the residue interaction network that could detect densely interconnected modules in local networks. The map of the cluster analysis suggested that compared with the wild-type *Cb*DPEase (**Figures 5A,B**), the highly interconnected nodes included the substrate, the functional residues (E150$_{CcDPEase}$/E152$_{CbDPEase}$), and the mutant sites (I68 and P109) that could form a larger and denser cluster in the Y68I/G109P mutant (**Figures 5C,D**).

It worth to noting that through the network analysis of the representative structures from the latter two equilibrium trajectories (**Figures S10**, **S11**), we found that for the wild-type *Cb*DPEase and Y68 and G109 mutant, the differences in the number of cliques and communities, the trend of the hubs, and the intensities of the clusters were highly compatible with the above conclusions. Overall, site-directed mutagenesis of Y68 and G109 of the *Cb*DPEase may induce more stable interactions than that of the wild-type *Cb*DPEase in the local network, which corresponded to the high level of cooperativity that may affect the catalytic activity of *Cb*DPEase.

Enhancement of Protein-Ligand Interactions With Site-Directed Mutagenesis of Y68 and G109

Reliable free energy calculations based on the first two equilibrium trajectories were performed to assess the quantitative effects of the binding affinities between the protein and the ligand. As can be seen in **Table 2**, ΔG_{bind} of D-fructose bound to the Y68I/G109P mutant was −13.23 kcal/mol, which was nearly 11 kcal/mol lower than that of wild-type complex (−2.49 kcal/mol). It meant that the binding affinities of the ligand in the Y68I/G109P mutant were stronger than that in the wild-type *Cb*DPEase, which were in accord with the experiment results (the k_{cat}/K_m in the Y68I/G109P mutant increased by 1.2-fold compared with the wild-type) (Zhang et al., 2016). And the binding free energies were driven by the electrostatic interaction (ΔE_{ele}), polar solvation (ΔG_{GB}), the *vdW* interaction (ΔE_{vdW}), nonpolar solvation (ΔG_{SA}), and entropy (–TΔS). To be specific, the contributions of ΔE_{ele} to the binding energies varied substantially in the wild-type *Cb*DPEase (−59.39 kcal/mol) and Y68I/G109P mutant (−72.82 kcal/mol), which were favorable to binding. And the value of conformational entropy (–TΔS) reached 19.95 kcal/mol and 19.46 kcal/mol in the wild-type *Cb*DPEase and Y68I/G109P mutant respectively, which were always unfavorable to binding. Moreover, there was almost no considerable difference in the ΔE_{vdW}, ΔG_{GB}, and ΔG_{SA} between the wild-type *Cb*DPEase and Y68I/G109P mutant. This showed that the *vdW* energy, polar solvation energy, and nonpolar solvation energy did not significantly contribute to the disparity of binding free energy. The free energy contributions were decomposed to the binding free energy in individual residues (**Figure S12**), which indicated that contributions in E152, E264, D185, and H211 of the Y68I/G109P mutant were far lower than that of the wild-type *Cb*DPEase corresponding the active site in *Cb*DPEase.

Furthermore, we applied the second free energy calculation for snapshot structures extracted from the latter two equilibrium trajectories, the results in **Table S1** indicated that the binding free energy of the Y68I/G109P mutant lower than that of the wild-type *Cb*DPEase. And the contributions of non-polar part, polar part, and conformational entropy to the binding free energies were in line with previous results. Conclusion could be drawn from those results that the combination of the D-fructose and the Y68I/G109P mutant was more stable than that of the wild-type *Cb*DPEase.

Identification of Pathways in Equilibration Trajectories

To select the best tunnel candidates for ASMD simulations and identify the discrepancy in representational tunnels between the wild-type *Cb*DPEase and Y68I/G109P mutant, Caver Analyst 2.0 was utilized for the analysis of 1,500 snapshots from equilibration trajectories after MD simulations. All Tunnels Heat Maps (ATH) (Byška et al., 2015), representing the whole tunnels in the molecular dynamics, showed that a tunnel with most frequent occurrence of the Y68I/G109P mutant was detected in tunnel 1 (**Figure S13B**) with the larger bottleneck (the color encodes the

FIGURE 6 | 3D visualization of the two candidate tunnels (tunnel 1) of the **(A)** wild-type *Cb*DPEase and **(B)** Y68I/G109P mutant. The detailed exploration of the tunnel bottleneck contour (collar) over time for the **(C)** wild-type *Cb*DPEase, and **(D)** Y68I/G109P mutant.
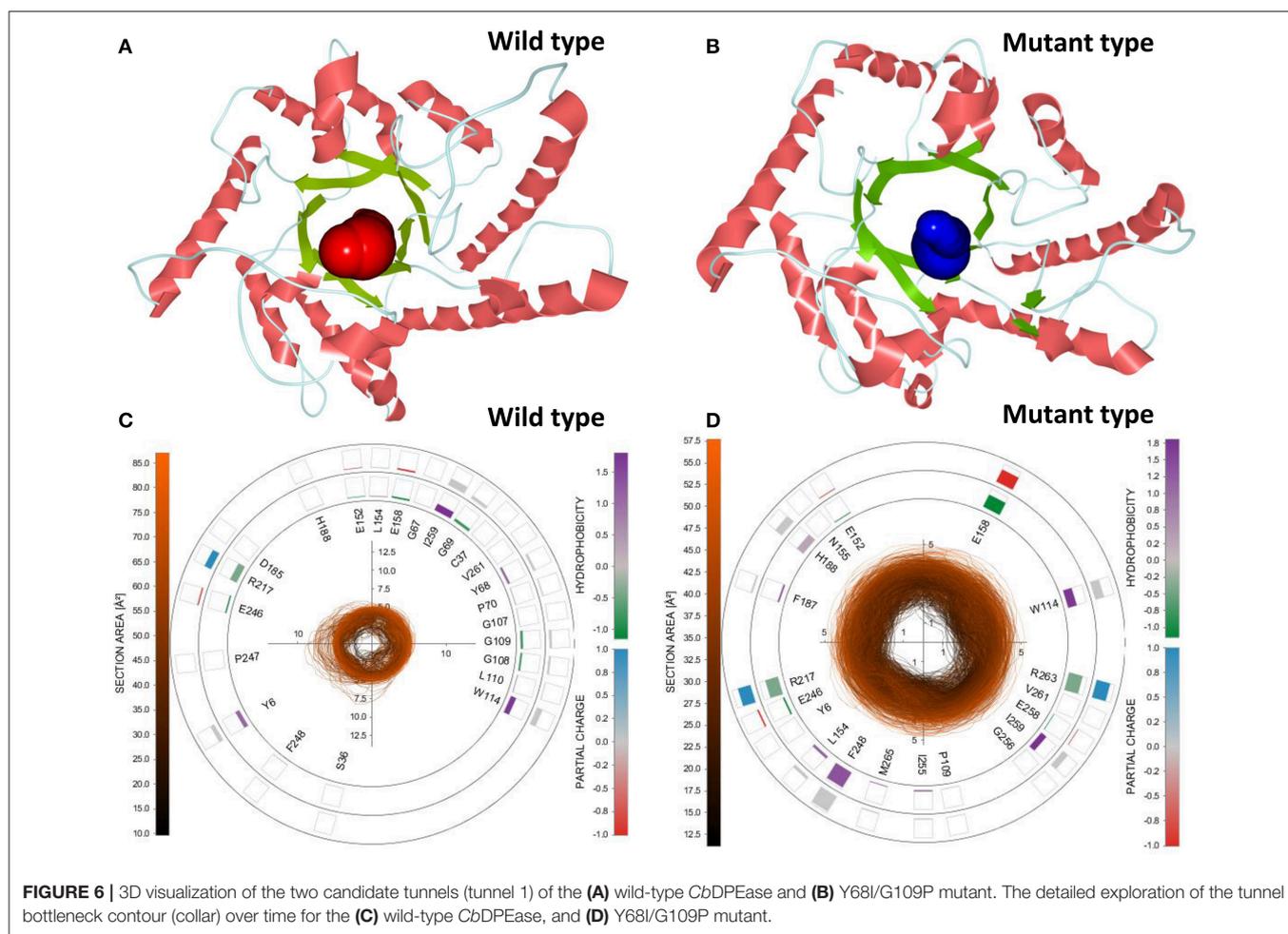
bottleneck size) compared with that of the wild-type *Cb*DPEase (**Figure S13A**). Single Tunnel Heat Maps (STH) (Býška et al., 2015) in **Figures S13C,D** represented the most accessible tunnels (tunnel 1) of the wild-type *Cb*DPEase and Y68I/G109P mutant over time, incorporating the individual rectangles colored by the width of the tunnel. It indicated that tunnel 1 in the Y68I/G109P mutant was longer and wider than the wild-type *Cb*DPEase. The results of ATH and STH manifested that the best tunnel candidate selected in the Y68I/G109P mutant (tunnel 1) may facilitate the unbinding of D-fructose more easily than that in the wild-type *Cb*DPEase. After that, the 3D visualization of the two candidate tunnels (tunnel 1) of the wild-type *Cb*DPEase and Y68I/G109P mutant were shown in **Figures 6A,B**, and the detailed exploration of the tunnel bottleneck and surrounding amino acids over time were presented in 2D collars (**Figures 6C,D**). From the collars, the shape and the area of the bottleneck of the main tunnel in the Y68I/G109P mutant was more stable and wider than that of the wild-type *Cb*DPEase. The surrounding amino acids displayed around the contour demonstrated the frequency they participated in the tunnel boundary in the light of the hydrophobicity and the partial charge of atoms. It was observed that W114, E158, R217, and R263 in the Y68I/G109P mutant had a bigger impact on the tunnel than that

in the wild-type *Cb*DPEase, which enabled the substrate of the outside solvent to easily pass through the tunnel into the buried active pocket of the Y68I/G109P mutant.

Analysis of Unbinding Pathways of the Wild-Type *Cb*DPEase and Y68i/G109p Mutant by ASMD Simulation and PMF Calculations

To explore the protein-ligand interactions and the affinity of the active sites for the D-fructose during the dissociation of ligand in the wild-type *Cb*DPEase and Y68I/G109P mutant, we conducted ASMD simulations on the wild-type *Cb*DPEase and Y68I/G109P mutant bound to D-fructose. Here, PMF profile displayed the energy changes for the departure of the D-fructose from the wild-type *Cb*DPEase and Y68I/G109P mutant, which amounted to approximately 13 and 22 kcal mol^{-1} respectively (**Figure 7**). In the Y68I/G109P mutant, the higher energy barrier should be transferred to totally dissociate of D-fructose from the tunnel. In other words, the D-fructose had stronger binding interactions with the Y68I/G109P mutant than wild-type *Cb*DPEase.

Furthermore, the hydrogen bond interactions between the protein and the ligand were detected. It assisted in the

FIGURE 7 | PMF profiles along the reactions coordinate for wild-type (green) and Y68I/G109P mutant type (pink).

TABLE 3 | The occupancies of hbond interaction occurred between protein and substrate (Fructose) for the Wild-type and the Y68I/G109P Mutant type.

System	Donor	Acceptor	Probability%
WT CbDPEase	FRU: O3	GLU158: OE2	18.17
	FRU: O1	GLU152: OE2	7.67
	FRU: O1	GLU158: OE2	5.17
	FRU: O1	GLU158: OE1	5.17
	FRU: O4	GLU158: OE2	5.58
	HIS188: NE2	FRU: O2	5.67
	FRU: O6	GLU34: OE1	16.33
	GLY109: N	FRU: O2	8.08
Y68I/G109P CbDPEase	FRU: O6	GLU158: OE2	20.92
	FRU: O5	GLU158: OE2	17.83
	FRU: O5	GLU158: OE1	11.33
	FRU: O1	GLU158: OE1	5.67
	FRU: O4	GLU34: OE1	12.25
	FRU: O1	GLU34: OE2	11.50
	FRU: O3	GLU34: OE1	7.83
	FRU: O1	GLU246: OE2	6.00
	FRU: O5	ILE68: O	7.83
	TYR6: OH	FRU: O6	9.67
	FRU: O6	TYR6: OH	9.67
	TRP114: NE1	FRU: O3	8.92
	FRU: O3	TRP114: NE1	8.92
	FRU: O4	TYR157: O	7.17
	FRU: O3	GLU152: OE2	5.17

understanding of the dissociation pathways and provides some vital clues for improving the catalytic efficiency by the site-directed mutagenesis of Y68 and G109. As shown in **Table 3**, in the Y68I/G109P mutant, the likelihood of hydrogen bond between D-fructose and tunnel residues was higher than that of the wild-type. Especially in the Y68I/G109P mutant, E246 coordinated with the Co^{2+} could make stronger hydrogen bond interactions with D-fructose, which may help a proton divert from C3 in D-fructose (Chan et al., 2012). Hereafter, certain hydrophobic residues (e.g., Y6, W114) and polar residues (e.g., E34 and E158) developed previously as crucial active sites for the substrate binding (Chan et al., 2012) were more likely to form hydrogen bond interactions with D-fructose of the Y68I/G109P mutant. During the dissociation, the number of hydrogen bonds that changed between D-fructose and tunnel residues along the reaction coordinate were shown in **Figure 8**. Interestingly, there was a sudden increase in the amount of hydrogen bonds in the Y68I/G109P mutant after 14 Å along the reaction coordinate, resulting in the sharp increase of free energy value. Thereafter, the number of hydrogen bonds continued to be more than that of the wild-type CbDPEase. Hence, the stronger protein-substrate interactions (hydrogen bond) in the Y68I/G109P mutant may facilitate the approach of the substrate (D-fructose) to the active pocket and induce a more efficacious catalytic reaction.

Now, the protein-ligand interactions for the wild-type CbDPEase and Y68I/G109P mutant along the unbinding pathways were generated from protein subnetworks and displayed in the protein structure (**Figures 9**, **10**). Possible interactions type occurred in the subnetwork were interatomic contact (blue/green) (cnt), hydrogen bonds (red) (hbond), the overlap (yellow) (ovl), and the combined (black) (combi).

The detailed analysis along the reaction coordinate for the unbinding pathway of the wild-type CbDPEase was displayed in **Figure 9**. First, for the initial coordinate (RC = 4 Å), the representative subnetwork generated from the wild-type CbDPEase (**Figure 9A**) showed that D-fructose was tightly fixed

FIGURE 8 | The number of hydrogen bonds change between D-fructose and tunnel residues along the reaction coordinate in the (A) wild-type CbDPEase and (B) Y68I/G109P mutant.

by hydrogen bond interactions with E34, Y68, G108, and G109. Also, it was fastened by functional residues (Y6, E34, G67, Y68, G69, G108, G109, E152, M183, H211, and E246) through the interatomic contact and combined interactions. Then at 7.5 Å, free energy value was experiencing a sharp increase due to the stronger hydrogen bond interactions

FIGURE 9 | Representative structures in the dissociation process of the D-fructose from the wild-type *Cb*DPEase during ASMD simulations. The detailed location of **(A–H)**, along the dissociation process at 4, 7.5, 10, 13, 15, 17, 20, and 24 Å, respectively. The residues forming hydrogen bonds (red dashed line) in ligand dissociation are highlighted in green. The residues participating in the other interactions are highlighted in blue. The D-fructose is highlighted in yellow. Subnetwork analyses of the D-fructose are displayed as two-dimensional view. Possible interactions type occurred in ligand dissociation are interatomic contact (blue/green), hydrogen bond (red), overlap (yellow), and combined (black).

between the D-fructose and the residues in active pockets (E34, G108, G109, E152, E158, and E246) than at 4 Å stage (**Figure 9B**). With the movement of the D-fructose, at 10∼16 Å, relatively stable hydrogen bond interactions resulted in a steady increase in the free energy. Nevertheless, the residues that formed hydrogen bonds with the substrate were constantly varying. For example, residues G109, E152 and H188 at 10 Å (**Figure 9C**), residues H188, E158, and E246 at 13 Å (**Figure 9D**), and residues E158, Q251, and R263 at 15 Å (**Figure 9E**) all altered frequently. Thereafter, the interconnections including the hydrogen bonds between tunnel residues and D-fructose

declined rapidly after ∼16 Å (**Figures 9F–H**), giving rise to the sharp decrease in free energy value. It was worth noting that E158, as an important tunnel residue, could continue to generate the hydrogen bond with the D-fructose during 13–17 Å, which was line with the tunnel analysis results and the hydrogen bond probability. Finally, at ∼28 Å, the D-fructose completely departed from the unbinding pathways of the wild-type *Cb*DPEase and the curves of PMF tended to be flat.

The D-fructose dissociation from the Y68I/G109P mutant was shown in **Figure 10**. At 4 Å along the reaction coordinate

FIGURE 10 | Representative structures in the dissociation process of the D-fructose from the Y68I/G109P mutant during ASMD simulations. The detailed location of **(A–H)**, along the dissociation process at 4, 7.5, 10, 13, 15, 18, 25, and 26Å, respectively. The residues forming hydrogen bonds (red dashed line) in ligand dissociation are highlighted in green. The residues forming Donor-π interactions (purple dashed line) in ligand dissociation are highlighted in orange. The residues participating in the other interactions are highlighted in cyan. The D-fructose is highlighted in yellow. Subnetwork analyses of the D-fructose are displayed as two-dimensional view. Possible interactions type occurred in ligand dissociation are interatomic contact (blue/green), hydrogen bond (red), overlap (yellow), and combined (black).

(**Figure 10A**), D-fructose was tightly fixed by hydrogen bond interactions with E34, I35, I68, G107, G108, P109, and E246, which were stronger than the wild-type *Cb*DPEase at the identical coordinate. Moreover, D-fructose was also fastened by functional residues in active sites (Y6, E34, I35, G67, G69, G107, G108, E152, and E246) and mutated residues (I68 and P109) through the interatomic contact and combined interactions. Then at 7.5 Å, the D-fructose was wrapped by hydrogen bond interactions with E34, G108, E152, and E246, which were weaker than the wild-type *Cb*DPEase, resulting in the decreased free

energy value (**Figure 10B**). Noticeably, at the same coordinate, the interatomic and combined interactions in the Y68I/G109P mutant were much smaller than that of the wild-type. At 10 Å, the regeneration of interatomic contact, combined with interactions and hydrogen bond interactions in Y6 and I68 (**Figure 10C**), was characterized by a sudden increase in free energy value. At 13 Å, D-fructose could merely form hydrogen bond interactions with one residue of E246, and the interatomic contact and combined interactions between the substrate and the tunnel residues were also decreased. Thus, the free energy located at a platform

FIGURE 11 | (A) The movements of W114 and F248 of the wild-type *Cb*DPEase at 13 Å (light green) and 24 Å (green), of the Y68I/G109P mutant at 13 Å (light pink) and 24 Å (pink). **(B)** The dihedral changes of F248 of the wild-type *Cb*DPEase and Y68I/G109P mutant during ASMD simulations. **(C)** The dihedral changes of W114 of the wild-type *Cb*DPEase and Y68I/G109P mutant during ASMD simulations. **(D)** The distance between the centroid of benzene rings of W114 and F248 in the wild-type *Cb*DPEase and Y68I/G109P mutant.

level (**Figure 10D**). After this plateau period of free energy, a rapid increase occured because of the strong hydrogen bond interactions between the substrate and the tunnel residues. For instance, D-fructose could form hydrogen bonds with Y6, E246 and R263 at 15 Å (**Figure 10E**). After that, at 18 Å, D-fructose formed certain new hydrogen bonds with W114, E158, H188, and F248 (**Figure 10F**). With the movement of the D-fructose, the substrate formed two donor-π interactions with F248 and W114. As illustrated in **Figures 11A–C**, the rotation motions of aromatic rings of F248 and W114 along the dissociation in the wild-type *Cb*DPEase and Y68I/G109P mutant were quantified. It was apparent that the benzene ring of the F248 in the wild-type *Cb*DPEase flipped in a great fluctuant way. But in the Y68I/G109P mutant, the dihedral angle of the F248 was stable at around 0°. Similarly, for the dihedrals angle of W114, the angle change of the wild-type was also larger than that of the mutant, and the dihedral angle of the mutant was steady at about 140°. From the structure diagram (**Figure 11A**), we could clearly see that F248 and W114 located at the entrance tunnel, which could act as the gatekeeper to regulate the opening of tunnel. The planes of benzene rings of F248 and W114 in the Y68I/G109P mutant could be continuously parallel to the stretching direction of the D-fructose, giving rise to the open of the gate of "entrance" in the tunnel. Such open state was favorable for the D-fructose to form the stable donor-π interactions with the benzene rings.

Moreover, in the Y68I/G109P mutant, the distance between F248 and W114 in the center of the benzene rings was larger than that of the wild-type, which conformed to the open state of "entrance" in the tunnel (**Figures 11A,D**). Then, at 25 Å, the free energy value had a rapid increase because of the strong hydrogen bond between D-fructose and residues W114, V116, Y157, E158, and R263 (**Figure 10G**). Lastly, the free energy was increased continually until the D-fructose was fully dissociated after 28 ns.

We could come to the following conclusions that D-fructose had the stronger binding affinity with the Y68I/G109P mutant than with the wild-type *Cb*DPEase. Moreover, analyses of interactions could enable us to explore the binding/unbinding pathways and provided vital explanations to the catalytic mechanism of *Cb*DPEase at the molecular structure level.

CONCLUSIONS

The catalytic efficiency of the *Cb*DPEase has been enhanced via using the site-directed mutagenesis (Y68I/G109P) technique. To comprehend the structure-activity relationship in the wild-type *Cb*DPEase and Y68I/G109P mutant, we applied molecular modeling methods with protein structure networks for the substrate D-fructose bound to the wild-type *Cb*DPEase

and Y68I/G109P mutant, correspondently. Protein structure network analyses and network-based centrality detected a great connectivity between D-fructose and the Y68I/G109P mutant through the high betweenness of several functional residues in the global interaction network. Moreover, during the dissociation of the D-fructose from the Y68I/G109P mutant, the planes of benzene rings of F248 and W114 could be continuously parallel to the stretching direction of the D-fructose. It was equivalent to opening the gate of "entrance" in the tunnel to trigger the formation of the stable donor-π interactions between the D-fructose and the benzene rings around 18 Å. Furthermore, the stronger substrate-protein interactions for the Y68I/G109P mutant were detected in comparison with the wild-type CbDPEase, which was in accordance with the binding free energy and PMF results. We could draw a conclusion that site-directed mutagenesis of Y68 and G109 would improve the substrate-binding affinity of CbDPEase and the opening "entrance" should be responsible for the outstanding catalytic efficiency of the Y68I/G109P mutant.

AUTHOR CONTRIBUTIONS

JZ wrote and revised this paper. YL and JW prepared the tables. ZY made the **Supplementary Materials**. YL provided some revision advice. YT and WH provided the ideas and modified the papers.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2018.00437/full#supplementary-material

REFERENCES

Afach, G., Kawanami, Y., Cheetangdee, N., Fukada, K., and Izumori, K. (2008). Lipase-catalyzed synthesis of d -psicose fatty acid diesters and their emulsification activities. *J. Am. Oil Chem. Soc.* 85, 755–760. doi: 10.1007/s11746-008-1242-x

Atilgan, A. R., Akan, P., and Baysal, C. (2003). Small-world communication of residues and significance for protein dynamics. *Biophys. J.* 86, 85–91. doi: 10.1016/S0006-3495(04)74086-2

Bowie J.U., Luthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170. doi: 10.1126/science.1853201

Bureau, H. R., Hershkovits, E., Quirk, S., and Hernandez, R. (2016). Determining the energetics of small β-sheet peptides using adaptive steered molecular dynamics. *J. Chem. Theory Comput.* 12, 2028–2037. doi: 10.1021/acs.jctc.5b01110

Byška, J., Juršík, A., Gröller, M. E., Viola, I. and Kozlíková, B. (2015). MoleCollar and tunnel heat map visualizations for conveying spatio-temporo-chemical properties across and along protein voids. *Comput. Graph. Forum* 34, 1–10. doi: 10.1111/cgf.12612

Case, D., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K., et al. (2005). The amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688. doi: 10.1002/jcc.20290

Chan, H. C., Zhu, Y., Hu, Y., Ko, T. P., Huang, C. H., Ren, F., et al. (2012). Crystal structures of d-psicose 3-epimerase from *Clostridium cellulolyticum* H10 and its complex with ketohexose sugars. *Protein Cell* 3, 123–131. doi: 10.1007/s13238-012-2026-5

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., et al. (2012). CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* 8:e1002708. doi: 10.1371/journal.pcbi.1002708

Daly, E. M., and Haahr, M. (2009). Social network analysis for information flow in disconnected delay-tolerant MANETs. *IEEE Trans. Mobile Comput.* 8, 606–621. doi: 10.1109/TMC.2008.161

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092. doi: 10.1063/1.464397

del Sol, A., Fujihashi, H., and O'Meara, P. (2005). Topology of small-world networks of protein–protein complex structures. *Bioinformatics* 21, 1311–1315. doi: 10.1093/bioinformatics/bti167

Delano, W. (2002). PyMOL molecular graphics system: an open-source molecular graphics tool. *CCP4 Newslett Protein Crystallogr.* 40, 82–92.

Deuflhard, P., Hermans, J., Leimkuhler, B., Mark, A. E., Reich, S., and Robert, D. (1999). *Computational Molecular Dynamics: Challenges, Methods, Ideas.* Berlin; Heidelberg: Springer.

Doncheva, N. T., Klein, K., Domingues, F. S., and Albrecht, M. (2011). Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.* 36, 179–182. doi: 10.1016/j.tibs.2011.01.002

Fiser, A., and Sali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol.* 374, 461–491. doi: 10.1016/S0076-6879(03)74020-8

Fukada, K., Ishii, T., Tanaka, K., Yamaji, M., Yamaoka, Y., Ken-Ichi, K., et al. (2010). Crystal structure, solubility, and mutarotation of the rare monosaccharide D-psicose. *Bull. Chem. Soc.* 84:678. doi: 10.1246/bcsj.20100148

Granström, T. B., Takata, G., Tokuda, M., and Izumori, K. (2004). Izumoring: a novel and complete strategy for bioproduction of rare sugars. *J. Biosci. Bioeng.* 97, 89–94. doi: 10.1016/S1389-1723(04)70173-5

Harrach, M. F., and Drossel, B. (2014). Structure and dynamics of TIP3P, TIP4P, and TIP5P water near smooth and atomistic walls of different hydroaffinity. *J. Chem. Phys.* 140:174501. doi: 10.1063/1.4872239

Hayashi, N., Iida, T., Yamada, T., Okuma, K., Takehara, I., Yamamoto, T., et al. (2010). Study on the postprandial blood glucose suppression effect of D-psicose in borderline diabetes and the safety of long-term ingestion by normal human subjects. *Agri. Biol. Chem.* 74, 510–519. doi: 10.1271/bbb.90707

Hossain, A., Yamaguchi, F., Matsunaga, T., Hirata, Y., Kamitori, K., Dong, Y., et al. (2012). Rare sugar D-psicose protects pancreas β-islets and thus improves insulin resistance in OLETF rats. *Biochem. Biophys. Res. Commun.* 425, 717–723. doi: 10.1016/j.bbrc.2012.07.135

Hossain, M. A., Kitagaki, S., Nakano, D., Nishiyama, A., Funamoto, Y., Matsunaga, T., et al. (2011). Rare sugar D-psicose improves insulin sensitivity and glucose tolerance in type 2 diabetes otsuka long-evans tokushima fatty (OLETF) rats. *Biochem. Biophys. Res. Commun.* 405, 7–12. doi: 10.1016/j.bbrc.2010.12.091

Hou, T., Wang, J., Li, Y., and Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* 51, 69–82. doi: 10.1021/ci100275a

Hou, T., Wang, J., Li, Y., and Wang, W. (2015). Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J. Comput. Chem.* 32, 866–877. doi: 10.1002/jcc.21666

Humphrey, W. F., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph. J. Mol. Graph.* 14, 27–28. doi: 10.1016/0263-7855(96)00018-5

Iida, S., Mashimo, T., Kurosawa, T., Hojo, H., Muta, H., Goto, Y., et al. (2016). Variation of free-energy landscape of the p53 C-terminal domain induced by acetylation: enhanced conformational sampling. *J. Comput. Chem.* 37, 2687–2700. doi: 10.1002/jcc.24494

Iida, T., Hayashi, N., Yamada, T., Yoshikawa, Y., Miyazato, S., Kishimoto, Y., et al. (2010). Failure of d-psicose absorbed in the small intestine to metabolize into energy and its low large intestinal fermentability in humans. *Metab. Clin. Exp.* 59, 206–214. doi: 10.1016/j.metabol.2009.07.018

Itoh, H., Okaya, H., Khan, A. R., Tajima, S., Hayakawa, S., and Izumori, K. (1994). Purification and characterization of D-tagatose 3-epimerase from pseudomonas sp. ST-24. *J. Agri. Chem. Soc.* 58, 2168–2171.

Izumori, K., Khan, A. R., Okaya, H., and Tsumura, T. (2008). A new enzyme, D-ketohexose 3-epimerase, from pseudomonas sp. ST-24. *J. Agri. Chem. Soc.* 57, 1037–1039. doi: 10.1271/bbb.57.1037

Jarzynski, C. (1997). A nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* 78, 2690–2693. doi: 10.1103/PhysRevLett.78.2690

Jia, M., Mu, W., Chu, F., Zhang, X., Jiang, B., Zhou, L. L., et al. (2014). A D-psicose 3-epimerase with neutral pH optimum from *Clostridium bolteae* for D-psicose production: cloning, expression, purification, and characterization. *Appl. Microbiol. Biotechnol.* 98, 717–725. doi: 10.1007/s00253-013-4924-8

Kang, Y. B., Haghighi, P. D., and Burstein, F. (2014). CFinder: an intelligent key concept finder from text for ontology development. *Expert Syst. Appl.* 41, 4494–4504. doi: 10.1016/j.eswa.2014.01.006

Kim, K., Kim, H. J., Oh, D. K., Cha, S. S., and Rhee, S. (2006). Crystal structure of d-Psicose 3-epimerase from *Agrobacterium tumefaciens* and its complex with true substrate d-fructose: a pivotal role of metal in catalysis, an active site for the non-phosphorylated substrate, and its conformational changes. *J. Mol. Biol.* 361, 920–931. doi: 10.1016/j.jmb.2006.06.069

Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/S0021889892009944

Lu, Q., and Luo, R. (2003). A Poisson–Boltzmann dynamics method with nonperiodic boundary condition. *J. Chem. Phys.* 119, 11035–11047. doi: 10.1063/1.1622376

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255

Matsuo, T., Baba, Y., Hashiguchi, M., Takeshita, K., Izumori, K., and Suzuki, H. (2001). Less body fat accumulation with D-psicose diet versus D-fructose diet. *J. Clin. Biochem. Nutri.* 30, 55–65. doi: 10.3164/jcbn.30.55

Matsuo, T., Baba, Y., Hashiguchi, M., Takeshita, K., Izumori, K., and Suzuki, H. (2015). Dietary D-psicose, a C-3 epimer of D-fructose, suppresses the activity of hepatic lipogenic enzymes in rats. *Asia Pacific J. Clin. Nutri.* 10, 233–237. doi: 10.1046/j.1440-6047.2001.00246.x

Matsuo, T., and Izumori, K. (2006). Effects of dietary D-psicose on diurnal variation in plasma glucose and insulin concentrations of rats. *Biosci. Biotechnol. Biochem.* 70, 2081–2085. doi: 10.1271/bbb.60036

Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* 7:339. doi: 10.1186/1471-2105-7-339

Morris, J., Huang, C., Babbitt, P., and Ferrin, T. (2007). structureViz: linking Cytoscape and UCSF Chimera. *Bioinformatics* 23, 2345–2347. doi: 10.1093/bioinformatics/btm329

Mu, W., Chu, F., Xing, Q., Yu, S., Zhou, L., and Jiang, B. (2011). Cloning, expression, and characterization of a d-Psicose 3-epimerase from *Clostridium cellulolyticum* H10. *J. Agric. Food Chem.* 59, 7785–7792. doi: 10.1021/jf201356q

Murao, K., Yu, X., Cao, W. M., Imachi, H., Chen, K., Muraoka, T., et al. (2007). d-Psicose inhibits the expression of MCP-1 induced by high-glucose stimulation in HUVECs. *Life Sci.* 81, 592–599. doi: 10.1016/j.lfs.2007.06.019

Murata, A., Sekiya, K., Watanabe, Y., Yamaguchi, F., Hatano, N., Izumori, K., et al. (2003). A novel inhibitory effect of D-allose on production of reactive oxygen species from neutrophils. *J. Biosci. Bioeng.* 96, 89–91. doi: 10.1016/S1389-1723(03)90104-6

Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to search for similarities in amino acid sequence of 2 proteins. *J. Mol. Biol.* 48, 443–453. doi: 10.1016/0022-2836(70)90057-4

Nicolaï, A., Delarue, P., and Senet, P. (2013). Decipher the mechanisms of protein conformational changes induced by nucleotide binding through free-energy landscape analysis: ATP binding to Hsp70. *PLoS Comput. Biol.* 9:e1003379. doi: 10.1371/journal.pcbi.1003379

Ochiai, M., Nakanishi, Y., Yamada, T., Iida, T., and Matsuo, T. (2013). Inhibition by dietary D-psicose of body fat accumulation in adult rats fed a high-sucrose diet. *J. Agri. Chem. Soc.* 77, 1123–1126. doi: 10.1271/bbb.130019

Ochiai, M., Onishi, K., Yamada, T., Iida, T., and Matsuo, T. (2014). D-psicose increases energy expenditure and decreases body fat accumulation in rats fed a high-sucrose diet. *Int. J. Food Sci. Nutri.* 65, 245–250. doi: 10.3109/09637486.2013.845653

O'Driscoll, P., Merényi, E., Karmonik, C., and Grossman, R. (2014). SOM and MCODE methods of defining functional clusters in MRI of the brain. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2014, 734–737. doi: 10.1109/EMBC.2014.6943695

Ozer, G., Quirk, S., and Hernandez, R. (2012). Adaptive steered molecular dynamics: validation of the selection criterion and benchmarking energetics in vacuum. *J. Chem. Phys.* 136:215104. doi: 10.1063/1.4725183

Palla, G., Derényi, I., Farkas, I., and Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435, 814–818. doi: 10.1038/nature03607

Pallottino, S. (2010). Shortest-path methods: complexity, interrelations and new propositions. *Networks* 14, 257–267. doi: 10.1002/net.3230140206

Pei, J., and Grishin, N. V. (2001). AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* 17, 700–712. doi: 10.1093/bioinformatics/17.8.700

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Robb, M., Scuseria, G. E., Ehara, M., Honda, Y., Scalmani, G., Iyengar, S., et al. (2009). *Gaussian 09, Revision, A.*02. Wallingford, CT.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* Wallingford, CT, 13, 2498–2504. doi: 10.1101/gr.1239303

Shen, M. Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524. doi: 10.1110/ps.062416606

Shintani, T., Sakoguchi, H., Yoshihara, A., Izumori, K., and Sato, M. (2017). d-Allulose, a stereoisomer of d-fructose, extends Caenorhabditis elegans lifespan through a dietary restriction mechanism: a new candidate dietary restriction mimetic. *Biochem. Biophys. Res. Commun.* 493, 1528–1533. doi: 10.1016/j.bbrc.2017.09.147

Sousa da Silva, A. W., and Vranken, W. F. (2012). ACPYPE-AnteChamber PYthon Parser interfacE. *BMC Res. Notes* 5:367. doi: 10.1186/1756-0500-5-367

Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C., and Stuckey, P. J. (2011). Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics* 27, 3315–3316. doi: 10.1093/bioinformatics/btr575

Sun, H., Li, Y., Shen, M., Tian, S., Xu, L., Pan, P., et al. (2014a). Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Phys. Chem. Chem. Phys.* 16, 22035–22045. doi: 10.1039/C4CP03179B

Sun, H., Li, Y., Tian, S., Xu, L., and Hou, T. (2014b). Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys. Chem. Chem. Phys.* 16, 16719–16729. doi: 10.1039/C4CP01388C

Takata, M. K., Yamaguchi, F., Nakanose, K., Watanabe, Y., and Hatano, N. (2005). Neuroprotective effect of D-Psicose on 6-hydroxydopamine-induced apoptosis in rat pheochromocytoma(PC 12)Cells. *J. Biosci. Bioeng.* 100, 511–516. doi: 10.1263/jbb.100.511

Tripathi, S., Srivastava, G., and Sharma, A. (2016). Molecular dynamics simulation and free energy landscape methods in probing L215H, L217R and L225M $\hat{I}^2$I-tubulin mutations causing paclitaxel resistance in cancer

cells. *Biochem. Biophys. Res. Commun.* 476, 273–279. doi: 10.1016/j.bbrc.2016. 05.112

Trott, O., and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi: 10.1002/jcc.21334

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174. doi: 10.1002/jcc.20035

Word, J. M., Lovell, S. C., Labean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., et al. (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* 285, 1711–1733. doi: 10.1006/jmbi.1998.2400

Yoshida, H., Yamada, M., Nishitani, T., Takada, G., Izumori, K., and Kamitori, S. (2007). Crystal structures of D-tagatose 3-epimerase from *Pseudomonas cichorii* and its complexes with D-tagatose and D-fructose. *J. Mol. Biol.* 374, 443–453. doi: 10.1016/j.jmb.2007.09.033

Zhang, W., Fang, D., Zhang, T., Zhou, L., Jiang, B., and Mu, W. (2013). Characterization of a metal-dependent D-psicose 3-epimerase from a novel strain, *Desmospora* sp. 8437. *J. Agri. Food Chem.* 61, 11468–11476. doi: 10.1021/jf4035817

Zhang, W., Jia, M., Yu, S., Zhang, T., Zhou, L., Jiang, B., et al. (2016). Improving the thermostability and catalytic efficiency of the d-Psicose 3-epimerase from clostridium bolteae ATCC BAA-613 using site-directed mutagenesis. *J. Agric. Food Chem.* 64, 3386–3393. doi: 10.1021/acs.jafc.6b 01058

Conflict of Interest Statement: YL, JW, and YT were employed by company COFCO (Jilin) Bio-Chemical Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.