Acta Pharmacologica Sinica 2017: 1673–1682
 1671-4083/17
www.nature.com/aps

Original Article

Conformation and dynamics of the C-terminal region in human phosphoglycerate mutase 1

Shi-en LIU[1, 3], Jun-chi HU[1, 3], Hao ZHANG[1, 3], Pan XU[1, 3], Wei WAN[1, 3], Ming-yue ZHENG[1, 3], Kun-qian YU[1, 3], Hong DING[1], Hua-liang JIANG[1, 3], Lu ZHOU[2, *], Cheng LUO[1, 3, *]

[1]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; [2]School of Pharmacy, Fudan University, Shanghai 201203, China; [3]University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

Phosphoglycerate mutase 1 (PGAM1), an important enzyme in glycolysis, is overexpressed in a number of human cancers, thus has been proposed as a promising metabolic target for cancer treatments. The C-terminal portion of the available crystal structures of PGAM1 and its homologous proteins is partially disordered, as evidenced by weak electron density. In this study, we identified the conformational behavior of the C-terminal region of PGAM1 as well as its role during the catalytic cycle. Using the PONDR-FIT server, we demonstrated that the C-terminal region was intrinsically disordered. We applied the Monte Carlo (MC) method to explore the conformational space of the C-terminus and conducted a series of explicit-solvent molecular dynamics (MD) simulations, and revealed that the C-terminal region is inherently dynamic; large-scale conformational changes in the C-terminal segment led to the structural transition of PGAM1 from the closed state to the open state. Furthermore, the C-terminal segment influenced 2,3-bisphosphoglycerate (2,3-BPG) binding. The proposed swing model illustrated a critical role of the C-terminus in the catalytic cycle through the conformational changes. In conclusion, the C-terminal region induces large movements of PGAM1 from the closed state to the open state and influences cofactor binding during the catalytic cycle. This report describes the dynamic features of the C-terminal region in detail and should aid in design of novel and efficient inhibitors of PGAM1. A swing mechanism of the C-terminal region is proposed, to facilitate further studies of the catalytic mechanism and the physiological functions of its homologues.

Keywords: phosphoglycerate mutase 1 (PGAM1); human cancers; metabolic target; C-terminal region; conformational changes; molecular dynamics simulation

Acta Pharmacologica Sinica (2017) 38: 1673–1682; doi: 10.1038/aps.2017.37; published online 27 Jul 2017

Introduction

Altered metabolism has recently been recognized as a nearly universal trait of cancer cells[1]. The best-known phenotype of upregulated metabolic profiles in cancer is aerobic glycolysis, known as the Warburg effect, which describes the preference of glycolysis and increased lactate production, regardless of oxygen availability[2, 3]. PGAM1, an important enzyme in glycolysis, is markedly upregulated in a number of human cancers, such as hepatocellular carcinoma[4], colorectal cancer[5] and urothelial bladder cancer[6]. Overexpression of PGAM1 is related to the loss of TP53 in cancer cells due to the negative regulation of *PGAM* gene expression by TP53. Hitosugi and co-workers[7] have reported that increased PGAM1 decreases the level of 3-phosphoglycerate (3-PG), thus leading to the disinhibition of 6-phosphogluconate dehydrogenase in the oxidative pentose phosphate pathway (PPP), whereas the increase in 2-phosphoglycerate (2-PG) activates 3-phosphoglycerate dehydrogenase, thereby steering 3-PG in glycolysis for serine synthesis. Inhibition of PGAM1 expression by shRNA or small molecules attenuates cancer cell proliferation and tumor growth. Hitosugi and co-workers have demonstrated the mechanism by which the abnormal expression of PGAM1 benefits growing tumors. Together, these studies suggest that PGAM1 is a promising metabolic target for human cancer treatments.

PGAM1 catalyzes isomerization between 3-PG and 2-PG, which occurs at the branch point during the glycolytic/gluconeogenic pathway and serine biosynthesis. This enzyme is cofactor-dependent and requires 2,3-BPG as a phosphate group donor to initialize the reaction in the first phosphorylation stage during catalysis. 2,3-BPG is also an intermediate of

*To whom correspondence should be addressed.
E-mail cluo@mail.shcnc.ac.cn (Cheng LUO);
zhoulu@fudan.edu.cn (Lu ZHOU)
Received 2017-02-14 Accepted 2017-03-09

the reaction process. This enzyme and its homologues have been extensively investigated by using basic enzyme experiments and crystallographic studies and determining disease associations. A ping-pong reaction mechanism of PGAM has been proposed several decades ago[8-12]. In brief, a phosphorus group is covalently linked to histidine 11 (H11) of PGAM1 from 2,3-BPG or other phosphorus group donors. The phosphorus group of phospho-H11 is then transferred to C-2 of 3-PG, thus promoting the formation of 2,3-BPG after 3-PG localizes at the catalytic site of PGAM1. When 2,3-BPG is created in the active site, it may change its orientation and facilitate the re-phosphorylation of H11 with the C-3 phosphate group by 2,3-BPG. Finally, PGAM1 reverts to its activated state and releases the product 2-PG. However, the molecular mechanisms of 2,3-BPG reorientation within the catalytic pocket and how H11 phosphorylation makes PGAM1 more activated remain elusive.

The structural details of PGAM1 are a prerequisite for understanding its catalytic mechanism and may guide suitable chemical entity design for cancer therapy. The available PGAM1 structures reveal a dimerization mode and contain a conserved folding type among other members of the PGAM superfamily[13-16]. However, the residues beyond 234 of the C-terminal domain are disordered, as evidenced by weak electron density. Among the PGAM family members, only two structures, *Escherichia coli* (PDB code: 1e58)[14, 17] and human bisphosphoglycerate mutase (BPGM, PDB code: 2hhj)[18, 19], contain a relatively well-ordered C-terminal region. Because the conformations of the C-terminal portion of these two structures are clearly different, they represent two achievable states corresponding to the closed and semi-closed forms of PGAM1. Because of the different conformations and insufficient electron densities in previous structures, the C-terminal region of PGAM1 is believed to be conformationally variable and may play a crucial role in substrate entrance and product release during the catalytic cycle. Lysine residues of the C-terminal domain are conserved and may interact with substrates or residues in the catalytic cavity. The mutase activity of PGAM from *Saccharomyces cerevisiae* is lost after removal of the last seven residues of the C-terminal region by limited proteolysis, although the phosphatase activity is not affected[20]. The *in vivo* evidence from Kowalski *et al*[21] indicates that the C-terminal region of PGAM is regulated by the concentration of exogenous lactate, thus modulating the effectiveness of the glycolytic enzyme complex. Accordingly, the C-terminal region is likely to play an important role during the catalytic cycle of PGAM1, although the detailed mechanism remains to be further investigated.

In this article, we elucidated the role of the C-terminal domain in PGAM1 by performing microsecond (μs)-scale MD simulations and demonstrate the flexible nature of the C-terminal segment. We characterized the structural transition of PGAM1 from the closed state to the open state, which is induced by large motions of the C-terminal region. The electrostatic potentials of the open and closed states were calculated and revealed the positively charged catalytic pocket of PGAM1. Concurrently, the C-terminal segment was determined to interact with residues located in the substrate/cofactor binding pocket and to affect 2,3-BPG binding. The C-terminal segment helps to maintain the closed conformation of PGAM1 and encloses the catalytic pocket, thereby facilitating 2,3-BPG binding and further catalysis. As a case study on a member of the PGAM family, the results presented here underscore the flexible nature of the C-terminal tail and demonstrate the structural transition of PGAM1 from the closed state to the open state. The electrostatic calculations should help guide the design of novel PGAM1 inhibitors. The proposed swing mechanism of the dynamic C-terminal region improves understanding of the PGAM catalytic cycle.

Materials and methods

Model construction and preparation

Owing to the absence of a C-terminal 10-residue tail in the crystal structures of human PGAM1, we chose two structures (PDB entry: 1e58, 2hhj) with a well-ordered C-terminal region as templates for homology modeling. Both structures show reasonable sequence identity (no less than 50%) to the human PGAM1 sequence. Crystal coordinates of 1e58 and 2hhj were obtained from the Protein Data Bank (PDB). A homology modeling approach was used to generate three-dimensional (3D) models of human PGAM1 by using Modeller v14[22]. Two apo-monomer structures of PGAM1 were generated on the basis of the abovementioned crystal structures. The sequence (Entry Number: P18669) of human PGAM1 containing 254 amino acids was retrieved from UniProtKB[23] (http://www.uniprot.org/). During the model building process, 60 comparative models were initially generated and then ranked by using discrete optimized protein energy (DOPE) assessment scores and probability density functions. The model with the best performance was selected for further evaluation to determine the compatibility of various structural parameters with Procheck[24]. Subsequently, four simulation systems (two apo models and two holo models with cofactor 2,3-BPG) were designed on the basis of the model structures of PGAM1 and the BPGM complex structures. Before simulations, the protonated states of the titratable residues from the models were set according to predicted pKa values with Propka[25, 26] at a neutral bulk pH. The crystal structure of 2,3-BPG was obtained from PDB entry 2hhj. The parameters for 2,3-BPG were produced by antechamber on the basis of its crystal coordinates. The protonated state of 2,3-BPG was predicted with Epik in the Schrodinger suite[27]. The RESP charges of 2,3-BPG were derived with antechamber[28] and Gaussian 09 software (Gaussian, Inc)[29] at the HF/6-31 (d, p) level. Tleap in AmberTools 13 was used to construct all simulation systems, and the amber03[30] force field and general Amber force field (GAFF)[31] were used to generate the parameters for simulations. In all simulation systems, the protein was submerged in explicit TIP3P water in cubic boxes with an extra 10 Å extension along each axis of the protein. The net charge of the system was neutralized by adding a suitable number of counterions.

Classical molecule dynamics (CMD) simulations

Acpype[32] was used to convert the amber topology into Gromacs format. All CMD simulations were performed with the Gromacs molecular dynamics package version 4.5.7[33, 34] using the amber03 force field[30]. The particle mesh Ewald[35] (PME) method was used to account for the long-range electrostatics. A 10 Å cutoff was applied for van der Waals forces and short-range electrostatic interactions. The periodic boundary condition (PBC) was used at 300 K and 1 atm. During the CMD runs, the integration time step was set to 2 fs. The LINCS algorithm[36] was applied to constrain the length of the hydrogen bonds. To avoid unfavorable steric clashes, the starting model structure of PGAM1 was energy-minimized by using the steepest-descents algorithm. First, the protein/protein complex in the system was constrained to equilibrate the solvent. Second, the protein was released to minimize the whole simulation system. The 2,3-BPG in holo systems remained fixed during this procedure. Then, the protein complex was maintained without any constraints, and the whole holo system was minimized. Third, each system was gradually heated to 300 K in the canonical (NVT) ensemble, and this was followed by a 10 ns equilibration of the whole system under isothermal isobaric (NPT) conditions. The equilibrated systems for PGAM1 models were subjected to the production period of simulation. During the production period, the Lennard-Jones potential cutoff was set at 0.9 nm. The solvent, ions and protein/protein complex were separately coupled in a thermal bath by using a modified Berendsen thermostat method[37] with a coupling time of 0.1 ps. The isotropic pressure coupling approach with a coupling time of 1 ps was used for the equilibration and production runs. Coordinates were saved every 10 ps throughout the production runs.

Trajectory analysis

The root-mean-square deviation (RMSD) is a critical indicator of trajectory stability. Owing to large fluctuations of the C-terminal segment, the Cα of other domains in PGAM1 were selected for RMSD calculations by using the starting structure as a reference. To demonstrate the highly flexible portions, the root-mean-square fluctuations (RMSFs) were calculated per residue during the simulations. Other common structural analyses, such as radius of gyration (Rg) and hydrogen bonds determined. The secondary structure elements of each frame in the trajectory were identified by using the DSSP program[38] as previously described[39]. All desired analyses were performed using the Gromacs package. The molecular viewers VMD and Pymol were used to analyze the protein structure.

Principal component analysis (PCA)

PCA was performed to detect the primary motions of flexible regions in PGAM1 by using the positional covariance matrix of the atomic coordinates and its eigenvectors. The principles of PCA used for MD simulations have been described in detail in previous publications[40, 41]. The PCA applied in this study was performed by using the protocol[42] in the Gromacs molecular dynamics package tools. The motions of the protein were analyzed by projecting the top three eigenvectors, including the maximum motions. PCA results were visualized using the Normal Mode Wizard (NMWiz) plugin[43] of VMD.

Free energy landscape (FEL) analysis

The FEL of a protein allows us to explore the conformational transition near the native state structures. Here, we use the sampled conformations in apo systems to show the states during the transition pathway. FEL analysis was carried out using a protocol described in previous publications[40, 41].

Results and discussion

Published crystal structures and the intrinsic sequence of PGAM suggest a flexible C-terminal region

A representative set of available 3D structures of human PGAM1 and its homologous proteins were aligned for comparison (Figure 1A). We found further evidence that the C-terminal tail (residues 234-254) of human PGAM1 structures (PDB code: 1yfk, 1yjx, 4gpi and 4gpz) lacks interpretable electron density. Several homologous structures have a similar lack of density , except for structures of phosphoglycerate mutase from *E. coli* (PDB code: 1e58, 1e59) and human BPGM (PDB code: 2a9j, 2f90, 2h4x, 2h4z, 2h52, 2hhj and 3nfy). A protein sequence alignment of the C-terminal region was performed (Figure 1B). The C-terminal region is considered a comparatively conserved domain, thus indicating the conservation of function among a set of orthologous proteins. These results directly suggest the highly dynamic nature of the C-terminal region. We hypothesized that the C-terminal segment is of great importance by virtue of its substantial function-related conformational change. Walter *et al*[20] have highlighted the crucial role of the C-terminal region in *S cerevisiae* through limited proteolysis and site-directed mutagenesis assays. Their results, as well as reported structural information, indicate that the C-terminal tail should have a tendency to be intrinsically disordered. In the next step, we determined whether the C-terminal tail is natively unstructured because of its specific amino acid sequence. The complete amino acid sequence of human PGAM1 was scored by a protein disorder predictor, PONDR-FIT (http://www.disprot.org/pondr-fit.php)[44, 45]. Residues with a VSl2B score[44] larger than 0.5 were recognized as disordered. The results clearly showed that the amino acids of the C-terminal segment were highly disordered (Figure 1C). Additionally, to confirm the sequence-dependent conformational variations of the isolated C-terminus, MC simulations were carried out in Campari software[46]. The MC simulations showed that the C-terminal region undergoes extensive conformational changes covering previously reported configurations, such as that from 2hhj and 1e58 (Figure S1). These preliminary analyses add to current understanding of the conformational flexibility of the C-terminal domain in human PGAM1.

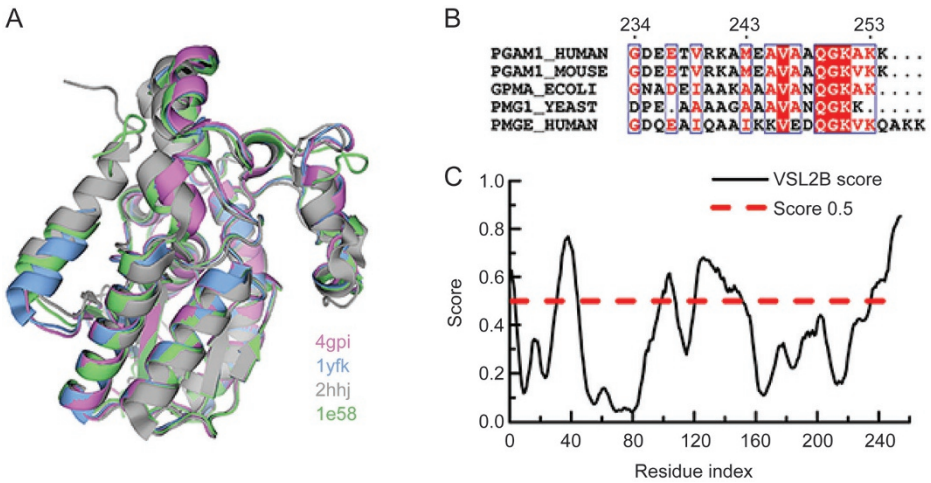

Figure 1. Superimposed structures of PGAM and sequence-based prediction of intrinsically disordered regions. (A) A cartoon representation showing the 3D structure alignment of the available crystal structures of PGAMs; structures of human PGAM1: magenta (PDB code: 4gpi), light marine (PDB code: 1yfk); the structure of human BPGM: gray (PDB code: 2hhj); the structure of *Escherichi coli* PGAM: green (PDB code: 1e58). The image was created with Pymol. (B) C-terminal sequence alignment of PGAMs from human, mouse, *Escherichia coli*, and *Saccharomyces cerevisiae* and human BPGM. Motif QGK is the most highly conserved motif within the C-terminal region. Identical residues are shown in red. Similar residues are highlighted in red and boxed in blue. (C) PGAM1 with four disordered regions at residues 31-44, residues 99-117, residues 121-151 and residues 235-254. Residues with a prediction score above 0.5 are predicted as disordered.

Intrinsic flexibility of the C-terminus from MD simulations

The sequence of the PGAM1 C-terminal region reflects its function-related dynamic behavior, as indicated by the PONDR-FIT prediction results and MC simulations. Despite the ergodicity of conformational space in the MC simulations, reliable interactive configurations cannot be obtained from the isolated C-terminus without other PGAM1 domains. MD simulations are routinely used computational tools for studying the structural flexibility of macromolecules, as shown previously[40, 47-53]. To determine the actual dynamic behavior of the C-terminal segment within its surrounding environment, we first performed MD simulations on the apo-PGAM1 models in aqueous solution. Two available homologous structures (PDB code: 1e58 and 2hhj) were selected to construct PGAM1 models, because these two homologous crystal structures share high sequence similarity with PGAM1 and contain a relatively complete C-terminal region that adopts distinguished conformations. Thus, we built two systems to solve the two apo model structures in water, namely, 1E58-apo and 2HHJ-apo (Table S1). The model building procedures and system preparations are described in the Methods section. For each system, a 1.5-μs CMD simulation was carried out. The structural features of PGAM1 were analyzed using the Gromacs build-in tools as previously described. During the simulations, the 1E58-apo system reached a stable condition up to 300 ns according to the RMSD diagram of the C-α backbone of the protein (the C-terminal region was not included) over time (Figure S2). However, the RMSD of the 2HHJ-apo model without the C-terminal region reached a relatively stable state after 1200 ns. The RMSF values of the two models demonstrated that the C-terminus undergoes large fluctuations during simulations (Figure S3A). Moreover, the central kink and back helix regions (residues 133–137) were detected as fluctuating regions, as were the helix region (residues 12-23), which have previously been reported to be flexible[16]. These conformational changes may be involved in the interaction networks induced by motions of the C-terminal tail. The secondary structure of the C-terminal segment was inspected for a detailed analysis of its organizational diversity during the simulations. The results of both systems indicated the flexible nature of the C-terminal tail in PGAM1. There was no fixed secondary structure in either system for residues 246 to 254. However, in contrast with the 1E58-apo system, the 2HHJ-apo system contained more helical regions (residues 243–245) than did the 1E58-apo system until 1200 ns. During a simulation time from 1200 ns to 1500 ns, both C-terminal regions reached the same state of equilibrium with a similar percentage of helical elements, in agreement with the crystal structure of human PGAM1 (PDB code: 1yfk, 1yjx, 4gpi and 4gpz), in which the C-terminal segment is presented as a helical element between residues 236-243 and lacks electron density beyond residue methionine 243. These simulations further suggest the flexible nature of the C-terminal region in PGAM1. Thus, we concluded that the C-terminal region of PGAM1 is indeed a flexible part that exhibits diverse conformational changes.

Interconversions between the open and closed state

Significant motions of the C-terminal region should reflect its related biological functions. To extract the large-scale motion modes and explore their functional relevance, we applied PCA of the MD trajectories across the two apo systems. We selected the top three principal components (Figure 2A–C), which separately accounted for 44.3%, 33.8% and 4.1% of the overall motions in the 2HHJ-apo system. The first component

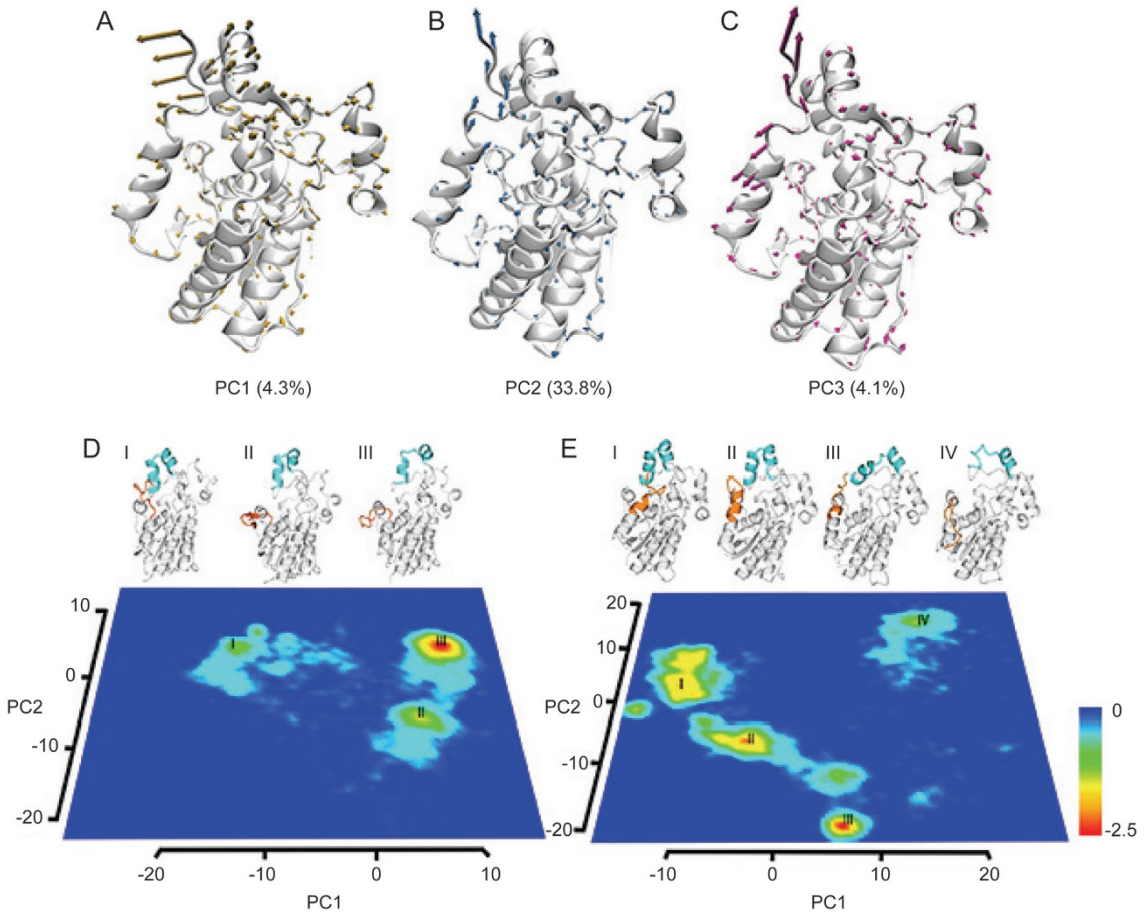

Figure 2. The C-terminal region is involved the primary motions and transition pathway of PGAM1 during the MD simulations. (A-C) Motions corresponding to PC1-PC3 of PGAM1 in the 2HHJ-apo system. PC1-PC3 directions are shown by yellow, blue and magenta arrows. (D, E) Bottom, energy landscape for the conformational transition pathway along projections onto the PC1 and PC2 of systems 1E58-apo and 2HHJ-apo; top, representative snapshots of the local energy minima during simulations of the two apo systems.

(PC1) indicated that the C-terminal region gradually extends into the bulk solvents and moves away from the central kink region. Simultaneously, the central kink region moves opposite from the motion direction of the C-terminus. Nevertheless, PGAM1 displays a common closed conformation in the presence of 2,3-BPG. According to the second component (PC2), the C-terminal region shows a strong tendency for the system toward the open state, which includes a more disordered C-terminal region, as compared with the closed state of PGAM1. In the third component (PC3), the central kink region attempts to restore its orientation to the closed state. These motions reveal new conformational space of the structural transitions. The FEL was calculated according to the plot selecting PC1 and PC2 as the reaction coordinates (Figure 2D, 2E). The FELs of the two apo models provide a direct view of the intrinsically disordered C-terminal region and a process of state conformational exchange between the open (I in Figure 2D, 2E) and closed (III in Figure 2D and IV in Figure 2E) state. Our MD simulations of the PGAM1 systems probed the open forms of PGAM1, which have not previously been identified. These observations also address the extensive conformational transition from the catalytically active (closed) state of PGAM1 to an open state. The contour map of free energy estimation indicates that the open state is the most frequently distributed state among the conformational states from apo-PGAM1 simulations (II and III in Figure 2D; IV in Figure 2E). It also shows that the open state should be the native conformation when PGAM1 is in an unbound state, thereby explaining the disordered tail in the crystal structures of PGAM1. As described in Figure 2D, the initial structure of the 1E58-apo model has a less closed conformation than that in the 2HHJ-apo model. Notably, the C-terminal tail (residues 243-254) reshapes its coordinates and orientation (II in Figure 2D). The conformational changes in the central kink region are coupled with fluctuations of the C-terminal tail. Consequently, the motions of these portions uncover the substrate-binding pocket, thus serving to recruit substrates or release products. Moreover, the information provided by the free energy landscape also reflects the ability of PGAM1 to sustain the open/closed conformation. Replica exchange molecular dynamics (REMD) simulations were performed to enhance the conformational sampling of PGAM1. The results reproduced the large-

magnitude functional motions from CMD simulations and emphasized the elastic characteristics of the C-terminal region in PGAM1 (Figure S4).

We also calculated the electrostatics of the open/closed state by using APBS (see Figure S5). When the C-terminal region covered the catalytic pocket and PGAM1 was in the closed conformation, the majority of the basic residues located in the catalytic pocket were wrapped by the C-terminal domain. Nevertheless, the basic residues were exposed to solvents in the open conformation (Figure S5B). The open pocket with high positive electrostatics would quickly attract the negatively charged substrate 3-PG, which contains a phosphate group. After the substrate binds to basic residues, such as K100, R116 or R117, the C-terminal region rescues its conformation in the closed state and maintains a necessary biochemical environment for enzyme catalysis. Several inhibitors have been reported to target PGAM1[7, 54-56]. MJE3 covalently inactivates PGAM1 at K100. However, it has been speculated that PGMI-004A binds exclusively to an allosteric pocket on PGAM1. Owing to the positive potential of the catalytic pocket and knowledge of the flexible C-terminus, more attention should be paid to determine the regular function of PGAM1. Only an electronegative molecule could occupy the substrate/cofactor binding pocket, thereby blocking access of the substrate to the catalytic pocket. Some non-specific polyanionic compounds have been described as mimics of the substrate that collapse the phospho-enzyme intermediate[56]. Therefore, analogs of the substrate or cofactor may act as effective inhibitors. Moreover, the design of a novel allosteric or covalent inhibitor is a promising strategy to mediate the biological activity of PGAM1.

Key residues involved in the dynamic mechanisms of PGAM1

To explore the mechanisms underlying the dynamic nature of PGAM1, we carefully examined all snapshots of trajectories from MD simulations. We found that the coordinates of K100 were associated with the conformational change in the central kink region. We then calculated the distance between K100 and H11. H11 is in the center of the catalytic pocket; therefore, the position of H11 was chosen as the reference coordinate set. We observed that the distance from the snapshots over the simulation time was larger than that from the initial model structures (Figure 3A). At the end of the simulations, the semi-open form of the central kink region had an increased kink angle compared with the initial model structures. The movement of this part further exposes the substrate-binding cavity composed of basic amino acid residues and may be the key process underlying product release. R116 and R117 are two flexible residues that are involved in the binding of cofactor 2,3-BPG and substrate 3-PG[17, 19]. Owing to their critical roles in the catalytic cycle, detailed hydrogen bond analyses were performed to determine whether they are affected by other domains. R116 forms hydrogen bonds with its adjacent residues, such as K113 and V112. Moreover, K100 was also found to contact R116 through hydrogen bonds. We observed that R117 interacts with E203 and E207. In the MD trajectory of the 1E58-apo system, these hydrogen bonds are fleetingly present throughout most of the simulation time but especially before 400 ns. However, during the simulations of the 2HHJ-apo system, hydrogen bonds between R117-E203/E207 are present only before 800 ns (Figure 3B). When R117 is captured by E203 or E207, it cannot facilitate recognition and binding to the cofactor or substrate. These interactions inhibit a substrate from entering the enzyme's active site as well as product release. Residue K254 forms hydrogen bonds with R117 within a short period, thus indicating the interplay between the C-terminal region and the substrate-binding pocket.

Holo system analysis

The apo systems provided with abundant, intrinsically dynamic information on PGAM1 in the unbound state. It is unclear whether more dynamic features are induced when the cofactor binds in the catalytic pocket of PGAM1. Specific questions regarding the underlying mechanism of critical residues for cofactor binding and catalytic activity remain unknown. To understand the conformational changes and the precise molecular mechanisms involved, we built two holo systems, 1E58-holo and 2HHJ-holo (Table S1), to investigate the effects introduced by the intermediate product, 2,3-BPG. We speculated that such a strategy (including the intermediate) may be necessary to obtain new information with respect to the interactions between the C-terminal region and the catalytic pocket of PGAM1. Details of model construction can be found in the Methods section. To determine the stability of PGAM1 in the four simulated systems, we compared their gyration radius changes over simulation time. The compactness of protein structures from the 2HHJ-holo system was homogeneous. The C-terminal region underwent large motions in the two 1E58-based systems. In the 2HHJ-apo system, the gyration radius was similar to that in the 2HHJ-holo system, which underwent large-amplitude fluctuations during the simulation and reached the same level of the two 1E58-based systems. We then clustered the trajectories of all holo systems and found three major conformations of the complex structures (Figure 4B–D). As mentioned above, a large conformational transition was observed in the 1E58-holo system. The representative conformation observed during the early simulation time is described in Figure 4B. The C-terminal tail generated the open state conformation, as demonstrated in Figure 4C, even in the presence of 2,3-BPG. This configuration reveals that the C-terminal tail is not necessary for the phosphatase activity of PGAM1, as corroborated by a previous study[20]. The structure of the cluster center in the compact 2HHJ-holo system is shown in Figure 4D. This structure shows few differences with the complex structure of human BPGM.

Furthermore, the occupancy of hydrogen bonds present in the 2,3-BPG binding pocket was analyzed during the simulation period. Key residues that contact 2,3-BPG were identified (Figure 4E-F). On the basis of the hydrogen bond occupancy, we determined that 2,3-BPG has dissimilar binding modes in the two holo systems. R116, R117 and K100 play dominant roles in the interaction

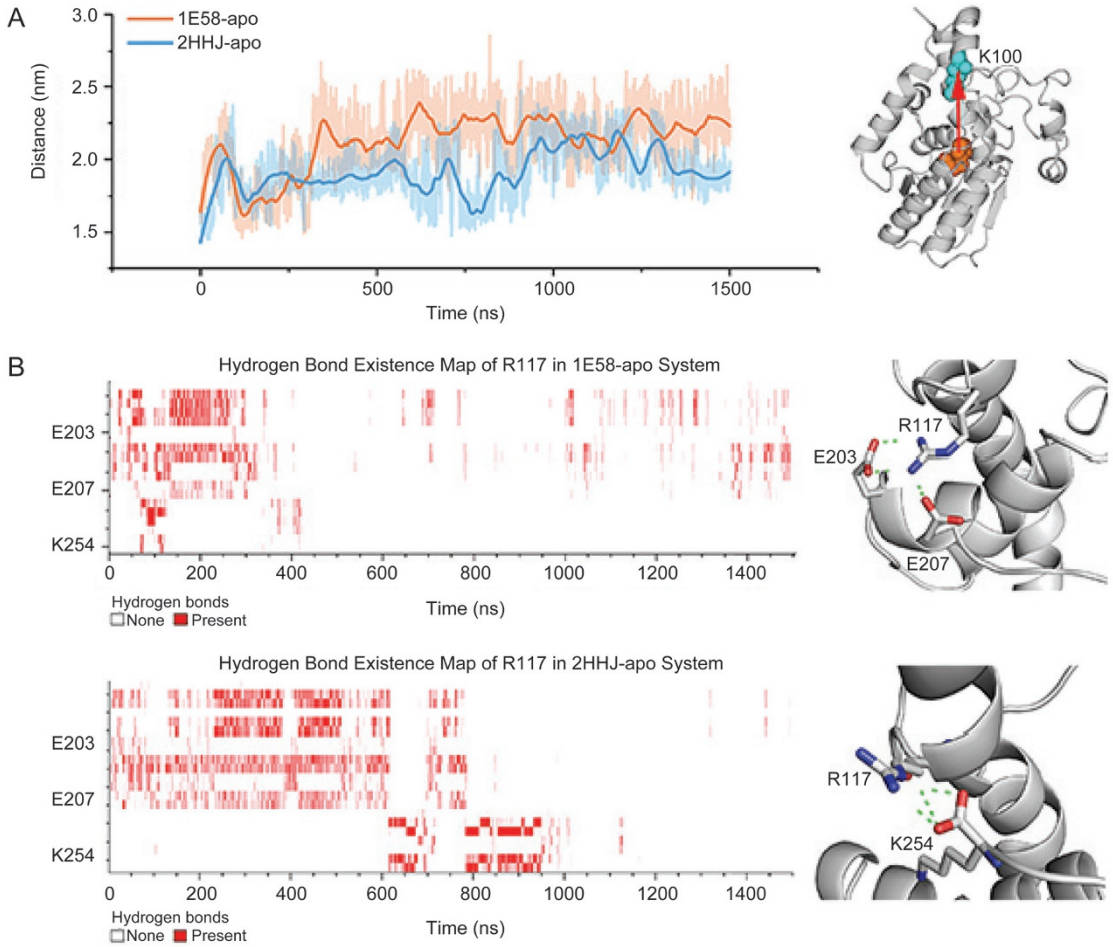

Figure 3. Dynamic features of key residues during MD simulations. (A) Left, distance between K100 and H11 in the 1E58-apo (orange) and 2HHJ-apo (cyan) systems versus simulation time. Time-smoothed values of distance are shown on the light background (raw data), 100 raw data points were averaged at a time; right, distance between K100 (cyan) and H11 (orange). (B) Left, time evolution of the hydrogen bond existence map of R117 in the 1E58-apo (top) and 2HHJ-apo (bottom) systems. Rows correspond to various hydrogen bonds formed. A hydrogen bond is defined when the acceptor atom lies within 3.5 Å of the donor atom in a cone of 120°. Right, R117, and its related residues are shown as sticks with hydrogen bonds depicted as green dashed lines.

with 2,3-BPG in both systems, with an occupancy rate as high as 35%. Their positive electrical potential stabilizes the negatively charged carbonyl oxygen and phosphate oxygen in 2,3-BPG, which is the direct reason for their role during substrate/cofactor binding. Hydrogen bonds between 2,3-BPG and arginine 10 (R10) appeared frequently in the trajectory of the 1E58-holo system; however, these interactions were not often observed in the 2HHJ-holo system. This phenomenon indicated that 2,3-BPG slightly shifts outward when PGAM1 is in the open state. In the case of the 2HHJ-holo system, we found that glutamine 249 (Q249) is the critical anchor for the interaction between 2,3-BPG and the C-terminal region. Q249 may interact directly with 2,3-BPG via the formation of a hydrogen bond, thus prompting partial portion adjustments of the catalytic pocket. Q249 can also form hydrogen bonds with R116. Additionally, Q249 is a perfectly conserved site in the homologous protein (Figure 1B), thus suggesting that Q249 may be crucial for enzyme function. In summary, the C-terminal region of PGAM1 exhibits remarkable conformational freedom. We confirmed that the flexible C-terminal region serves as a modulator affecting the binding mode of the substrate or intermediate product.

Swing model of the C-terminal region during the catalytic cycle

On the basis of the evidence in this study and previous reports[20], we propose the following swing model to describe the functional role of the C-terminal region within the catalytic pathway (Figure 5). The C-terminal region reorganizes its secondary structure and exhibits large-scale conformational transitions (swing) between the open and closed configurations in the apo state. The dynamics of other domains (the kink region and its adjacent helix region) are accompanied by movement in the C-terminus. In the 1E58-holo system, however, 2,3-BPG can localizes in the binding cavity without large motions, even

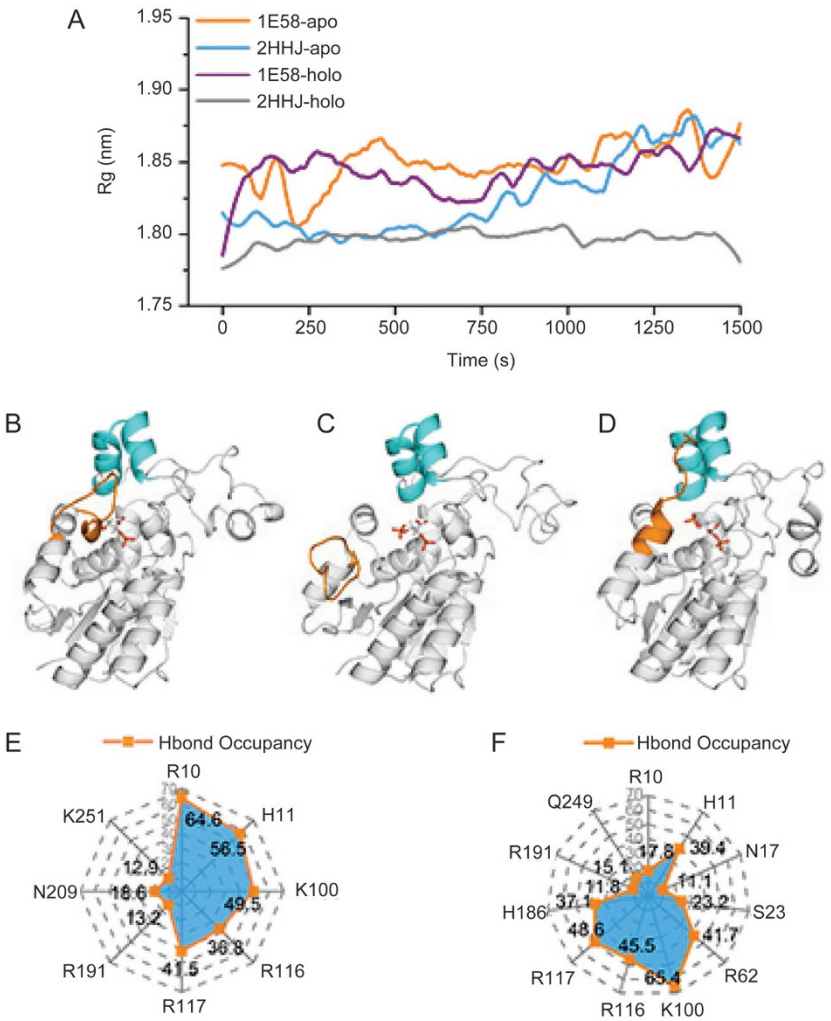

Figure 4. Structural characterization of the C-terminal region and its influence on 2,3-BPG binding in two holo systems. (A) Time evolution of Rg in 1E58-apo (orange), 2HHJ-apo (cyan), 1E58-holo (purple) and 2HHJ-holo (gray) systems. (B-D) Typical conformations of PGAM1 with various C-terminal tails in the holo systems. 2,3-BPG is indicated by sticks. The central kink region is colored cyan, and the C-terminal tail is colored orange. (E-F) Hydrogen bond occupancy between 2,3-BPG and key residues for substrate/cofactor binding in 1E58-holo (E) and 2HHJ-holo (F) systems. A hydrogen bond is defined when the acceptor atom lies within 3.5 Å of the donor atom in a cone of 120°.

when the C-terminal region is in the open state, thus indicating the minimal role of the C-terminal domain in phosphatase activity. Additionally, the 2HHJ-holo simulation system indicates that the C-terminal region is primarily invoked as a means to constrain structural flexibility to stabilize the binding mode of the intermediate so that R116, R117 and other basic residues interact favorably with 2,3-BPG. Overall, we observed the dynamic motions of the C-terminal region of PGAM1 and described its role in different stages of catalysis by using computational methods.

Conclusions

Crystal structures of the C-terminal region of PGAM1 do not show a fixed 3D structure, and residues within the C-terminal tail are strongly predicted to be natively disordered, according to PONDR-FIT. Here, different conformations of the C-terminal region were explored by using MC simulations. The results are in agreement with the available structure of the C-terminal region. We also validated the flexible nature of the C-terminal region and revealed the transition pathway of PGAM1 from the closed state to the open state, by using CMD simulations. The open or closed state of the C-terminal region may be involved in events associated with substrate binding or product release during catalysis. On the basis of the MC and CMD simulations, which were operated with different force fields, we found that the C-terminal region is inherently plastic, in a manner dependent on its amino acid sequence. Thus, the dynamic features of the C-terminus are not the result of force field deficiencies.

Considering the relationship between the C-terminal region and the catalytic core, we performed CMD simulations of complex systems. In the 1E58-holo system, the coil of the C-terminal tail underwent large-amplitude fluctuations and ultimately generated the open state confor-

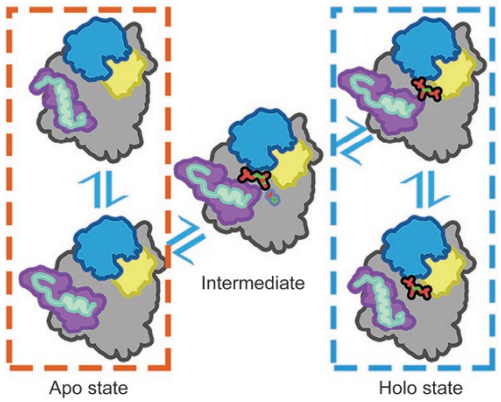

Figure 5. Swing model of the C-terminal region during the catalytic cycle. The surface area of the C-terminal region is colored purple, and its secondary scaffold is shown in light green. The kink region is colored light blue, and its adjacent helix region (residues 12–23) is filled with yellow. Other parts of PGAM1 are shown in gray. 3-PG, phosphohistidine 11 and 2,3-BPG are represented as sticks.

mation, even in the presence of 2,3-BPG. In contrast, most conformations in the 2HHJ-holo system were structurally stable, in accordance with the crystal structure of the BPGM complex in the closed state, thus demonstrating that the C-terminal region may play a different role in the presence of cofactor 2,3-BPG. Moreover, some residues of the C-terminal region were determined to interfere with substrate/cofactor binding, in agreement with previous reports.

Furthermore, it is difficult to identify small-drug molecules that target the binding cavity, owing to its physicochemical properties. Hence, an allosteric inhibitor or a covalent modulator may provide a new direction for drug design against PGAM1. In summary, this simulation study sheds new light on the spontaneous interactions of the flexible C-terminus and active pocket of PGAM1 and its homologous proteins. Future efforts will focus on the process of 2,3-BPG repositioning (i.e., transferring the C3 phosphate to H11) by using MD simulations.

Acknowledgements

The authors wish to thank Dr Jun-yan LU for advice on simulation design and data analyses, as well as for editing the manuscript. We gratefully acknowledge financial support from the Ministry of Science and Technology of China (No 2015CB910304 to Hua-liang JIANG), the National Natural Science Foundation of China (No 21472208 and 81625022 to Cheng LUO, No 21210003 and 81230076 to Hua-liang JIANG, 21472026 to Lu ZHOU), the State Key Laboratory of Toxicology and Medical Countermeasures, Academy of Military Medical Science (No TMC201505 to Cheng LUO), the Shanghai Municipal Committee of Science and Technology (No 14XD1400300) and the program for Shanghai Rising Star (No 15QA1400300). The computation resources were supported by the Computer Network Information Centre, Chinese Academy of Sciences and the Guangdong Supercomputing Centre (No nsfc2015-446 to Cheng LUO).

Author contribution

Cheng LUO and Lu ZHOU designed the study; Shi-en LIU and Cheng LUO performed the simulations, analyzed the data and wrote the paper; Jun-chi HU, Hao ZHANG, Wei WAN, Pan XU, Ming-yue ZHENG, Kun-qian YU, Hong DING and Hua-liang JIANG discussed and edited the manuscript. All authors contributed to writing the manuscript and have approved the final version of the manuscript.

Supplementary information

Supplementary information is available at the website of Acta Pharmacologica Sinica.

References

1 Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. Cell Metab 2016; 23: 27–47.
2 Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer 2004; 4: 891–9.
3 Hitosugi T, Chen J. Post-translational modifications and the Warburg effect. Oncogene 2014; 33: 4279–85.
4 Ren F, Wu H, Lei Y, Zhang H, Liu R, Zhao Y, *et al*. Quantitative proteomics identification of phosphoglycerate mutase 1 as a novel therapeutic target in hepatocellular carcinoma. Mol Cancer 2010; 9: 81.
5 Liu L, Wang SA, Zhang QL, Ding YQ. Identification of potential genes/proteins regulated by Tiam1 in colorectal cancer by microarray analysis and proteome analysis. Cell Biol Int 2008; 32: 1215–22.
6 Peng XC, Gong FM, Chen Y, Qiu N, Cheng K, Tang J, *et al*. Proteomics identification of PGAM1 as a potential therapeutic target for urothelial bladder cancer. J Proteomics 2016; 132: 85–92.
7 Hitosugi T, Zhou L, Elf S, Fan J, Kang HB, Seo JH, *et al*. Phosphoglycerate mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. Cancer Cell 2012; 22: 585–600.
8 Grisolia S, Cleland WW. Influence of salt, substrate, and cofactor concentrations on the kinetic and mechanistic behavior of phosphoglycerate mutase. Biochemistry 1968; 7: 1115–21.
9 Britton HG, Carreras J, Grisolia S. Mechanism of yeast phosphoglycerate mutase. Biochemistry 1972; 11: 3008–14.
10 Britton HG, Clarke JB. Mechanism of the 2,3-diphosphoglycerate-dependent phosphoglycerate mutase from rabbit muscle. Biochem J 1972; 130: 397–410.
11 Rose ZB, Hamasaki N, Dube S. The sequence of a peptide containing the active site phosphohistidine residue of phosphoglycerate mutase from chicken breast muscle. J Biol Chem 1975; 250: 7939–42.
12 Rose ZB, Dube S. Rates of phosphorylation and dephosphorylation of phosphoglycerate mutase and bisphosphoglycerate synthase. J Biol Chem 1976; 251: 4817–22.
13 Winn SI, Watson HC, Harkins RN, Fothergill LA. Structure and activity of phosphoglycerate mutase. Philos T Roy Soc B 1981; 293: 121–30.
14 Bond CS, White MF, Hunter WN. High resolution structure of the phosphohistidine-activated form of Escherichia coli cofactor-dependent phosphoglycerate mutase. J Biol Chem 2001; 276: 3247–53.
15 Wang YL, Wei ZY, Liu L, Cheng ZJ, Lin YJ, Ji FY, *et al*. Crystal structure of human B-type phosphoglycerate mutase bound with citrate. Biochem Biophys Res Commun 2005; 331: 1207–15.
16 Hitosugi T, Zhou L, Fan J, Elf S, Zhang L, Xie J, *et al*. Tyr26 phosphorylation of PGAM1 provides a metabolic advantage to

tumours by stabilizing the active conformation. Nat Commun 2013; 4: 1790.
17 Bond CS, White MF, Hunter WN. Mechanistic implications for *Escherichia coli* cofactor-dependent phosphoglycerate mutase based on the high-resolution crystal structure of a vanadate complex. J Mol Biol 2002; 316: 1071–81.
18 Wang Y, Wei Z, Bian Q, Cheng Z, Wan M, Liu L, *et al.* Crystal structure of human bisphosphoglycerate mutase. J Biol Chem 2004; 279: 39132–8.
19 Wang Y, Liu L, Wei Z, Cheng Z, Lin Y, Gong W. Seeing the process of histidine phosphorylation in human bisphosphoglycerate mutase. J Biol Chem 2006; 281: 39642–8.
20 Walter RA, Nairn J, Duncan D, Price NC, Kelly SM, Rigden DJ, *et al.* The role of the C-terminal region in phosphoglycerate mutase. Biochem J 1999; 337: 89–95.
21 Kowalski W, Nocon D, Gamian A, Kolodziej J, Rakus D. Association of C-terminal region of phosphoglycerate mutase with glycolytic complex regulates energy production in cancer cells. J Cell Physiol 2012; 227: 2613–21.
22 Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993; 234: 779–815.
23 UniProt: a hub for protein information. Nucleic Acids Res 2015; 43: D204–12.
24 Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck–a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26: 283–91.
25 Olsson MH, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput 2011; 7: 525–37.
26 Sondergaard CR, Olsson MH, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa Values. J Chem Theory Comput 2011; 7: 2284–95.
27 Schrödinger Release 2015-4: Epik, Schrödinger, LLC, New York, NY, 2015.
28 Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 2006; 25: 247–60.
29 Gaussian 09. Wallingford, CT, USA: Gaussian, Inc. 2009.
30 Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003; 24: 1999–2012.
31 Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004; 25: 1157–74.
32 Sousa da Silva AW, Vranken WF. ACPYPE–Antechamber python parser interface. BMC Res Notes 2012; 5: 367.
33 Berendsen HJC, Vanderspoel D, Vandrunen R. Gromacs - a message-passing parallel molecular-dynamics implementation. Comput Phys Commun 1995; 91: 43–56.
34 Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model 2001; 7: 306–17.
35 Darden T, York D, Pedersen L. Particle Mesh Ewald–an N.Log(N) method for ewald sums in large systems. J Chem Phys 1993; 98: 10089–92.
36 Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 1997; 18: 1463–72.
37 Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. Molecular-dynamics with coupling to an external bath. J Chem Phys 1984; 81: 3684–90.
38 Kabsch W, Sander C. Dictionary of protein secondary structure-pattern-recognition of hydrogen-bonded and geometrical features. Biopolymers 1983; 22: 2577–637.
39 Wang YY, Li L, Chen TT, Chen WY, Xu YC. Microsecond molecular dynamics simulation of A beta(42) and identification of a novel dual inhibitor of A beta(42) aggregation and BACE1 activity. Acta Pharmacol Sin 2013; 34: 1243–50.
40 Du Y, Yang H, Xu Y, Cang X, Luo C, Mao Y, *et al.* Conformational transition and energy landscape of ErbB4 activated by neuregulin1beta: one microsecond molecular dynamics simulations. J Am Chem Soc 2012; 134: 6720–31.
41 Ye F, Zhang J, Liu H, Hilgenfeld R, Zhang R, Kong X, *et al.* Helix unfolding/refolding characterizes the functional dynamics of Staphylococcus aureus Clp protease. J Biol Chem 2013; 288: 17643–53.
42 Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. Proteins 1993; 17: 412–25.
43 Bakan A, Meireles LM, Bahar I. ProDy: protein dynamics inferred from theory and experiments. Bioinformatics 2011; 27: 1575–7.
44 Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK. Exploiting heterogeneous sequence properties improves prediction of protein disorder. Proteins 2005; 61: 176–82.
45 Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. BBA-Proteins Proteom 2010; 1804: 996–1010.
46 Vitalis A, Pappu RV. Methods for Monte Carlo simulations of biomacromolecules. Annu Rep Comput Chem 2009; 5: 49–76.
47 Li J, Du Y, Liu X, Shen QC, Huang AL, Zheng MY, *et al.* Binding sensitivity of adefovir to the polymerase from different genotypes of HBV: molecular modeling, docking and dynamics simulation studies. Acta Pharmacol Sin 2013; 34: 319–28.
48 Liu GX, Tan JZ, Niu CY, Shen JH, Luo XM, Shen X, *et al.* Molecular dynamics simulations of interaction between protein-tyrosine phosphatase 1B and a bidentate inhibitor. Acta Pharmacol Sin 2006; 27: 100–10.
49 Xiong B, Huang XQ, Shen LL, Shen JH, Luo XM, Shen X, *et al.* Conformational flexibility of beta-secretase: molecular dynamics simulation and essential dynamics analysis. Acta Pharmacol Sin 2004; 25: 705–13.
50 De Vivo M, Masetti M, Bottegoni G, Cavalli A. Role of molecular dynamics and related methods in drug discovery. J Med Chem 2016; 59: 4035–61.
51 Lu J, Hu L, Cheng J, Fang D, Wang C, Yu K, *et al.* A computational investigation on the substrate preference of ten-eleven-translocation 2 (TET2). Phys Chem Chem Phys 2016; 18: 4728–38.
52 Yang LL, Yang DH, de Graaf C, Moeller A, West GM, Dharmarajan V, *et al.* Conformational states of the full-length glucagon receptor. Nat Commun 2015; 6: 7859.
53 Yuan S, Chan HC, Vogel H, Filipek S, Stevens RC, Palczewski K. The molecular mechanism of P2Y1 receptor activation. Angew Chem Int Ed Engl 2016; 35: 10331–5.
54 Evans MJ, Morris GM, Wu J, Olson AJ, Sorensen EJ, Cravatt BF. Mechanistic and structural requirements for active site labeling of phosphoglycerate mutase by spiroepoxides. Mol Biosyst 2007; 3: 495–506.
55 Evans MJ, Saghatelian A, Sorensen EJ, Cravatt BF. Target discovery in small-molecule cell-based screens by in situ proteome reactivity profiling. Nat Biotechnol 2005; 23: 1303–7.
56 Puyau PD, Perie JJ. Synthesis of substrate analogues and inhibitors for the phosphoglycerate mutase enzyme. Phosphorus Sulfur 1997; 129: 13–45.