A computational study on role of 6-(hydroxymethyl)-3-[3,4,5-trihydroxy-6-[(3,4,5-trihydroxyoxan-2-yl)oxymethyl]oxan-2-yl]oxyoxane-2,4,5-triol in the regulation of blood glucose level

Karthikeyan Muthusamy* and Gopinath Krishnasamy

Department of Bioinformatics, Alagappa University, Science Block, Karaikudi, 630 004 Tamil Nadu, India

Communicated by Ramaswamy H. Sarma

(Received 6 September 2015; accepted 21 November 2015)

6-(hydroxymethyl)-3-[3,4,5-trihydroxy-6-[(3,4,5-trihydroxyoxan-2-yl)oxymethyl]oxan-2-yl]oxyoxane-2,4,5-triol (SID 242078875) was isolated from the fruits of Syzygium densiflorum Wall. ex Wight & Arn (Myrtaceae), which has been traditionally used in the treatment of diabetes by the tribes of The Nilgiris, Tamil Nadu, India. In this study, reverse pharmacophore mapping approach and text-based database search identified the dipeptidyl peptidase-IV, protein-tyrosine phosphatase 1B, phosphoenolpyruvate carboxykinase, glycogen synthase kinase-3β and glucokinase as potential targets of SID 242078875 in diabetes management. Further, molecular docking was performed to predict the binding pose of SID 242078875 in the active site region of the target protein. In addition, dynamic behaviour and stability of protein–ligand complexes were observed for a period of 50 ns through molecular dynamics simulation.

Keywords: antidiabetic property; molecular docking; molecular dynamics; natural product; SID 242078875; Syzygium densiflorum

Introduction

Diabetes mellitus is one of the largest chronic endocrine metabolic disorder worldwide; it is debilitating and often life threatening with an increasing prevalence rate of 8.3% (IDF, 2014). Investigation of medicinal plants is an emerging research area due to their lesser side effects (Osadebe, Odoh, & Uzor, 2014). Recently, natural sources have been used for the treatment of diabetes throughout the world. A large number of medicinal plants that possess highly diversified chemical compounds with hypoglycaemic activity (Onkaramurthy et al., 2013). A wide variety of natural and naturally derived compounds are in the clinical phase. Among the FDA-approved diabetes medicines, 15 anti-diabetic drugs – voglibose, acarbose, miglitol, exenatide, tipramycin acetate, liraglutide, nateglinide, tolrestat, epalrestat, troglitazone, rosiglitazone maleate, sitagliptin, vildaglaptin, saxagliptin and alogliptin benzoate – were categorized under natural products. These include natural products, naturally derived compounds that mimic natural products, and synthetic compounds with natural product pharmacophore (Newman & Cragg, 2012). However, use of natural products in drug discovery has decreased due to technical barriers such as high-throughput screening assays against disease targets (Harvey, Edrada-Ebel, & Quinn, 2015).

Diabetes target proteins have been identified by various approaches such as chemical proteomics, comprehensive phenotypic screening, computational inference methods, direct biochemical methods, expression profiling, gene knockouts studies, genetic interactions analysis, genome-wide genetic assays, high-throughput functional genomics, integrated network-based approach, microarray gene expression analysis, PageRank algorithm and pathway analysis (Emig et al., 2013; Grolmusz, 2015; Gu, Chen, Yuan, & Xu, 2013; Haanstra & Bakker, 2015; Schenone, Dančík, Wagner, & Clemons, 2013; Walke et al., 2001; Wang, Zhao, Shang, & Xia, 2014). Various online resources have been updating new novel protein targets and offer target information to the scientific community. At present, following resources provide information about the proteins associated with diabetes: Potential drug target database (Gao et al., 2008), T2D-Db (Agrawal et al., 2008), T1Dbase (Burren et al., 2011), T2D@ZJU (Yang et al., 2013), The Human Diabetes Proteome Project (Schvartz et al., 2015; Topf et al., 2013) and diabetes associated proteins database (DAPD) (Gopinath, Jayakumararaj, & Karthikeyan, 2015). Identification of specific molecular targets of the drug is crucial for drug innovation, personalized medicine and rational use of drugs. Also, it has a significant impact on clinical trials and treatment with reduction in
side effects (França, 2015; Liu, Constantinides, & Li, 2014). However, target identification from thousands of molecules is still a laborious, time-consuming and cost-intensive process (Kore, Mutha, & Antre, 2012). Over the past decades, computer-aided drug design approaches have accelerated the early-stage pharmaceutical research in target identification and validation. Computational tools have been reported as an alternative to experimental methods in predicting target protein for small molecules/plant-derived molecules (Bashar, Khalipha, & Chowdhury, 2012; Zheng et al., 2013). Computational tools employ ligand and structure-based approaches for target prediction. Reverse/inverse docking uses a structures-based approach to identify the target protein, where a query compound is docked against a panel of protein targets or receptors using scoring functions (McPhillie, Cain, Narramore, Fishwick, & Simmons, 2015). Various online sources – INVDOCK (Chen & Ung, 2001), Target Fishing Dock (Li et al., 2006) and IDTarget (Wang, Chu, Chen, & Lin, 2012) – have employed reverse/inverse docking approaches to identify the target protein (McPhillie et al., 2015). The above-mentioned server uses docking program and ligand–protein interaction energy to assess the docking results (Zheng, Chen, & Lu, 2011). DOCK scoring, ITScore and robust AutoDock4 scoring functions are in use for ranking the target protein (Kharkar, Warrier, & Gaud, 2014). Besides, parallel strategies have been used to improve the execution of inverse docking (McPhillie et al., 2015; Vasseur et al., 2013). The recently developed VinaMPI program decreases the time required for completing the screening process (Kharkar et al., 2014). In addition, ligand-based methods use the concept of shape similarity, where chemical compounds of similar shape tend to have similar biological activities (McPhillie et al., 2015; Vasseur et al., 2013). Ligand-based strategies utilize 2D fingerprints, 3D similarity search and reverse pharmacophore mapping for target prediction. Reverse pharmacophore and reverse docking have proved to be efficient and cost-effective target identification approaches that increase the chances of success in the drug discovery process (Liu et al., 2010).

Structure-based pharmacophore determines the chemical features with their spatial relationships in the ligand-binding site. Chemical features reflect the interactions that play a key role in the modulation of protein function by interacting with crucial residues (Wolber, Seidel, Bendix, & Langer, 2008). PharmMapper server identifies the target protein by reverse pharmacophore screening against more than 1500 drug targets from PharmTargetDB (Liu et al., 2010). Also, this server provides function and disease-relevant information for the identified target proteins. The efficiency of drug discovery relies on the knowledge of protein function and regulation in the pathogenesis of a disease or metabolic disorder (Burbaum & Tobal, 2002). Thus, additional information on signalling and metabolic pathway(s) associated with diseases would provide a basic understanding towards the role of cellular proteins in the pathogenesis of the disease.

The present investigation was aimed at identifying the diabetes target for 6-(hydroxymethyl)-3-[3,4,5-trihydroxy-6-[3,4,5-trihydroxyoxan-2-yl]oxymethyl]oxan-2-yl]oxyoxane-2,4,5-triol (SID 242078875) (Figure 1). The SID 242078875 was isolated from the ethanolic extract of Syzygium densiflorum Wall. ex Wight & Arn (Myrtaceae) fruits that exhibited antidiabetic property (Gopinath, David Raj, Nagarajan, & Karthikeyan, 2015). Further, it is necessary to determine the therapeutic role of the isolated compound to correlate with the antidiabetic effect of ethanolic extract of S. densiflorum fruits. In this work, we used PharmMapper server to obtain the human target hit for SID 242078875. Subsequently, the identified targets were screened manually against DAPD (Gopinath et al., 2015) to identify the obtained diabetes targets along with their pathway information. Further, molecular docking, binding free energy calculation, quantum mechanics/molecular mechanics (QM/MM) calculations and molecular dynamics (MD) simulation were performed to understand the mechanism of SID 242078875 in regulation of the blood glucose level.

**Materials and methods**

**Screening of potential diabetes targets for SID 242078875**

Target identification is one of the main steps in the determination of the therapeutic role of a new molecule (Hughes, Rees, Kalindjian, & Philpott, 2011). The current investigation was aimed at understanding the

![Figure 1. Structure of 6-(hydroxymethyl)-3-[3,4,5-trihydroxy-6-[3,4,5-trihydroxyoxan-2-yl]oxymethyl]oxan-2-yl]oxyoxane-2,4,5-triol (SID 242078875).](image-url)
antidiabetic effect of SID 242078875 by identifying the target protein. In this study, PharmMapper server was used to predict the disease targets from PharmTargetDB. PharmTargetDB has pharmacophore models, which describes the binding modes of known drug at the binding sites of protein–ligand complexes that were extracted from various databases. PharmMapper flexibly aligns the submitted small molecule onto each pharmacophore model of proteins in the target list and calculated the fit values between the small molecule and the pharmacophores. Further, it presents the aligned pose with the corresponding pharmacophore model and lists the candidate targets based on the fit values. Finally, PharmMapper provides the top N hits of the ranking list, from which the user may select potential target candidates (Liu et al., 2010; Xie, Kinnings, Xie, & Bourne, 2012).

In the present screening, “conformational generation” and “human protein targets” options were selected. Conformational generation allows the PharmMapper server to generate conformers for query compound by in-house multi-objective evolution algorithm-based conformational analysis program (Liu et al., 2009). Further, “Perform GA Match” with default parameters was selected from the “Advanced Options” menu. Genetic algorithm (GA) performs post-optimization of the alignment in order to maximize the fit values. As a result, PharmMapper server listed human targets that have fit value score higher than the cut-off (2.0) and the screened targets were used for further investigation. In addition, the text-based search was employed to screen the obtained human targets against DAPD (www.mkarthikeyan.bioinfoau.org/dapd). The screened diabetes target molecules were further investigated by molecular docking and MD simulation methods.

**Protein and ligand preparation**

The crystal structure of screened target proteins were obtained from Protein Data Bank (Berman et al., 2000). The structures were prepared and optimized by the addition of hydrogens, rectifying the missing atoms in side chain and loop region. Furthermore, the protein(s) was refined by fixing overlapping hydrogens through H-bond optimization and terminal flips, or by minimization. The restrained minimization was employed using Optimized Potentials for Liquid Simulations (OPLS)-2005 force field (Jorgensen, Maxwell, & Tirado-Rives, 1996). Minimization was performed until the average root mean square deviation (RMSD) of the non-hydrogen atoms reached 0.3 Å (Tripathi & Singh, 2014). The overall protein preparation was performed using glide protein preparation wizard in Maestro (Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015). The SID 242078875 was prepared by LigPrep (LigPrep, version 3.3, Schrödinger, LLC, New York, NY, 2015.), which fixed the bond orders, generated ionization states and tautomers. The possible tautomeric and ionization states of compound were generated at the pH range of 7.0 ± 2.0 and have been rigorously adjusted via Epik program rather than separate ionizer and tautomerizer treatments. The optimized structures were minimized with the OPLS-2005 force field.

**Molecular docking analysis**

Molecular docking simulation was performed in order to find the probable binding mode of SID 242078875. The prepared target proteins were subjected to grid generation to ensure that possible active sites were not missed (Schrödinger, 2015). The co-crystallized ligand was considered as the reference molecule and a grid-enclosing box was centred at the co-crystallized ligand. The grid box was generated around the ligand-binding site of the screened targets. The position of grid box is set as ATZ axis with radius 2.0 Å and the van der Waals radii of receptor atoms as 1.0 Å with a partial charge cut-off 0.25 Å to soften the potential for non-polar part of receptor. The SID 242078875 was docked onto the ligand-binding site of the screened targets using glide extra-precision (XP) docking (Glide, version 6.6, Schrödinger, LLC, New York, NY, 2015). Glide score (a modified and extended version of the empirically base function), Glide energy (modified Coulomb–van der Waals interaction energy), hydrogen bond interaction and hydrophobic interactions were considered to investigate the therapeutic effect of SID 242078875.

**QM-polarized ligand docking**

The lowest-energy docked complex was further subjected to Quantum Polarized Ligand Docking (QPLD). The QPLD protocol was employed in docking calculations rather than usual fixed charges assigned by the OPLS-AA force field. It improves the partial charges on the ligand atoms by replacing them with charges derived from quantum mechanical calculations. In this hypothesis, the protein is considered as MM region and ligand as QM region (Selvaraj, Singh, & Singh, 2014). When there are covalent connections between the QM and MM regions, it uses frozen localized molecular orbitals along the covalent bonds to construct an interface between the two regions that improves the binding pose and accuracy (Singh & Muthusamy, 2013). The atomic charge was calculated from the electrostatic potential energy surface of the ligand, which is generated from a single-point calculation using density functional theory (DFT) employing 6-31G*/LACVP* basis set, B3LYP density functional and “Ultrafine” SCF accuracy level (iaec = 1, iaecsf = 2) (Cho, Guallar, Berne, & Friesner, 2005). Glide then re-docks each ligand with updated charges
and returns the most energetically favourable pose. The QPLD protocol utilizes Glide (Glide, version 6.6, Schrödinger, LLC, New York, NY, 2015), Jaguar (Jaguar, version 8.7, Schrödinger, LLC, New York, NY, 2015) and QSite (QSite, version 6.6, Schrödinger, LLC, New York, NY, 2015). Jaguar is a high-performance *ab initio* quantum chemical program for both gas and solution phase simulations and specializes in fast electronic structure predictions for molecular systems. It performs DFT calculations and computes molecular orbitals, electron density and electrostatic potential for the molecules (Bochevarov et al., 2013). QSite is a QM/MM program that applies QM to the reactive centre of a protein active site and MM to the rest of the system. Its accuracy allows detailed understanding of reactions involving proteins, making it a powerful tool for lead optimization (Friesner & Guallar, 2005).

**E-pharmacophore generation**

The pharmacophore overlay of SID 242078875, co-crystallized “experimental” pose, and most active compounds for each target was generated from the protein–ligand complex with Phase (Phase, version 4.2, Schrödinger, LLC, New York, NY, 2015). Structure-based pharmacophore or energy-optimized pharmacophores (E-pharmacophore) approach is based on mapping of the energetic terms from the Glide XP score function onto atom centres. Structure-based pharmacophore combines the speed of pharmacophore screening with the energetic binding terms from Glide XP (Loving, Salam, & Sherman, 2009; Salam, Nuti, & Sherman, 2009).

**Prime MM-GBSA free energy calculations**

Prime/MM-GBSA approach (Prime, version 3.9, Schrödinger, LLC, New York, NY, 2015) was used to calculate the binding free energy. The docked poses obtained from QPLD were minimized using the local optimization feature in Prime. OPLS-AA 2005 force field was used to calculate the protein–ligand complex energies in GB/SA continuum solvent model (Du et al., 2011; Lyne, Lamb, & Saeh, 2006). Prime uses a surface generalized Born model employing a Gaussian surface instead of a van der Waals surface for better representation of the solvent-accessible surface area (SASA). The following equation was used to calculate the binding free energy

\[
\Delta G_{\text{bind}} = \Delta E + \Delta G_{\text{solv}} + \Delta G_{\text{SA}}
\]  

(1)

\[
\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}}
\]  

(2)

where \(E_{\text{complex}}, E_{\text{protein}}\) and \(E_{\text{ligand}}\) are the minimized energies of protein–inhibitor complex, protein and inhibitor, respectively.

\[
\Delta G_{\text{solv}} = G_{\text{solv (complex)}} - G_{\text{solv (protein)}} - G_{\text{solv (ligand)}}
\]  

(3)

where \(G_{\text{solv (complex)}}, G_{\text{solv (protein)}}, G_{\text{solv (ligand)}}\) are the solvation free energies of protein–inhibitor complex, protein and inhibitor, respectively.

\[
G_{\text{SA}} = G_{\text{SA (complex)}} - G_{\text{SA (protein)}} - G_{\text{SA (ligand)}}
\]  

(4)

where \(G_{\text{SA (complex)}}, G_{\text{SA (protein)}}, G_{\text{SA (ligand)}}\) are the surface area energies for protein–inhibitor complex, protein and inhibitor, respectively.

**QM/MM calculations**

The QPLD complexes were subjected to QM/MM calculations with QSite (QSite, version 6.6, Schrödinger, LLC, New York, NY, 2015) employing hybrid DFT with B3LYP/6-31G++** basis set (Becke, 1993; Binkley, Pople, & Hehre, 1980; Lee, Yang, & Parr, 1988; Pietro et al., 1982). The use of B3LYP hybrid functional is most popular in the calculation of properties and reactions of organic molecules (McCarren, 2009; Yanai, Tew, & Handy, 2004). It provides good results with more accuracy with respect to geometrical parameters and energy calculation (Hopmann & Himö, 2010). The cubic unit cell was neutralized and solvated with explicit water by adding simple point charge. Further, each system was neutralized

**MD simulation**

MD simulation was performed to ensure the stability of SID 242078875 in the binding cavity of the screened target proteins. MD simulation was carried out using GROMACS 4.6.3 package with the GROMOS96 (53a6) force field. The GROMOS96 (53a6) was used here to prepare the protein topology (Oostenbrink, Soares, Van der Vegt, & Van Gunsteren, 2005; Oostenbrink, Villa, Mark, & Van Gunsteren, 2004). The SID 242078875 topology was obtained from PRODRG server (Schüttelkopf & Van Aalten, 2004). The cubic unit cell was defined and solvated with explicit water by adding simple point charge. Further, each system was neutralized...
by adding either Na\(^+\) or Cl\(^-\) counter ions. The whole system was subjected to energy minimization with the help of 50,000 minimization steps and a tolerance of 1000 kJ/mol/nm. Electrostatic and van der Waals (vdW) interactions were applied based on the particle mesh Ewald method with a cut-off distance of 1.0 nm for short-range neighbour list and 1.0 nm for Coulomb cut-off and vdw interactions (Kawata & Nagashima, 2001). The LINCS (A linear constraint solver for molecular simulations) and SETTLE (An analytical version of the SHAKE and RATTLE algorithm for rigid water models) algorithms were used to constrain the bond angles and geometry of the water molecules (Hess, Bekker, Berendsen, & Fraaije, 1997; Miyamoto & Kollman, 1992). V-rescale coupling was used to regulate the temperature (weak coupling method). The Parrinello–Rahman coupling was employed to maintain a constant temperature of 300 K and a constant semi-isotropic pressure of 1 bar with a coupling time of 2 fs and the coordinates were saved (Martoňák, Laio, & Parrinello, 2003). The system was equilibrated with NVT and NPT simulations. The dynamic behaviour of protein–ligand complex was observed in an equilibrated system for 50 ns MD run with a time step of 2 fs.

**Results and discussion**

**Screening of diabetes targets for SID 242078875**

In this study, potential disease targets for SID 242078875 were predicted using PharmMapper server. PharmMapper automatically finds the best mapping poses of SID 242078875 against all the pharmacophore models in PharmTargetDB through a “reverse” pharmacophore mapping approach. It listed out 897 best mapping poses that belong to 319 human disease targets. Of the 319 human targets retrieved from PharmMapper, 5 targets such as dipeptidyl peptidase-IV (DPP-IV), protein-tyrosine phosphatase 1B (PTP1B), phosphoenolpyruvate carboxykinase (PEPCK), glycogen synthase kinase-3β (GSK-3B) and glucokinase (GK) were recognized as potential diabetes targets while performing text-based manual search of 319 targets in DAPD (Figure 2). Modulating these protein functions balance the blood glucose level and thus help in diabetes treatment (Figure 3, see supplementary data). Pathways associated with the screened targets are depicted in Table 1.

Further, pharmacophore features for SID 242078875 were generated with Phase module (Phase version 4.2, Schrödinger, LLC, New York, NY, 2015) and the alignment between the molecule (SID 242078875) and receptor-based pharmacophore models of potential targets are shown in Figure 4. The receptor-based pharmacophore feature of complex crystal structures is shown in Table 2. The alignment (Figure 4(b)-(f)) shows the selected pharmacophore features for the interaction of the drug molecule with its respective potential target.

**Molecular docking analysis**

X-ray crystal structures of DPP-IV (PDB id: 4PNZ), PTP1B (PDB id: 4I8N), PEPCK (PDB id: 1M51), GSK-3B (PDB id: 4AFJ) and GK (PDB id: 3FR0) were obtained from Protein Data Bank and prepared using protein preparation wizard. The active sites were determined to all the five targets by generating the grid with reference to its co-crystal ligands that has been reported as a potent

---

*Figure 2. Workflow of screening diabetes target by reverse pharmacophore and database search approaches.*
inhibitor/activator to the respective target. Extra precision glide docking was performed to predict the binding pose of SID 242078875 with screened targets. The predicted binding pose was obtained through molecular docking. Glide score, glide energy, hydrogen bond and hydrophobic interactions were considered for evaluating the binding affinities and the molecular basis of the interaction with SID 242078875 (Table 3). The docking scores ranged from $-14.88$ to $-10.337$ kcal/mol and Glide energies ranged from $-53.857$ to $-42.572$ kcal/mol, revealing better binding affinities for SID 242078875 with the screened targets. The root mean square deviation (RMSD) between predicted binding pose of SID 242078875 and the crystallographic coordinates of DPP-IV, PTP1B, PEPCK, GSK-3B and GK was found to be $0.37$, $0.79$, $0.47$, $0.72$ and $0.57$ Å, respectively. An RMSD of $2$ Å is considered as the cut-off for correct docking. Herein, the observed RMSD values suggest the reliability of the Glide XP docking mode in reproducing the experimentally observed binding mode (Tripathi & Singh, 2014). In addition, hydrogen bond interaction with the active residues supports potent antagonist activity of SID 242078875 with DPP-IV, PTP1B, PEPCK, and GSK-3B; agonist activity for GK (Figure S1).

**QM/MM docking analysis**

Accurate binding pose of SID 242078875 with the screened targets were predicted using QPLD and the obtained glide score, glide energy, hydrogen bond and hydrophobic interactions were tabulated (Table 4). The electronic charges play a key role in the protein–ligand interaction. XP docking obtains the charges from OPLS force field, while QPLD uses mixed QM/MM methods to compute the ligand charge distribution that increase the accuracy of docking. The fixed charges of ligand obtained from the OPLS MM force field was replaced by QM force field in the protein atmosphere. Hence, polarization effects in the quantum region significantly contribute to accurate interactions of protein–ligand complexes (Cho et al., 2005).

![Figure 3. Schematic representation of role of identified targets (DPP-IV, PTP1B, PEPCK, GSK-3B, and GK) in the pathogenesis of diabetes.](image-url)
In this study, variation in the glide score, hydrogen bond number and distances were observed between the XP docking and QPLD (Tables 3 and 4). Protein–ligand complexes obtained by XP-docking and QPLD were superimposed to check the binding mode of the SID 242078875 in the binding cavity of screened targets. Superimposed complexes clearly show that SID 242078875 is fit into the binding cavity through the interaction(s) with active site residues (Figure 5). Moreover, correlation between Glide XP and QPLD scores for isolated compound yielded a statistically significant correlation coefficient of .702. Glide score of SID 242078875-DPP-IV and SID 242078875-GK are found to be comparatively less in the QPLD; however, these complexes showed more hydrogen bond interactions than the XP docking complexes (Figure S2). Further, protein–ligand interaction analysis provides more insights into the mechanism of SID 242078875 to determine its therapeutic effect.

**Molecular recognition analysis**

Molecular recognition plays a key role in promoting fundamental biomolecular events such as drug–protein, enzyme–substrate and drug–nucleic acid interactions. The detailed understanding of drug interactions with their target protein/nucleic acid may provide a conceptual framework for determining the desired potency and therapeutic role of the drug (Mohan, Gibbs, Cummings, Jaeger, & DesJarlais, 2005). The therapeutic effect of drug molecule depends on the ability to interact with a particular binding site on target protein to exert a desired biologic effect. Ligands that share favourable interactions will exert the similar therapeutic effect on the target protein (Kortagere & Ekins, 2010). Therefore, the therapeutic role of novel compounds can be determined by analysing the favourable interaction at the binding site of the target protein. The common residues that form hydrogen bond interaction in the protein–ligand complexes of XP docking and QPLD were noted for molecular recognition analysis.

The DPP-IV binding site is highly druggable, which includes catalytic triad of Ser630 with Asp708 and His740, hydrophobic S1 pocket (Tyr631, Val656, Trp659, Tyr662, Tyr666 and Val711), hydrophobic S2 pocket (Arg125, Phe357, Arg358, Tyr547, Pro550 and Asn710) and the N-terminal recognition region (Glu205, Glu206 and Tyr662). The hydrophobic S1 pocket and the Glu205/Glu206 dyad were regarded as crucial molecular anchors for DPP-IV inhibition. The DPP-IV inhibitors showed strong interaction with the Glu205, Glu206 and residues in the hydrophobic S1 pocket of DPP-IV (Table S1) (Guasch et al., 2012). Herein, the SID 242078875–DPP-IV complex showed hydrogen bond interaction with Glu205, Glu206 and Tyr666, which are located in the hydrophobic S1 pocket and catalytic region (Figure 5(a)). Thus, it is expected that SID 242078875 can act similarly as the reported potent DPP-IV inhibitor.
The active site region of PTP1B includes PTP loop (Ser216, Ala217, Gly218, Ile219, Gly220 and Arg221), WPD loop (Asp181 and Phe182), substrate recognition loop (Asp181 and Phe182), substrate recognition loop (Lys36, Val49 and Lys120) and secondary binding site (Tyr20, Arg24, His25, Phe52 and Arg254) (Montalibet et al., 2006). The WPD hydrophobic region maintains the
normal catalytic activity of PTP1B, PTP and WPD regions are regarded as attractive hydrophobic pockets (Brandão, Johnson, & Hengge, 2012; Li et al., 2014). Potent PTP1B inhibitors were found to have interaction with PTP and WPD loop region of PTP1B (Table S2). Herein, the SID 242078875 binds with PTP and WPD loops by interacting with Ser216, Gly220, Arg221 and Phe182 (Figure 5(b)) and it theoretically supports PTP1B inhibitory activity of SID 242078875.

It is reported that PEPCK-binding pocket consists of Phe517, Phe525 and Phe530 that form the walls of binding pocket (Dunten et al., 2002). The residues Phe530 and Asn533 are found to have a role in the catalytic mechanism of PEPCK (Carlson & Holyoak, 2009). The reported potent inhibitors are found to be placed between the Phe517 and Phe530 residues of the enzyme (Table S3). The SID 242078875–PEPCK complex showed interactions with Ala287, Asn292, Trp527, Pro528, Phe530 and Asn533 (Figure 5(c)). Herein, the interaction with Phe530 and Asn533 theoretically supports the inhibitory activity of SID 242078875 on PEPCK.

GSK-3B consists of ATP site (Leu132, Tyr134, Val135, Pro136 and Arg141), Axin-binding site (Lys85, Asp133, Val135, Lys183 and Asp200) and priming site Arg96 (Arg180, Ser203, Lys205 and Val214) (Phukan, Babu, Kannoji, Hariharan, & Balaji, 2010). Besides, it is also reported that the residues Asp133, Val135, Glu137, Arg141, Gln185, Asp200 and Arg220 are responsible for strong interactions with the binding pocket of GSK-3B.

Figure 5. Hydrogen bonding interactions formed in XP docking (Left) and QPLD (Right) complexes. Superimposed binding pose of SID 242078875 in the active site of DPP-IV (a), PTP1B (b), PEPCK (c), GSK-3B (d), and GK (e) (Middle). Common hydrogen bonding residues are highlighted as a red circle, which has a major role in the modulating protein function during hyperglycemia.
The potent GSK-3B inhibitors are found to have interactions with the ATP binding site (Table S4). Herein, the Val135 and Asp200 were noted as the common interacting residues in the SID 242078875–GSK-3B complex of XP docking and QPLD (Figure 5(d)) and it is expected that the SID 242078875 can block GSK-3B by acting as ATP-competitive inhibitors.

The small molecule GK activators bind to the allosteric site and stabilize the high-glucose affinity conformations and increase the affinity of GK for glucose (Ratcliffe, 2011). This region contains hydrophobic amino acids (Val62, Ile159, Val452 and Val255) and polar amino acid residue (Arg63), which forms the hydrophobic pocket for GK activators (Kumari & Li, 2008). All the reported potent GK agonists showed interaction with Arg63 of GK (Table S5). Similarly, the SID 242078875-GK complex has also showed interaction with Arg63 (Figure 5(e)) and thus, theoretically supports glucokinase as the agonist activity of SID 242078875.

In addition to interaction analysis, co-crystal ligands were re-docked with the respective target protein to check the docking score and interactions. Re-docking analysis shows that SID 242078875 and co-crystal ligands shares most similar interactions (Tables S1–S5). Docking scores of SID 242078875 and co-crystal ligands were compared and represented as a bar graph (Figure S3). As shown in Figure S3, SID 242078875 showed better score compared to co-crystal ligands of DPP-IV inhibitors (Figure S3(a)), PTP1B inhibitors (Figure S3(b)), PEPCK inhibitors (Figure S3(c)), GSK-3B inhibitors (Figure S3(d)) and GK agonists (Figure S3(e)). However, a few of the PTP1B inhibitors and GK agonists showed better score than the SID 242078875.

Pharmacophore overlay of SID 242078875, co-crystallized “experimental” pose, and most active compounds was generated for DPP-IV, PTP1B, PEPCK, GSK-3B and GK (Figure 6(a)–(c)). This pharmacophore overlay clearly represents pharmacophore features such as hydrogen bond donors and acceptors largely contribute in the binding of SID 242078875 and most active compounds in the active site region of its respective target.

**Binding energy calculations**

The obtained QPLD posture was minimized through local optimization feature in Prime, and the complex energies are calculated in OPLS-AA force field and GBSA continuum solvent medium. MM/PBSA combines an explicit molecular mechanical model with a continuum method for the solvation free energy. This method is more efficient than the traditional free energy method (Cozzini et al., 2004). The molecular interaction mechanism of protein and ligand is based on the theory of second law of thermodynamics. Thus, it is desirable to perform lead optimization on the basis of binding free energy, which in turn provides information regarding thermodynamic behaviour of the ligand necessary for drug design (Kitamura et al., 2014).

![Figure 6](image_url)

*Figure 6. Pharmacophore overlay of SID 242078875 (a), co-crystallized “experimental” pose (b), and most active compounds (c) with DPP-IV, PTP1B, PEPCK, GSK-3B, and GK.*
Table 5. Binding free energy calculation results for SID 242078875 bound with DPP-IV, PTP1B, PEPCK, GSK-3, and GK.

| Target name | $\Delta G_{\text{coulomb}}$ | $\Delta G_{\text{vdW}}$ | $\Delta G_{\text{covalent}}$ | $\Delta G_{\text{li}}$ | $\Delta G_{\text{Hbond}}$ | $\Delta G_{\text{bind}}$ |
|-------------|-----------------------------|-------------------------|-------------------------------|------------------------|--------------------------|------------------------|
| DPP-IV      | $-52.17$                    | $-22.96$                | $4.68$                        | $-26.21$               | $-4.12$                  | $-42.86$               |
| PTP1B       | $-73.04$                    | $-12.37$                | $6.00$                        | $-39.41$               | $-4.18$                  | $-68.54$               |
| PEPCK       | $-36.40$                    | $-28.52$                | $2.98$                        | $-39.09$               | $-2.75$                  | $-75.87$               |
| GSK-3B      | $-28.63$                    | $-29.51$                | $3.85$                        | $-28.63$               | $-3.39$                  | $-66.98$               |
| GK          | $-45.42$                    | $-23.92$                | $8.11$                        | $-32.01$               | $-3.23$                  | $-53.03$               |

* Coulomb energy.
* Van der Waals energy.
* Covalent energy (internal energy).
* Lipophilic energy.
* Hydrogen-bonding energy.
* Binding free energy.

All energies are in kcal/mol.

Figure 7. 3D molecular electrostatic potential contour map of SID 242078875 in complex with DPP-IV (a), PTP1B (b), GK (c), GSK-3B (e), and PEPCK (e). The map was calculated using density functional theorem at B3LYP/6-31G++**. The rainbow colours in the map denote the electrostatic potential energy around the molecules. Electrostatic potential of the SID 242078875 is increases in the order of Red < Yellow < Green < Cyan < Blue < Purple. The map represents that electrophilic and nucleophilic attacks favours the hydrogen bonding interactions. The red region denotes the preferred site for electrophilic attack indications and purple/blue represents the preferred site for nucleophilic attack.
obtained through QPLD are represented in Table 5. Calculated energy component shows that Coulomb energy ($\Delta G_{\text{Coulomb}}$) and covalent energy ($\Delta G_{\text{covalent}}$) are the most favourable contributors to ligand binding, whereas van der Waals energy ($\Delta G_{\text{vdW}}$), lipophilic energy ($\Delta G_{\text{lipo}}$) and hydrogen bonding energy ($\Delta G_{\text{Hbond}}$) had also shown effective contribution to ligand binding but comparatively less than the Coulomb energy and covalent energy. The $\Delta G_{\text{bind}}$ values ranged from $-75.87$ to $-42.86$ kcal/mol and show better binding affinities for the SID 242078875 in the active site of screened targets.

**QM/MM calculations**

The molecular electrostatic potential, HOMO, LUMO, HOMO–LUMO energy gap (HLG) and QM/MM energy were calculated to study the stereoelectronic complementarities of SID 242078875 and the most active compounds by B3LYP/6-31G++** basic set using Jaguar module. The QM/MM energy showed that the ligand is effective in discerning correct binding poses, which is calculated as the Coulomb–van der Waals force of the complex and estimated from the electrostatic potential energy of the ligand. The value of the electrostatic potential was mapped onto an electron density isosurface corresponding to the overall molecular size of SID 242078875. The predicted electron density map shows the electron-rich and -poor regions that present the value of the electrostatic potential in addition to structure. Generally, a negative potential surface is a delineated region in the ligand for electrophilic attack, while a positive surface area defines nucleophilic attack. The electrostatic potential of SID 242078875 at a particular surface area was represented with rainbow colours to designate the electrostatic potential value (Figure 7). Herein, red represents a large negative potential value, while purple represents a large positive potential value and orange, yellow and green represent intermediate values of the electrostatic potential. The electrostatic potential map of GET 49 with DPP-IV (Figure 7(a)), PTP1B (Figure 7(b)), PEPCK (Figure 7(c)), GSK-3B (Figure 7(d)) and GK (Figure 7(e)) clearly shows the electron donor and acceptor regions of SID 242078875. The hydrogen bond interactions were formed between the SID 242078875 and targets due to electrophilic and nucleophilic attacks.

The molecular orbital energies at ground state (HOMO) and first excited state (LUMO) were calculated to characterize the reactive sites and substituent influence on the electronic structure of the compounds. The HOMO value of the SID 242078875 represents the ability to donate the electron and the LUMO value represents the ability to accept the electron (Figure 8). The energy difference of HOMO–LUMO was calculated as the band gap energy. The calculated values were compared with the most active compounds. The HOMO–LUMO energy gap measures the electron conductivity. Interestingly, the HOMO–LUMO energy gap of SID 242078875 was found to be lower than the most active compounds (Table 6). Decrease in the energy gap corresponds to the stability and possibility of a compound to form the interaction between the ligand and receptor (Reddy & Singh, 2015).

**MD simulation**

The flexibility and binding affinity of SID 242078875 with receptors (DPP-IV, PTP1B, PEPCK, GSK-3B and GK) for a period of 50 ns specific time period were studied through MD simulation. The total energy, potential

![Figure 8. HOMO and LUMO surfaces of SID 242078875 in complex with DPP-IV (a), PTP1B (b), PEPCK (c), GSK-3B (d), and GK (e).](image-url)
energy, pressure and temperature (Figures S4–S7) of the equilibrated system have been illustrated in the supplementary contents. The observed trajectories show that the MD simulation system was stable during the whole simulation period. The average values of total energy, potential energy, pressure and temperature were depicted in Table 7. The deviation in the system values was represented as standard deviation that shows the system was stable during the whole simulation period. The total energy and potential energy varies among the complex because of the variation in size and residues. The RMSD, root mean square fluctuations (RMSF), SASA, radius of gyration (Rg) and number of hydrogen bonds (H bond) were calculated to evaluate the stability and dynamic behaviour of protein–ligand complex.

The backbone RMSD of five docked complexes with respect to the initial conformation was calculated as a function of time. The conformational stability of the protein–ligand complex during the simulation period was assessed by the calculated backbone RMSD (Figure 9(a)). The SID 242078875 has been bound with five different receptors that differ in size and secondary structure, hence the initial run to obtain the equilibration phase shows variation for each complex and later it maintains the stable protein–ligand complex in equilibrated condition. The RMSD plot of five complexes show that the trajectories of SID 242078875–DPP-IV are more stable from 20 to 50 ns with slight decrease in RMSD during 40–45 ns; SID 242078875–PTP1B was stable from 32 to 50 ns; SID 242078875–PEPCK was

### Table 6. Calculated quantum chemical properties for SID 242078875 bound with DPP-IV, PTP1B, PEPCK, GSK-3 and GK.

| Target name | Ligands | HOMO (eV) | LUMO (eV) | HLG (eV) | QM/MM energy (kcal/mol) | Most negative electrostatic potential (kcal/mol) | Most positive electrostatic potential (kcal/mol) |
|-------------|---------|-----------|-----------|----------|------------------------|-----------------------------------------------|-----------------------------------------------|
| DPP-IV      | SID 242078875 | .027      | .221      | −.194    | −1847.43               | −50.39                                        | 119.19                                        |
|             | Active compound | .088      | .237      | −.149    | −1854.44               | −41.40                                        | 64.00                                         |
| PTP1B       | SID 242078875 | −.238     | −.003     | −.234    | −1818.09               | −50.98                                        | 66.33                                         |
|             | Active compound | −.183     | −.001     | −.182    | −2024.15               | −163.68                                       | 20.31                                         |
| PEPCK       | SID 242078875 | −.172     | .051      | −.223    | −1838.85               | −58.64                                        | 75.85                                         |
|             | Active compound | −.138     | .017      | −.155    | −1359.25               | −59.94                                        | 59.17                                         |
| GSK-3B      | SID 242078875 | −.357     | −.143     | −.213    | −1818.59               | −55.14                                        | 90.71                                         |
|             | Active compound | −.328     | −.168     | −.16     | −1071.05               | −41.31                                        | 44.39                                         |
| GK          | SID 242078875 | .152      | .374      | −.222    | −1827.76               | −58.09                                        | 115.89                                       |
|             | Active compound | .202      | .349      | −.147    | −1775.22               | −53.69                                        | 54.65                                         |

*Highest occupied molecular orbital energy.

| Energy (eV) | Pressure (eV) | Temperature (eV) |
|-------------|--------------|-----------------|
| DPP-IV      | −1,265,888.97 ± 1523 | −1,544,040.86 ± 1274.50 |
| PTP1B       | −379,586.90 ± 839.02  | −462,022.39 ± 700.60  |
| PEPCK       | −1,160,718.78 ± 1458  | −1,415,494.72 ± 1223.10 |
| GSK-3B      | −443,807.56 ± 898    | −541,908.62 ± 755.70  |
| GK          | −905,419.05 ± 1272.6  | −1,101,440.21 ± 1065.06 |

The values are expressed as mean ± standard deviation.
stable from 5 to 50 ns. SID 242078875–GSK-3B was stable from 23 to 50 ns; the SID 242078875–GK was stable during 19 to 25 ns and it has increased without much deviation, further it maintained the stability from 45 ns. The average RMSD of all the five complexes are shown in Figure 10, which explains the RMSD mean variation in the entire period of MD simulation and the last 10 ns (40–50 ns) of MD simulation. The fluctuation in the RMSD was represented as the standard deviation (SD). The overall SD of all complexes show the deviation in the initial run, while RMSD mean variation during the last 10 ns (40–50 ns) is much less than the overall RMSD mean variation. This confirmed the ligand stability in the last 10 ns, and it indicates the stability of the SID 242078875 with the target proteins (Figure 9(b)).

The flexibility of each residue in protein–ligand complex was evaluated by RMSF. The RMSF of hydrogen bond interacting residues of DPP-IV, PTP1B, PEPCK, GSK-3B and GK were observed to understand the mobility of residues during the whole simulation period. In SID 242078875–DPP-IV, the residues Glu205, Glu206 and Tyr666 were found to be the most important for DPP-IV inhibitory activity. RMSF of Glu205, Glu206 and Tyr666 are .17, .18 and .13 nm, respectively. Nearby residues in the binding cavity also showed very less fluctuation around ~1.5 nm (Figure 11(a)). The RMSF of residues Arg24, Asp48, Ser216, Gly220 and Arg221 in SID 242078875–PTP1B were observed at .65, .32, .11, .1 and .12 nm, respectively; it is found that RMSF for Arg24 and Asp48 are comparatively increased, because these residues lie next to the loop region (Arg24) and in the middle of the loop region (Asp48) (Figure 11(b)). In SID 242078875–PEPCK, the RMSF values for Ala287, Asn292, Trp527, Pro528, Phe530 and Asn533 were found to be .25, .16, .18, .19, .45 and .21 nm, respectively. All the interacting residues showed less fluctuation except for the loop residue Phe530 (Figure 11(c)). The SID 242078875–GSK-3B showed very less RMSF for both interacting residues Val135 (.14 nm) and Asp200 (.11 nm) (Figure 11(d)). In SID 242078875–GK, the interacting residues Tyr61, Arg63, Asp158 and Lys459 exhibit .19, .21, .24 and .3 nm, respectively. A slight increase in the RMSF is due to its position in the loop region (Arg63, Asp158) and terminal region (Lys459) (Figure 11(e)). Overall, RMSF of all five complexes were showed deviation only in the loop and terminal regions, which excludes the active site residues. However, very few residues were observed in the loop region, which has strong hydrogen bond with the SID 242078875. The RMSF of active site residues was found to be less and it confirms the protein–ligand stability in active sites of all screened targets; however, less fluctuation in the active site residues depict the atomic fluctuations of SID 242078875 and activeness inside the binding site pocket.

Total SASA of the five complexes were calculated to evaluate the overall change in the shape of receptors during the simulation (Figure 12). Further, average total SASA was calculated to estimate the SASA during simulation period and calculated values were expressed as

![Figure 9](image9.png)  
**Figure 9.** Time-dependent backbone RMSD plot of DPP-IV, PTP1B, PEPCK, GSK-3B, and GK in complex with SID 242078875 for entire simulation period (a) and the last 10 ns (40–50 ns) (b).

![Figure 10](image10.png)  
**Figure 10.** The pictorial representation of average means RMSD using bars for the entire simulation time period (50 ns) and the last 10 ns (40–50 ns). The error bars represents the corresponding standard deviations.
a mean ± standard deviation. The standard deviation represents the variation in the total SASA. The average total SASA for SID 242078875–DPP-IV, SID 242078875–PTP1B, SID 242078875–PEPCK, SID 242078875–GSK-3B, and SID 242078875–GK complexes were 344.04 ± 4.55, 170.17 ± 3.29, 295.96 ± 3.89, 194.03 ± 3.56 and 231.67 ± 5.3. Herein, all the five receptors belong to various families and differ in sequence length and size of the drug-binding cavity. Thus, the average total SASA values vary according to their protein size. However, all the complexes have shown very less deviation during MD simulation, which indicate that the ligand was bound tightly in the binding pocket and did not induce any significant change in the total SASA upon binding of SID 242078875 to the target protein.

The compactness of protein–ligand complex was measured by radius of gyration (Rg) of protein and was plotted against simulation time (Figure 13), which provides the structural changes in the protein–ligand complexes during simulation time. The Rg values of SID 242078875–DPP-IV, SID 242078875–PTP1B, SID

![Figure 11. The RMSF of the residues in DPP-IV (a), PTP1B (b), PEPCK (c), GSK-3B (d), and GK (e) in complex with SID 242078875. The RMSFs are measured using the Cα of each residue from trajectories of the entire simulations period that represents the deviation in the residue position which corresponds to the reference position.](image-url)
Figure 12. SASA of DPP-IV, PTP1B, PEPCK, GSK-3B, and GK in complex with SID 242078875.

Figure 13. The radius of gyration of DPP-IV, PTP1B, PEPCK, GSK-3B, and GK in complex with SID 242078875.

Figure 14. Hydrogen bond analysis of DPP-IV, PTP1B, PEPCK, GSK-3B, and GK in complex with SID 242078875. Vertical axis represents the number of hydrogen bond interaction between target proteins and SID 242078875.
242078875–PEPCK, SID 242078875–GSK-3B and SID 242078875–GK showed fluctuation between 2.6 and 2.76 nm with mean of 2.69 ± 0.1, 1.9–1.99 with mean of 1.96 ± 0.1, 2.43–2.53 with mean of 2.5 ± 0.1, 2.11–2.2 with mean of 2.15 ± 0.1 and 2.23–2.45 with the mean of 2.32 ± 0.4, respectively. Overall, the Rg values of SID 242078875–DPP-IV and SID 242078875–GK showed slightly higher fluctuations with standard deviation of .03 and .04 nm, respectively, it maintained the compactness after 45 ns, while the other three complexes maintain stability from the initial run of the MD simulation. The number of hydrogen bonds between receptor and ligand was calculated during simulation and plotted against time (Figure 14); it perceives that the SID 242078875 has high potential towards the acceptor–donor relations for strong interaction between protein and ligand. These acceptor–donor relation results with hydrogen bond interactions throughout the 50 ns of timescale confirmed the presence of SID 242078875 in the binding cavity of DPP-IV, PTP1B, PEPCK, GSK-3B and GK. Overall, the RMSD, RMSFs, SASA, radius of gyration and number of hydrogen bonds confirm the stability of SID 242078875 in the binding cavity of DPP-4, PTP1B, PEPCK, GSK-3B and GK.

Conclusion
Target identification is one of the main steps in determining the therapeutic role of a new molecule. In this study, we have used reverse pharmacophore mapping and databases search approaches to identify the protein targets for SID 242078875. The reverse pharmacophore screening retrieved a sum of 897 pharmacophore poses of 325 human protein targets for the query SID 242078875. The reverse pharmacophore targets for the query SID 242078875. Subsequent text-based search with DAPD identified five diabetes targets such as dipeptidyl peptidase-4, protein-tyrosine phosphatase 1B, phosphoenolpyruvate carboxykinase, glycogen synthase kinase-3β and glucokinase. Further, it was examined with molecular docking, molecular recognition analysis, binding free energy calculation, QM/MM calculations and MD. Overall, the series of computational tools explore that SID 242078875 has the potential to modulate the function of DPP-IV, PTP1B, PEPCK, GSK-3B and GK for balancing the blood glucose level. Thus, the SID 242078875 may be tested as a potential lead molecule for the treatment of diabetes.

Supplemental material
The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2015.1124289.

Acknowledgement
The authors gratefully acknowledge the infrastructure facilities provided by the Department of Bioinformatics, Alagappa University, Post Graduate Diploma in Structural Pharmacognomies.

Disclosure statement
The authors declare no competing financial interest.

Funding
This work was supported by the UGC, Government of India [number F.14-13/2013 (Inno/ASIST) dated 30.03.3013.

ORCID
Karthikeyan Muthusamy http://orcid.org/0000-0001-7755-3381
Kopinath Krishnasamy http://orcid.org/0000-0003-0063-1905

References
Agrawal, S., Dimitrova, N., Nathan, P., Udayakumar, K., Lakshmi, S. S., Siriam, S., … Sengupta, U. (2008). T2D-Db: An integrated platform to study the molecular basis of type 2 diabetes. BMC Genomics, 9, 320. doi:10.1186/1471-2164-9-320
Bashar, A., Khalipha, R., & Chowdhury, M. R. (2012). Use of computer in drug design and drug discovery: A review. International Journal of Pharmaceutical and Life Sciences, 1(2), 1–21. doi:10.3329/ijpls.v1i2.12955
Becke, A. D. (1993). Density–functional thermochemistry. III. The role of exact exchange. The Journal of Chemical Physics, 98, 5648-5652. doi:10.1063/1.464913
Berman, H. M., Westbrook, J., Feng, Z., Gililand, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The protein data bank. Nucleic Acids Research, 28, 235–242. doi:10.1093/nar/28.1.235
Binkley, J. S., Pople, J. A., & Hehre, W. J. (1980). Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. Journal of the American Chemical Society, 102, 939-947. doi:10.1021/ja00523a008
Bochevarov, A. D., Harder, E., Hughes, T. F., Greenwood, J. R., Braden, D. A., Philipp, D. M., … Friesner, R. A. (2013). Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. International Journal of Quantum Chemistry, 113, 2110–2142. doi:10.1002/qua.24481
Brandão, T. A. S., Johnson, S. J., & Hengge, A. C. (2012). The molecular details of WPD-loop movement differ in the protein-tyrosine phosphatases YopH and PTP1B. Archives of Biochemistry and Biophysics, 525, 53–59. doi:10.1016/j.abb.2012.06.002
Burbay, J., & Tobal, G. M. (2002). Proteomics in drug discovery. Current Opinion in Chemical Biology, 6, 427–433. doi:10.1016/S0929-8665(02)00102-4
Burren, O. S., Adlem, E. C., Achuthan, P., Christensen, M., Coulson, R. M. R., & Todd, J. A. (2011). T1DBase: Update 2011, organization and presentation of large-scale data sets for type 1 diabetes research. Nucleic Acids Research, 39 (Suppl. 1), D997–D1001. doi:10.1093/nar/gkq912
Carlson, G. M., & Holyoak, T. (2009). Structural insights into the mechanism of phosphoenolpyruvate carboxykinase catalysis. *Journal of Biological Chemistry*, 284, 27037–27041. doi:10.1074/jbc.R109.040568

Chen, Y. Z., & Ung, C. Y. (2001). Prediction of potential toxicity and side effect protein targets of a small molecule by a ligand–protein inverse docking approach. *Journal of Molecular Graphics & Modelling*, 20, 199–218. doi:10.1016/S1093-2647(00)01099-7

Cho, A. E., Guillar, V., Berne, B. J., & Friesner, R. (2005). Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *Journal of Computational Chemistry*, 26, 915–931. doi:10.1002/jcc.20222

Cozzini, P., Fornabaio, M., Marabotti, A., Abraham, D. J., Kellogg, G. E., & Mozarelli, A. (2004). Free energy of ligand binding to protein: Evaluation of the contribution of water molecules by computational methods. *Current Medicinal Chemistry*, 11, 3093–3118. doi:10.2174/0929867043363929

Du, J., Sun, H., Xi, L., Li, J., Yang, Y., Liu, H., & Yao, X. (2011). Molecular modeling study of checkpoint kinase 1 inhibitors by multiple docking strategies and prime/MM-GBSA calculation. *Journal of Computational Chemistry*, 32, 2800–2809. doi:10.1002/jcc.21859

Dunten, P., Belunis, C., Crowther, R., Holfelder, K., Kammlott, U., Levin, W., ... Werther, S. J. (2002). Crystal structure of human cytosolic phosphoenolpyruvate carboxykinase reveals a new GTP-binding site. *Journal of Molecular Biology*, 316, 257–264. doi:10.1016/S0022-2836(01)003564

Emig, D., Iliiev, A., Pong, O., Lancaster, L., Bureeva, S., Nikolsky, Y., & Bessarabanov, M. (2013). Drug target prediction and repositioning using an integrated network-based approach. *PLoS ONE*, 8, e60618. doi:10.1371/journal.pone.0060618

França, T. C. C. (2015). Homology modeling: An important tool for the drug discovery. *Journal of Biomolecular Structure and Dynamics*, 33, 1780–1793. doi:10.1080/07391102.2014.971429

Friesner, R. A., & Guallar, V. (2005). Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annual Review of Physical Chemistry*, 56, 389–427. doi:10.1146/annurev.physchem.56.040904.101643

Gu, Z., Li, H., Zhang, H., Liu, X., Luo, X., ... Jiang, H. (2008). PDTD: A web-accessible protein database for drug target identification. *BMC Bioinformatics*, 9, 104. doi:10.1186/1471-2105-9-104

Gopinath, K., David Raj, C., Nagarajan, S., & Karthikeyan, M. (2015). Antidiabetic, antihyperlipidemic and antioxidant activity of Syzygium densiflorum fruits in streptozotocin and nicotinamide induced diabetic rats. Submitted.

Gopinath, K., Jayakumararaj, R., & Karthikeyan, M. (2015). DAPD: A knowledgebase for diabetes associated proteins. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 12, 604–610. doi:10.1109/TCBB.2014.2359442

Grolmusz, V. I. (2015). Identifying diabetes-related important protein targets with few interacting partners with the PageRank algorithm. *Royal Society Open Science*, 2, 140252-1–140252-9. doi:10.1098/rsos.140252

Guasch, L., Ojeda, M. J., González-Abuin, N., Sala, E., Cereto-Massagué, A., Mulero, M., ... Pujadas, G. (2012). Identification of novel human dipeptidyl peptidase-IV inhibitors of natural origin (part I): Virtual screening and activity assays. *PLoS ONE*, 7, e44971. doi:10.1371/journal.pone.0044971

Gu, J., Chen, L., Yuan, G., & Xu, X. (2013). A drug-target network-based approach to evaluate the efficacy of medicinal plants for type II diabetes mellitus. *Evidence-Based Complementary and Alternative Medicine: eCAM*, 203614, 1–7. doi:10.1155/2013/203614

Haanstra, J. R. & Bakker, B. M. (2015). Drug target identification through systems biology. *Drug Discovery Today: Technologies*, 15, 17–22. doi:10.1016/j.ddct.2015.06.002

Harvey, A. L., Edrada-Ebel, R., & Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery*, 14, 111–129. doi:10.1038/nrd4510

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18, 1463–1472. doi:10.1002/jcc.5870181121

Hopmann, K. H., & Himo, F. (2010). Quantum chemical modeling of enzymatic reactions – Applications to epoxide-transforming enzymes. In Newnes (Ed.), *Comprehensive natural products II* (pp. 719–747). Philadelphia: Elsevier. doi:10.1016/B978-080045382-8.00160-X

Hughes, J. P., Rees, S. S., Kalindjian, S. B., & Philipk, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, 162, 1239–1249. doi:10.1111/j.1476-5381.2010.01127.x

IDF. (2014). *IDF diabetes atlas* (6th ed.) Belgium: International Diabetes Federation.

Jensen, F. (Ed.) (2006). Basis sets. In *Introduction to computational chemistry* (2nd ed., pp. 192–232) Hoboken, NJ: Wiley.

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118, 11225–11236. doi:10.1021/ja9621760

Kawata, M., & Nagashima, U. (2001). Particle mesh Ewald method for three-dimensional systems with two-dimensional periodicity. *Chemical Physics Letters*, 340, 165–172. doi:10.1016/S0009-2614(01)00393-1

Kharkar, P. S., Warrier, S., & Gaud, R. S. (2014). Reverse docking: A powerful tool for drug repositioning and drug rescue. *Future Medicinal Chemistry*, 6, 333–342. doi:10.4155/fmc.13.207

Kitamura, K., Tamura, Y., Ueki, T., Ogata, K., Noda, S., Himeno, R., & Chuman, H. (2014). Binding free-energy calculation is a powerful tool for drug optimization: Calculation and measurement of binding free energy for 7-azaindole derivatives to glycogen synthase kinase-3β. *Journal of Chemical Information and Modeling*, 54, 1653–1660. doi:10.1021/ci400719v

Kore, P., Mutha, M., & Antre, R. (2012). Computer-aided drug design: An innovative tool for modeling. *Open Journal of Medicinal Chemistry*, 2, 139–148. doi:10.4236/ojmc.2012.24017

Kortagere, S., & Ekins, S. (2010). Troubleshooting computational methods in drug discovery. *Journal of Pharmacological and Toxicological Methods*, 61, 67–75. doi:10.1016/j.jpharm.2010.02.005

Kramer, T., Schmidt, B., & Lo Monte, F. (2012). Small-molecule inhibitors of GSK-3: Structural insights and their application to Alzheimer’s disease models. *International Journal of Alzheimer’s Disease*, 2012, 1–32. doi:10.1155/2012/381029
Antidiabetic therapeutic potential Syzygium densiflorum

Kumari, V., & Li, C. (2008). Comparative docking assessment of glucokinase interactions with its allosteric activators. Current Chemical Genomics, 2, 76–89. doi:10.2174/1875973080802010076

Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical Review B, 37, 785–789. doi:10.1103/PhysRevB.37.785

Li, H., Gao, Z., Kang, L., Zhang, H., Yang, K., Yul, K., ... Jiang, H. (2006). TarFisDock: A web server for identifying drug targets with docking approach. Nucleic Acids Research, 34(Web. SERV. ISS.), W219–W224. doi:10.1093/nar/gkl1114

Li, S., Zhang, J., Lu, S., Huang, W., Geng, L., Shen, Q., & Zhang, J. (2014). The mechanism of allosteric inhibition of protein tyrosine phosphatase 1B. PLOS ONE, 9, e97668. doi:10.1371/journal.pone.0097668

Liu, C., Constantinides, P. P., & Li, Y. (2014). Research and development in drug innovation: Reflections from the 2013 bioeconomy conference in China, lessons learned and future perspectives. Acta Pharmaceutica Sinica B, 4, 112–119. doi:10.1016/j.apsb.2014.01.002

Liu, X., Bai, F., Ouyang, S., Wang, X., Li, H., & Jiang, H. (2009). Cyndi: A multi-objective evolution algorithm based method for bioactive molecular conformational generation. BMC Bioinformatics, 10, 101. doi:10.1186/1471-2105-10-101

Liu, X., Ouyang, S., Yu, B., Liu, Y., Huang, K., Gong, J., ... Jiang, H. (2010). PharmMapper server: A web server for potential drug target identification using pharmacophore mapping approach. Nucleic Acids Research, 38(Suppl. 2), 5–7. doi:10.1093/nar/gkq300

Loving, K., Salam, N. K., & Sherman, W. (2009). Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. Journal of Computer-Aided Molecular Design, 23, 541–554. doi:10.1007/s10822-009-9268-1

Lyne, P. D., Lamb, M. L., & Saehe, J. C. (2006). Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. Journal of Medicinal Chemistry, 49, 4805–4808. doi:10.1021/jm060522a

Martonák, R., Laio, A., & Parrinello, M. (2003). Predicting crystal structures: The Parrinello–Rahman method revisited. Physical Review Letters, 90, 075503. doi:10.1103/PhysRevLett.90.075503

McCarren, P. R. (Ed.). (2009). Assessment of DFT methods in the study of transition metals and organometallic compounds. In Computational investigations of organometallic and S-nitrosothiol reaction mechanisms (pp. 151–180). Ann Arbor, Michigan: ProQuest.

McPhailie, M. J., Cain, R. M., Naramore, S., Fishwick, C. W. G., & Simonds, K. J. (2015). Computational methods to identify new antibacterial targets. Chemical Biology & Drug Design, 85, 22–29. doi:10.1111/cbdd.12385

Miyamoto, S., & Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. Journal of Computational Chemistry, 13, 952–962. doi:10.1002/jcc.540130805

Mohar, V., Gibbs, A. C., Cummings, M. D., Jaeger, E. P., & Desjarlais, R. L. (2005). Docking: Successes and challenges. Current Pharmaceutical Design, 11, 323–333. doi:10.2174/1381612053382106

Montalibet, J., Skorey, K., McKay, D., Scapin, G., Asante-Appiah, E., & Kennedy, B. P. (2006). Residues distant from the active site influence protein-tyrosine phosphatase 1B inhibitor binding. Journal of Biological Chemistry, 281, 5258–5266. doi:10.1074/jbc.M511546200

Newman, D. J., & Cragg, G. M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. Journal of Natural Products, 75, 311–335. doi:10.1021/np200906s

Oktaramurthy, M., Veerapur, V. P., Thippeswamy, B. S., Madhusudana Reddy, T. N., Rayappa, H., & Badami, S. (2013). Anti-diabetic and anti-cataract effects of Chromolaena odorata Linn.; in streptozotocin-induced diabetic rats. Journal of Ethnopharmacology, 145, 363–372. doi:10.1016/j.jep.2012.11.023

Oostenbrink, C., Soares, T. A., Van der Vegt, N. F. A., & Van Gunstersen, W. F. (2005). Validation of the 53A6 GROMOS force field. European Biophysics Journal, 34, 273–284. doi:10.1007/s00249-004-0448-6

Oostenbrink, C., Villas, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. Journal of Computational Chemistry, 25, 1656–1676. doi:10.1002/jcc.20090

Osadbe, P., Odoe, E., & Uzor, P. (2014). Natural products as potential sources of antidiabetic drugs. British Journal of Pharmaceutical Research, 4, 2075–2095. doi:10.9734/BJPJR/2014/8382

Phukan, S., Babu, V., Kannoji, A., Hariharan, R., & Balaji, V. (2010). GSK3β: Role in therapeutic landscape and development of modulators. British Journal of Pharmacology, 160, 1–19. doi:10.1111/j.1476-5381.2010.00661.x

Pietro, W. J., Franc, M. M. Hehre, W. J., Defrees, D. J., Pople, J. A., & Binkley, J. S. (1982). Self-consistent molecular orbital methods. 24. Supplemented small split-valence basis sets for second-row elements. Journal of the American Chemical Society, 104, 5039–5048. doi:10.1021/ja00383a007

Ratcliffe, A. J. (2011). The drug discovery and development of kinase inhibitors outside of oncology. In R. A. Ward & F. Goldberg (Eds.), Kinase drug discovery (pp. 218–243). Cambridge: Royal Society of Chemistry. doi:10.1039/9781849733557-00218

Reddy, K., & Singh, S. (2015). Insight into the binding mode between N-methyl pyridimiones and prototype foamy virus integrase-DNA complex by QM-polarized ligand docking and molecular dynamics simulations. Current Topics in Medicinal Chemistry, 15, 43–49. doi:10.2174/156802661566150112110948

Salam, N. K., Nuti, R., & Sherman, W. (2009). Novel method for generating structure-based pharmacophores using energetic analysis. Journal of Chemical Information and Modeling, 49, 2356–2368. doi:10.1021/ci900212v

Schenone, M., Dančík, V., Wagner, B. K., & Clemons, P. A. (2013). Target identification and mechanism of action in chemical biology and drug discovery. Nature Chemical Biology, 9, 232–240. doi:10.1038/nchembio.1199

Schrödinger. (2015). Receptor grid generation. In Glide 6.6 user manual (pp. 25–44).

Schütte-Kopf, A. W., & Van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein–ligand complexes. Acta Crystallographica Section D: Biological Crystallography, 60, 1355–1363. doi:10.1107/S0907444904011679

Schwartz, D., Bergsten, P., Baek, K., De La, A. B., Cantley, J., & Dayon, L. (2015). The human diabetes proteome project
Selvaraj, C., Singh, P., & Singh, S. K. (2014). Molecular modeling studies and comparative analysis on structurally similar HTLV and HIV protease using HIV-PR inhibitors. *Journal of Receptors and Signal Transduction*, 34, 361–371. doi:10.1016/j.trprot.2014.03.001

Singh, K. D., & Muthusamy, K. (2013). Molecular modeling, quantum polarized ligand docking and structure-based 3D-QSAR analysis of the imidazole series as dual AT1 and ETA receptor antagonists. *Acta Pharmacologica Sinica*, 34, 1592–1606. doi:10.1038/aps.2013.129

Tripathi, S. K., & Singh, S. K. (2014). Insights into the structural basis of 3,5-diaminoindazoles as CDK2 inhibitors: Prediction of binding modes and potency by QM-MM interaction, MESP and MD simulation. *Molecular BioSystems*, 10, 2189. doi:10.1039/C4MB00077C

Vasseur, R., Baud, S., Steffenel, L. A., Vigouroux, X., Martiny, L., Krajecki, M., & Dauchez, M. (2013). Parallel strategies for an inverse docking method. In *Proceedings of the 20th European MPI Users’ Group Meeting on – EuroMPI ’13* (p. 253). New York, NY: ACM Press. doi:10.1145/2488551.2488584

Walke, D. W., Han, C., Shaw, J., Wann, E., Zambrowicz, B., & Sands, A. (2001). *In vivo* drug target discovery: Identifying the best targets from the genome. *Current Opinion in Biotechnology*, 12, 626–631. doi:10.1016/S0958-1669(01)00271-3

Wang, J.-C., Chu, P.-Y., Chen, C.-M., & Lin, J.-H. (2012). iTarget: A web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. *Nucleic Acids Research*, 40, W393–W399. doi:10.1093/nar/gks496

Wang, Q., Zhao, Z., Shang, J., & Xia, W. (2014). Targets and candidate agents for type 2 diabetes treatment with computational bioinformatics approach. *Journal of Diabetes Research*, 2014(Gse 38642), 763936-1–763936-8. doi:10.1155/2014/763936

Wolber, G., Seidel, T., Bendix, F., & Langer, T. (2008). Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discovery Today*, 13, 23–29. doi:10.1016/j.drudis.2007.09.007

Xie, L., Kinnings, S. L., Xie, L., & Bourne, P. E. (2012). Predicting the polypharmacology of drugs: Identifying new uses through chemoinformatics, structural informatics, and molecular modeling-based approaches. In M. J. Barratt & D. E. Frail (Eds.), *Drug repositioning* (pp. 163–205). Hoboken, NJ: Wiley. doi:10.1002/9781118274408.ch7

Yanai, T., Tew, D. P., & Handy, N. C. (2004). A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters*, 393, 51–57. doi:10.1016/j.cplett.2004.06.011

Yang, Z., Yang, J., Liu, W., Wu, L., Xing, L., Wang, Y., Cheng, Y. (2013). T2D@ZJU: A knowledgebase integrating heterogeneous connections associated with type 2 diabetes mellitus. *Database, 2013*, bat052. doi:10.1093/database/bat052

Zheng, M., Liu, X., Xu, Y., Li, H., Luo, C., & Jiang, H. (2013). Computational methods for drug design and discovery: Focus on China. *Trends in Pharmacological Sciences*, 34, 549–559. doi:10.1016/j.tips.2013.08.004

Zheng, R., Chen, T., & Lu, T. (2011). A comparative reverse docking strategy to identify potential antineoplastic targets of tea functional components and binding mode. *International Journal of Molecular Sciences*, 12, 5200–5212. doi:10.3390/ijms12085200