Supporting Material

Multimodal Generative Neural Networks and Molecular Dynamics based Identification of PDK1 PIF-pocket modulators

Kailasam N. Vennila* [a] and Kuppanagounder P. Elango [b]

Free energy metadynamics of PDK1

The advanced free energy calculation of protein-ligand binding reaction known as metadynamics was applied to observe the different protein conformations upon ligand binding. Here, a starting structure for the Metadynamics analysis was fetched out from molecular docking analysis which possess energetically more favourable conformation has been subjected to study the formation of complex, stability and interaction mode by using Desmond in Schrodinger Suite. The complex was first soaked in an orthorhombic water box containing 9613 water molecules (TIP3P water model) totalling to 33581 atoms, all atom force field (OPLS-2005) and 0.15 M NaCl were added to the system to mimic the physiological system. After the energy minimization, simulations were carried out at 300 K and 1 bar pressure [1]. The simulation was carried out for 200 ns at NPT ensemble and the trajectory files were collected at every successive 200 ps as snapshots. The two collective variables used for metadynamics simulation are: (i) CV1, αB helix end to end distance and ii) CV2, distance between Gly rich loop and Asp205 which describes the kinase hinge motion. Finally, the setup was minimized and pre-equilibrated using the default relaxation routine [1-4].

Thousand frames were collected during 200 ns and the Root mean square fluctuations, Root mean square deviation, protein-ligand interaction and energy potential were acquired. The Root Mean Square Deviation (RMSD) measured the average change in displacement of a

selection of atoms (Protein/ligand) for a particular frame with respect to the first reference frame.

The free energy landscape obtained is given in Figure S1 in which three energy minima were obtained. One of the minima are well populated with closed structures with hall mark salt bridge between Lys111-Glu130 and hinge distance < 20 Å. The other minimum was populated with open structures of hinge distance > 25 Å without the salt bridge. These free energy landscape with least energy of -8 to -9.8 kcal/mol discriminates the equilibrium transition between open and closed state of PDK1 upon ligand binding. In addition to these two minima, the third one also corresponds to open structure but with very much disturbed N lobe and unwind αB helix in the PIF pocket.

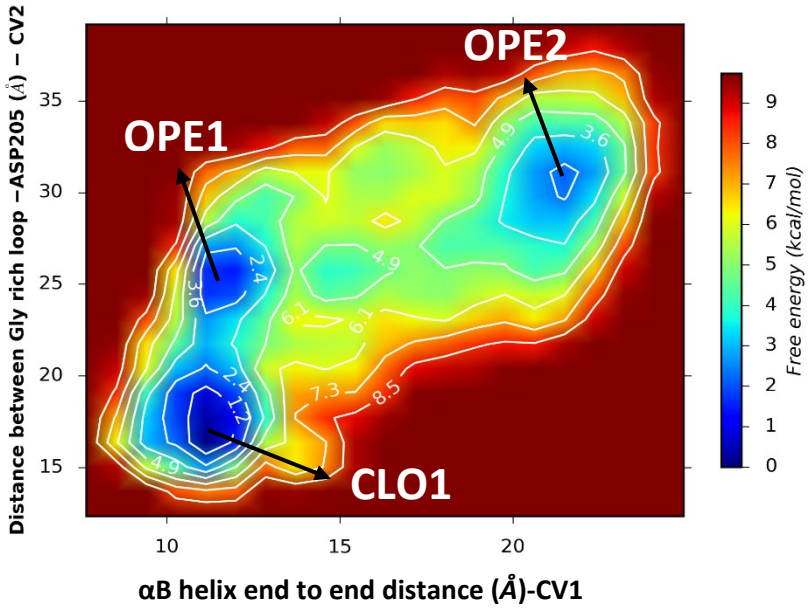
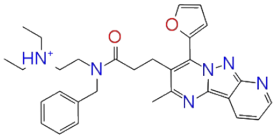
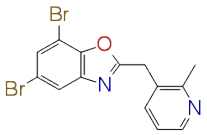
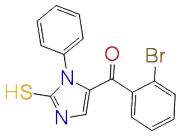
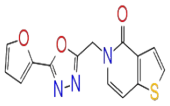
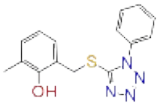
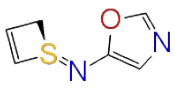
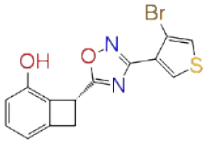
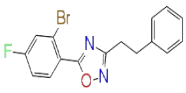
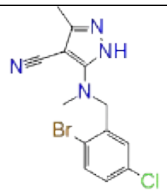
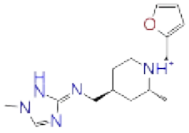
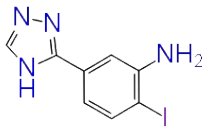
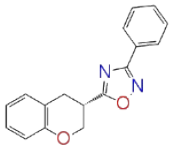
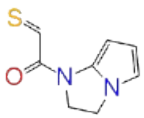
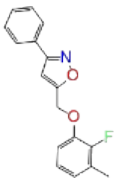
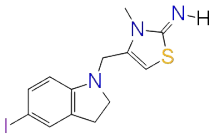
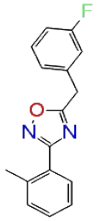
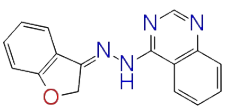


Figure S1: Free energy surface calculated from metadynamics of PDK1 depicting three energy minima without off setting to zero.

Analysis of generated ligand datasets

The datasets generated by LiGANN, a structure-based drug generator using multimodal generative models for five model systems were analysed to find out any repetition of ligands among the five sets. For this process, the 20 ligands obtained all the filtering stages were checked and they were found to be distributed as given in Table T1.

5LVO	4RRV	OPE1	CLO1	OPE2
 Score: -18.552 ClogP: 3.732 QED: 0.328	 Score: -20.038 ClogP: 4.647 QED: 0.677	 Score: -20.524 ClogP: 4.155 QED: 0.655	 Score: -23.144 ClogP: 2.754 QED: 0.571	 Score: -24.262 ClogP: 2.969 QED: 0.759
 Score: -18.791 ClogP: 1.635 QED: 0.682	 Score: -22.896 ClogP: 3.954 QED: 0.771	 Score: -23.976 ClogP: 4.423 QED: 0.721	 Score: -18.5 ClogP: 3.642 QED: 0.949	 Score: -22.132 ClogP: 0.126 QED: 0.852
 Score: -18.073 ClogP: 1.658 QED: 0.745	 Score: -19.327 ClogP: 3.455 QED: 0.739	 Score: -18.623 ClogP: 0.834 QED: 0.653	 Score: -21.673 ClogP: 4.368 QED: 0.727	 Score: -19.197 ClogP: 2.733 QED: 0.725
 Score: -21.373 ClogP: 3.775 QED: 0.746	 Score: -21.405 ClogP: 2.838 QED: 0.764			

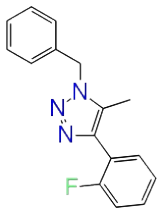
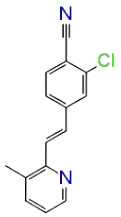
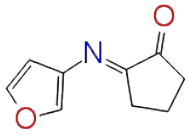
 Score: -24.096 ClogP: 3.441 QED: 0.734	 Score: - -18.76 ClogP: 4.086 QED: 0.835			
 Score: -27.555 ClogP: 2.105 QED: 0.703				

Table T1: Distribution of identified hit ligands among the chosen five model systems

References:

1. S. Ponkarpagam, G. Mahalakshmi, K. N. Vennila, & K. P. Elango. *J. Biomol. Struct. Dyn.* 2021, 1–11.
2. J. O. Schulze, G. Saladino, K. Busschots, S. Neimanis, E. Süß, D. Odadzic, et al., Cell Chem. Biol. 2016, 23, 1193.
3. W. Liu, P. Li, & Y.Mei. *J Mol Model*, 2019, *25*(7), 187.
4. S. Ponkarpagam, K. N. Vennila, & K. P. Elango. *J. Biomol. Struct. Dyn.* 2022, 1–9.