International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
(A Peer Reviewed, Referred, Indexed and Open Access Journal)
DOI: 10.22192/ijarbs Coden: IJARQG (USA) Volume 9, Issue 5 -2022

Research Article

Special Issue on Potential Applications of Bioinformatics in Biological Sciences - PABBS 2022

DOI: http://dx.doi.org/10.22192/ijarbs.2022.09.05.01.011

In silico analysis of polystic kidney disease using molecular docking and virtual screening

R. Priya1*, M.Vinoth[2], Mahendran Radha[3]
Department of Bioinformatics, Vels institute of Science and Technology & Advanced Studies, Chennai.
*Corresponding e-mail : *priya.sls@velsuniv.ac.in*

Abstract

A designed to perform blind docking at predicted sites, instead of the entire surface of a protein. Therefore, the first step is to detect putative binding sites (cavity detection). Since the ligand binding sites are usually larger cavities, we select several top cavity according to cavity size for further analysis (cavity sorting). Then, we calculate the docking center and adjust the docking box size. These parameters are required for molecular docking with AutoDock Vina (center and size). After the docking process finished, the bound poses are reranked according to the docking score (Dock and Rerank). The first conformation is considered as the best binding pose and the corresponding site is the optimal binding site for the query ligand.

As the number of elucidated protein structures is rapidly increasing, the growing data call for methods to efficiently exploit the structural information for biological and pharmaceutical purposes. Given the three-dimensional (3D) structure of a protein and a ligand, predicting their binding sites and affinity are a key task for computer-aided drug discovery. To address this task, a variety of docking tools have ben developed. Most of them focus and sizes with a novel curvature-based cavity detection approach, and performs docking with a popular docking program, AutoDock Vina. Whose root mean square deviation (RMSD) were within 2Å from the Xraypose, which outperformed the state-of-the-art blind docking tools in our benchmark tests. CB-Dock offers an interactive 3D visualization 3D visualization of results.

Virtual molecular screening is used to dock small-molecules libraries to a macromolecules in order to find lead compounds with desire biological function. This in silico method is well known for its application in computer-aided drug design. This chapter describes how to perform small-molecule virtual screening by docking with PyRX, which is open-source software with an intuitive user interface that runs on all major operating systems. Specific steps for using PyRx, as well as considerations for data preparation, docking and analysis, are also described.

Keywords: Virtual molecular screening, Computer-aided drug design, Molecular docking, AutoDock, Vina, Open babel, PyRx.

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is one of the most common monogenic disorders in humans, affecting 1 in 1000 individuals. Its major manifestation is progressive cystic dilatation of the renal tubules, leading to renal failure in half of affected individuals by age 50. ADPKD is also associated with hepatic, pancreatic, and splenic cysts, cardiac valve abnormalities, and an increased incidence of cranial aneurysms and subarachnoid hemorrhage. Despite intensive investigation, the underlying biochemical defect in ADPKD remains unknown. A series of apparently unrelated abnormalities has been detected at the cellular and tissue levels both in ADPKD and in other forms of renal cystic disease. The most carefully documented of these findings are abnormalities in the composition of the tubular basement membrane, proliferation of tubular epithelial cells, and a reversal of the normal polarized distribution of cell membrane proteins. Phenotypically indistinguishable forms of ADPKD are caused by mutations in three separate Two of these loci, PKD1 and PKD2, have been mapped to the short arm of chromosome 16 and chromosome 4, respectively. The third locus has not been mapped. Mutations in PKDI account for approximately 90% of ADPKD cases. This locus previously had been mapped to a gene-rich 500 kb interval in band 16p13.3 that includes the TSC2 locus for tuberous sclerosis (TS). Some TS patients are known to develop renal cystic lesions that resemble those of ADPKD, which led investigators to examine families with TS for positional segregation of ADPKD.

Autodock:

Protein–ligand docking has been widely used to predict binding modes and affinities of ligands. Protein–ligand docking is a powerful tool for computer-aided drug discovery (CADD). Most docking tools require the ligand binding region in advance to search for the most energy favourable binding mode. The binding region is usually represented as a cubic box, so its size and center are critical for accurate docking because it defines the boundaries of the conformational sampling space. In many application scenarios, the binding regions are unknown. To identify potential interactions between a given protein and a ligand then perform several rounds of protein–ligand docking to obtain the final result. we described a new blind docking tool, named CB-Dock, which focuses on enhancing the docking accuracy. CB-Dock predicts binding regions of a given protein, calculates the centers and sizes with a curvature-based cavity detection approach, and performs docking with the state-of-the-art docking software Auto dock Vina CB-Dock also ranks the binding modes according to Vina scores and provides an interactive 3D visualization of the binding modes.

Virtual screening:

Virtual screening, also called *in silico* screening, can be viewed as a funnel approach where one or more computational methods are used to select, from a pull of candidate molecules (usually in a molecular database), a subset of compounds for experimental validation. One of the major goals is to increase the probabilities of identifying active compounds. Depending on the information available for the system, the search can be performed using structure-based methods such as moleculardocking, or ligand-based methods such as similarity searching.

Thus far, virtual screening has proven to be quite successful in identifying small molecule. A primary example is the discovery of docking-based virtual screening of the NCI database. A second example is the identification also using docking-based screening of the NCI database.

Protein intraction:

As interactions between proteins represent such a crucial component for modern biology, STRING is by far not the only online resource dedicated to this topic. Apart from the primary databases that hold the experimental data in this field and hand-curated databases serving expert annotations a number of resources take a meta-analysis

approach, similar to STRING. Within this wide variety of online resources and databases dedicated to interactions, STRING specializes in three ways:

- It provides uniquely comprehensive coverage, with >1000 organisms, 5 million proteins and >200 million interactions stored;
- It is one of very few sites to hold experimental, predicted and transferred interactions, together with interactions obtained through text mining;
- It includes a wealth of accessory information, such as protein domains and protein structures, improving its day-to-day value for users.

Materials and Methods

Materials:

1.**PDB**-PROTEIN DATA BANK is a database for the three-dimensional structural data of large biological molecules, such as protein and nucleic acids.

2.**PUBCHEM**-This is a database of chemical molecules and their activities against biological assays.

3.**QED**- Speed of drug discovery is very necessary today to meet the new drugs in market with increase in the incidences of diseases and drug resistance strains.

4.**NCBI**- National Service for Biological Information it is an houses a series of databases relevant to biotechnology and biomedicine and is an important resource for bioinformatics tools and services.

5.**UNIPROT**- The mission of UNIPROT is to providethe scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

6.**MEGA**-MOLECULAR EVOLUTIONARY GENETICS analysis is computer software for conducting statistical analysis of molecular evolution and for constructing phylogenetic trees.

7.**STRING**-STRING is a database of known and predicted protein – protein interactions.

8.**CB DOCK**-CB-DOCKis a protein – ligand docking method which automatically indentifies the binding sites, calculates the center and size, customizes the docking box size according to the query ligands.

9.**OPENBABEL**- It's a tool to convert most of the chemical file format, analyze or store data from molecular modelling.

10.**PYRX**- Is virtual screening software for computational drug discovery that can be used to screen libraries of compounds against potential drugs targets.

Methods:

Protein Structure and Sequence Repositories:

PDB is the most comprehensive repository of structure data for biological macromolecules. The repository contains the primary structure and secondary structure information along with the atomic coordinates of a constituent atoms of biomolecule. It also contains corresponding experimental data. PDB101, an education portal of PDB provides detailed information about the PDB. As of 27th September 2017, PDB contains structure data for 133,920 Biological Macromolecular Structures. On an average, the length of proteins ranges between 100 and 300 residues. However, there are big proteins containing 1000 or more residues as well as small proteins with at most 30 residues.

Compound identification

PUBCHEM is an open archive accepting information from many sources about a given molecule, it is imperative to provide the end-user with an aggregated view of all that is known for a single chemical structure. PubChem Compound records are derived summaries that give users access to a rich set of related content.

Compound records contain unique chemical structures extracted from contributed Substance records through a process called ‘standardization’. Each Compound record points to at least one Substance record. In contrast, a Substance record might have no derived Compound record if the structure cannot be standardized or is missing.

In silico drug likeness

The concept of drug-likeness, established from the analyses of the physiochemical properties and structural features of existing small organic drugs or/and drug candidates, has been widely used to filter out compounds with undesirable properties, especially poor ADMET (absorption, distribution, metabolism, excretion, and toxicity) profiles. Here, we summarize various approaches for drug-likeness evaluations, including simple rules/filters based on molecular properties/structures and quantitative prediction models based on sophisticated machine learning methods, and provide a comprehensive review of recent advances in this field. Moreover, the strengths and weaknesses of these approaches are briefly outlined. Finally, the drug-likeness analyses of natural products and traditional Chinese medicines are discussed.

Molecular Evolutionary Relationship of protein

The Molecular Evolutionary Genetics Analysis (MEGA) software analysis of homologous gene sequences either from multigene families or from different species with a special emphasis on inferring evolutionary relationships and patterns of DNA and protein evolution. In addition to the tools for statistical analysis of data, MEGA provides many convenient facilities for the assembly of sequence data sets from files or web-based repositories, and it includes tools for visual presentation of the results obtained in the form of interactive phylogenetic trees and evolutionary distance matrices.

Protein- protein interaction

A protein's functions using protein-protein interaction data and the functional annotations of its interaction protein partners. For each function of interest and a protein, predict the probability that the protein has that function using. Unlike in other available approaches for protein annotation where a protein has or does not have a function of interest, that give a probability for having the function. This probability indicates how confident we are about the prediction. We apply our method to predict cellular functions (43 categories including a category "others") for yeast proteins defined in the Yeast Proteome Database, using the protein-protein interaction data from the Munich Information Center for Protein Sequences. That our approach outperforms other available methods for function prediction based on protein interaction data.

Protein-ligand docking with binding sites.

The three-dimensional (3D) structure of a protein and a ligand, predicting their binding sites and affinity are a key task for computer-aided drug discovery. To address this task, a variety of docking tools have been developed. Most of them focus on docking in the preset binding sites given by users. To automatically predict binding modes without information about binding sites, we developed a user-friendly blind docking web server, named CB-Dock, which predicts binding sites of a given protein and calculates the centers and sizes with a novel curvature-based cavity detection approach, and performs docking with a popular docking program, Auto dock Vina.

Virtual screening

Virtual molecular screening is used to dock small molecule libraries to a macromolecule in order to find lead compounds with desired biological function. This in silico method is well known for its application in computer-aided drug design. In this chapter, we describe how to perform small molecule virtual screening by docking with PyRx, which is opensource software with intuitive user

interface that runs on all major operating systems. The steps for using PyRx, as well as considerations for data preparation, docking, and data analysis.

Results and Discussion

Top section of the Search Results page after a term search for polycystic kidney disease. Only a single hit is shown here for brevity.

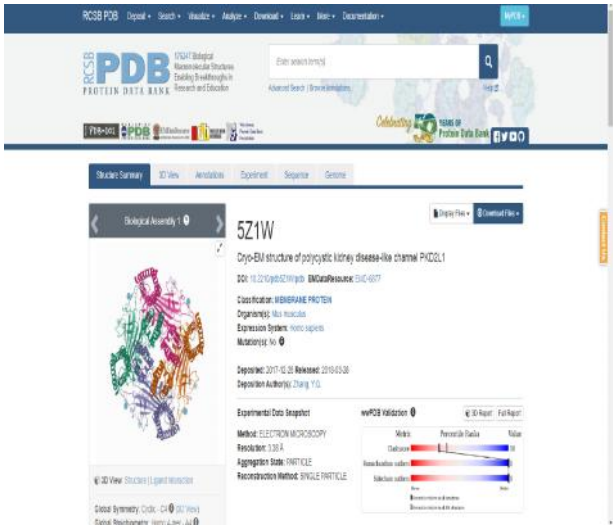
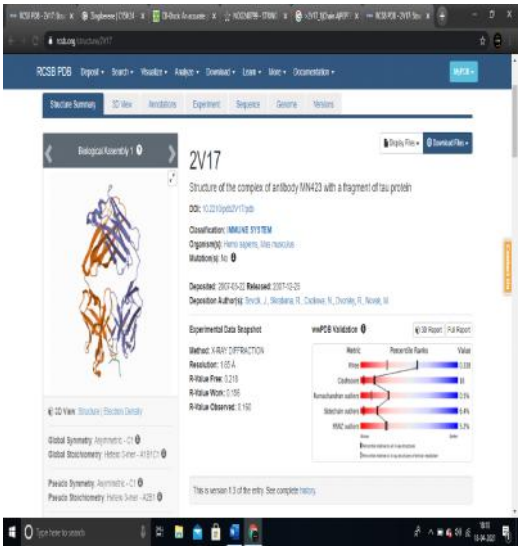


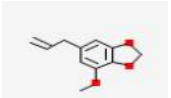
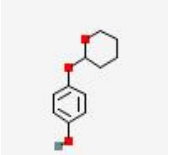
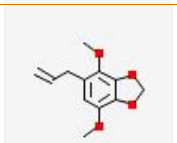
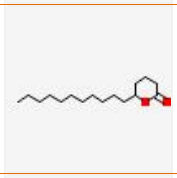
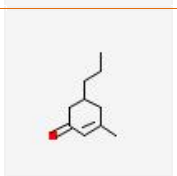
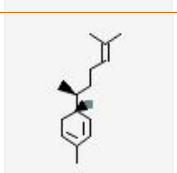
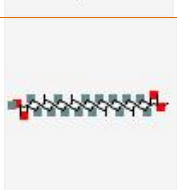
Fig ~1.1 Protein data bank (5Z1W)



Fig~ 1.2Protein data bank (2V17)

List of compounds:

Ligand is an ion or molecule (functional group) that binds to a central metal atom to form a coordination complex. The bonding with the metal generally involves formal donation of one or more of the ligand's electron pairs

S. NO	LIGAND NAME	PUBCHEM ID	MOLECULE FORMULA	STRUCTURE OF LIGAND
1	Benzyl(diphenyl)arsane	238198	$C_{19}H_{17}As$	
2.	Myristicin	4276	$C_{11}H_{12}O_3$	
3.	Deoxyarbutin	11745519	$C_{11}H_{14}O_3$	
4.	Apiole	10659	$C_{12}H_{14}O_4$	
5.	Tetrahydro-6-undecyl-2H-pyran-2-one	110976	$C_{16}H_{30}O_2$	
6.	Silicon dioxide	24261	O_2Si	
7.	3-Methyl-5-propyl-2-cyclohexen-1-one	3782139	$C_{10}H_{16}O$	
8.	Psyllium	249815196		
9.	Zingiberene	92776	$C_{15}H_{24}$	
10.	beta-Bixin	6376436	$C_{25}H_{30}O_4$	

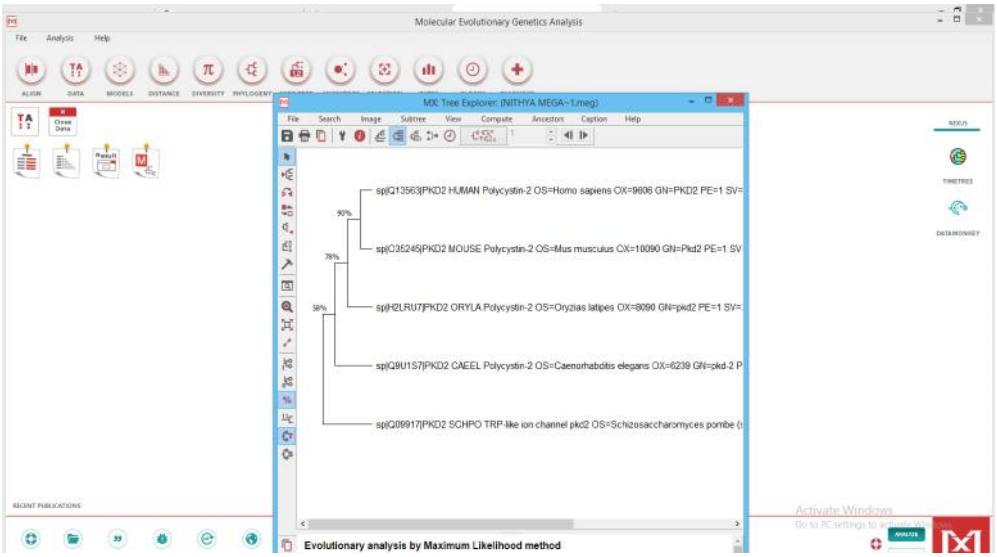
Insilco drug likeness :

Drug-likeness rules are set of guidelines for the structural properties of compounds, used for fast calculation of drug-like properties of a molecule. These guidelines are not absolute, nor are they intended to form strict cutoff values for which property values are drug-like and which are not drug-like.

S.NO	COMPOUND NAME	PREDICITION (DRUG LIKE – NONLIKE)
1	Benzyl(diphenyl)arsane	Drug like
2	Myristicin	Drug like
3	Deoxyarbutin	Drug like
4	Apiole	Drug like
5	Tetrahydro-6-undecyl-2h-pyran-2-one	Non drug like
6	Silicon dioxide	Non drug like
7	3-methyl-5-propyl-2-cyclohexen-1-one	Non-drug like
8	Psyllium	Non drug like
9	Zingiberene	Drug like
10	Beta-bixin	Non drug like

Molecular Evolutionary Relationship of protein:

Timetree inferred in MEGAand shown in the Tree Explorer, where it is displayed with divergence times and their respective 90% confidence intervals. A scale bar for absolute divergence times is shown.



Protein- protein interaction:

The STRING network view. Combined screenshots from the STRING website, which has been queried with a subset of proteins belonging to two different protein complexes in yeast (the COP9 signalosome, as well as the proteasome). Colored lines between the proteins indicate the various types of interaction evidence. Protein nodes which are enlarged indicate the availability of 3D protein structure information. Inset top right: for each protein, accessory information is available which includes annotations, cross-links and domain structures. Inset bottom right: the same network is shown after the addition of a user-configurable In this case, the payload corresponds to color-coded protein abundance information, and reveals systematic differences in the expression strength of both complexes .

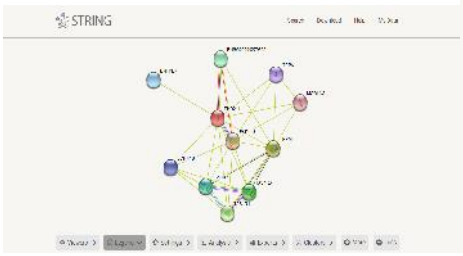


Fig.1 protein- 5Z1W

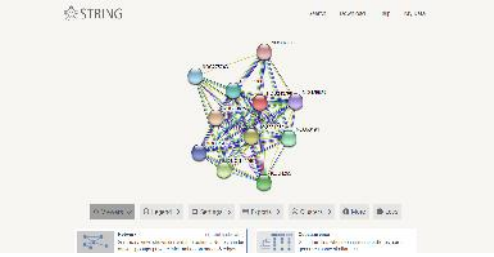


Fig.2 protein- 2V17

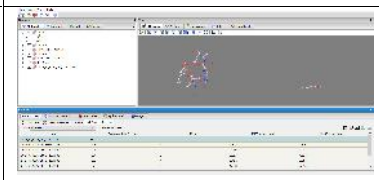
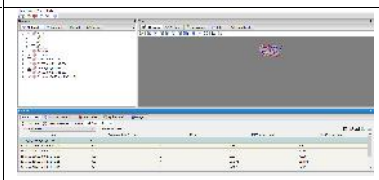
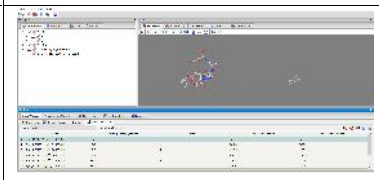
Protein-ligand docking with binding sites(Cd-dock).

predicts the binding activities of proteins to compounds and calculates the center and size of the cavity. It is also integrated with AutoDock Vina and has been carefully optimized.

S.no	Ligand	Protein	Binding energy	Figure of binding molecule
1	4276		-5.9	
2	10659	2V17	-6.2	
3	11745519		-6	

Virtual screening

The calculations are done, results will be populated as seen in the below table with the Binding Affinity (kcal/mol) values. More negative the binding affinity better the orientation of the ligand in the binding site.

S.no	Ligand	Protein	Binding infinity	Figure of VS
1	4276		-2.6	
2	10659	2V17	-5.5	
3	11745519		-3.0	

Conclusion

The molecular study myristicin docking of polycystic kidney disease (PKD) protein revealed that Myristicin are having good interaction in favourabformle pose with PKD which was explained by lowest binding energy, strong bond length and -2.6 no of interactions with active site of PKD molecule Thus it can be concluded that some could be used as a template for the future development through modification or derivatization to design more potent therapeutic agents. Compounds synthesized if properly changed into therapeutic agent can be used for POLYCYSTIC KIDNEY DISEASE the possibility of success of this approach in identification of true positive inhibitors was 100%, because it was able to screen all the three other known inhibitors as true positive.

References

1. Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. Biophys Rev. 2017;9:91–102. 2. Yuriev E, Holien J, Ramsland PA. Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. J Mol Recognit. 2015;28:581–604. 3. Meiler J, Baker D. ROSETTALIGAND: protein-small molecule docking with full sidechain flexibility. Proteins. 2006;65:538–48
2. Ng, R. (2008) Drugs: From Discovery to Approval, 2nd edn. Wiley-Blackwell, Hoboken, New Jersey.
3. Swinney, D.C., and Anthony, J. (2011) How were new medicines discovered? Nat Rev Drug Discov. 10, 507-19.

4. Jacob, R.B., Andersen, T., and McDougal, O.M. (2012) Accessible high-throughput virtual screening molecular docking software for students and educators. PLoS Comput Biol. doi:10.1371/journal.pcbi.1002499
5. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bon- field, J., Burton, J., Connell, M., Copsey, T., Cooper, J., et al. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome Ill of C. elegans. Nature 368, 32-38.
6. Vinoth, M., and Natarajan, B., "Isolation, characterization, antibiological and anticancer activity of Tetrahydrobenzochromone glycosides from leaf of Melochia corchorifolia." Rasayan Journal of Chemistry, 14(2), pp. 768-777, 2021.
7. Vinoth, M., and Natarajan, B., "Isolation, structural elucidation of flavonolignan derivative from Melochia corchorifolia - Evaluation of their biological, anticancer activity, and molecular docking studies." Rasayan Journal of Chemistry, 14(3), pp. 166-1679, 2021.
8. Vinoth, M., and Natarajan, B., and Shanmuga Sundaram, C., "Characterization and evaluation ethyl acetate extract of Melochia corchorifolia leaf- anticancer antibiological and molecular docking studies on breast cancer estrogen receptor." Indian Journal of Chemical Technology, 28(2), pp. 188-196, 2021.
9. Vinoth, M., and Natarajan, B., "Evaluation of phytochemical, anthelmintic activity and Antidiabetic activity-Melochia corchorifolia ethanol extract." International Journal of Research in Pharmaceutical Sciences, 12(1), pp. 1-6, 2021.

Access this Article in Online	
Quick Response Code	Website: www.ijarbs.com Subject: Bioinformatics
DOI:10.22192/ijarbs.2022.09.05.01.011	

How to cite this article:
R. Priya, M.Vinoth, Mahendran Radha. (2022). *In silico* analysis of polystic kidney disease using molecular docking and virtual screening. Int. J. Adv. Res. Biol. Sci. 9(5): Special Issue 1: 98-107.
DOI: http://dx.doi.org/10.22192/ijarbs.2022.09.05.01.011