European Journal of Chemistry

Check for updates

View Journal Online
View Article Online

In silico screening for the interaction of small molecules with their targets and evaluation of therapeutic efficacy by free online tools

Praveen Kumar [1], Nayak Devappa Satyanarayan [2], Subba Rao Venkata Madhunapantula [3], Hulikal Shivashankara Santhosh Kumar [4] and Rajeshwara Achur [1,*]

[1] *Department of Biochemistry, Kuvempu University, Jnana Sahyadri 577451, Shimoga, Karnataka, India*
praveenbiochemku@gmail.com (P.K.), rajachur@gmail.com (R.A.)
[2] *Department of Pharmaceutical Chemistry, Kuvempu University, Post Graduate Centre, Kadur 577458, Chikkamagaluru Dist., Karnataka, India*
satya1782005@gmail.com (N.D.S.)
[3] *Center of Excellence in Molecular Biology and Regenerative Medicine Laboratory, Department of Biochemistry, Jagadguru Sri Shivarathreeshwara Medical College, Academy of Higher Education and Research, Mysuru, Karnataka, 570015, India*
madhunapantulas@yahoo.com (S.R.V.M.)
[4] *Department of Biotechnology, Kuvempu University, Jnana Sahyadri 577451, Shimoga, Karnataka, India*
sk.genesan@gmail.com (H.S.S.K.)

* *Corresponding author at: Department of Biochemistry, Kuvempu University, Jnana Sahyadri 577451, Shimoga, Karnataka, India.*
e-mail: anr@kuvempu.ac.in (R. Achur).

REVIEW ARTICLE

doi 10.5155/eurjchem.11.2.168-178.1962

Received: 24 January 2020
Received in revised form: 17 April 2020
Accepted: 21 April 2020
Published online: 30 June 2020
Printed: 30 June 2020

KEYWORDS

CASTp
ADMET
STRING
ChemAxon
Doxorubicin
MTiAutoDock

ABSTRACT

Pharmaceutical chemistry deals with the process of isolating organic compounds from natural sources or chemically synthesizing them in order to explore potential drugs. Drugs are small molecules, used to prevent or treat various diseases. Of several lead molecules, only few of them reach clinical trial phases and emerge as effective drugs, whereas the majority will be eliminated at different stages. On the other hand, due to the lack of proper identification of their pharmacokinetic properties and biological potential, many small molecules fail to reach this stage. This could be because of the fact that it is either time consuming and costly or there is full of uncertainty due to lack of analyses that are necessary for the confirmation. In the post-genomic era, computational methods have been implemented in almost all stages of drug research and development owing to the drastic increase in the available knowledge about small molecules and the target biomacromolecule. This includes identifying the suitable and specific targets for drug candidates, lead discovery, lead optimization and ultimately preclinical phases. In this context, numerous websites have become highly valuable and influence the drug development and discovery process. Here, we have attempted to bring together some of the online computational approaches and tools that are available to facilitate research efforts in the field of drug discovery and drug design. The output information from these tools is extremely helpful in selecting and deciding about the future direction or specific path needed to be followed by the researchers. These computational methods are indeed help to focus the intended research in the right direction. As detailed in this review, the information provided about the servers and methods should be useful throughout the process of screening of synthesized or chemical database originated small molecules to find the appropriate targets along with the active sites without depending on any commercial tools or time-consuming and costly assays. It should however be remembered that the bioinformatics-based prediction cannot completely replace the wet lab data of chemical compounds or specific assays.

Cite this: *Eur. J. Chem.* **2020**, *11(2)*, 168-178 **Journal website**: www.eurjchem.com

1. Introduction

Molecular modeling is a computer-based tool for the creation, representation and transformation of small molecule drug candidates into active or efficient drugs. The search for a potential drug begins from a large repository of synthetic molecules which undergoes the elimination process at different steps of drug discovery process. Thus, the number of bioactive drug molecules will be very small as compared to a vast collection of candidate compounds [1].

In computational drug discovery, there are several parameters that are needed to be evaluated to check drug likeliness of molecules, as shown in Figure 1. These include structural, molecular and physico-chemical properties of a small molecule, such as their molecular weight, hydrogen bonding ability, polarity, lipophilicity and structural adaptation to increase their drug likeliness. In addition, prediction of toxicity and their binding capacity with the appropriate target are also important parameters. All these multiple criteria should be fulfilled by a molecule for it to qualify as a hit or lead candidate in the area of computational drug design pipeline.

It is very important that drug-like properties of compounds as identified in the virtual screening must agree with the laboratory assays involving the specific target in order to become a potent drug.

http://dx.doi.org/10.5155/eurjchem.11.2.168-178.1962

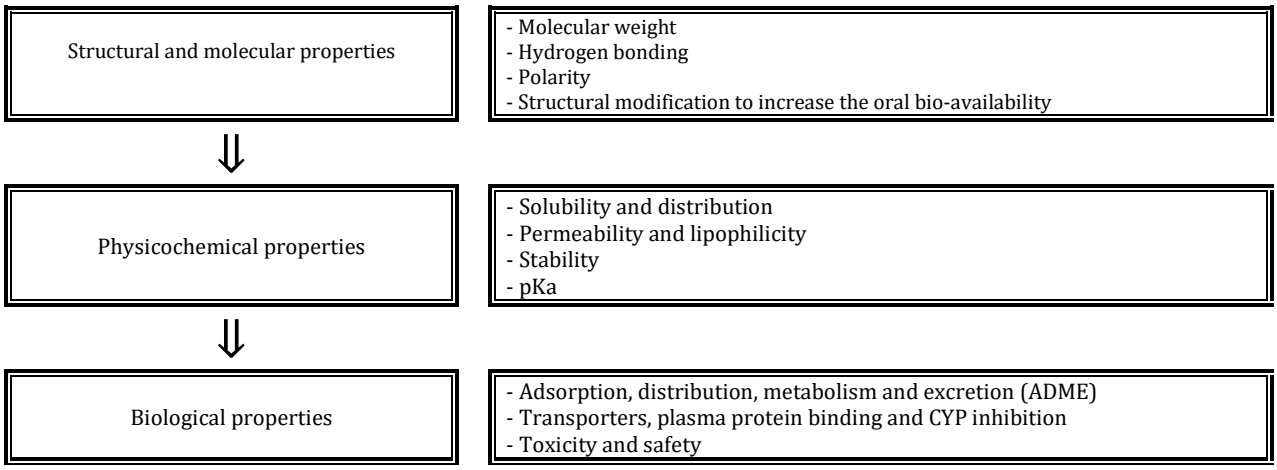

Figure 1. Schematic representation of pharmaceutical profiling and drug discovery stages and their associated parameters.

A drug target is usually a main molecule that involves particular metabolic or signaling pathways that are unique to a state of the disease.

In this manuscript, we have presented a compilation of some of the freely accessible online resources that are frequently used in drug discovery. For the purpose of demonstration, a common drug molecule such as doxorubicin has been selected from ZINC database, which is a collection of commercially produced substances designed for virtual screening. The specific characteristic properties of this drug molecule were evaluated using the available online tools. The target was identified, binding pockets analyzed, virtual testing and compound profiling were performed to highlight the importance of *in silico* drug design tools. Although various docking algorithms-based tools are available online, an understanding of each method's advantages and limitations is of fundamental importance in designing successful strategies and obtaining the expected results. Thus, the aim of this study was to compare some of the currently available online accessible tools for the pharmacokinetic evaluation and molecular docking studies utilized in drug discovery and medicinal chemistry.

2. Online resources for drug discovery and development

Over the last few years, plenty of free online tools are available to encourage and promote drug development research. These sites assist withdrawing and converting of the chemical structures into appropriate formats, including MDL MOL, SMILES, SDF and several other formats which are compatible inputs for the selected web tools. Marvin Suite is a smart online toolkit designed to help editing the skeleton of a molecule (https://academia.chemaxon.com) [2]. This program has a function that automatically detects the structure of a chemical molecule from any input format. ChemDraw JS, a web version of ChemDraw helps to build JavaScript and HTML5 web applications (https://chemdrawdirect.perkinelmer.cloud/rest/) [3]. All kinds of scientific template inputs can be made into a better presentation in this web tool with modeling of small molecules. MolView v2.4 is a program that allows the user to analyze and display molecular structure in any format [4]. This program can read small molecules in any format including PDB structure files such as American Standard Code for Information Interchange (ASCII). The structure can further be screened and analyzed by showing hydrogen bonds, constructing Ramachandran plots for favorable residues, marking the atoms, calculating the distance between atoms and finding adjacent atoms. This program also helps to carry out structural analysis by updating their database interlinked with PubChem molecules, RCSB protein database and crystallographic database. By the help of these tools, once the lead molecule is in hand, it is very important to find out ways to transform these molecules to improve the desired pharmacological properties. The reason why many drug molecules fail to develop into lead molecules is because of exhibition of weak drug-like properties which could be improved by increasing its binding affinity to a receptor or target. For this, it is critical to know the structural or molecular, physicochemical and biological properties of any potential small molecule in the process of drug discovery (Figure 1).

3. Structural and molecular properties

3.1. Molecular weight

The molecular weight of an organic compound plays an important role as it should follow the Lipinski rule of 5 according to which smaller molecules are better as the dispersal is directly affected [5]. Most of the existing organic molecules that are available in the library and approved by the FDA are having molecular weight in the range of 200 and 600 Daltons. Specifically, the majority of drug molecules belong to <500 Daltons group [6].

3.2. Hydrogen bond

The hydrogen bond is a unique descriptor because hydrogen is the only atom that can carry a positive charge at physiological pH while remaining covalently bonded in a molecule. Both intra-molecular and inter-molecular hydrogen bonds play a crucial role in the drug-receptor interaction. Intra-molecular bonding is a very essential property of a molecule that may have a significant effect on the lead modification approach. In computational approaches (Molecular docking), hydrogen bonds are very much significant to maintain the structural components such as α-helix and the β-sheet conformations of the protein or receptor. The ΔG° or free energy for hydrogen bonding occurs between -1 to -7 kcal/mol. The binding affinity increases by order of magnitude per hydrogen bond [7]. Proteins or receptors that are composed of various groups such as NH and OH can donate hydrogen bond and CO and other groups can act as acceptors [8]. Thus, the hydrogen bonds have significant biochemical importance in drug likeliness of a molecule. An example for the hydrogen bond interaction between the ligand and the surface of a receptor is shown in Figure 2.

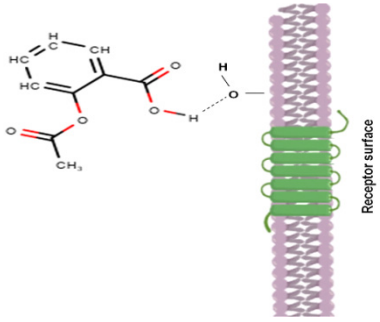

Figure 2. Example of hydrogen bonding interaction of aspirin with the receptor surface. The wavy line represents the receptor surface and the dotted line indicates the hydrogen bond.

3.3. Polarity

Polarity is the most important parameter to determine the ability of a small molecule to be an effective drug. The extent of Total Polar Surface Area (TPSA) relies on the sum of hydrogen donor and acceptor count in the structure of a molecule which is an important predictor of good oral bioavailability. The TPSA value should be in the range of 20 to 130 Å for a molecule to be a drug. The TPSA is a descriptor that is shown to correlate well with passive molecular transport through the membrane of receptors. Hence, the sum of polar atoms in a molecule allows for the prediction of transport properties of drugs [9].

3.4. Structural modification of molecules for better oral bioavailability

About half of the drug candidates do not reach the clinical trial phases due to lack of exhibition of acceptable pharmacokinetic properties in animal model studies [10]. Thus, a major challenge in organic synthesis is the structural modification of lead compounds. This should not only modify the structure of starting material but also increases its biological activity [11]. During modeling and visualization of small molecules using online tools such as Molinspiration (https://www.molinspiration.com/) [12], one can find the pharmacokinetic violations which are needed to be corrected. In order to make a potent molecule, several lead modification approaches are required such as homologation, chain branching, fingerprint analysis of a molecule, ring-chain transformations and bio-isosterism to increase the potency of a molecule.

4. Physicochemical properties

4.1. Solubility and distribution

The insolubility, poor solubility and poor permeability issues are the main reasons to classify them as inactive drugs during the development and design of a particular drug candidate [13]. It is therefore important to determine these pharmacokinetic and physicochemical properties associated with a drug molecule and receptor. Poor or weak solubility can often be a more limiting factor in drug development. The lipophilicity (log P) and Lipinski's analysis of compounds are very critical for those molecules to be destined for drug development or used as drugs [14]. Having a soluble molecule greatly facilitates many drug development activities [15]. Once the drug molecule gets absorbed in the Gastro intestinal-tract (GI-tract) or through some other administration route, it will pass through the capillaries and dispersed into the tissues of the body [16]. The amount of drug in the body is described by the term "volume of the drug distributed" which denotes the sum of the drug in the body (mg) divided by the concentration of the drug in plasma (mg/L). The drug distribution largely relies on the size of the molecule, log P, H bonding, polarity etc. In this context, the drug molecules designed to be effective on tissues and organs should not be able to cross the blood-brain barrier (BBB) to avoid adverse effects on the CNS [17].

4.2. Permeability and lipophilicity

The permeability is the ability of a drug molecule to diffuse passively across the cell membrane, enter the phospholipid bilayer and from the cells apical side to the basal lateral side through a cytoplasmic aqueous phase or along the lipid membranes of the cell [18]. The permeability of a drug molecule across biological membranes depends upon the diffusion coefficient and lipophilicity property. If a drug molecule has poor water solubility (high lipophilicity) that can lead to a limiting factor in oral bioavailability, and highly lipophilic compounds are metabolized easily or it can bind to plasma proteins [7]. However, low lipophilicity is more typically problematic because that leads to poor permeability across the membrane and thus increased lipophilicity leads to improved physicochemical properties. The log P values, as shown in Figure 3, denote the *in vivo* lipophilicity which is derived from the 1-octanol/water partition model [19]. If the log P value is greater than 5.5, then the molecule will be highly lipophilic and therefore it is very important to be aware of choosing the solvent to obtain the log P data.

4.3. Stability

The storage as well as *in vivo* stability of a drug molecule is an important criterion to be fulfilled by the successful lead molecule. If a molecule is photosensitive or air-sensitive, any exposure can lead to decomposition or alteration of the functional groups of the molecule. Any potential drug molecule having weak alkaline pH value of stability that ranges between 7.20 and 7.80 can countervail the acid toxins *in vivo.* This type of property of a molecule helps to remove the toxicants from the body to keep the body fluid augment digestion and nutrient absorption [20]. The highly stable small molecule is useful for the medicinal chemists as their toolbox for drug discovery. Due to lack of stability of a drug molecule, its biological potential can be disrupted. A specific reaction of a drug molecule could lead to putrefaction which includes oxidation and hydrolysis. All the potential stability issues of a drug candidate can be screened *in silico* and could be solved [21].

4.4. pKa

The acid-base dissociation constant (pKa) of a drug candidate is an important parameter that influences many biopharmaceutical properties. The pH of medium interacts differently with the ionizable groups having variable pKa values.

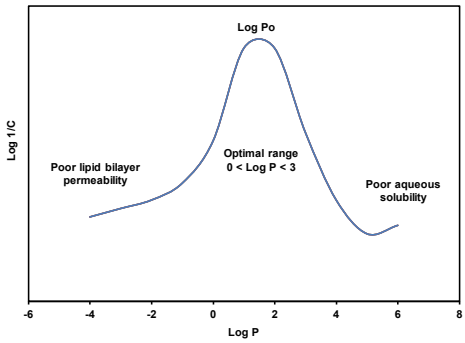

Figure 3. Example for the maximum drug-likeness model score of a standard drug with partition co-efficient and concentration to produce standard biological effect. Log 1/C is the concentration of compound required to produce a standard response in a given time, log P is the logarithm of molecules partition coefficient between 1-Octanol and water, log P_o is the logarithm of optimum partition coefficient for biological activity.

The pKa distribution of drug molecules can impact in two different channels. One is structure related activity of functional groups of the drug molecule and the ideal range of pKa values that they span. The other factor concerns the biological targets of these compounds that are designed to hit [22]. The pKa value of a drug molecule can impact lipophilicity, solubility, receptor binding ability and membrane permeability, which in turn directly affect the pharmacokinetic properties.

5. Biological properties

For depicting the screening of a small molecule and their target evaluation of biological properties, we selected doxorubicin as a drug molecule, the structure of which was downloaded from the ZINC database (ZINC3918087) (https://zinc.docking.org/substances/home/) [23]. The therapeutic screening of this drug molecule was performed and the structure was visualized by using online tool Chemaxon (https://marvinjs-demo.chemaxon.com/latest/demo.html). The generated images of the molecule are as shown in Figure 4.

5.1. Absorption, distribution, metabolism, excretion and toxicity (ADMET)

The absorption, distribution, metabolism, excretion and toxicity (ADMET) data can be obtained in three ways. First, by a series of cell-based assays (*in vitro*), the second is from the computational approaches or by *in silico* prediction. In the third category, the predictive models have been improved that can eventually influx in the process of drug discovery and drug design to replace *in vitro* assays and/or *in vivo* experiments. The *in vitro* and *in vivo* experiments are generally time-consuming, expensive and the most important thing is that the drug candidate has to be procured to carry out these experiments. By using *in silico* approach, one can first derive a potent promising molecule for further *in vitro* and *in vivo* based assays. The idea behind the rule of 5 (ADMET) is to find the biological potential of a drug candidate. According to ADMET descriptors, the drug molecules should pass through all the pipelines of every descriptor without violating rule of 5. This is very important for their successful biological activity when considering the administration of drugs orally or from any route with real therapeutic potential. The results from the rule of 5 are only indicative, which means that a drug molecule violating the rule of 5 does not mean they have completely poor bioavailability. It should be viewed more like a quantitative predictor as it alerts that poor absorption or permeability is possible, and thus, helps to pay special attention towards preparing potential drug molecule having less toxicity and more bioavailability to overcome the pitfalls in drug properties [24].

The absorption occurs across different channels such as passive transcellular and paracellular absorption, carrier-mediated absorption, and receptor-mediated endocytosis in the biological membrane [25]. The polar compounds absorption takes place by the paracellular absorption which occurs by diffusion through the tight junctions between the cells [26]. The most important physicochemical properties for passive absorption of a drug candidate from the Gastro intestinal (GI) tract are its aqueous solubility and lipophilicity. The P-glycoprotein (P-gp) acts as an efflux pump in order to push the molecules out of the membrane. Hence, measuring the P-gp substrate through *in silico* will help the prediction of drug likeliness of a molecule. The distribution of entirely administered test dose of a drug in the body (mg) will be divided equally in the plasma [27]. When a drug molecule is administrated by any route, it may permeate through the BBB or it may undergo human intestinal absorption (HIA). Further, the drug's biodegradability, Caco-2 permeability, AMES toxicity and carcinogenicity need to be evaluated. All these descriptors can easily be predicted by the online web server swissADME (http://www.swissadme.ch/) [28]. This is a free web tool which enables the computation of key properties such as physicochemical, pharmacokinetic, drug-like and related parameters (Permeability, lipophilicity, solubility, pKa and absorption). Based on the results obtained for each molecule, one can predict the bioavailability, plasma-protein binding, metabolism and drug-drug interaction of a likely drug candidate of any type of small molecules that are synthesized chemically or extracted from any natural product source. In the present example, within this filtration criterion, doxorubicin has been selected as a drug candidate for the next step. The ADMET results of doxorubicin molecule are shown in Table 1.

5.2. Prediction of oral toxicity of small molecules

Determining the toxicity of drug molecules is necessary to identify their harmful effects on humans, animals, plants, or the environment. It is also one of the main decisive factors in drug design. ProTox-II is a free web server to predict the toxicity of any small molecule or chemicals by entering the Pubchem name or in Canonical SMILES format of the molecule. The organ toxicity and toxicity endpoints such as carcinogenicity, immunotoxicity, mutagenicity and cytotoxicity of the molecule can be estimated by using the online tool (http://tox.charite.de/protox_II/index.php?site=compound_input) [34]. Computational prediction of all these important adverse characteristics of a molecule is very critical in suggesting new research directions and provides recommendations for designing novel *in silico* models.

Table 1. *In silico* ADMET prediction of doxorubicin by using six different online tools *.

ADMET profile of the doxorubicin using different type of web tools						
Web servers / URL	**HIA (Probability)**	**BBB**	**CYP inhibition/ substrate**	**AMES toxicity**	**Carcinogenicity**	**LD_{50} in rat (mol/kg)**
admetSAR [29] (http://lmmd.ecust.edu.cn)	-	-	Substrate	AMES toxic	Non-carcinogen	2.6644
SWISS ADME (http://www.swissadme.ch/)	Low	No	Substrate	-	-	-
scbdd.com [30] (http://admet.scbdd.com/)	0.019	0.015	Substrate	0.866	-	3.224
pkCSM [31] (http://biosig.unimelb.edu.au/pkcsm/prediction)	49.703	-1.688	No	-	-	2.698
vNN-ADMET [32] (https://vnnadmet.bhsai.org/vnnadmet/login.xhtml)	-	-	Substrate	Yes	-	-
preADMET [33] (https://preadmet.bmdrc.kr/)	31.952	0.032	Weak substrate	non-mutagen	Negative	128

* HIA: Human intestinal absorption; BBB: The blood-brain barrier; CYP: Cytochrome P450; AMES: Mutagenic effect; LD_{50}: Median lethal dose.

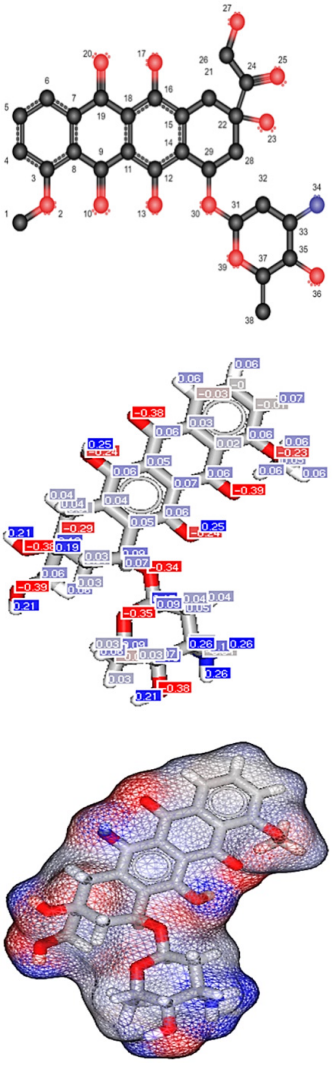

Figure 4. Structure of doxorubicin as a model ligand (a). The images (b) and (c) are the three-dimensional images of doxorubicin shown with its charge and polarity.

5.3. Calculation of molecular properties of doxorubicin

In the present model, doxorubicin as the ligand molecule has been subjected to molecular properties prediction using seven different freely accessible online tools to evaluate its drug-likeness. The online tools used are Molinspiration, Molview, mcule.com, SWISSADME, proTOX, ZINC and DRUGBANK [35]. All these seven tools gave similar results and among these, six of them gave highly comparable output data as shown in Table 2.

5.4. CYP inhibition

The kinetic profile of a molecule indicates the extent of binding interaction between the compound and its target as well as rate of binding and dissociation from the target. The tools also estimate or predict the metabolism property of cytochromes P450 (CYPs) and amongst these, the three isoforms (2C9, 2D6, and 3A4), that are mainly expressed in the gut and liver [36].

Table 2. Molecular property predictions of doxorubicin using online *in silico* tools.

Types of web servers/tools	x log P	H bond donors	H bond acceptor	tPSA	MW	Rotatable bonds
ZINC	0.57	8	12	208.00	544.533	5
Molinspiration	0.57	-	-	206.08	543.520	5
SWISSADME	1.27	6	12	206.70	543.520	5
Molview	-	6	12	-	543.525	-
DRUGBANK	0.92	6	12	206.00	-	5
mcule.com (https://mcule.com/apps/property-calculator/)	0.70	6	12	206.07	543.510	5
proTOX (http://tox.charite.de/protox_II/index.php?site=home)	-	0	12	206.07	543.520	5

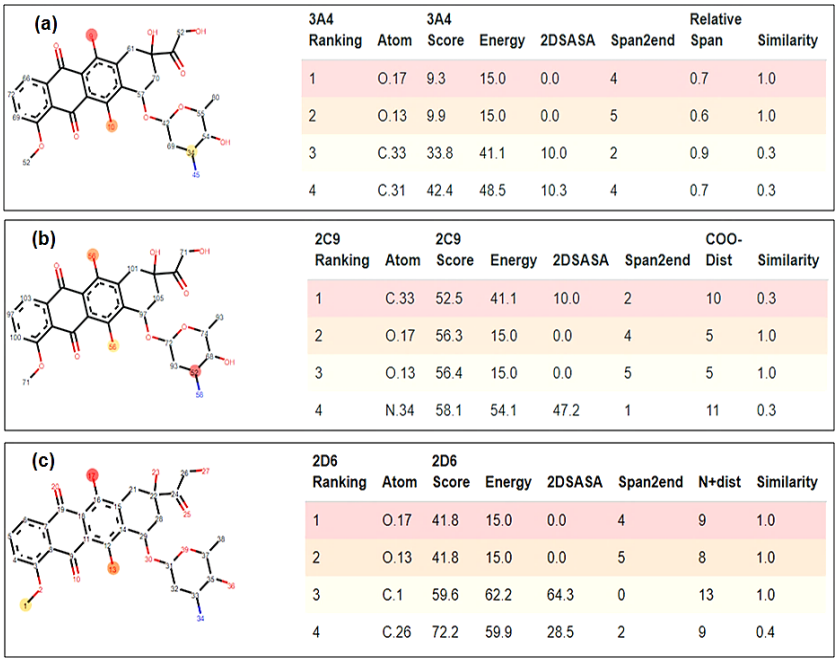

Figure 5. Prediction of top three best ranking atoms of doxorubicin involved in the CYP inhibition by using SMARTCyp online tool. (a) Predicted SMARTCyp score of the enzyme 3A4, (b) Predicted SMARTCyp score of the enzyme 2C9 and (c) Predicted SMARTCyp score of the enzyme 2D6.

In this context, a publicly available web server SMARTCyp (https://smartcyp.sund.ku.dk/mol_to_som) can help the researchers to predict the sites of cytochrome P450-mediated metabolism of drug-like molecules and it displays the predicted sites that are metabolized by the cytochrome P450 3A4 isoform [37]. The advantage of SMARTCyp is that the result generated exactly predicts the site of metabolism directly from the structure of a drug molecule, without the need for the generation of 3D structures [38]. The result of predicted CYPs metabolism property of doxorubicin molecule is shown as indicated by various parameters. The findings are displayed in the form of a structure and in a Table for doxorubicin. Both the structure and the Table shows three highest ranking atoms. The top three ranking atoms of the doxorubicin molecule are highlighted in Figure 5. The atoms in the table are ranked by the score, the lowest score resulting in the least rank, and accordingly the highest probability of being a site of metabolism. The comparison is expressed by a value between 0 (low) to 1 (high) showing the greatest probability of being a site of metabolism. Although the top three sites are colored to encourage recognition, this does not mean that there is a cutoff and that there are only three probable locations as the ranking matters. It shows how close the atom in the molecule is to an atom in a fragment on which Discrete Fourier Transform (DFT) has calculated the activation energies.

6. Receptor selection for molecular docking with doxorubicin

DNA topoisomerases are essential for DNA replication, transcription, recombination, repair, and mitosis by introducing transient single-strand breaks (SSBs) or double-strand breaks (DSBs) in the DNA to adjust its topology in the cell. Doxorubicin initially docked into the same active domains of the target enzyme of interest. Doxorubicin is a broad-spectrum anticancer drug that targets Top II through stabilizing the topoisomerase-DNA complex, thus leading to the accumulation of DNA breaks and eventually cancer cell death [39]. Here, we selected the receptor 1ZXM for docking with doxorubicin ligand. The receptor 1ZXM was downloaded from the RCSB (https://www.rcsb.org/) protein data bank and visualized in the same web server to see its three-dimensional structure (Figure 6) [40].

6.1. Active sites pocket of receptor and its drug ability predictions by CASTp and PDBsum web servers

Computed Atlas of Surface Topography of proteins (CASTp) is a web server interlinked with the PDB website at the San Diego Supercomputing Center (http://sts.bioe.uic.edu/castp/index.html?201l) [41]. This is a user-friendly site wherein a four-letter code (PDB-ID) can be entered to get the display list of relevant PDB structures. This is done before molecular docking to find the active sites that are compatible with the ligand with better binding energy and interaction. In this, CASTp server identifies all the active pockets and measures the volume and area of each pocket as well as the size of mouth openings of individual pockets, which helps to assess the accessibility of binding sites to various ligands and substrates [42].

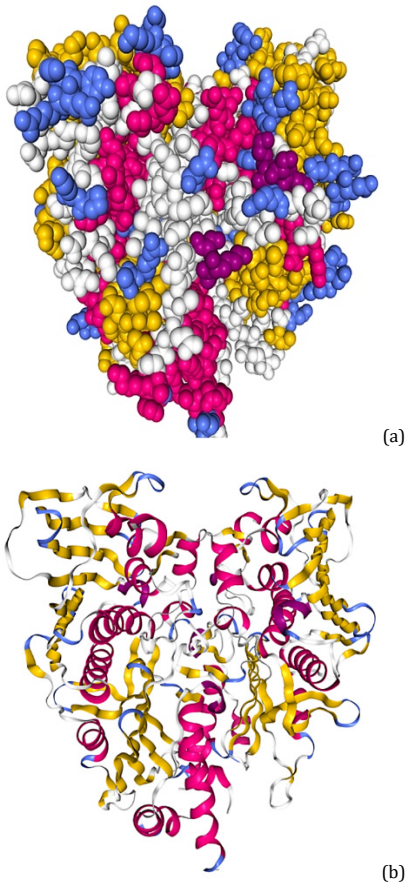

Figure 6. Structural presentation of human Topoisomerase II (PDB: 1ZXM). (a) Surface structure of Topo II and (b) Ribbon structure of Topo II.

To get the topographic computation, the input can be in the form of a PDB structure / four letters PDB ID. Further, one can submit the structures directly to get a customized evaluation. For pre-computed results, a default probe radius of 1.4 Å is used, which is the standard value for computing the solvent accessible surface area. For the customized computation requests, users can specify any probe radius as desired [43]. For the receptor 1ZXM, we analyzed the active site by using the CASTp server and the obtained results are as shown in Figure 7. Alternatively, there is another freely accessible online database called PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/cgibin/pdbsum/GetPage.pl?pdbcode=index.html) to analyze the complete details of a protein. This database provides a summary of the molecules in each PDB file (i.e. proteins, nucleic acids, ligands, water molecules and metals) together with various analyses of their structural features [44,45]. A calculation based on Ramachandran plot quality and the side chain properties for the main-chain of the 1ZXM was performed by using PROCHECK (sub-application of PDBdsum) which were found to be normal and the tested parameters were included in the standard deviation of the chi angles [46]. The enlarged secondary structure and Ramachandran plot for 1ZXM predicted by PDBsum web server generated are shown in Figure 8.

6.2. Determination of interaction partners of 1ZXM by STRING database

The aim of Search Tool Retrieval of Interacting Genes/Proteins (STRING) database is to collect, score, and incorporate all publicly available sources of knowledge on protein-protein interaction and complement these with computational predictions which helps to create a global network that is comprehensive and objective, including both direct (physical) and indirect (functional) interaction. The STRING is a publicly available free database which has been used to analyze protein interactions of 1ZXM (http://string-db.org/) [47]. The STRING database has an advantage that it predicts most aggregate information of the protein-protein associated clusters. In the present example, the PPI network of genes was constructed with the STRING database and the interaction with a combined score >0.4 was considered statistically significant. The value obtained here is statistically very significant as TPX2 has shown the highest score of 0.997 (Figure 9).

7. Molecular docking and visualization of doxorubicin interaction profile with 1ZXM receptor

The docking analysis carried out using the DNA TOP2A (PDB:1ZXM) has shown very high binding affinity with the standard anticancer drug doxorubicin as revealed by three different web servers such as SwissDock [48], mcule [49] and MTiAutoDock [50]. Visualization of binding interactions of doxorubicin upon docking onto receptor 1ZXM in top scored poses obtained by the mcule and MTiAutodock are as shown in Figures 10 and 11. Once the ligand structure and appropriate receptor are uploaded, the web servers splits the ligand file in different conformations and sets up the grid box in the active sites of residues for better interactions and the data is generated for the visualization. In the present example, the binding energy values of doxorubicin ranges between -6.8 to -9.11 as indicated by SwissDock.

Table 3. Docking interaction score and clustering results of doxorubicin with 1ZXM by SWISSDOCK.

Receptor	No. of Swiss Dock clusters	Cluster rank	Full fitness (kcal/ mol)	Estimated ΔG (Kcal/mol)
1ZXM	29 (257 runs)	1	-4590.14	-7.92
		2	-4589.66	-8.95
		3	-4586.9	-6.8
		4	-4585.05	-9.11
		5	-4585.48	-7.24

Table 4. Docking interaction score with best three docking poses of doxorubicin and 1ZXM by mcule server.

No	Docking pose	Docking score
1	#1	-8.9
2	#2	-8.4
3	#3	-8.3
4	#4	-8.2

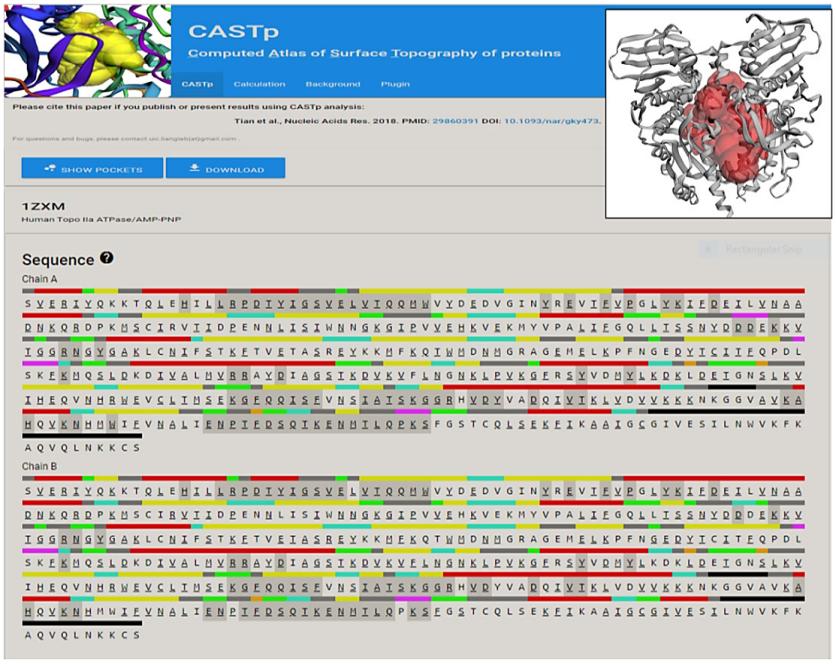

Figure 7. Active sites prediction for the receptor (1ZXM) side chains A and B by using CSATp web server.

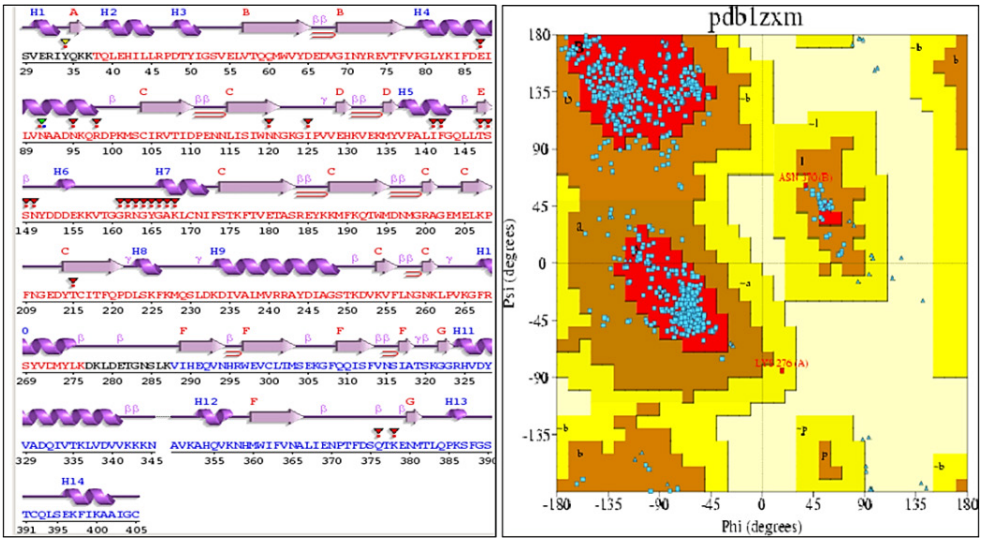

Figure 8. Prediction by Ramachandran plot and enlarged secondary structure of 1ZXM by PDBsum web server.

The mcule server provides four best docking poses based on top four docking scores with best poses of binding affinity and interaction as shown in Figure 10 (Pose 1= -8.9, Pose 2= -8.4, Pose3= -8.3 and Pose4= -8.2). In the case of MTiAutoDock server, which is implemented in AutoDock 4.2.6, it generated 10 different conformations of doxorubicin. The data generated from the above docking servers for the ligand doxorubicin and target receptor 1ZXM is shown in Tables 3-5.

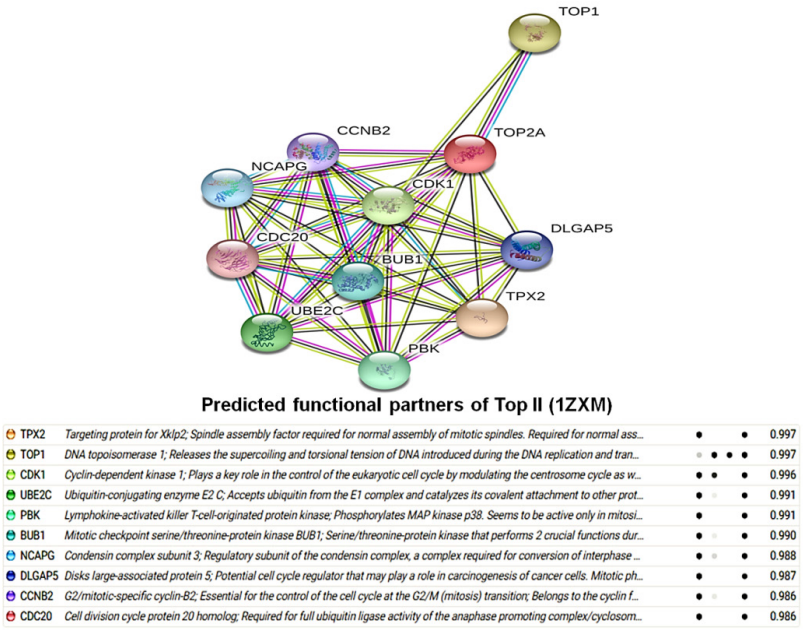

Figure 9. Snapshot of a typical output from a STRING Database. Topo II isomerase (1ZXM) and its interaction partners are shown here.

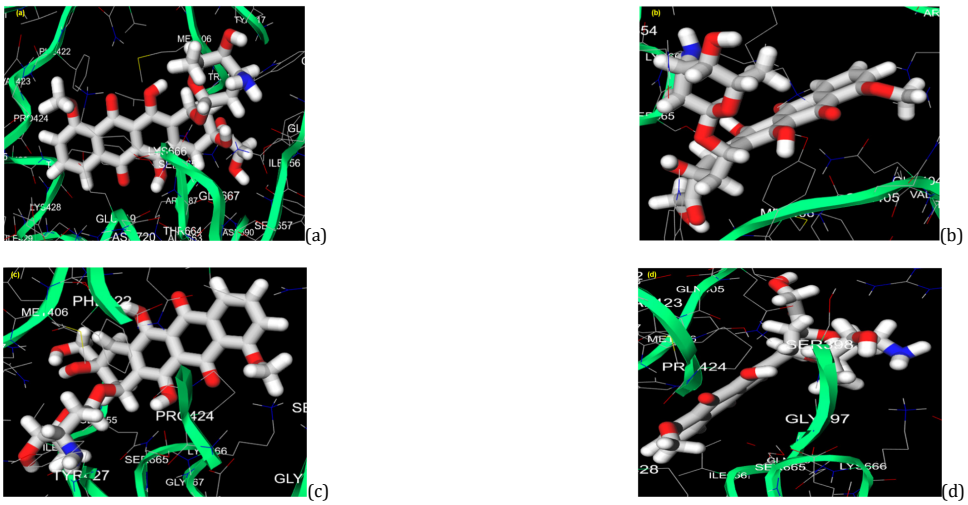

Figure 10. The best four pose molecular visualization for the binding of doxorubicin with the receptor 1ZXM DNA topoisomerase II with top-ranked multi-conformational view. (a), (b), (c) and (d) are the different docking poses generated by the server mcule.

7.1. Results of docking interaction of doxorubicin with1ZXM receptor by SwissDock, mcule and MTiAutoDock

The molecular docking study with DNA TOP2A (PDB: 1ZXM) revealed that doxorubicin has shown high binding affinity as indicated by all the three different online docking servers. The SwissDock docking tool focused on ligand-protein interaction after uploading the protein and ligand structure files. The required structure input format of the receptor could be protein PDB code, name of the protein, sequence, or URL etc. The ligand input could be ZINC substance ID, name of the ligand or category (like scaffolds or side chains), or URL etc. For the modified ligands and receptors, the structure needs to be uploaded separately. As indicated above, the docking studies of doxorubicin was performed with mcule, SwissDock and MTiAutodock web servers to understand the ligand's interaction with Topo II (1ZXM). The results of docking analyses with docking score are shown in Tables 3-5.

8. Conclusions

In the discovery and optimization of hit molecules, computational approach using online web servers and databases have become an essential component of current research. These are highly useful in the preparation of ligands to enable pharmacokinetic predictions and analysis of appropriate receptors with the active pocket residues to find the specific target and validation among many others. Many of these methods and databases are now available freely online for the researchers. These tools provide needed support for the drug discovery tasks including the ability to find a new drug molecule, poly-pharmacology, drug-drug and drug-protein interaction, prediction of active sites and very valuable aggregate information of the protein-protein associated clusters. There are still gaps that are needed to be improved with regard to enhancing guidance to users and screening methods for combining receptor flexibility and best scoring schemes, better selection of binding pocket in a receptor and

Table 5. Ten different conformational interactions of doxorubicin with its target receptor 1ZXM as provided by MTiAutoDock server.

Ligand	All poses concatenated file (pdbqt)	All poses concatenated file (mol2)	Pose	Energy	Rotatable bonds
Doxorubicin	ligands_in1_all_poses.pdbqt	ligands_in1_all_poses.mol2	ligands_in1_1.pdbqt	-5.94	2
			ligands_in1_2.pdbqt	-5.93	2
			ligands_in1_3.pdbqt	-5.93	2
			ligands_in1_4.pdbqt	-5.73	2
			ligands_in1_5.pdbqt	-5.54	2
			ligands_in1_6.pdbqt	-5.51	2
			ligands_in1_7.pdbqt	-5.45	2
			ligands_in1_8.pdbqt	-5.37	2
			ligands_in1_9.pdbqt	-5.28	2
			ligands_in1_10.pdbqt	-5.24	2

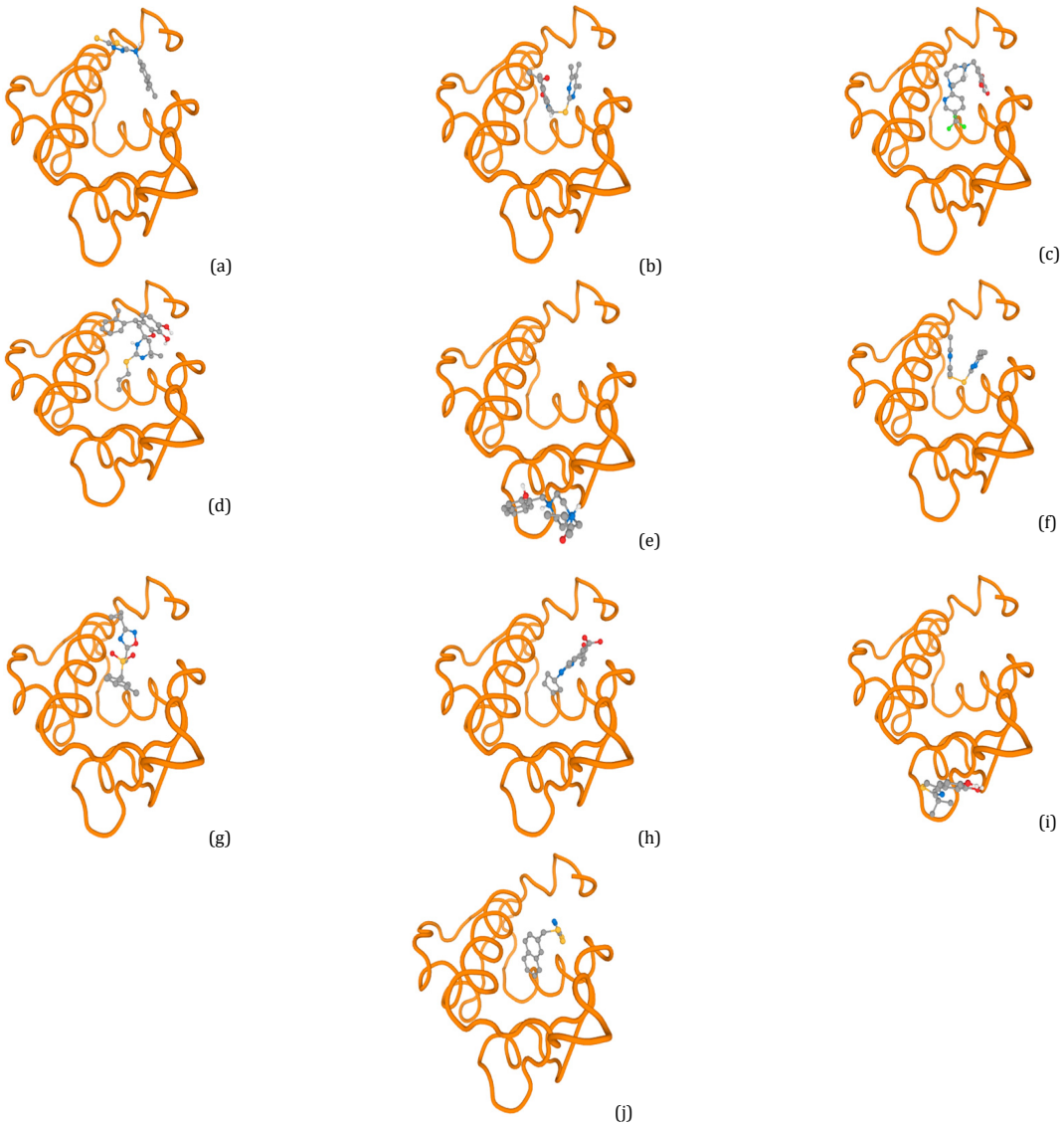

Figure 11. Visualization of binding interactions of doxorubicin upon docking onto receptor 1ZXM in top 10 scored poses. (a-j) are different conformation of binding interaction between the ligand doxorubicin and the receptor (1ZXM).

developing user-friendly workflows. Yet, the freely available and online accessible web servers and databases have already contributed enormously to do faster, better and cheaper identification of new drug candidates either synthesized or extracted from natural sources.

In this present compilation, by taking doxorubicin as a model ligand and its 1ZXM target receptor, we have shown the utility of some of the free online available tools in the area of *in silico* drug design.

Disclosure statement

Conflict of interests: The authors report no conflicts of interest in this work.

Author contributions: Rajeshwara Achur and Praveen Kumar conceptualized the study, performed the review of literature and wrote the manuscript. The manuscript contents were designed and compiled in consultation and inputs from Nayak Devappa Satyanarayan, Subba Rao Venkata Madhunapantula and Hulikal Shivashankara Santhosh Kumar.

Ethical approval: The present study did not require any type of ethical approval.
Sample availability: Not applicable.

ORCID

Praveen Kumar
http://orcid.org/0000-0002-8034-6348
Nayak Devappa Satyanarayan
http://orcid.org/0000-0001-4511-3749
Subba Rao Venkata Madhunapantula
http://orcid.org/0000-0001-9167-9271
Hulikal Shivashankara Santhosh Kumar
http://orcid.org/0000-0003-0371-3535
Rajeshwara Achur
http://orcid.org/0000-0001-9867-8917

References

[1]. Fang, Y.; Ding, Y.; Feinstein, W. P.; Koppelman, D. M.; Moreno, J.; Jarrell, M.; Ramanujam, J.; Brylinski, M. *PLoS ONE* **2016**, *11(7),* e0158898.
[2]. Bode, J. W.; Reactor ChemAxon Ltd. Retrieved December 15, 2019, from https://chemaxon.com/
[3]. Cousins, K. R. Chemdraw Ultra 12. 0. Retrieved December 17, 2019 from https://chemdrawdirect.perkinelmer.cloud/js/sample/index.html
[4]. Smith, T. J. J. *Mol. Graphics* **1995**, *13(2),* 122-125.
[5]. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv Drug Deliv. Rev.* **2001**, *46(1-3),* 3-26.
[6]. Duffy, F. J.; Devocelle M.; Shields, D. C. Computational Approaches to Developing Short Cyclic Peptide Modulators of Protein-Protein Interactions. In: Zhou P., Huang J. (eds) Computational Peptidology. Methods in Molecular Biology, vol 1268. Humana Press, New York, NY, 2015.
[7]. Silverman, R. B.; Holladay, M. W. The organic chemistry of drug design and drug action. Academic press, ISBN 978-0-12-382030-3, 2016.
[8]. Kostal, J. Computational Chemistry in Predictive Toxicology: Status Quo et Quo Vadis? In Advances in Molecular Toxicology, Ed. J. Fishbein, Elsevier, San Diego, 2016, vol. 10, pp. 139-186.
[9]. Ertl, P.; Rohde, B.; Selzer, P. *J. Med. Chem.* **2000**, *43(20),* 3714-3717.
[10]. Edwards, R. A.; Olsen, G. J.; Maloy, S. R. *Trends Microbiol.* **2002**, *10(2)*, 94-99.
[11]. Murtaza, S.; Altaf, A. A.; Hamayun, M.; Iftikhar, K.; Tahir, M. N.; Tariq, J.; Faiz, K. *Eur. J. Chem.* **2019**, *10(4)*, 358-366.
[12]. Cheminformatics, M. Calculation of molecular properties and bioactivity score, Retrieved November 26, 2019, from http://www.molinspiration.com
[13]. Norinder, U.; Bergstrom, C. A. *ChemMedChem* **2006**, *1(9)*, 920-937.
[14]. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45(12)*, 2615-2623.
[15]. Ritchie, T. J.; Macdonald, S. J.; Peace, S.; Pickett, S. D.; Luscombe, C. N. *Med. Chem. Comm.* **2013**, *4(4)*, 673-680.
[16]. Thakker, D. R.; Hoyer, D.; Cryan, J. F. *Pharmacol. Therapeut.* **2006**, *109(3)*, 413-438.
[17]. Alavijeh, M. S.; Chishty, M.; Qaiser, M. Z.; Palmer, A. M. *Neurotherapeutics* **2005**, *2(4)*, 554-571.
[18]. Baringhaus, K. H.; Matter, H.; Stengelin, S.; Kramer, W. *J. Lipid Res.* **1999**, *40*, 2158-2168.
[19]. Kwon, Y., Handbook of essential pharmacokinetics, pharmaco-dynamics and drug metabolism for industrial scientists. New York: Kluwer Academic/Plenum Publishers, 2001.
[20]. Jiang, P.; Li, Q.; Lu, J. *IOP Conf. Ser.: Earth Environ. Sci.* **2018**, *178*, 012020.
[21]. Kerns, E. H.; Di, L., Drug-like properties: concepts, structure design and methods from ADME to toxicity optimization, *Academic press,* ISBN 978-0-12-369520-8, 2008.
[22]. Manallack, D. T. *Perspec. Med. Chem.* **2007**, *1*, 25-48.
[23]. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inform. Model.* **2005**, *45(1)*, 177-182.
[24]. Van De Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. *J. Med. Chem.* **2001**, *44(9)*, 1313-1333.
[25]. Chan, O. H.; Stewart, B. H. *Drug Discov. Today* **1996**, *1(11)*, 461-473.
[26]. Harris, D. A. *Trends Biochem. Sci.* **1993**, *18(9)*, 354.
[27]. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53(7)*, 2719-2740.
[28]. Daina, A.; Michielin, O.; Zoete, V. *Sci. Reports* **2017**, *7*, 42717.
[29]. Yang, H.; Lou, C.; Sun, L.; Li, J.; Cai, Y.; Wang, Z.; Li, W.; Liu, G.; Tang, Y. *Bioinformatics* **2019**, *35(6)*, 1067-1069.
[30]. Dong, J.; Wang, N. N.; Yao, Z. J.; Zhang, L.; Cheng, Y.; Ouyang, D.; Lu, A. P.; Cao, D. S. *J. Cheminfor.* **2018**, *10(1)*, 29.
[31]. Pires, D. E.; Blundell, T. L.; Ascher, D. B. *J. Med. Chem.* **2015**, *58(9)*, 4066-4072.
[32]. Schyman, P.; Liu, R.; Desai, V.; Wallqvist, A. *Frontiers Pharmacol.* **2017**, *8*, 889, 1-14.
[33]. Lee, S. K.; Park, S. H.; Lee, I. H.; No, K. T. PreAD-MET Ver. v2. 0, Seoul, Korea, Bioinformatics and Molecular Design Research, BMDRC, 2007.
[34]. Banerjee, P.; Eckert, A. O.; Schrey. A. K.; Preissner, R. *Nucleic Acids Res.* **2018**, *46(W1)*, W257-W263.
[35]. Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C; Sayeeda, Z.; Assempour, N. *Nucleic Acids Res.* **2018**, *46(D1)*, D1074-1082.
[36]. Dalvie, D.; Kalgutkar, A. S.; Chen, W. *Drug Metabol. Rev.* **2015**, *47(1)*, 56-70.
[37]. Rydberg, P.; Gloriam, D. E.; Olsen, L. *Bioinformatics* **2010**, *26(23)*, 2988-2989.
[38]. Rydberg, P.; Gloriam, D. E.; Zaretzki, J.; Breneman, C.; Olsen, L. *ACS Med. Chem. Let.* **2010**, *1(3)*, 96-100.
[39]. Zhang, Y.; Chen, G.; Ma, H.; Guo, M. *Molecules* **2019,** *24(1),* 166, 19.
[40]. Rose, P. W.; Prlic, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Costanzo, L. D.; Duarte, J. M.; Dutta, S.; Feng, Z.; Green, R. K.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y. P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. *Nucleic Acids Res.* **2017**, *45(D1)*, D271-D281.
[41]. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. *Nucleic Acids Res.* **2018**, *46(W1)*, W363-W367.
[42]. Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. *Nucleic Acids Res.* **2006**, *34(Suppl2)*, W116-W118.
[43]. Tian, W.; Chen, C.; Liang, J. *Biophys. J.* **2018**, *114(3), (Suppl1)*, 50a.
[44]. Laskowski, R. A.; Hutchinson, E. G.; Michie, A. D.; Wallace, A. C.; Jones, M. L.; Thornton, J. M. *Trends Biochem. Sci.* **1997**, *22(12)*, 488-490.
[45]. Basta, H. A.; Sgro, J. Y.; Palmenberg, *A. C. Virology.* **2014**, *448*, 176-184.
[46]. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallog.* **1993**, *26(2)*, 283-291.
[47]. Mering, C. V.; Huynen, M.; Jaeggi, D.; Schmidt, S.; Bork, P.; Snel, B. *Nucleic Acids Res.* **2003**, *31(1)*, 258-261.
[48]. Grosdidier, A.; Zoete, V.; Michielin, O. *Nucleic Acids Res.* **2011**, *39(Suppl2)*, W270-W277.
[49]. Kiss, R.; Sandor, M.; Szalai, F. A. *J. Cheminform.* **2012**, *4*, P17.
[50]. Labbe, C. M.; Rey, J.; Lagorce, D.; Vavrusa. M.; Becot, J.; Sperandio. O.; Villoutreix, B. O.; Tuffery, P.; Miteva, M. A. *Nucleic Acids Res.* **2015**, *43(W1)*, W448-W454.