Biotechnology and Genetic Engineering Reviews

ISSN: 0264-8725 (Print) 2046-5556 (Online) Journal homepage: http://www.tandfonline.com/loi/tbgr20

Bioinformatics approaches for new drug discovery: a review

Kullappan Malathi & Sudha Ramaiah

To cite this article: Kullappan Malathi & Sudha Ramaiah (2018): Bioinformatics approaches for new drug discovery: a review, Biotechnology and Genetic Engineering Reviews, DOI: 10.1080/02648725.2018.1502984

To link to this article: https://doi.org/10.1080/02648725.2018.1502984

Published online: 31 Jul 2018.

BIOTECHNOLOGY AND GENETIC ENGINEERING REVIEWS
https://doi.org/10.1080/02648725.2018.1502984

Taylor & Francis
Taylor & Francis Group

ARTICLE

Bioinformatics approaches for new drug discovery: a review

Kullappan Malathi and Sudha Ramaiah

Medical & Biological Computing Laboratory, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, India; Department of Biosciences, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu

ABSTRACT

Prolonged antibiotic therapy for the bacterial infections has resulted in high levels of antibiotic resistance. Initially, bacteria are susceptible to the antibiotics, but can gradually develop resistance. Treating such drug-resistant bacteria remains difficult or even impossible. Hence, there is a need to develop effective drugs against bacterial pathogens. The drug discovery process is time-consuming, expensive and laborious. The traditionally available drug discovery process initiates with the identification of target as well as the most promising drug molecule, followed by the optimization of this, *in-vitro, in-vivo* and in pre-clinical studies to decide whether the compound has the potential to be developed as a drug molecule. Drug discovery, drug development and commercialization are complicated processes. To overcome some of these problems, there are many computational tools available for new drug discovery, which could be cost effective and less time-consuming. *In-silico* approaches can reduce the number of potential compounds from hundreds of thousands to the tens of thousands which could be studied for drug discovery and this results in savings of time, money and human resources. Our review is on the various computational methods employed in new drug discovery processes.

ARTICLE HISTORY
Received 27 January 2016
Accepted 18 July 2018

KEYWORDS
Drug discovery; antibiotic resistance; bacterial infections; Molecular docking; Molecular modelling

Introduction

The control of pathogenic infections is severely hampered by the constant rise in the number of pathogenic micro-organisms which are highly resistant to several antibiotics (Holmberg, Solomon, & Blake, 1987). Drug-resistant infections will tend to increase the morbidity as well as the period of hospitalization. Hence, there arises a need to design and develop better drug candidates to meet this challenge. The traditionally available drug discovery and development approach is a manifold process and a highly time-consuming one. These multi-step processes are; (i)

CONTACT Sudha Ramaiah sudhaanand@vit.ac.in Department of Biosciences, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India

identification of an essential disease target, that will be used to identify a specific therapeutic ligand (0.5–1 year); (ii) validation of the target (0.5–1 year); (iii) assay development; searching for an 'ideal' drug against the disease specific target (0.5–1 year); (iv) lead compound identification; the use of chemical libraries to test for activity against the target protein (0.5–1 year); (v) Optimization of the lead candidate (2–3 years); (vi) Pre-clinical development includes *in-vivo* and *in-vitro* studies to test the tolerability as well as the side effects of the compound (1 year); (vii) The industry will file for Investigational New Drug (IND) with Food and Drug Administration (FDA in the USA); (viii) Clinical development studies comprise three phases of clinical trials. Phase I clinical trials are conducted on 20–100 healthy volunteers to check absorption, distribution, metabolism and excretion (ADME) levels of the drug (2 years). Phase II clinical trials are conducted with 100–500 volunteers for a few months to determine the final dosage required (2 years). Phase III trials are conducted with 1000–5000 volunteers to confirm the safety and effectiveness of the drug (2 years); (ix) NDA has to be filed for FDA review; (x) Finally the drug will be approved for marketing (Figure 1). Drug discovery and the development process include the investment of major capital ranging from millions to billions of dollars per successful drug (Paul et al., 2010). Hence, it has become a challenge for the pharmaceutical industries to develop innovative drugs. Despite major Research and Development (R&D) investments and technical advancements, the number of new drug discovery applications approved by the FDA per year is low (Mullard, 2011). The increase in large capital investments and decline in FDA approval is known as 'the innovation gap' (Earm & Earm, 2014). To overcome some of these problems, more sophisticated *in-silico* approaches are being used for discovery of new drugs. Most of the *in-silico* methods are carried out alongside *in-vitro* and *in-vivo* data for concordant results (Noori & Spanagel, 2013). These computational techniques play a significant role in the pharmacology sectors that cover worldwide drug development through the use of various software and tools. Our review article examines the various *in-silico* methods involved in the new drug discovery process (Figure 2).

In-silico approaches for drug designing

Molecular modelling

In structure-based drug design, protein-ligand binding plays a crucial role. Hence, there is a need for a three-dimensional (3D) structure of the protein. For our computational drug discovery process, first the 3D structure of the protein is searched for in an appropriate Protein

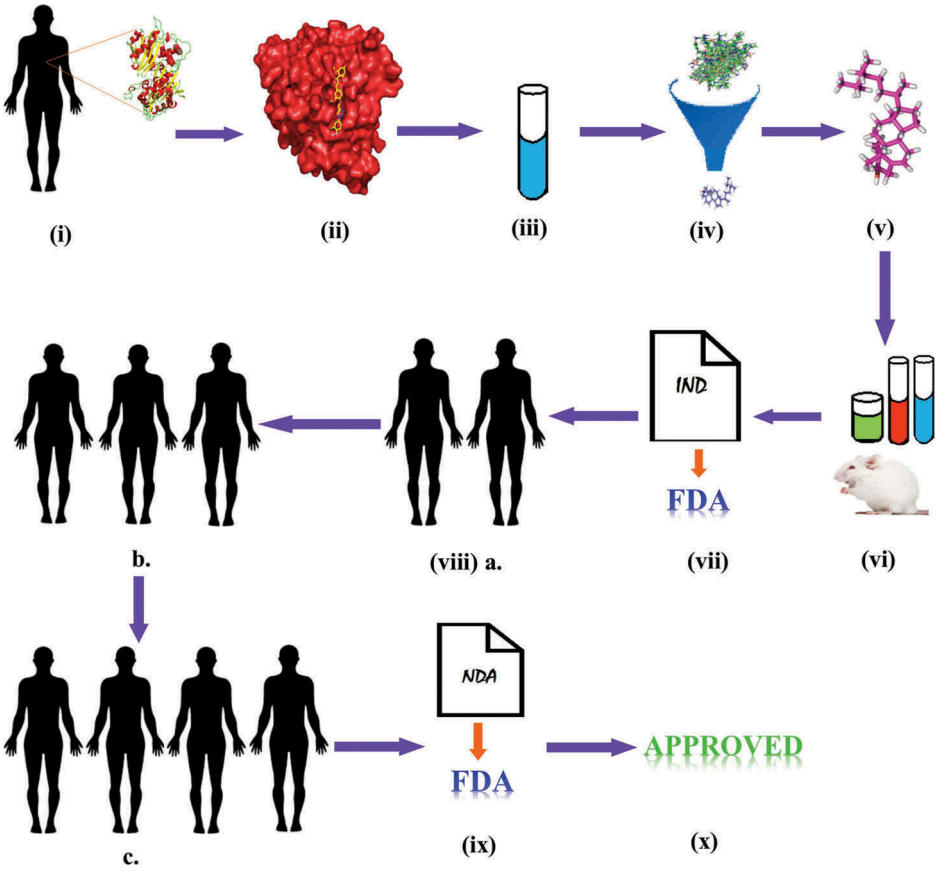

Figure 1. Drug discovery process. (i) Identification of disease and specific drug target (ii) Validation of the target (iii) Assay development (iv) Lead identification against the target (v) Optimization of the lead candidate (vi) Pre-clinical development including *in-vivo* and *in-vitro* studies (vii) File for Investigational New Drug (IND) with Food and Drug Administration (FDA; USA) (viii) Clinical development studies, a. phase I clinical trials (20 – 100 healthy volunteers) b. Phase II clinical trials (100–500 volunteers) c. Phase III trials (1000 to 5000 volunteers) (ix) NDA filed for FDA review (x) Drug approved for marketing.

data bank (PDB) (Berman et al., 2000). PDB is the protein structural database, where all the known 3D structures of the proteins are deposited (http://www.rcsb.org/). All these 3D structures are experimentally determined by two major techniques, namely: X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR). Experimental determination of structure is difficult and very few structures are available in the PDB database. Hence, there is also a need to model a protein from existing data and sequence. The *in silico* approach offers a 'homology-based modelling' method for protein modelling. This modelling approach works on the basis of sequence similarity in evolutionarily equivalent proteins. Homology modelling involves the following steps.

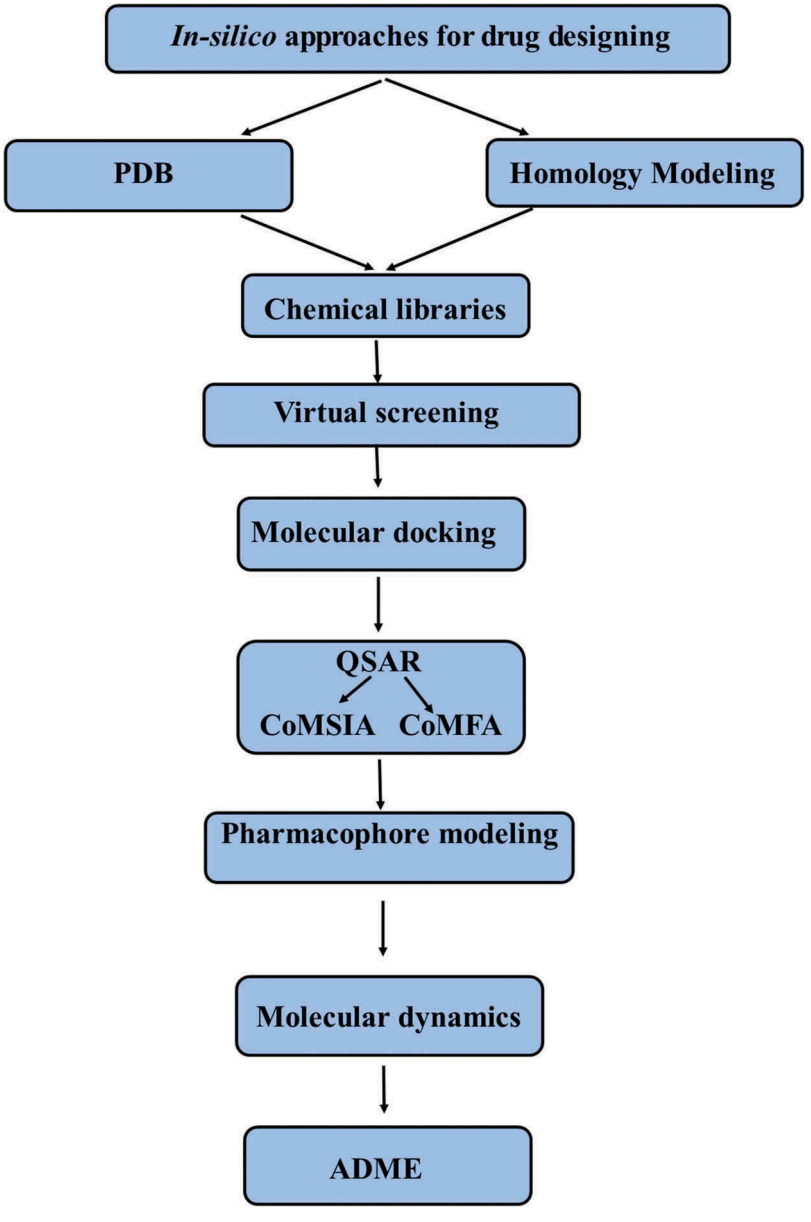

Figure 2. *In-silico* approaches for drug designing.

Recognition of template and sequence alignment

The foremost step in homology modelling is template recognition. This includes querying unknown amino acid sequences against already existing ones in the protein data bank in order to identify the homologous sequences, whose structure is already resolved. These homologues sequences are identified by similarity searches using programmes such as BLAST (Basic local alignment search tool) for the amino acid sequence alignment.

Model building

Building of the model can follow a number of methods: a) spatial restraint, b) rigid-body assembly, c) segment matching and d) artificial evolution.

Refinement modelling

Model-building methods involve the addition, deletion and substitution of amino acid residues. This model refinement involves loop modelling and side-chain modelling. This model refinement is based on molecular dynamics, Monte Carlo methods and genetic algorithms. The modelled structures are energy minimized by using certain force fields such as OPLS, AMBER, MM3 and CHARMM22.

Loop modelling

Insertion and substitution of amino acid residues into the sequences of homologous proteins are known as loops, which tend to be the variable portion of the protein. In template-based loop modelling, the missing loop region sequence segment is searched in template proteins for modelling. If that particular sequence is not available in the template then the loop is modelled through non-template based methods (*de novo* methods).

Side chain modelling

Modelling of side chains involves the substitution of side chains on the backbone structure. The modelled side chains are analyzed through their root mean square deviation (RMSD) values.

Validation of modelled protein structure

After the homology modelling, it is necessary to validate the protein to check the accuracy of the modelling. The stereochemical nature of the protein can be investigated by WHATCHECK, WHAT IF, VADAR and PROCHECK. The Ramachandran plot evaluation is carried out to check the torsion angles of main chain, ϕ and ψ. In a scatter plot, if a greater percentage of residues are found in the allowed region than in the disallowed zone, it is a good sign of the structural feasibility of the protein.

Programs available for homology modelling are MODELLER, SWISS PDB VIEWER, SWISS MODEL and COMPOSER (Sali & Blundell, 1993).

The homology modelling tool MODELLER 9.16 was used to build the OXA-143 Carbapenem Hydrolyzing class D β-lactamase (CHDL) with the help of the template OXA-24 CHDL and it was validated (Figure 3(a,b)). In this study, the OXA-143 CHDL structure was developed to analyze the resistance mechanism of imipenem in OXA-143 CHDL expressing *Acinetobacter baumannii* (Malathi, Anbarasu, & Ramaiah, 2017; Wei et al., 2016).

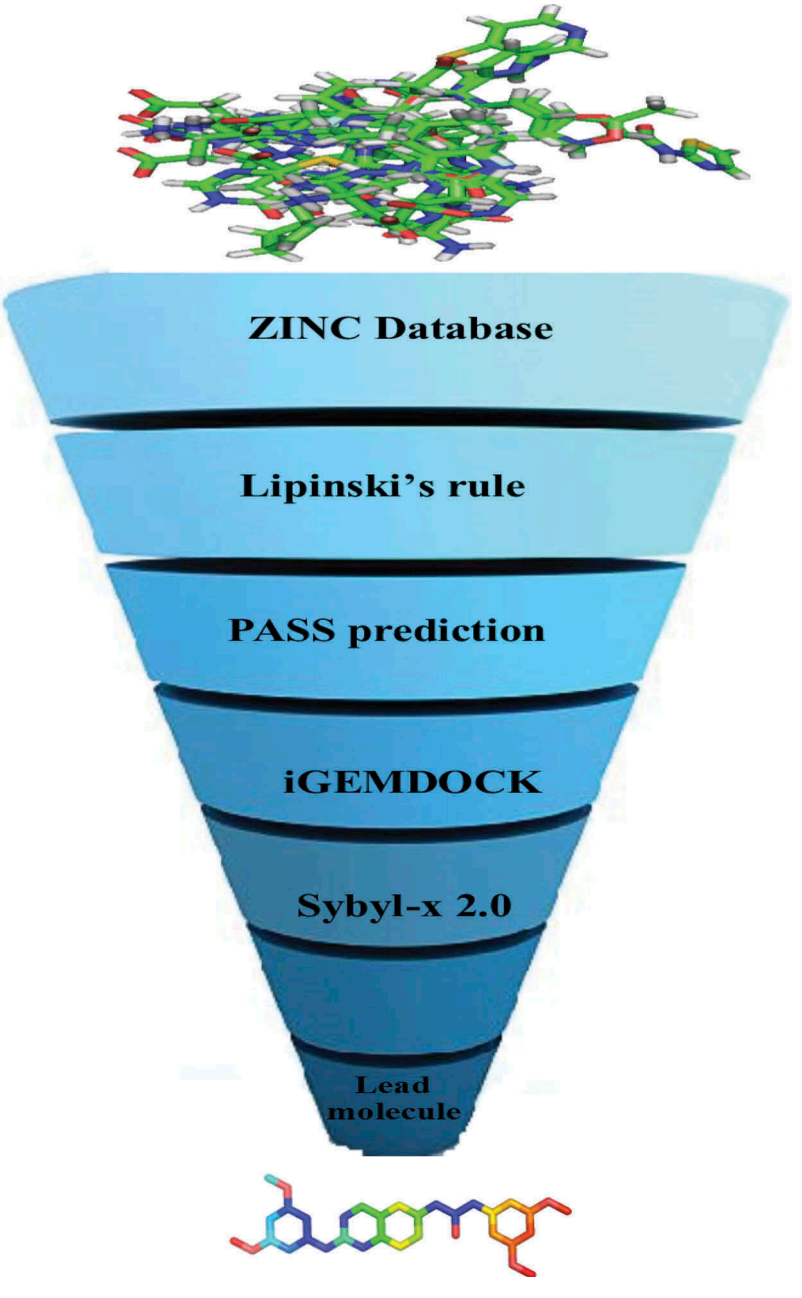

Figure 3. Techniques used in virtual screening process.

Small molecule databases

Chemical libraries offer an extraordinary diversity and purity of screening compounds which have been used for more than two decades to identify novel and drug-like properties in drug discovery.

Most commonly used chemical libraries are the NCBI PubChem, ZINC Database, DrugBank and ChEMBL database (Kim et al., 2016)

Virtual screening

Virtual screening is the commonly used *in-silico* approach for the lead identification process, which searches for an active lead from the chemical database. Virtual screening approaches are of two types namely structure-based and ligand-based. Structure-based virtual screening involves the use of the 3D structure of target protein to screen against the compounds present in the chemical libraries through molecular docking studies. Structure-based virtual screening methods follow the docking procedure to identify the active lead based on the binding affinity of the compound with the target and the functional scores.

The structure-based virtual screening procedure was employed in a study to screen the potential inhibitors for OXA-10 ESBL expressing *P. aeruginosa* against the millions of compounds present in the ZINC database (Malathi & Ramaiah, 2016). It was also utilized to identify novel inhibitors for Penicillin binding protein 2a (PBP2a) of ceftaroline resistant methicillin-resistant Staphylococcus aureus (MRSA). The Dock blaster server was used for the purpose of virtual screening (Irwin et al., 2009; Lavanya, Ramaiah, & Anbarasu, 2016). The different methodologies involved in virtual screening process are represented in Figure 4.

Molecular docking

Molecular docking studies are used to understand the interaction between a ligand and a target protein at the atomic level. These molecular docking

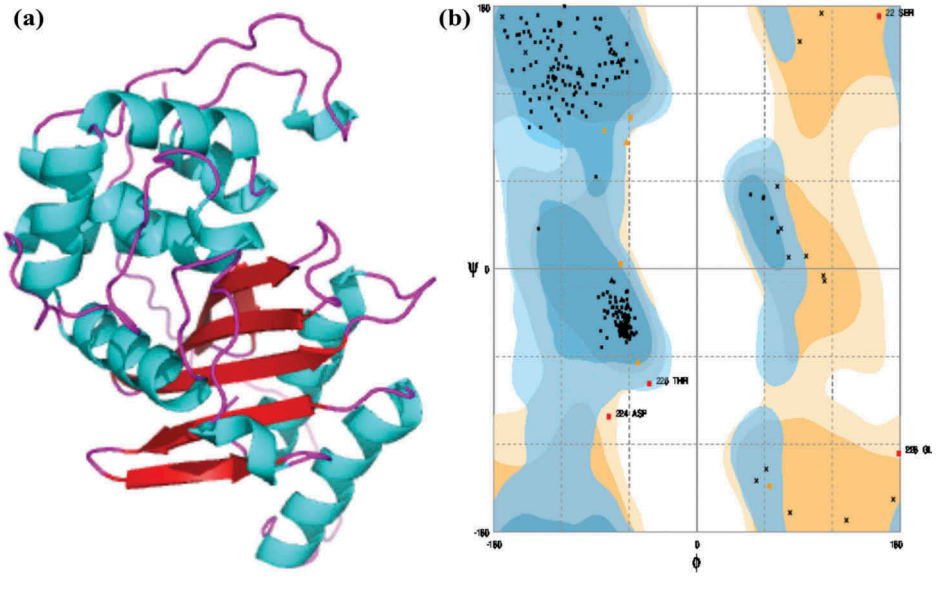

Figure 4. Homology modelling of protein. (a) OXA-143 CHDL structure (b) Ramachandran plot.

studies have two goals: prediction of protein-ligand binding orientation as well as the estimation of the interaction. Before the start of the docking process, it is necessary to identify the active site of the protein where the ligands can bind. For the prediction of active sites, many online programs are available like CASTp, Q-SiteFinder, LigA Site and MetaPocket (Laurie & Jackson, 2005). A molecular docking study without the prediction of the active site is known as a 'blind docking'(Malathi, Anbarasu & Ramaiah, 2016).

Searching algorithms

The sampling algorithms are available in currently used docking software. Those searching algorithms are (i) Matching algorithms—DOCK, LibDock and SANDOCK (ii) Incremental construction method—DOCK 4.0, FlexX and Hammerhead (iii) Multiple Copy Simultaneous Search (iv) LUDI (v) Monte Carlo—AutoDock—ICM, QXP (vi) Genetic algorithms—AutoDock, GOLD and DARWIN.

Scoring functions

The main goal of scoring is to estimate the binding affinity of a compound with the protein and rank the complex against other candidates. These scoring functions are classified as 'empirical', 'force field' and 'knowledge-based'. The empirical scoring involves the conversion of protein-ligand binding energy into many energy components. All of these components are multiplied and summed for the final score. Force field functional scorings estimate the binding affinity by the calculation of electrostatic and van der Waals interactions. Leonard-Jones potential energy was used for the van der Waals interactions and Columbic potential energy formulation was used for the electrostatic interactions. Knowledge-based scoring functions are emerging as an alternate method of grading protein-ligand binding affinities along with 3D structures. It is a statistical method that utilizes the structural information of protein-ligand complexes deposited in the databases to develop potentials of mean force (PMF). The calculation of binding energy through knowledge-based scoring functions is computationally a very simple method (Muegge & Martin, 1999).

Software widely used for molecular docking studies are Surflex-dock, AutoDock, GOLD Glide, FlexX, DOCK, HADDOCK (Allen et al., 2015 ; Dominguez, Boelens & Bonvin, 2003).

Sybyl-x 2.0, a surflex-dock software was used in many studies that explore the resistance mechanism of antibiotics in different classes of β-lactamases expressing pathogenic bacteria and the screening for inhibitors of β-lactamases (Kumar, Anbarasu, & Ramaiah, 2014; Malathi & Ramaiah, 2016). The interaction of class D β-lactamase OXA-10 ESBL and the antibiotic imipenem is given in Figure 5(a,b). It was also used in the

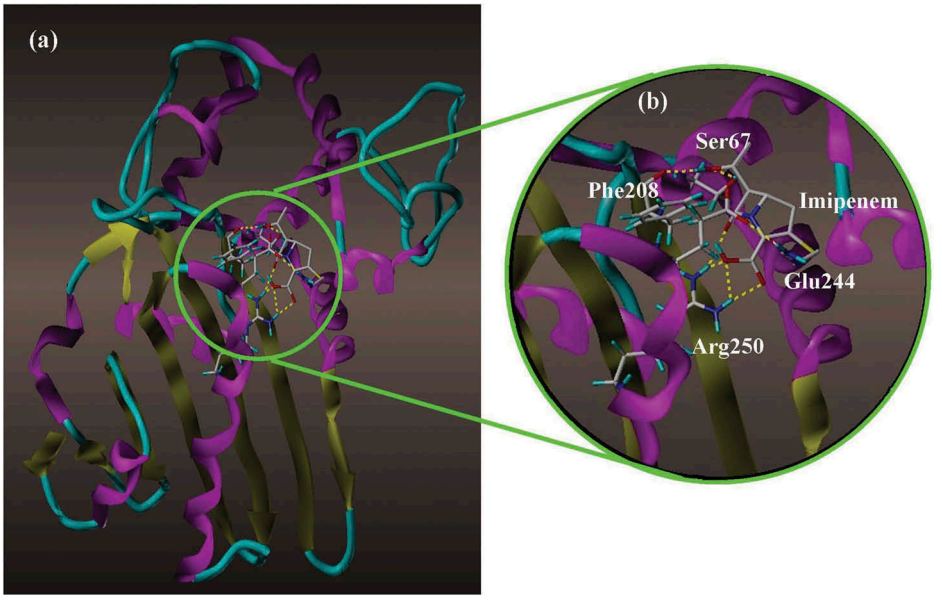

Figure 5. Molecular docking results for OXA-10 ESBL and imipenem. (a) Interaction of OXA-10 ESBL and imipenem (b) Hydrogen bond interaction between the imipenem and active site residues of the OXA-10 ESBL.

screening of anti-malarial compounds for the dihydroorotate dehydrogenase of *Plasmodium falciparum* from chalcone and flavone hybrids (Thillainayagam, Malathi, & Ramaiah, 2017). The inhibitor screening of MurB, a cell wall protein of *Vibrio cholera*, against phytocompounds used the surflex-dock package (Ragunathan, Malathi, & Anbarasu, 2017).

QSAR

Ligand-based virtual screening involves the use of QSAR studies. The QSAR method involves the development of mathematical models to correlate relationships between the biological function and physicochemical characteristics. In the drug discovery process, QSAR methods are used for the lead optimization, a major step in drug discovery. In 3D-QSAR, Comparative molecular field analysis (CoMFA) and Comparative molecular similarity indices analysis (CoMSIA) are the two techniques developed for ligand-based drug design.

CoMFA

For a CoMFA study, a set of derivatives of bioactive compounds were first selected, these compounds having different substitutions. All of these compounds were divided into a 70% training set and 30% test set. For QSAR performance, various commercial softwares are available. The

regression models are created with the training set compounds and the predicting ability of the generated 3D QSAR models are estimated by test sets. A lattice box was arranged with all these structures with a regular substructure. The docked high energy compound is selected as a template for prediction. Powell gradient and Gasteiger-Huckel charges (Gasteiger & Marsili, 1980) were used to optimize the energy. The electrostatic and steric fields were calculated with columbic potential energy and Lennard-Jones potential energy, respectively. The CoMFA descriptors are calculated by the sp^3 probe. For regression analysis, the partial least square method was used. The generated CoMFA models were cross-validated using the LOO method. The steric and electrostatic contour plots play an essential role in the drug design process. These contour plots relate to the variation of the molecular fields and changes in biological activity.

CoMSIA

CoMSIA is a more advanced method of CoMFA, which has fewer limitations. Molecular similarity indices of CoMSIA are estimated from the SEAL similarity, as descriptors. This SEAL similarity method utilizes the electrostatic, steric, hydrogen bonding and hydrophobic descriptors. Similarities of the molecules are calculated by sp^3 probes with +1 charge and +1 hydrophobicity. For the estimation of similarity, the distances between the molecule atom and the atom of the carbon probe are calculated. To delineate the distance between these two atoms and to estimate the molecular as well as the physicochemical properties, Gaussian functions are utilized. Gaussian functions producing slopes are not as steep as in CoMFA, which is produced by Lennard-Jones and Columbic potentials. In the CoMSIA method, the generated contours will indicate the favoured or unfavoured regions in the ligand binding areas. Software used for QSAR studies are Sybyl-X 2.0 and E-Dragon.

The COMFA study was carried out on quinolinyl and aryl chalcone derivatives with the help of the Sybyl-x 2.0 module to identify the antimalarial activity of compounds (Thillainayagam, Anbarasu, & Ramaiah, 2016; Thillainayagam et al., 2015). Partial least square methods are applied to correlate the COMFA models with the pIC50 values. The steric and electrostatic contour plots of the most active compound XIII ((2E)-3-(4-methoxyphenyl) −1-(4-[1,2,4]-triazol-1-yl-phenyl) prop-2-en-1-one) is illustrated in Figure 6(a,b). Three-dimensional steric (Green and yellow polyhedrons) and electrostatic (red and blue polyhedrons) contour plots offer valuable information for the modification or proposal of novel ligands.

The CoMFA and CoMSIA analyses were carried out on pyrrole hydrazine derivatives that act on enoyl-acyl carrier protein reductase, to reveal potential anti-mycobacterium activity (More et al., 2014).

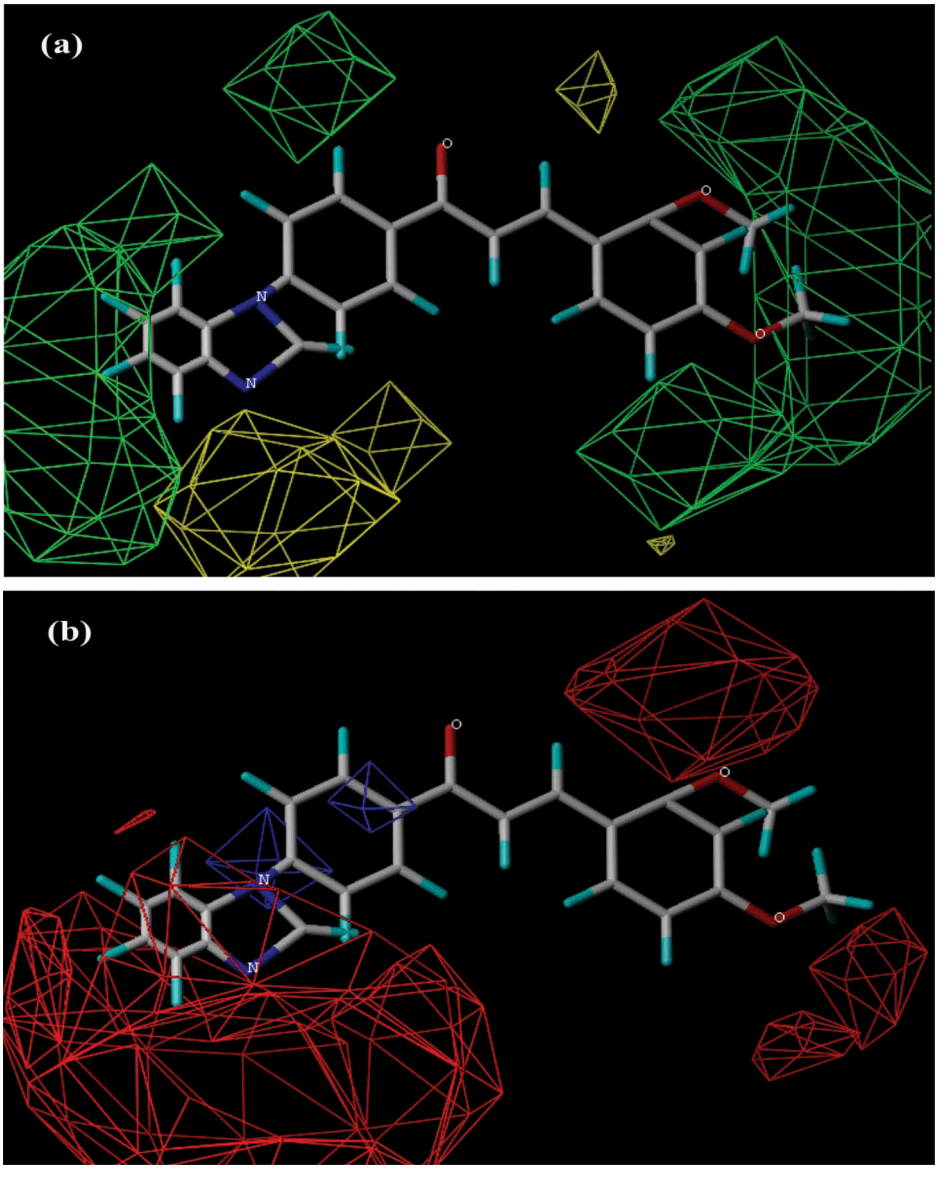

Figure 6. COMFA contour plots of compound XIII ((2E)-3-(4-methoxy phenyl) −1-(4-[1,2,4] – triazol −1- yl-phenyl) prop-2-en-1-one). (a) Steric contour plots—green polyhedrons indicates the favoured region, yellow polyhedrons indicates the disfavoured region (b) Electrostatic contour plots—red polyhedrons indicates the negatively charged groups, blue polyhedrons indicates the positively charged groups.

Pharmacophore modelling

Pharmacophore modelling is known as the set of molecular characteristics essential for the macromolecular recognition of ligands that trigger a biological reaction. Pharmacophores are modelled with the features such as aromatic, hydrophobic, hydrogen bond acceptor,

hydrogen bond donor, anion and cation residues. Pharmacophore models are used as a query in virtual screening for the identification of molecules from the chemical library. Pharmacophore modelling is of two types namely structure based and ligand-based. In structure-based pharmacophore modelling, it is totally dependent on the 3D molecular structure of the target protein that has been resolved by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. For structure-based pharmacophore modelling the active site as well as the spatial relationships are analysed for the certain chemical characteristics that are complementary to that of the binding ligand. Combination of 3D conformation of the protein as well as the essential characteristics of the ligand will lead to the selective identification of predicted hits from large chemical libraries. The target proteins undergo structural changes upon binding of ligand. Hence it is necessary to incorporate structural flexibility during the modelling to increase the hit rate. Ligand-based modelling is used in the absence of the 3D structure of target protein. In this situation, structure-based screening is impossible. The generation of pharmacophore from the ligand is completely based on the common features of pharmacophore and QSAR pharmacophore generation. In common feature pharmacophores, two different chemical structures will share the same features at the same position. These chemical features are extracted from the set of familiar ligand structures. The common features include hydrophobic interaction, hydrogen bonding and electrostatic interactions. In ligand-based pharmacophore modelling, the construction of pharmacophore models have two major steps: i) Ligands in the training set are arranged with a conformational space to create ligand conformational flexibility and (ii) aligning set of ligands in the training set to find common chemical features from all ligands.

Software used for pharmacophore modelling are HipHop, DISCO, HypoGen and PHASE (Martin, 2000).

The ligand-based pharmacophore search was carried out in a study to find non-ATP competitive inhibitors for the mechanistic or mammalian target of rapamycin (mTOR). The spatial arrangement of ligand and protein model was generated with hydrophobic interactions of residues C5, C19, C21, C43, C45 and C49 of rapamycin to construct a model by the ZincPharmer platform. It has resulted in eight new inhibitors with better activity (Kist, Timmers, & Caceres, 2018).

Molecular dynamics (MD)

Computer-aided simulation of molecular systems involves the solution of Newton's equation for motion. The molecular system simulation has

results in the trajectories, where all the molecular properties of the target protein or protein and ligand complexes are calculated. The MD simulations are carried out to understand the stability of the free protein or protein-ligand complexes. The docked best binding energy protein-ligand complex is subjected to MD simulation. Initially the protein topology is obtained by standard parameters using gromacs or LEap programme. The ligand topology is generated by the online server PRODRG program (Schuttelkopf & van Aalten, 2004). Various force fields like Amber and Gromos are used for the protein topology generation. The protein along with the ligand complexes are kept in a cubic box and solvated with simple point charge water. There must be a distance between the cubic box and protein-ligand complex. Since proteins are negatively charged, the system is neutralized with counter ions like Na^+ or Cl^-. The systems are energy minimized for 1000 steps by steepest descent algorithm. Next, at 300 K position restraining simulations are carried out for 100 pS under constant volume and temperature dynamics (NVT) and followed by production simulation for 100 pS under pressure and temperature dynamics (NPT) at 300 K. Linear Constraint Solver algorithm (Hess, Bekker, Berendsen, & Fraaije, 1997) was used for the bond length restraining. The Particle Mesh Ewald Method was used for the estimation of electrostatic interactions. Finally, each complex is subjected to MD simulation for particular ps. After the simulation, the trajectories were analysed and the Xmgrace tool was used for graph generation (Turner, 2005). The root mean square deviation (RMSD), root mean square fluctuation (RMSF), Radius of gyration (Rg) and Intermolecular hydrogen bond formations were analyzed to check the stability of the complex.

Software packages used for MD simulations are GROMACS, NAMD, AMBER and CHARMM (van der Spoel et al., 2005).

To understand the stability of the protein-ligand complex and free protein, molecular dynamics are carried out for a point mutant (SHV-E166A in *Klebsiella pneumoniae*), double mutant (toho-1-R274N/R276N in *Escherichia coli*) and triple mutant (toho-1-E166A/R274N/R276N in *Escherichia coli*) systems of class A β-lactamases (Kumar, Lavanya, Anbarasu, & Ramaiah, 2014). In an investigation of the anti-dyslipidemic property of the plant compounds against HMG-CoA reductase, the stability of the rutin-HMG CoA complex was analyzed through the molecular dynamics simulation (Suganya, Nandagopal, & Anbarasu, 2017). The RMSD and RMSF plot of Epicatechin-HMG CoA complex and free HMG CoA is portrayed in Figure 7(a,b). These plots reveal the stability of the Epicatechin-HMG CoA complex rather than the free HMG CoA.

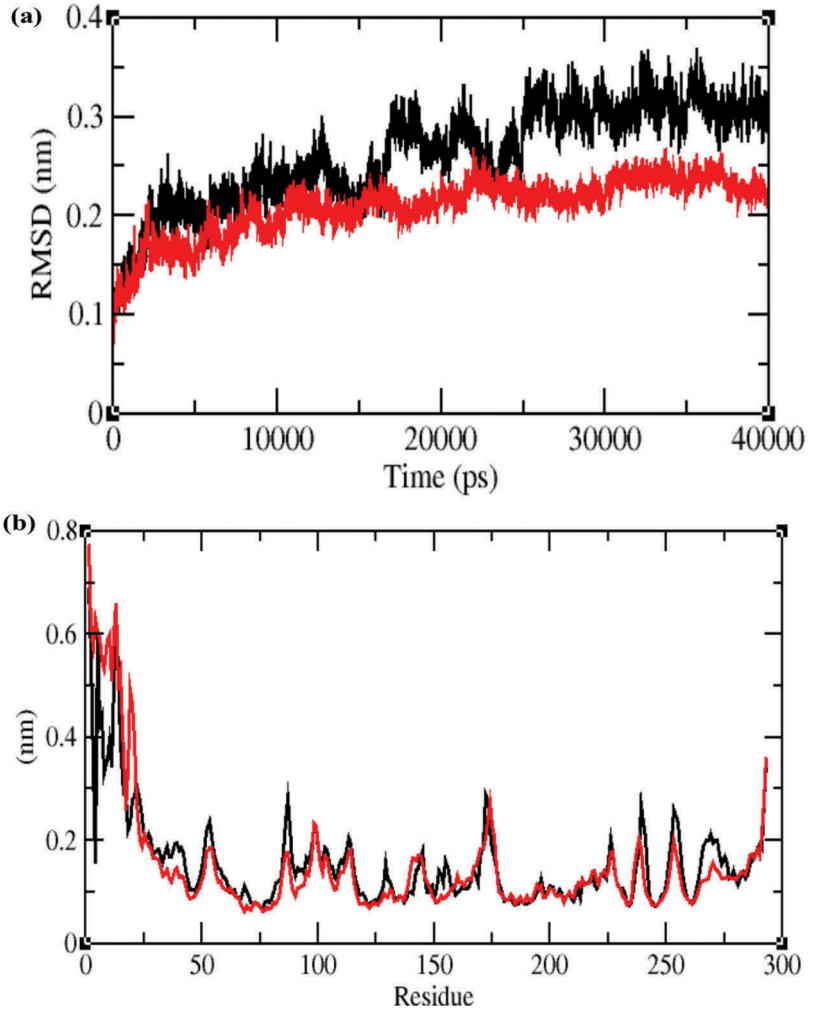

Figure 7. Molecular dynamic simulation results. (a) RMSD of Cα atoms—HMG CoA-Epicatechine complex (red) and HMG CoA (black) (b) RMSF of residues—HMG CoA-Epicatechine complex (red) and HMG CoA (black).

ADME prediction

ADME predictions play an important role in the drug discovery process. Most of the drugs fail in clinical trials due to the poor pharmacokinetic and toxicity predictions. The traditionally available methods take a longer time than for ADME prediction. The toxicity as well as the molecular property prediction will lead to either the success or failure of the drug in clinical trials. Lipinski's rule of five plays an essential role in property prediction (Gimenez, Santos, Ferrarini, & Fernandes, 2010). This rule accepts compounds with one violation. This rule of five states that molecular weight (MW) of the ligand should not exceed 500 Daltons, H-bond donor (HBD) must be less than 5, H-bond acceptor (HBA) must be below

10, the milogP value should be below 5 and the number of rotatable bonds must be below 10. The compounds that satisfy these rules will be orally available for humans.

Software available for ADMET predictions are admetSAR, PreADMET, Qikprop, Volsurf, Molinspiration, PASS (Cheng et al., 2012; Lagunin, Stepanchikova, Filimonov, & Poroikov, 2000).

The tool admetSAR was used to screen the ADMET properties of the compounds extracted from *Carica papaya*. This study was carried out to analyze the anti-dengue activity of the compounds isolated from *Carica papaya* against the NS2B-NS3 protease of dengue 2 virus (DENV-2) (Senthilvel et al., 2013). The Molinspiration tool was used in the virtual screening process to filter the imipenem analogues based on the Lipinski's rule of five. This study involves the search of novel inhibitors for OXA-10 ESBL from ZINC database (Malathi & Ramaiah, 2016).

Conclusion

The traditionally available drug discovery process involves large capital investments and takes a long time for drug development. Due to the limitations and the longer period of drug development, there is a need for new techniques in the drug discovery process. In our review article, we have discussed various *in-silico* methods such as molecular modelling, virtual screening, molecular docking, pharmacophore modelling, molecular dynamics and ADMET prediction tools that can be used to make new drug discovery more efficient.

Acknowledgments

SR gratefully acknowledges the Indian Council of Medical Research (ICMR), Government of India Agency for the research Grant (IRIS ID: 2014-0099). KM thanks ICMR for Senior Research Fellowship (IRIS ID: 2015-25750). The authors would also like to thank the management of VIT for providing the necessary facilities.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

SR gratefully acknowledges the Indian Council of Medical Research (ICMR), Government of India Agency for the research Grant [IRIS ID: 2014-0099]. KM thanks ICMR for Senior Research Fellowship [IRIS ID: 2015-25750].

Notes on contributors

Malathi Kullappan is an ICMR-SRF at Vellore Institute of Technology (VIT), Vellore, India. She pursues her Ph.D degree at VIT under the guidance of Dr. Sudha Ramaiah. She has received her Master of Philosophy in Biochemistry from Thiruvalluvar University, Vellore, India. After completing her masters in Biochemistry, she has been a research scientist at Brainwave Biosolutions Limited, Chennai. Her research interests include designing of drugs for the antibiotic resistant microorganisms through the molecular modelling.

Dr. Sudha Ramaiah is an Associate Professor in Department of Biosciences at Vellore Institute of Technology. She completed her Ph.D in Bioinformatics at Vellore Institute of Technology. She has received her Master of Philosophy in Biochemistry from Pondicherry University and her postgraduate studies at University of Madras, Chennai. Her research interest lies in the area of Biochemistry, Bioinformatics and Infectious Diseases. She has collaborated actively with researchers in several other disciplines and has more than 70 publications. She has received the grants from Indian Council of Medical Research and Department of Science and Technology, Government of India.

ORCID

Kullappan Malathi http://orcid.org/0000-0003-2981-6751

References

Allen, W. J., Balius, T. E., Mukherjee, S., Brozell, S. R., Moustakas, D. T., Lang, P. T., … Rizzo, R. C. (2015). DOCK 6: Impact of new features and current docking performance. *Journal of Computational Chemistry*, *36*, 1132–1156.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*, 235–242.

Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., … Tang, Y. (2012). admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. *Journal of Chemical Information and Modeling*, *52*, 3099–3105.

Dominguez, C., Boelens, R., & Bonvin, A. M. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*, 1731–1737.

Earm, K., & Earm, Y. E. (2014). Integrative approach in the era of failing drug discovery and development. *Integrative Medicine Research*, *3*, 211–216.

Gasteiger, J., & Marsili, M. (1980). Iterative partial equalization of orbital electronegativity – A rapid access to atomic charges. *Tetrahedron*, *36*, 3219–3228.

Gimenez, B. G., Santos, M. S., Ferrarini, M., & Fernandes, J. P. (2010). Evaluation of blockbuster drugs under the rule-of-five. *Pharmazie*, *65*, 148–152.

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.

Holmberg, S. D., Solomon, S. L., & Blake, P. A. (1987). Health and economic impacts of antimicrobial resistance. *Reviews of Infectious Diseases*, *9*, 1065–1078.

Irwin, J. J., Shoichet, B. K., Mysinger, M. M., Huang, N., Colizzi, F., Wassam, P., & Cao, Y. (2009). Automated docking screens: A feasibility study. *Journal of Medicinal Chemistry*, *52*, 5712–5720.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., & Bryant, S. H. (2016). PubChem substance and compound databases. *Nucleic Acids Research*, *44*, D1202–1213.

Kist, R., Timmers, L. F. S. M., & Caceres, R. A. (2018). Searching for potential mTOR inhibitors: Ligand-based drug design, docking and molecular dynamics studies of rapamycin binding site. *Journal of Molecular Graphics and Modelling*, *80*, 251–263.

Kumar, K. M., Anbarasu, A., & Ramaiah, S. (2014). Molecular docking and molecular dynamics studies on β-lactamases and penicillin binding proteins. *Molecular BioSystems*, *10*, 891–900.

Kumar, K. M., Lavanya, P., Anbarasu, A., & Ramaiah, S. (2014). Molecular dynamics and molecular docking studies on E166A point mutant, R274N/R276N double mutant, and E166A/R274N/R276N triple mutant forms of class A β-lactamases. *Journal of Biomolecular Structure and Dynamics*, *32*, 1953–1968.

Lagunin, A., Stepanchikova, A., Filimonov, D., & Poroikov, V. (2000). PASS: Prediction of activity spectra for biologically active substances. *Bioinformatics*, *16*, 747–748.

Laurie, A. T., & Jackson, R. M. (2005). Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*, *21*, 1908–1916.

Lavanya, P., Ramaiah, S., & Anbarasu, A. (2016). A molecular docking and dynamics study to screen potent anti-staphylococcal compounds against ceftaroline resistant MRSA. *Journal of Cellular Biochemistry*, *117*, 542–548.

Malathi, K., Anbarasu, A., & Ramaiah, S. (2016). Ethyl Iso-allocholate from a medicinal rice karungkavuni inhibits dihydropteroate synthase in escherichia coli: A molecular docking and dynamics study. *Indian Journal of Pharmaceutical Sciences*, *78*, 780–788.

Malathi, K., Anbarasu, A., & Ramaiah, S. (2017). Exploring the resistance mechanism of imipenem in carbapenem hydrolysing class D beta-lactamases OXA-143 and its variant OXA-231 (D224A) expressing Acinetobacter baumannii: An in-silico approach. *Computational Biology and Chemistry*, *67*, 1–8.

Malathi, K., & Ramaiah, S. (2016). Molecular docking and molecular dynamics studies to identify potential OXA-10 extended spectrum β-Lactamase non-hydrolysing inhibitors for pseudomonas aeruginosa. *Cell Biochemistry and Biophysics*, *74*, 141–155.

Martin, Y. C. (2000). DISCO: What we did right and what we missed. In O. F. Gunner (Ed.), *Pharmacophore perception, development, and use in drug design* (pp. 49–68). La Jolla: International University Line.

More, U. A., Joshi, S. D., Aminabhavi, T. M., Gadad, A. K., Nadagouda, M. N., & Kulkarni, V. H. (2014). Design, synthesis, molecular docking and 3D-QSAR studies of potent inhibitors of enoyl-acyl carrier protein reductase as potential antimycobacterial agents. *European Journal of Medicinal Chemistry*, *71*, 199–218.

Muegge, I., & Martin, Y. C. (1999). A general and fast scoring function for protein-ligand interactions: A simplified potential approach. *Journal of Medicinal Chemistry*, *42*, 791–804.

Mullard, A. (2011). FDA drug approvals. *Nature Reviews Drug Discovery*, *10*, 82–85.

Noori, H. R., & Spanagel, R. (2013). In silico pharmacology: Drug design and discovery's gate to the future. *Silico Pharmacology*, *1*, 1.

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, *9*, 203–214.

Ragunathan, A., Malathi, K., & Anbarasu, A. (2017). MurB as a target in an alternative approach to tackle the *Vibrio cholerae* resistance using molecular docking and simulation study. *Journal of Cellular Biochemistry*, *119*, 1726–1732.

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, *234*, 779–815.

Schuttelkopf, A. W., & van Aalten, D. M. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*, 1355–1363.

Senthilvel, P., Lavanya, P., Kumar, K. M., Swetha, R., Anitha, P., Bag, S. …., & Anbarasu, A. (2013). Flavonoid from Carica papaya inhibits NS2B-NS3 protease and prevents Dengue 2 viral assembly. *Bioinformation*, *9*, 889–895.

Suganya, S., Nandagopal, B., & Anbarasu, A. (2017). Natural Inhibitors of HMG-CoA Reductase-An Insilico approach through molecular docking and simulation studies. *Journal of Cellular Biochemistry*, *118*, 52–57.

Thillainayagam, M., Anbarasu, A., & Ramaiah, S. (2016). Comparative molecular field analysis and molecular docking studies on novel aryl chalcone derivatives against an important drug target cysteine protease in Plasmodium falciparum. *Journal of Theoretical Biology*, *403*, 110–128.

Thillainayagam, M., Malathi, K., & Ramaiah, S. (2017). In-Silico molecular docking and simulation studies on novel chalcone and flavone hybrid derivatives with 1, 2, 3-triazole linkage as vital inhibitors of *Plasmodium falciparum* dihydroorotate dehydrogenase. *Journal of Biomolecular Structure and Dynamics*, 1–17. doi:10.1080/07391102.2017.1404935

Thillainayagam, M., Pandian, L., Murugan, K. K., Vijayaparthasarathi, V., Sundaramoorthy, S., Anbarasu, A., & Ramaiah, S. (2015). In silico analysis reveals the anti-malarial potential of quinolinyl chalcone derivatives. *Journal of Biomolecular Structure and Dynamics*, *33*, 961–977.

Turner, P. J. (2005). *XMGRACE, Version 5. 1.19*. Beaverton, ORE: Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology.

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, *26*, 1701–1718.

Wei, L., Chintala, S., Ciamporcero, E., Ramakrishnan, S., Elbanna, M., Wang, J., & Pili, R. (2016). Genomic profiling is predictive of response to cisplatin treatment but not to PI3K inhibition in bladder cancer patient-derived xenografts. *Oncotarget*, *7*, 76374–76389.