

Virtual screening strategies: a state of art to combat with multiple drug resistance strains

Abstract

Introduction: The resistance among the bacterial world has been increasing against majority of antibiotics. An effective approach to combat with this situation is the use of potential inhibitors with proper combination of these antibiotics. Traditionally, high throughput screening is used to screen the potential drug for the selected target. However, this strategy is costly and time consuming.

Approaches: Thus, an alternative method is required for facilitating the optimal leads in early drug discovery process. Virtual screening strategies have also been successfully applied during the past decades in screening large chemical libraries with aim to identify the potential bioactive hits. Such techniques are cost-effective and reliable that can be applied to identify potential leads with improved hits rates.

Conclusion: In the present review, we have described some of the computational techniques that can be used for designing potential drug candidates against the multiple drug resistant strains. At the end some of the recent successful studies were also discussed.

Keywords: virtual screening, antibiotic resistance, docking, drug designing

Volume 2 Issue 2 - 2015

Mohd Danishuddin, Asad U Khan

Interdisciplinary Biotechnology Unit, Aligarh Muslim University, India

Correspondence: Asad U Khan, Interdisciplinary Biotechnology Unit, A.M.U., Aligarh 202002, India, Tel 00919837021912, Email asad.k@rediffmail.com

Received: December 19, 2014 | **Published:** March 24, 2015

Abbreviations: NDM-1, new Delhi metallo- β -lactamase-1; PBPs, penicillin binding proteins; MT, *Mycobacterium tuberculosis*; MDR, multidrug-resistant; WHO, world health organization; HTS, high throughput screening; EF, enrichment factor; SBVS, structure-based virtual screening; LBSV, ligand based virtual screening; QSAR, quantitative structure-activity relationships

Introduction

Bacterial resistance to major classes of antibacterial agents is becoming an increasing problem worldwide.¹⁻⁴ Gram-negative bacteria are strongly associated with the production of both chromosomal- and plasmid-encoded beta-lactamases. These enzymes exhibit broad-spectrum hydrolytic activity against most classes of beta-lactams. One of the most recently reported carbapenemase, New Delhi Metallo- β -lactamase-1 (NDM-1), producers have become resistant against almost all antibiotics.^{1,5} An alternative way to combat with this situation is to use the effective inhibitors in clinical settings with combination of different antibiotics.⁶ These inhibitors inhibit the beta-lactamases so that antibiotics can reach to penicillin binding proteins (PBPs), the target of beta-lactam antibiotics. Presently, three inhibitors named as clavulanic acid, sulbactam and tazobactam are used in combinations with different antibiotics to restore their potency. However, with several years of introduction, resistance of these inhibitors was observed against the bacterial strains in recent years^{7,8} and has become another challenge to treat serious infections in clinical settings.⁹ Inhibitor resistance poses a significant threat to many current antibiotic combinations, used against the beta-lactamase enzymes. Similarly, the emergence of multidrug-resistant (MDR) *Mycobacterium tuberculosis* (MT) strains is alarming and represents a worldwide health care problem.¹⁰ These resistances are attributed primarily to the accumulation of mutations in the drug target genes, lead either to an altered target.¹¹ A large number of mutations among clinical isolates have been identified that confer resistance against life saving drugs like, isoniazid, fluoroquinolones and rifampicin.¹²

Recently, world health organization (WHO) has reported that antibiotics resistance in bacteria become a worldwide threat to public health. For example, resistance to one of the most commonly used antibiotic (fluoroquinolones) used for the oral treatment of urinary tract infections caused by *E.coli* is getting common. Moreover, resistance against last resort of antibiotic for life threatening infections caused by common intestinal bacteria has spread to all regions of the world.^{13,14} Treatment failure to the last resort of treatment for gonorrhoea—third generation cephalosporins also been confirmed in Austria, Australia, Canada, France, Japan, Norway, Slovenia, South Africa, Sweden and the United Kingdom (WHO report 2014). Thus, there is urgent need to check the prevalence of resistant strains by development of new specific mechanism-based inhibitors. However, to bring a new drug to the market takes approximately 10-14 years to get approval of a single drug from FDA (Food and Drug Administration, USA) before it comes in the market with the expense of ~1 billion USD.^{15,16} Traditionally, high throughput screening (HTS) methods are routinely employed to screen hundreds of thousands of molecules against the targets for their biological activity.^{17,18} However, the success rate of this approach is very low as only few compounds screened from library were found to pass through drug development phases. Recently, computational approaches have been appreciated and become a crucial component of many drug discovery programmes. Their applications span almost all stages in the discovery and development pipeline from hit identification to lead optimization. These techniques are well-established tools in the modern drug discovery process and proven to be more efficient than the traditional way of drug discovery. One key approach structure based virtual screening remains a highly active area of research during past decade.¹⁹⁻²¹ In this article we have discussed the role of virtual screening approaches such as structure based and ligand based that may be used for the identification of novel specific inhibitors against the drug resistance bacteria or virus strains. Finally, some of the case studies that have successfully applied these strategies were also discussed.

Virtual screening

Virtual screening (VS) is a powerful technique for identifying hits as starting point for medicinal chemistry.²² It can be differentiated into three main approaches; a) structure-based virtual screening which requires knowledge of the 3D structure of the target; b) ligand-based virtual screening which requires knowledge of active ligands; and c) hybrid approach. Numbers of methods which use the ligand and target-based VS approaches are increasing at a rapid pace (Table 1).^{23–26} These methods are different in their respective docking algorithms, scoring functions as well as supported platforms.^{27–29} These methods have been examined for their docking, virtual screening and their capabilities for reproducing the experimental poses.^{30,31} A rational approach in early drug discovery is an appropriate selection of the disease target, mostly a protein or enzyme. The completion of the human genome project has provided with an abundance of potential

new drug targets and has a significant impact on the process of drug development. Another important requirement for VS is the preparation of compound library. From the standpoint of exploring chemical space, compound databases have provided the platform for several options, helpful in target specific compounds selection. Chemical space is understood as the ensemble of all organic molecules in the context of drug discovery and comprises millions of known molecules collected in public databases such as PubChem,³² CoCoCo,³³ ZINC³⁴ and DrugBank.³⁵ These databases contain the vast amount diverse molecules for library design and easier to optimize toward the identification of promising lead compounds. To overcome the limitations of structure and ligand based virtual screening there has been increasing interest in combining these methods as noticed in recent years.^{36,37} Such integrated or hybrid approaches address the exploitation of all available structural and chemical information in the search for new molecules and holds significant potential.

Table 1 Some of the available tools for protein-ligand docking. A not determined

S. no	Software	Algorithm	Reference	Citationa
Free software's				
1	SwissDock	Evolutionary algorithm	27	88
2	AutoDock Vina	Genetic algorithm	28	1741
3	AutoDock	Genetic algorithm	29	6098
4	DOCK	Lamarckian and Simulated annealing Shape Fitting	78	ND
Commercials				
7	GOLD	Genetic Algorithm	23	ND
8	LigandFit	Monte Carlo	26	626
9	Glide	Monte Carlo	76	2140
10	FlexX	Incremental construction	77	ND

Structure based virtual screening (SBVS)

Structure-based virtual screening (SBVS) has been widely applied in early-stage drug discovery.³⁸ For any successful outcome this technique can follow the steps as listed in Figure 1. Target identification or selection is the first vital step in the drug discovery pipeline. Improper selection of target may lead to false hits in the later stages of drug development. An increase of X-ray crystallography and nuclear magnetic resonance spectroscopy structures of targets in Protein Data Bank (PDB)³⁹ has further opened doors to structure-based virtual screening. One can find the in-depth structural details of these targets as well as their interactions with ligands in PDB. Target structure, either experimentally solved or homology modeled, and a library of small molecules are basic requirements for any SBVS protocol. Homology or comparative modeling is a technique use to predict protein structure based on the observation that proteins with similar sequences have similar structures. The process of homology or comparative modeling consists of the following steps: (1) Identification of known 3D structure(s) of a related protein with reasonable identity or similarity that can serve as template (2) Alignment of target and template protein sequences (3) model building (4) validation of the generated models.⁴⁰ The most crucial step in theoretical or computational modelling is the search of suitable template. The success of any modeled structure depends on this template. Structure that are >35% identical with the query sequence is better for the three dimensional modelling. Molecular docking is one of the most commonly used methods in structural

based virtual screening approach. It consists of sampling the ligand pose within the binding site of the target to form a stable complex. Significant outcomes of any docking depends highly on the capability of method in reproducing experimentally solved ligand poses and scoring functions to rank the binding poses. Scoring functions are fast approximate mathematical methods used to predict the binding affinity in between two molecules. Three types of scoring functions are currently applied in currently available docking programs. These scoring functions are based on: force field,^{41,42} empirical approaches^{43,44} and knowledge based.^{45,46} Discriminating or predictive power of these functions have high influence on an outcome of any docking protocol. An ideal scoring function has potential to generate the ranking for ligand set to separate active compounds from inactive ones. The performance of these scoring functions can be measured by using the matrices like ROC, AUC and Enrichment Factor (EF), which are frequently used to evaluate the performance of the ranking methods in virtual screening process.⁴⁷ A ROC curve is a graphical plot created by plotting the fraction of the true positive rate (TPR or Sensitivity) versus the false positive rate (FPR, or 1-specificity) at various thresholds. The AUC-ROC (Area Under Curve-ROC) can be used to quantify the overall quality of the plot. It is a way to measure that how randomly selected active molecules are ranked compared to randomly chosen inactive molecules. A perfect value will result in an area under the curve of 1, while a random scoring function will have an ROC-AUC of 0.5.^{48,49} However, even with the advancement and vast amount of methods available, we still lack the universal algorithm or method that can enhanced the performance of docking

based virtual screening performance. The main obstacles faced by docking and scoring functions are the receptor flexibility, solvation and entropic contributions. For proper fitting the ligand the active site of target often adapts some conformational changes upon ligand binding.⁵⁰ Such conformational changes have high impact on DBVS (docking based virtual screening) and still represents one of the greatest challenges for current docking programs. However, proper consideration of target flexibility can improve DBVS outcomes.⁵¹ Such challenges need to be address properly during the virtual screening protocols.

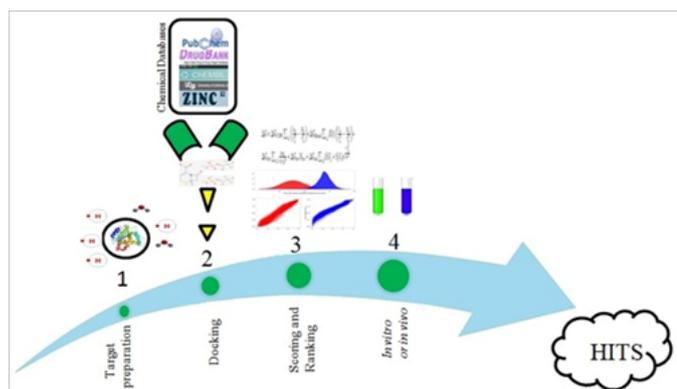


Figure 1 Depicts the steps involved in structure based virtual screening protocol.

Ligand based virtual screening (LBSV)

Ligand based virtual screening approaches utilize structure-activity data from a set of known actives molecules in order to identify likelihood drug candidates for experimental confirmation.⁵² Quantitative structure-activity relationships (QSAR), pharmacophore modeling, similarity or substructure searching and three-dimensional shape matching are some of the strategies that have been utilized in LBSV method. Quantitative Structure Activity Relationship (QSAR) is one of the frequently used approach in ligand based virtual screening. Generally, QSAR is used to study the structural or physiochemical relationship of active molecules with their biological targets.^{53–55} High quality data, diverse set compounds, appropriate descriptors, suitable mathematical algorithm and proper validation sets are required for the development of any effective and successful QSAR model. There are certain reports that showed the influence of these features on any model development.^{56–58} However, despite these challenges, these models are still prefer as they reduced the time, cost and false hit rates for any designed biological assay. Machine learning algorithms are among the most popular tools used to perform a robust and quantitative structure activity relationship modeling. These techniques applied to QSAR modeling are not only useful for virtual screening but also play an important role in predicting the parameters of pharmacological and pharmaceutical relevance. Different machine learning methods have been proposed with its own advantages and disadvantages. Some of these methods named as, Neural Network,⁵⁹ Support Vector Machines,⁶⁰ PLS⁶¹ and Decision Tree Classification.⁶² The use of these techniques in the chemistry field has increased in the last decades.^{63–66} They are applied for the calculation of the optimal distance between the descriptors of active and inactive compounds. The models developed by these algorithms have potential to discriminate the biologically active compounds from the inactive compounds for their likelihood of interacting with the target. These techniques are helpful in developing the effective prediction models and to discover

the optimal decision for your problem. The output may depend on the size of the dataset which is known to be the major drawback of these methods. Adequate amount of data size is necessary to get the optimal output. Another, major drawback of ligand based virtual screening is its dependency on existing ligands information as templates which limits the scaffold diversity as comparison to structure based virtual screening.

Hybrid approaches

The main challenge in drug discovery is not only to find the hits, but to produce more potent and selective compounds with desired biological activity. Currently available ligand and structure based virtual screening approaches faced several issues associated with their outcomes. Combination of these two methods can also be in order to enhance⁶⁷ the performance of the integrated protocol. However, not all of these combinations are successful. Proper choice of the methods for their integration have high impact on any significant outcome. Various methods for example; pharmacophore methods, fingerprint methods and approaches integrating docking with similarity-based methods are available that can be integrated with each others. Pharmacophore features represent the information from experimental solved proteins, ligands or protein–ligand complexes. The observed protein–ligand interactions are directly translated into pharmacophore features. However, in lack of any experimental solved protein–ligand complexes, pharmacophore model can be derived from docking into homology models or experimental solved protein structures.^{67,68} These protein–ligand pharmacophore models have been successfully applied in virtual screening process.⁶⁹ The interest in using these hybrid protocols is evident in the publications in past few years. Qing et al.,⁷⁰ proposed a hybrid virtual screening approach based on SVM classification and pharmacophore models for the identification of potential C5aR antagonists. The proposed hybrid approach showed the significantly increased in the hit rate and enrichment factor as compared with the individual method. Despite this, there are others several successful studies, where the hybrid protocol has showed the potential in significant number of outcomes. Complete presentation of these approaches is beyond the scope of this review, however herein, we have discussed some of the protocols successfully implemented in recent studies. By combining the molecular docking, pharmacophore and fingerprint-based 2D similarity, a novel two layer workflow was developed to enhance the virtual screening performance.⁷¹ By combining these methods the authors noticed the improvement in the performance of virtual screening process on DUD database. Recently, a hybrid protocol includes support vector machine, pharmacophore and docking-based virtual screening was used for retrieving the new Bruton's tyrosine kinase inhibitors from commercially available chemical databases.⁷² Performances of this hybrid approach showed the considerably increased the hit rate and enrichment factor by shortened the overall screening time as compared with the individual method or their combinations. This approach was used to screen several chemical databases including Specs (202,408 compounds) and Enamine (980,000 compounds) databases, from which thirty nine compounds were finally selected for experimental evaluation

Successful studies

Case study I

Two different virtual-screening approaches were applied in order to identify novel inhibitors for the *Mycobacterium tuberculosis* InhA (MtInhA) target.⁷³ First, a 3-D pharmacophore model was built based

on 36 available MtInhA crystal structures. The second approach consisted of using four well established docking programs, with different search algorithms, to compare the binding mode and score of the selected molecules from the aforementioned library. Finally, six ligands were selected for *in vitro* analysis. Three of these molecules presented a satisfactory inhibitory activity with IC₅₀ values in between the 24 μM to 83 μM. The best compound presented an uncompetitive inhibition mode to NADH and 2-trans-dodecenoyl-CoA substrates, with Ki values of 24 μM and 20 μM respectively.

Case study 2

Nan Li et al.,⁷⁴ screened the SPECS compound library by using *structural based virtual screening* to identify the potential inhibitors of the histidine kinase (HK) VicK protein from *S. pneumoniae*. Based on molecular diversity, shape complementarities, and the potential to form hydrogen bonds and hydrophobic interactions in the binding pocket, they identified a series of 105 compounds. Six of them were then validated *in vitro* and were found to inhibit the growth of *S. pneumoniae*. These compounds were found potential in decreasing the mortality of the mice infected by *S. pneumoniae*. Finally, the authors reported that these compounds were the first inhibitors of HK with antibacterial activity *in vitro* and *in vivo*, and were novel lead to be drugs that can helpful to combat pneumococcal infection.

Case study 3

A novel hybrid virtual screening protocol (comprised of pharmacophore based virtual screening, rigid and flexible docking based virtual screenings) was applied to identify the potential inhibitors for the bacterial enzyme named, Methionyl-tRNA Synthetase.⁷⁵ Specs (202,408 compounds) database was used for virtual screening process. Based on the outcome of this hybrid protocol, 15 hits were selected for the experimental studies. After experimental evaluation, two compounds with novel scaffold were reported.^{76–78}

Conclusion

Despite the advancement that drawn the positive impact on structure or ligand based virtual screening approaches, challenges associated with these methods remained unsolved. We are still lacking a cutting edge technology that can helpful in predicting the accurate binding pose of the compounds. However, careful database preparation, judicious methods choices with optimized parameters, use of proper positive controls can increases probability of any virtual screening experiments success. In this article, we have discussed some of the recent developments in virtual screening methods that can be utilized to find the potential leads molecules against the multidrug resistance strains. We have also highlighted the challenges associated with the VS techniques. At last some of the case studies applied these methods have also been discussed. We believed that these studies and their reported successful outcomes can play a significant role in planning the future experiments.

Acknowledgements

Author acknowledges DBT, government of India for the support and internal facilities of the department. This work was supported by internal funds of Biotechnology Unit, AMU and DBT grant, BT/PR11453/BID/07/296/2009, BT/HRD/NBA/34/01/2012 to AUK and BSR (Basic Science Research) Fellowship to Danishuddin.

Conflict of interest

The author declares no conflict of interest.

References

1. Khan AU, Nordmann P. Spread of carbapenemase NDM-1 producers: The situation in India and what may be proposed. *Scand J Infect Dis*. 2012;44:531–535.
2. Xia S, Fan X, Huang Z, et al. Dominance of CTX-M-type extended-spectrum β-lactamase (ESBL)-producing *Escherichia coli* isolated from patients with community-onset and hospital-onset infection in China. *PLoS One*. 2014;9(7):e100707.
3. Bush K. Alarming β-lactamase-mediated resistance in multidrug-resistant *Enterobacteriaceae*. *Curr Opin Microbiol*. 2010;13(5):558–564.
4. Kapil A. The challenge of antibiotic resistance: need to contemplate. *Indian J Med Res*. 2005;121(2):83–91.
5. Khan AU, Parvez S. Detection of blaNDM-4 in *Escherichia coli* from hospital sewage. *J Med Microbiol*. 2014;63:1404–1406.
6. Hickman RA, Hughes D, Cars T, et al. Cell-wall-inhibiting antibiotic combinations with activity against multidrug-resistant *Klebsiella pneumoniae* and *Escherichia coli*. *Clin Microbiol Infect*. 2014;20(4):267–273.
7. Papp-Wallace KM, Bethel CR, Distler AM, et al. Inhibitor resistance in the KPC-2 beta-lactamase, a preeminent property of this class A beta-lactamase. *Antimicrob Agents Chemother*. 2010;54(2):890–897.
8. Manageiro V, Ferreira E, Coughnoux A, et al. Characterization of the inhibitor-resistant SHV β-Lactamase SHV-107 in a clinical *Klebsiella pneumoniae* strain coproducing GES-7 enzyme. *Antimicrob Agents Chemother*. 2012;56(2):1042–1046.
9. Shlaes D M. New β-lactam-β-lactamase inhibitor combinations in clinical development. *Ann NY Acad Sci*. 2013;1277:105–114.
10. Meacci F, Orru G, Iona E, et al. Drug Resistance Evolution of a *Mycobacterium tuberculosis* Strain from a Noncompliant patient. *J Clin Microbiol*. 2005;43(7):3114–3120.
11. Ballif M, Harino P, Ley S, et al. Drug resistance-conferring mutations in *Mycobacterium tuberculosis* from Madang, Papua New Guinea. *BMC Microbiol*. 2012;12:191.
12. Rosales-Klitz S, Jureen P, Zalutskaya A, et al. Drug resistance-related mutations in multidrug-resistant *Mycobacterium tuberculosis* isolates from diverse geographical regions. *Int J Mycobacteriol*. 2012;1(3):124–130.
13. WHO (World Health Organization) Report 2014.
14. Khan AU, Nordmann P. NDM-1-producing *Enterobacter cloacae* and *Klebsiella pneumoniae* from diabetic foot ulcers in India. *J Med Microbiol*. 2012;61(Pt 3):454–456.
15. Xiang M, Cao Y, Fan W, et al. Computer-aided drug design: lead discovery and optimization. *Comb Chem High Throughput Screen*. 2012;15(4):328–337.
16. Macarron R, Banks MN, Bojanic D, et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov*. 2011;10(3):188–195.
17. Ou-Yang SS, Lu JY, Kong XQ, et al. Computational drug discovery. *Acta Pharmacol Sin*. 2012;33(9):1131–1140.
18. Khan SN, Islam B, Yennamalli R, et al. Characterization of doxorubicin binding site and drug induced alteration in the functionally important structural state of oxyhemoglobin. *J Pharm Biomed Anal*. 2008;48(4):1096–1104.
19. Zhuang C, Narayanapillai S, Zhang W, et al. Rapid identification of keap1-nrf2 small-molecule inhibitors through structure-based virtual screening and hit-based substructure search. *J Med Chem*. 2014;57(3):1121–1126.

20. Evers A, Klabunde T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J Med Chem.* 2005;48(4):1088–1097.
21. Good A. Structure-based virtual screening protocols. *Curr Opin Drug Disco Devel.* 2001;4(3):301–307.
22. Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem.* 2013;20(23):2839–2860.
23. Jones G, Willett P, Glen R C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol.* 1995;245(1):43–53.
24. Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol.* 1996;261(3):470–489.
25. Makino S, Kuntz ID. Automated flexible ligand docking method and its application for database search. *J Comput Chem.* 1997;18:1812–1825.
26. Venkatachalam CM, Jiang X, Oldfield T, et al. Ligand Fit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graphics Model.* 2003;21(4):289–307.
27. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 2011;39(Web server issue):W270–W277.
28. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455–461.
29. Morris GM, Goodsell DS, Halliday, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry.* 1998;19(14):1639–1662.
30. Kellenberger E, Rodrigo J, Muller P, et al. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins.* 2004;57(2):225–242.
31. Warren GL, Andrews CW, Capelli AM, et al. A critical assessment of docking programs and scoring functions. *J Med Chem.* 2006;49(20):5912–5931.
32. Wang Y, Xiao J, Suzek TO, et al. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* 2009;37:W623–W633.
33. Del Rio A, Barbosa AJ, Caporuscio F, et al. CoCoCo: a free suite of multi conformational chemical databases for high throughput virtual screening purposes. *Mol Biosyst.* 2010;6(11):2122–2128.
34. Irwin JJ, Sterling T, Mysinger MM, et al. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model.* 2012;52(7):1757–1768.
35. Wishart DS, Knox C, Guo AC, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 2008;36(Database issue):D901–D906.
36. McGaughey G, Sheridan R, Bayly C, et al. Comparison of topological, shape, and docking methods in virtual screening. *J Chem Inf Model.* 2007;47(4):1504–1519.
37. Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem.* 2013;20(23):2839–2860.
38. Lionta E, Spyrou G, Vassilatis DK, et al. Structure-based virtual screening for drug discovery: Principles, applications and recent advances. *Curr Top Med Chem.* 2014;14(16):1923–1938.
39. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235–242.
40. Martí-Renom MA, Stuart AC, Fiser A, et al. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct.* 2000;29:291–325.
41. Ewing TDJ, Kuntz DI. Critical evaluation of search algorithms for automated molecular docking and database screening. *J Comput Chem.* 1997;18(9):1175–1189.
42. Kramer B, Rarey M, Lengauer T. Evaluating of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins.* 1999;37(2):228–241.
43. Sotriffer CA, Sanschagrin P, Matter H, et al. SFCscore: Scoring functions for affinity prediction of protein-ligand complexes. *Proteins.* 2008;73(2):395–419.
44. Bohm HJ. The Development of a Simple Empirical Scoring Function to Estimate the Binding Constant for a Protein-ligand Complex of Known three-dimensional Structure. *J Comput Aided Mol Des.* 1994;8(3):243–256.
45. Sippl MJ. Calculation of Conformational Ensembles from Potentials of Mean Force. An approach to the Knowledge-based Prediction of Local Structures in Globular Proteins. *J Mol Biol.* 1990;213(4):859–883.
46. Thomas PD, Dill KA. Statistical potentials extracted from protein structures: how accurate are they? *J Mol Biol.* 1996;257(2):457–469.
47. Danishuddin M, Khan AU. Structure based virtual screening to discover putative drug candidates: Necessary considerations and successful case studies. *Methods.* 2015;71:135–145.
48. Truchon JF, Bayly CI. Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J Chem Inf Model.* 2007;47(2):488–508.
49. Zhao W, Hevener KE, White SW, et al. A statistical framework to evaluate virtual screening. *BMC Bioinformatics.* 2009;10:225.
50. Cheng T, Li Q, Zhou Z, et al. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *AAPS J.* 2012;14(1):133–141.
51. Rueda M, Bottegoni G, Abagyan R. Consistent improvement of cross-docking results using binding site ensembles generated with elastic network normal modes. *J Chem Inf Model.* 2009;49(3):716–725.
52. Venkatraman V, Pérez-Nueno VI, Mavridis L, et al. Comprehensive Comparison of Ligand-Based Virtual Screening Tools Against the DUD Data set Reveals Limitations of Current 3D Methods. *J Chem Inf Model.* 2010;50(12):2079–2093.
53. Myint K Z, Xie XQ. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. *Int J Mol Sci.* 2010;11(10):3846–3866.
54. Kruhlik NL, Benz RD, Zhou H, et al. (Q)SAR modeling and safety assessment in regulatory review. *Clin Pharmacol Ther.* 2012;91(3):529–534.
55. Esposito EX, Hopfinger AJ, Madura JD. Methods for applying the quantitative structure-activity relationship paradigm. *Methods Mol Biol.* 2004;275:131–214.
56. Gozalbes R, Doucet JP, Derouin F. Application of topological descriptors in QSAR and drug design: history and new trends. *Curr Drug Targets Infect Disord.* 2002;2(1):93–102.
57. Golbraikh A, Shen M, Xiao Z, et al. Rational selection of training sets for the development of validated QSAR models. *J Comput Aided Mol Des.* 2003;17(2–4):241–253.
58. Tropsha A, Gramatica P, Gombar, et al. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR model. *QSAR Comb Sci.* 2003;22:69–76.
59. Winkler DA. Neural networks as robust tools in drug lead discovery and development. *Mol Biotechnol.* 2004;27(2):138–168.
60. Burbidge R, Trotter M, Buxton B, et al. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput Chem.* 2001;26(1):5–14.

61. Dayal BS, MacGregor JF. Improved PLS algorithms. *J Chemometrics*. 1997;11:73–85.
62. Breiman L, Friedman J, Olshen R, et al. *Classification and Regression Trees*. CA: Wadsworth; 1984.
63. Burbidge R, Trotter M, Buxton B, et al. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput Chem*. 2001;26(1):5–14.
64. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5–32.
65. Salt DW, Yildiz N, Livingstone DJ, et al. The Use of Artificial Neural Networks in QSAR. *Pestic Sci*. 1992;36:161–170.
66. Liew CY, Pan C, Tan A, et al. QSAR classification of metabolic activation of chemicals into covalently reactive species. *Mol Divers*. 2012;16(2):389–400.
67. Drwal, MN, Griffith R. Combination of ligand- and structure-based methods in virtual screening. *Drug Discov Today Technol*. 2013;10(3):395–401.
68. Wilson, GL, Lill MA. Integrating structure-based and ligand-based approaches for computational drug design. *Future Med Chem*. 2011;3(6):735–750.
69. Arooj M, Sakkiah S, Kim S, et al. A combination of receptor-based pharmacophore modeling & QM techniques for identification of human chymase inhibitors. *PLoS One*. 2013;8(4):e63030.
70. Qing XY, Zhang CH, Lib LL, et al. Retrieving novel C5aR antagonists using a hybrid ligand-based virtual screening protocol based on SVM classification and pharmacophore models. *J Biomol Struct Dyn*. 2013;31(2):215–223.
71. Zhang Y, Yang S, Jiao Y, et al. An integrated virtual screening approach for VEGFR-2 inhibitors. *J Chem Inf Model*. 2013;53(12):3163–3177.
72. Wan HL, Wang ZR, Li LL, et al. Discovery of novel Bruton's tyrosine kinase inhibitors using a hybrid protocol of virtual screening approaches based on SVM model, pharmacophore and molecular docking. *Chem Biol Drug Des*. 2012;80(3):366–373.
73. Pauli I, dos Santos RN, Rostirolla DC, et al. Discovery of new inhibitors of Mycobacterium tuberculosis InhA enzyme using virtual screening and a 3D-pharmacophore-based approach. *J Chem Inf Model*. 2013;53(9):2390–2401.
74. Li N, Wang F, Niu S, et al. Discovery of novel inhibitors of Streptococcus pneumoniae based on the virtual screening with the homology-modeled structure of histidine kinase (VicK). *BMC Microbiol*. 2009;9:129.
75. Liu C, He G, Jiang Q, et al. Novel Hybrid Virtual Screening Protocol Based on Molecular Docking and Structure-Based Pharmacophore for Discovery of Methionyl-tRNA Synthetase Inhibitors as Antibacterial Agents. *Int J Mol Sci*. 2013;14:14225–14239.
76. Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid accurate docking and scoring. I Method and assessment of docking accuracy. *J Med Chem*. 2004;47(7):1739–1749.
77. Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*. 1996;261(3):470–489.
78. Moustakas DT, Lang PT, Pegg S, et al. Development and validation of a modular, extensible docking program: DOCK 5. *J Comput Aided Mol Des*. 2006;20(10–11):601–619.