Annals of Medicinal Chemistry

Research Article
Published: 26 Mar, 2019

Development of Consensus Scoring Functions for the Identification of Novel Beta-Lactamases Inhibitors by Virtual Screening

*Vieira TF, Magalhaes RP and Sousa SF**

Department of Biomedicine, University of Porto, Portugal

Abstract

Docking and Virtual Screening (VS) are important tools for identifying novel drug candidates that possess antagonistic activity for several enzymes or protein receptors of medical or biological importance. Docking predicts and ranks the conformations of a ligand in a specific target and VS is used to explore large virtual databases containing millions of molecules.

In this work, we report the optimization of a virtual screening protocol for the inhibition of β-lactamases. These are a class of enzymes that are associated with the development of resistance to β-lactam antibiotics. The results highlight the strengths and weaknesses of different scoring functions (AutoDock, Auto Dock Vina, GoldScore, ChemScore, CHEMPLP, and ASP) when dealing with this type of target. This allows us to perform large-scale VS campaigns in the search for novel inhibitors of β-lactamases. In general, the results have shown that the overall performance of the different scoring functions is comparable in discriminating between actives and decoys. However, the results can vary significantly with the type of target and ligand. ChemScore gives considerable better results, followed by Vina.

Introduction

Protein-ligand docking is a computational method that tries to predict and rank the conformation and orientation (also known as pose) of a ligand, typically a small organic molecule, within a specific binding site of a protein, normally an enzyme or receptor of therapeutic relevance [1-8]. It is often used alongside Virtual Screening (VS) as an integral part of the drug design and development process [1,5,9]. The goal of virtual screening is to evaluate large databases of molecular compounds to generate a much smaller set of potentially active drug candidates or lead compounds that can be later evaluated experimentally [10].

Despite its limitations, molecular docking is a quick and computationally non-expensive way to predict how a specific ligand can bind to a protein in the absence of an X-ray or NMR structure of the resulting complex. In the context of VS can also be used to discriminate ligands that bind or that are more likely to bind to a specific target from those that do not bind or that are less likely to bind. However, protein-docking is always a multistep process. On a first step, the search algorithm explores the different conformations and orientations of the ligand within the target. Then, the scoring function tries to identify the correct binding mode, evaluating the different possibilities [4,11].

Docking algorithms and software have been continually improving but obtaining accurate binding modes is still a challenging task [2,9]. The scoring function is still the limiting step in protein-ligand docking as it is often simplified and does not express the true complexity of protein-ligand interactions [12]. For a scoring function to weight all aspects involved in ligand binding, the computational costs would have to increase, making it an extremely time-consuming task particularly in a context of VS of large databases, due to the analysis of several binding modes required per compound [13].

The limitations in the scoring functions may affect the accuracy of docking results for a specific target but are normally even more significant in the context of VS. In VS, millions of compounds are docked against a given target; scoring metrics are applied to rank the compounds with the highest affinity towards the protein target. The result is normally a selection of 100 to 500 molecules that eventually proceed to experimental testing. VS is, therefore, an important tool in drug design as it

OPEN ACCESS

Correspondence:

Sousa SF, Department of Biomedicine, University of Porto, Portugal, E-mail: sergiofsousa@med.up.pt

Received Date: *27 Feb 2019*

Accepted Date: *22 Mar 2019*

Published Date: *26 Mar 2019*

Citation:

Vieira TF, Magalhaes RP, Sousa SF. Development of Consensus Scoring Functions for the Identification of Novel Beta-Lactamases Inhibitors by Virtual Screening. Ann Med Chem. 2019; 1(1): 1002.

limits the number of compounds that are experimentally evaluated [10,13]. Besides, the cost of performing screens computationally is significantly lower than using experimental procedures.

Due to the lack of accuracy in the scoring functions, the top solutions of a typical virtual screening campaign include a large number of compounds that when tested experimentally fail to bind strongly to the target (i.e. false positives). A worse problem is that normally many molecules that are strong ligands, fail to be identified as such in docking (i.e. false negatives), remaining in cognito among millions of other compounds in the database. This is particularly challenging because possible hit compounds may never reach the preliminary experimental studies, despite their high pharmacological potential [14]. Several strategies to minimize this problem have been described in the literature [5,10,13,15,16]. In general, however, it has been shown that the accuracy of different docking programs and scoring functions depends on the specific protein target under study and type of ligands. In fact, binding pockets in different proteins vary in terms of size, polarity, solvent accessibility, composition of amino acid residues and most notably number, type and orientation of specific interaction groups available [17]. On the other hand, ligand can differ appreciably in terms of size, length, polarity and hydrophobicity, number of rotatable bonds, number of possible hydrogen donors or acceptors, etc. The combination of these features results in that very different challenges can be offered to the docking program, and that the best solution to tackle one problem is not necessarily the best solution to handle another problem [17]. Hence, it has become necessary, when starting a VS campaign anchored in protein-ligand docking to assess the accuracy of different protein-ligand docking programs and scoring functions against that specific type of protein-target [17].

β-lactamases

β-lactamases area class of enzymes (EC3.5.2.6) produced by gram-negative bacteria responsible for the most widespread resistance mechanisms to β-lactam antibiotics such as penicillins, cephalosporins, cephamycins, and carbapenems. To inactivate the action of antibiotics, these enzymes break their structure, by hydrolyzing a four-atom ring, called the β-lactam. Breaking this ring open leads to a loss of antibacterial properties followed by antibiotic resistance [18,19].

Antibiotic resistance is an increasing threat in public heath, with high morbidity and mortality rates, mainly due to the inappropriate prescription and widespread use of antibiotic agents. Therefore, finding strategies against this resistance is of the outmost importance [20].

The first way that the pharmaceutical industry tried to overcome antibiotic resistance, consisted in the use of β-lactamase-stable antibiotics (i.e. antibiotic agents resistant to hydrolysis) and the concomitant use of selective β-lactamase inhibitors (BLIs) alongside β-lactam antibiotics (with the example of the association between amoxycillin and clavulanate, which was a success). However, bacteria are continuously evolving and adapting and are becoming resistant to the β-lactamase inhibitors as well, leading to Multi-Drug Resistant (MDR) and extensively-drug resistance (XDR) strains [21].

Bacteria are also producing, carbapenemases, enzymes that destroy almost all β-lactams. In this study, we focused on β-lactamases, but we also performed experimental and computational studies on carbapenem resistance [22].

New β-lactamase inhibitors are required. Virtual screening can be used to identify novel promising molecules for future experimental testing. However, to ensure accurate results it very important to careful understands the best scoring functions to treat this important therapeutic target.

So, in this study, six scoring functions from three of the most commonly used docking tools AutoDock4.2.6 [23], VINA [24] and GOLD [25] - were evaluated in terms of their ability to discriminate between real β-lactamase inhibitors from other molecules with similar physico-chemical properties that are non-binders (decoys). More details about the scoring functions used are described in Table 1 [23-28].

Autodock 4 is a docking program developed by Morris & co-workers at the Scripps Research Institute [23]. It is a very popular choice between users due to the good accuracy and high versatility shown. It is also often a first choice because it is free to academic users, which has contributed to its widespread use and high number of citations [2]. Autodock 4 offers a variety of search algorithms and a scoring function that is based on a linear regression analysis, the AMBER force field, and a large set of diverse protein-ligand complexes with known inhibition constants. The program can be used with a visual interface called Auto Dock Tools (ADT) which ensures an efficient analysis of the docking results.

Auto dock Vina is another popular docking program, developed by Trott & Olson at the Scripps Research Institute, La Jolla, California. It follows the success of previous Auto dock versions [24]. Vina is open source software and it is also very popular among users. It inherits some of the ideas and approaches of Auto Dock 4, but it features a new search algorithm and a hybrid scoring function, combining empirical and knowledge-based scoring function. It offers significant improvements in the average accuracy of the binding mode predictions. It is faster than Auto Dock and it is easy to use. Those are some of the reasons why it has become very popular.

GOLD is docking software created by collaboration between the University of Sheffield, GlaxoSmithKline and the Cambridge Crystallographic Data Centre (CCDC) and it is commercially available [25]. The program contains a Genetic Algorithm (GA) based search method; it allows full ligand flexibility, while ensuring partial protein flexibility, through protein side-chain and backbone flexibility for up to a maximum of ten user-defined residues. The GOLD scoring functions (GoldScore, ChemScore, the Astex Statistical Potential (ASP) and ChemPLP) are dimensionless and incorporate different factors and terms regarding H-bonding energy, free energy and ligand flexibility. However, for each of the scoring functions, the scale of the score gives a guide as to how good the pose is; the higher the score, the better the docking result is likely to be [27].

The results obtained here can provide important clues for future work. Namely when preparing extensive virtual screening studies to identify novel β-lactamase inhibitors for subsequent experimental testing.

Materials and Methods

Target selection, preparation and docking

One AmpC β-lactamase X-ray crystal structure was chosen from the PDB databank [29], as a starting point (1L2S) [30]. The structure is in complex with a non-covalent inhibitor allowing the identification and characterization of the binding pocket. After inspection of the

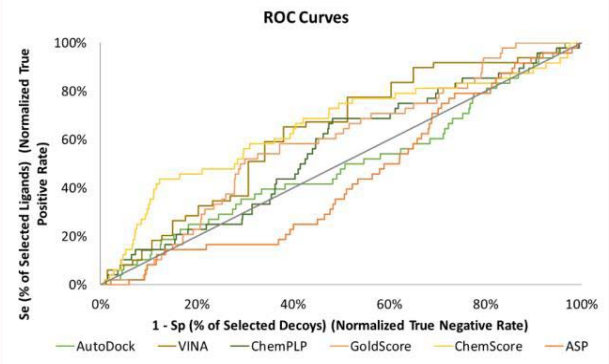

Figure 1: ROC Plots for the different scoring functions.

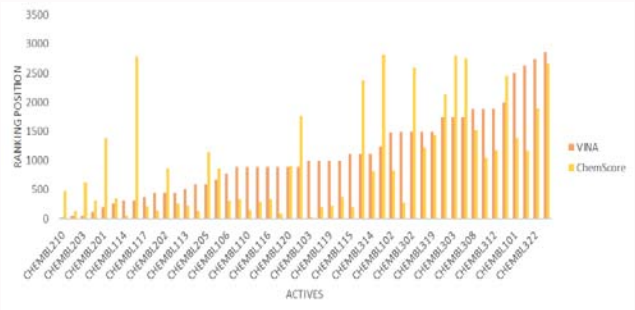

Figure 2: Bar chart showing the ranking position of all the active molecules obtained from Vina and ChemScore Virtual Screening assays.

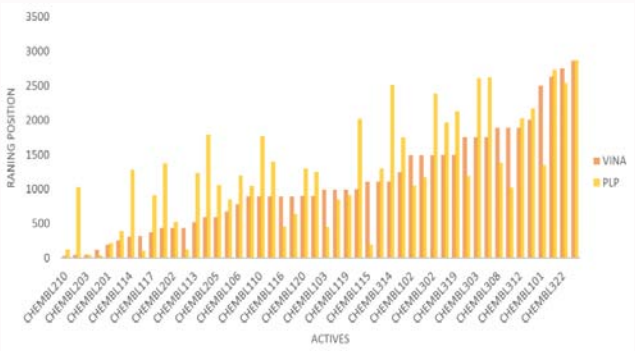

Figure 3: Bar chart showing the ranking position of all the active molecules obtained from Vina and CHEMPLP Virtual Screening assays.

target, the ligand was re-docked with AutoDock, Vina and Gold and the docking parameters were optimized to reproduce the known experimental binding poses.

For Vina, the parameters adjusted in the docking protocol included the box size and position, exhaustiveness and number of generated binding modes. For AutoDock, the parameters optimized also included the box size and position, as well as the number of GA runs, number of grid points and spacing, maximum number of generations, maximum number of energy evaluations and population size. For GOLD, and in order to guarantee consistency with the previous box sizes used for VINA and AutoDock, maximum attention was dedicated to the radius of the binding pocket.

Virtual screening of ligands and decoys

The Directory of Useful Decoys Enhanced (DUD-E) was used to test the 6 scoring functions and docking algorithms [31,32]. DUD-E is a database that containing 102 targets and average of 224 active compounds per target and is a powerful tool to help benchmark molecular docking programs because it provides challenging "decoys". Decoys are molecules that present similar physico-chemical properties to the active molecules but have different 2-D topology to be likely non-binders. In this database, there is an average of 50 decoys for each one of the active compounds.

The protocols optimized in A, for the different scoring functions, were applied to all the β-lactamase active ligands (48 molecules) and corresponding decoys (2332 molecules) contained in DUD-E dataset.

Evaluation of the virtual screening results

From the Virtual screening studies, ranked lists of ligands and decoys were prepared, based in the binding scores obtained with the six scoring functions. The different scoring functions were evaluated on their ability to discriminate active ligands from decoys. To compare their performance, the ranked lists were used to determine recognition metrics such us receiver operating characteristic (or ROC) plots and respective Area under the Curve (AUC) and enrichment factor.

Receiver Operating Characteristic (or ROC) plots expresses the True Positive Rate (TPR=TP/P) *vs.* the False Positive Rate (FPR=FP/N). TP is the number of True Positives, P is the total number of Positives (actives), FP is the number of False Positives, and N is the total number of Negatives (decoys). From ROC plots we also obtained another useful measure, the Area under the Curve (AUC). An increase in the AUC value in a ROC curve means that the discrimination between the true positive and the false positive poses is higher [33].

Another metric used was the enrichment factor at 1% (or EF1%). This is also a very helpful tool because it measures the number of active ligands recovered at 1% of the ligand/decoy database over the number of active ligands that should be expected at the same fraction of the database with random selection. The higher the EF at the top of the database (i.e. EF 1%), the better the scoring function [33]. This measure is very important in VS, as early recognition metric, especially in VS campaigns involving large databases of molecules, because only the top solutions identified by the scoring function can later evaluated visually and eventually reach experimental testing. In this study, EF 20% and EF max were also evaluated.

To assess the correlation between the 48 active molecules ranking position and type of scoring function, chemical and physical descriptors were calculated using MOE [34]. The main properties chosen were, Polarity (the sum of the atomic polarizabilities), F Charge (total charge of the molecule), Weight (molecular weight), Log P (o/w) (log of the octanol/water partition coefficient), Volume (van der Waals volume), Heavy Atoms (number of heavy atoms), Rotatable bonds (fraction of rotatable bonds), Acceptor (number of hydrogen bond acceptor atoms), Donor (number of hydrogen bond donor atoms), Area+; fractional ASA+ (water accessible surface area of all atoms with positive partial charge) calculated as ASA+ /ASA and Area-; fractional ASA- (water accessible surface area of all atoms with negative partial charge) calculated as ASA- /ASA.

Consensus scoring

It is well known that each scoring function has its strengths and its weaknesses. A way to improve the results beyond the level that can be given by a single scoring function it to use a methodology termed consensus scoring [10,17,35-37]. Consensus scoring functions combine different scoring functions, attributing different

weights to the contribution of different alternative, as to improve the discrimination between actives and non-actives in VS [17]. Several studies have demonstrated that the application of consensus scoring functions can decrease the number of false positives and false negatives and virtual screening (REFs).

In this work, the scoring functions with the best discriminatory ability were later combined to create a specific consensus scoring function with superior accuracy. This was evaluated by analysis of the early recognition metrics in the combined ranked list of actives/decoys.

Results and Discussion

Virtual screening results

Table 2 compares the performance of Vina, AutoDock, ChemPLP, GOLDScore, ChemScore and ASP indiscriminating β-lactamase ligands and decoys. From the results obtained, ChemScore, is the scoring function that provides the best early ligand/decoys discrimination, with an EF 1% of 2.14. All the other scoring functions fail to rank the active molecules on the top 1% of the database. Regarding EF20% and EF_{max}, Vina and ChemScore provide the best results.

The same tendency of results is reinforced when analyzing Figure 1 as the ROC plots of the different scoring function, represent the recovery rate of the real ligands as a function of the recovery rate of the decoys.

Another important parameter to address is the AUC value, illustrated in Figure 1 and presented in Table 2. It does confirm the superior performance of ChemScore, with an AUC value of 65.14. It proves that ChemScore has higher discrimination ability throughout the virtual screening, not only in the early recognition metrics (EF 1%). The scoring function with the second best AUC value is Vina (AUC=60.98), followed by GoldScore (AUC=58.98) and ChemPLP (AUC=55.89). ASP and Auto dock showed an inferior performance in early recognition metrics and present AUC values bellow 50, that is, bellow of a random statistic selection.

Upon a detailed analysis of Figure 1, it is evident that ChemScore provides the best ligand/decoys discrimination early on in the virtual screening; however it is not able to recover all the real ligands among the top 80% results. Vina, on the other hand, despite not being so effective in the early recognition metrics, is able to recover 90% of the active ligands among the top 70% of the dataset.

GoldScore is the only scoring function able to recognize all the active ligands, after exploring 83% of all the solutions.

Analyzing active ligands

Analyzing the VS results regarding the active molecules, Figure 2 shows us that different scoring functions rank the molecules in a different way. It is clear that active ligands that appear higher on the VS list with Vina are not the same as with ChemScore. Comparing the molecule CHEMBL210, for example, which is in the 30th position, and is the first active to be identified with Vina, actually is in the 488th position in the ChemScore ranking. The opposite is verified with molecule CHEMBL103. It is the active molecule that is ranked higher with ChemScore (26th position) and is placed the 992th position when using Vina.

Curiously, Vina and CHEMPLP exhibit similar discriminatory ability, as seen in Figure 3.

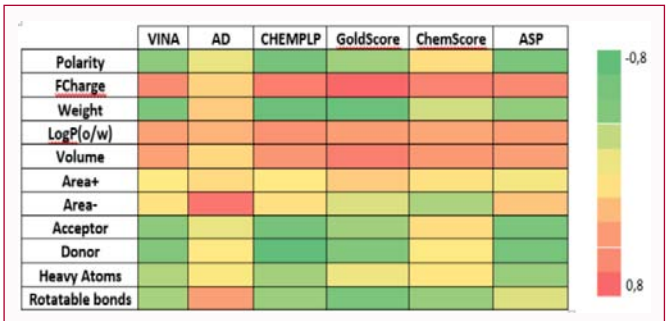

Figure 4: Heat map plots showing the correlation of the active molecule position on the VS ranking with their chemical and physical properties for the different scoring functions.

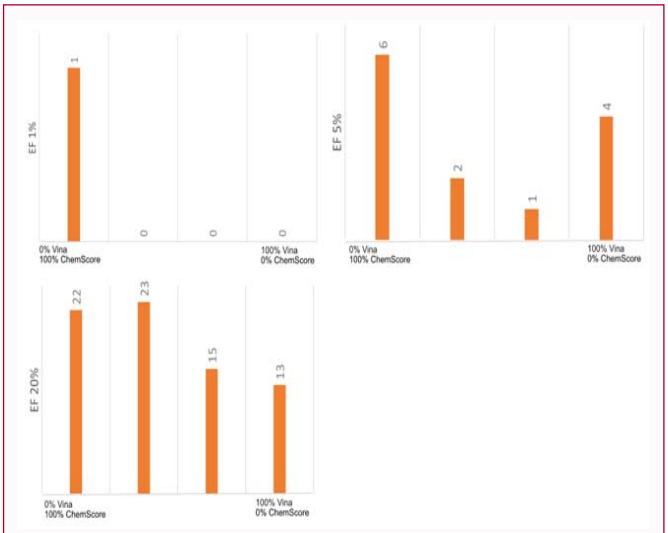

Figure 5: Consensus scoring.

In order to try to understand why different scoring functions evaluate active molecules in a different way, several chemical and physical properties of these ligands were calculated using MOE [34]. Figure 4 represents a heat map of the correlation between the ranking position of the active molecules and their chemical and physical properties, for the different scoring functions. Correlation values of +1 (darker red), imply that there is a perfect positive correlation and values of -1 stands for a perfect negative correlation (darker green). Values near 0 mean no correlation (yellow, orange or light green).

Parameters such as area+ and area-, do not seem to correlate with any scoring function. Charge and volume present a positive correlation. That is, increasing charge or molecule volume means an increase in the ranking position. For these parameters, there is a high positive correlation for all the scoring functions except for AutoDock.

Properties as polarity, molecular weight, acceptor, donor, presence of heavy atoms and fraction of rotatable bonds, exhibit a negative correlation with the ranking position, meaning that, when these parameters increase, the position in the ranking decreases. There is a high negative correlation for Vina, GoldScore, ChemPLP and ASP.

ChemScore and Vina, which were the scoring functions with the best discrimination ability, seem to have opposing results. When there is a high correlation with Vina, it is not so high for ChemScore. Vina, CHEMPLP and ASP, however, display similar correlation profiles.

Table 1: Summary of the scoring functions used in this study.

Program	Scoring Function	Classification	Ref.
AutoDock Vina	Vina	Hybrid (Empirical+knowledge-based)	[24]
AutoDock	AutoDock	Empirical	[23]
GOLD	GoldScore	Force field	[25]
	ChemScore	Empirical	[26]
	CHEMPLP	Empirical	[27]
	ASP	Knowledge-based	[28]

Table 2: Summary of results for all the different scoring functions studied.

	EF (1%)	EF (20%)	EF (Max)	AUC
Vina	0	1.43	3.92	60.98
AutoDock	0	1.25	1.69	49.58
ChemPLP	0	1.15	2.61	55.89
GoldScore	0	1.15	1.71	58.98
ChemScore	2.14	2.29	3.43	65.14
ASP	0	0.73	1.09	44.04

These results suggested that combining different scoring functions could be the best strategy to identify different β-lactamase ligands and that VINA and ChemScore can be used to complement each other is a VS campaign to identify novel ligands.

Consensus scoring

Analyzing the enrichment factors (1%, 5% and 20%) of the consensus scoring, (Figure 5) it is clear that the combination between ChemScore and Vina is not very effective, especially for EF 1%. The combination does improve a little bit when we analyze the EF 20%, but we are more interested in the early recognition metrics on the top 1% of the database.

These results seem to suggest that, in this case, combining the different scoring functions has no effect in the discriminating ability and it is more effective using them separately. ChemScore is, indeed, the scoring function that provides the best discriminating power on this particular target.

Conclusion

Comparing Docking software is difficult, binding sites and recognition processes have unique features that ultimately render protein-ligand interactions specific which might require a case by case analysis (analyzing specific targets or related receptors) [4].

The results presented provide important clues for preparing extensive virtual screening studies involving databases containing millions of compounds to identify novel β-lactamase inhibitors for subsequent experimental testing.

ChemScore (GOLD) was the scoring functioning that presented the best early discrimination results between actives and decoys in virtual screening for β-lactamases. Consensus Scoring did not provide better results. However, Vina and ChemScore can be used together to find more active molecules.

Analyzing the active molecules can be very helpful in finding common patterns in the molecules that fail to be identified in the top solutions in the virtual screening tests. This enables the choice of specific criteria to maximize their recovery in the top solutions. In the case of β-lactamases, Vina was shown to be the best choice in identifying more polar and charged molecules, while ChemScore was better in identifying molecules with more negative partial charges.

In general, these results provide important clues to the preparation of large Virtual Screening campaigns involving extensive chemical databases to identify novel possible active molecules against β-lactamases.

Acknowledgment

This work was supported by national funds from Fundação paraa Ciência e a Tecnologia (SFRH/BD/137844/2018 and IF/00052/2014) and co-financed by the ERDF under the PT2020 Partnership Agreement (POCI-01-0145- FEDER-007728).

References

1. Lohning AE, Levonis SM, Williams-Noonan B, Schweiker SS. A Practical Guide to Molecular Docking and Homology Modelling for Medicinal Chemists. Curr Top Med Chem. 2017;17(18):2023-40.
2. Sousa SF, Ribeiro AJ, Coimbra JT, Neves RP, Martins SA, Moorthy NS, et al. Protein-ligand docking in the new millennium--a retrospective of 10 years in the field. Curr Med Chem. 2013;20(18):2296-314.
3. Sousa FS, Fernandes PA, Ramos MJ. Protein - Ligand Docking: Current Status and Future Challenges. Proteins. 2006;65(1):15-26.
4. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004;3(11):935-49.
5. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;20(7):13384-421.
6. Kroemer RT. Structure-based drug design: docking and scoring. Curr Protein Pept Sci. 2007;8(4):312-28.
7. Taylor RD, Jewsbury PJ, Essex JW. A review of protein-small molecule docking methods. J Comput Aided. Mol Des. 2002;16(3):151-66.
8. Huang SY, Zou X. Advances and challenges in protein-ligand docking. Int J Mol Sci. 2010;11(8):3016-34.
9. Chen YC. Beware of docking! Trends Pharmacol Sci. 2015;36(2):78-95.
10. Cerqueira NM, Gesto D, Oliveira EF, Santos-Martins D, Brás NF, Sousa SF. et al. Receptor-based virtual screening protocol for drug discovery. Arch Biochem Biophys. 2015;582:56-67.
11. Muegge I, Rarey M. Small Molecule Docking and Scoring. In: Lipkowitz KB, Boyd DB, editors. Reviews in Computational Chemistry. New York: Wiley-VCH; 2001. p. 1-46.
12. Adeniyi AA, Soliman MES. Implementing QM in docking calculations: is it a waste of computational time? Drug Discov Today. 2017;22(8):1216-23.
13. Sousa SF, Cerqueira NM, Fernandes PA, Ramos MJ. Virtual screening in drug design and development. Comb Chem High Throughput Screen. 2010;13(5):442-53.
14. Shoichet BK. Virtual screening of chemical libraries. Nature. 2004;432(7019):862-65.

15. ZZhou, Cheng T, Li Q, Bryant SH, Wang Y. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. AAPS J. 2012;14(1):133-41.

16. Yuriev E. Challenges and advances in structure-based virtual screening. Future Med Chem. 2014;6(1):5-7.

17. Vieira TF, Magalhaes R, Sousa SF. Tailoring specialized scoring functions for more efficient virtual screening. Front Drug Chem Clin Res. 2019;2(1):1-4.

18. Powers RA, Morandi F, Shoichet BK. Structure-based discovery of a novel, non covalent inhibitor of AmpC beta-lactamase. Structure. 2002;10(7):1013-23.

19. Antipin RL, Beshnova DA, Petrov RA, Shiryaeva AS, Andreeva IP, Grigorenko VG, et al. Synthesis, SAR and molecular docking study of novel non-ß-lactam inhibitors of TEM type β-lactamase. Bio org Med. 2017;27(7):1588-92.

20. Yelin I, Kishony R. Antibiotic Resistance. Cell. 2018;172(5):1136-1136.e1.

21. Docquier JD, Mangani S. An update on β-lactamase inhibitor discovery and development. Drug Resist Updat. 2018;36:13-29.

22. Silva AP, Ramos IF, Santos RT, Moura D, Vieira TF, Sousa SF, et al. A flow cytometric and computational approaches to carbapenems affinity to the different types of carbapenemases. Front Microbiol. 2016;7:1259.

23. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J Comput Chem. 1998;19(14):1639-62.

24. 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):455-61.

25. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. J Mol Biol. 1997;267(3):727-48.

26. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Proteins. 2003;52(4):609-23.

27. Liebeschuetz JW, Cole JC, Korb O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Mol Des. 2012;26(6):737-48.

28. Mooij WTM, Verdonk ML. General and targeted statistical potentials for protein-ligand interactions. Proteins. 2005;61(2):272-87.

29. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol. 1977;112(3):535-42.

30. Powers RA, Morandi F, Shoichet BK. Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. Structure. 2002;10(7):1013-23.

31. Huang N, Shoichet BK, Irwin JJ. Benchmarking sets for molecular docking. J Med Chem. 2006;49(23):6789-801.

32. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. J Med Chem. 2012;55(14):6582-94.

33. Huang SY, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. Phys Chem Chem Phys. 2010;12(40):12899-908.

34. Chemical Computing Group ULC Molecular Operating Environment (MOE), 2013.08; 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2019.

35. Clark RD, Strizhev A, Leonard JM, Blake JF, Matthew JB. Consensus scoring for ligand/protein interactions. J Mol Graph Model. 2002;20(4):281-95.

36. Houston DR, Walkinshaw MD. Consensus docking: improving the reliability of docking in a virtual screening context. J Chem Inf Model. 2013;53(2):384-90.

37. Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem. 1999;42(25):5100-9.