Interdisciplinary Sciences: Computational Life Sciences (2019) 11:320–328
https://doi.org/10.1007/s12539-019-00327-w

REVIEW

An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking

Jin Li[1,2] · Ailing Fu[3] · Le Zhang[1,4,5,6]

Received: 4 November 2018 / Revised: 6 February 2019 / Accepted: 6 March 2019 / Published online: 15 March 2019

Abstract

Currently, molecular docking is becoming a key tool in drug discovery and molecular modeling applications. The reliability of molecular docking depends on the accuracy of the adopted scoring function, which can guide and determine the ligand poses when thousands of possible poses of ligand are generated. The scoring function can be used to determine the binding mode and site of a ligand, predict binding affinity and identify the potential drug leads for a given protein target. Despite intensive research over the years, accurate and rapid prediction of protein–ligand interactions is still a challenge in molecular docking. For this reason, this study reviews four basic types of scoring functions, physics-based, empirical, knowledge-based, and machine learning-based scoring functions, based on an up-to-date classification scheme. We not only discuss the foundations of the four types scoring functions, suitable application areas and shortcomings, but also discuss challenges and potential future study directions.

Keywords Molecular docking · Scoring function · Ligand pose · Binding affinity · Protein–ligand interaction

Abbreviations

SF	Scoring function
QM	Quantum mechanics
MM	Molecular mechanics
SVM	Support vector machine
RF	Random forest
ANN	Artificial neural network
DL	Deep learning
DNN	Deep neural networks

1 Introduction

Molecular docking [1] is a key tool in drug discovery and molecular modeling applications. The goal of ligand–protein docking is to explore the predominant binding mode(s) of a ligand when it binds to a protein with a known three-dimensional structure. In molecular docking, based on the protein structures, tens of thousands of possible poses of ligand are generated; the binding poses are evaluated by a scoring function (SF) [2], which guides and determines ligand poses. As the most important component of molecular docking, scoring functions have three major functions [3–5]: the first is to determine the binding mode and site of a ligand binding to a protein [6]; the second is to predict the absolute binding affinity between the protein and ligand in lead optimization [7, 8]; the third is virtual screening, which can identify the potential drug leads for a given protein target by searching a large ligand database [9, 10].

Previous research has classified the scoring functions into three major classes: force field, empirical and knowledge-based SFs [3–5]. However, since current research has greatly improved the studies for scoring functions [11, 12], especially in protein–ligand interactions, it is necessary for us to review the recent scoring functions for protein–ligand interactions with new classification scheme [13], which classifies the scoring functions into physics-based, empirical,

✉ Le Zhang
zhangle06@scu.edu.cn

1 College of Computer and Information Science, Southwest University, Chongqing 400715, China

2 School of Medical Information and Engineering, Southwest Medical University, Luzhou 646000, China

3 College of Pharmaceutical Sciences, Southwest University, Chongqing 400715, China

4 College of Computer Science, Sichuan University, Chengdu 610065, China

5 Medical Big Data Center, Sichuan University, Chengdu 610065, China

6 Zdmedical, Information Polytron Technologies Inc Chongqing, Chongqing 401320, China

knowledge-based and machine learning-based SFs. The first three classical scoring functions are classified according to the types of feature items, and they mainly use the linear regression method. The fourth type of scoring function incorporates the nonlinear regression machine-learning method. The four types of soring functions are described in Fig. 1. Here, we briefly introduce these SFs and discuss the scope of their applications.

2 Physics-Based Scoring Functions

Figure 2 describes the physics-based SFs including the scoring functions based on force field [14], solvation models [15, 16] and quantum mechanics methods [17, 18]. The classical force field-based SF computes the binding energy by accumulating the van der Waals and electrostatic interaction between the protein–ligand atom pairs (Eq. 1 of Fig. 2), which considers the contribution of enthalpy to energy [14]. Since it neglects entropy and solvent effect, the performance of the force field-based SF is not good enough [3]. Thus, the force field-based SF is improved by incorporating the torsion entropy [19] of ligands and the solvation/desolvation effect described by explicit [20–24] and implicit solvent models [25–27] (Eq. 2 of Fig. 2). However, the predictive accuracy

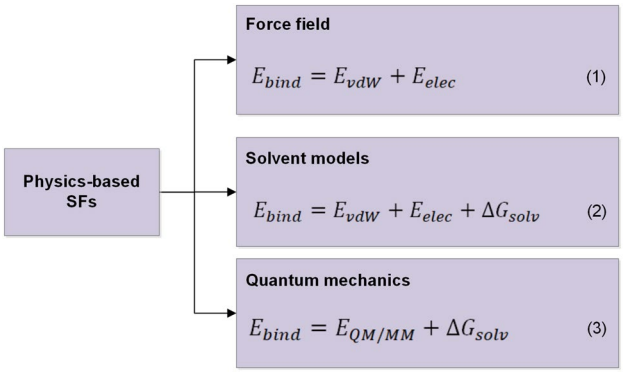
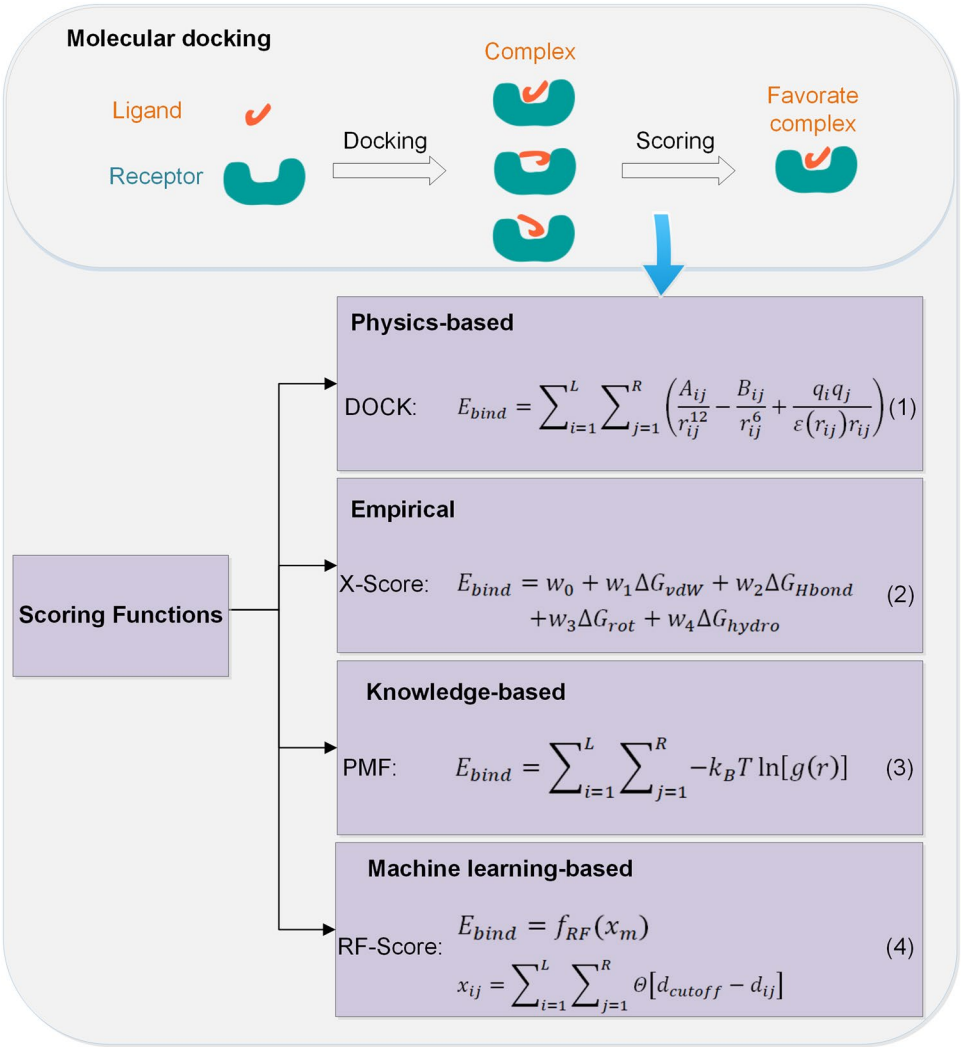


Fig. 2 The description for physics-based SFs

Fig. 1 Four categories of the scoring functions for protein–ligand docking



for the binding energy is significantly subjected to the functional form of the potential energy and related parameters that are hard to locate because this type of scoring function is based on the force field.

Therefore, recent studies have developed the SF based on quantum mechanics (QM) to address the challenges of covalent interactions, polarization, and charge transfer in docking [28–31]. However, the QM-based SF has greater accuracy and computational cost than the force field-based SF. For this reason, a hybrid quantum mechanical/molecular mechanics (QM/MM) approach (Eq. 3 in Fig. 2) [32] was developed to compromise the computational cost and predictive accuracy [17, 18]. How to speedup QM/MM computing [33] has become a hot research area. Though there are still many existing technical difficulties, QM has significant potential to replace the force field when evaluating the direct interactions between proteins and ligands, which represents the current novel research trend [32].

In general, physics-based SFs can directly compute the interactions between the atoms of protein and ligand. Physics-based SFs are appropriate to compute binding free energy between proteins and ligands with relatively greater predictive accuracy than other types of SFs due to the consideration of the enthalpy, solvation and entropy. However, MM or QM models of physics-based SFs are computationally expensive.

3 Empirical Scoring Functions

Empirical SFs [34–37] estimate the binding affinity of a complex by summing up the important energetic factors for protein–ligand binding, such as hydrogen bonds, hydrophobic effects, steric clashes, etc. We usually employ a training set with known binding affinities to optimize the weights of the energetic factors for empirical SFs by linear regression analysis [38]. An example of empirical SFs: X-score [35] can be written as Eq. 2 in Fig. 1.

Empirical SFs are comprised of two research directions. One direction is how to employ a large and high-quality training data set to optimize the protein–ligand structures; the other direction is how to choose appropriate energy terms by stepwise variables and systematic selection regarding the target protein [39–42]. Currently, empirical SFs are commonly employed by protein–ligand docking programs [43–45].

Although the empirical SFs decompose protein–ligand binding affinities into several individual energy terms, similar to physic-based SFs, they usually employ a flexible and intuitive functional form other than using the well-established models that physics-based SFs use. Because of their simple energy terms, these SFs are good at predicting binding affinity, ligand pose, and virtual screening with low computing cost [46], but they are poorly suited for describing the relationship between binding affinity and the crystal structures and they encounter double-counting problems [47].

4 Knowledge-Based Scoring Functions

Knowledge-based SFs [48, 49] derive the desired pairwise potentials from three-dimensional structures of a large set of protein–ligand complexes based on the inverse Boltzmann statistic principle. It is assumed that the frequency of different atom pairs in different distances is related to the interaction of two atoms and converts the frequency into the distance-dependent potential of mean force. Figure 3 describes the computational flow for knowledge-based SF.

The greatest advantage for knowledge-based SFs is compromising the computing cost and predictive accuracy compared with the physics-based and empirical SFs. However, it is difficult for knowledge-based SFs to locate the reference state. Currently, there are two classical strategies used to determine the reference state. One is approximating the reference state by the random distribution of atomic pairs in the training set [50–53]; the other is introducing the corrections

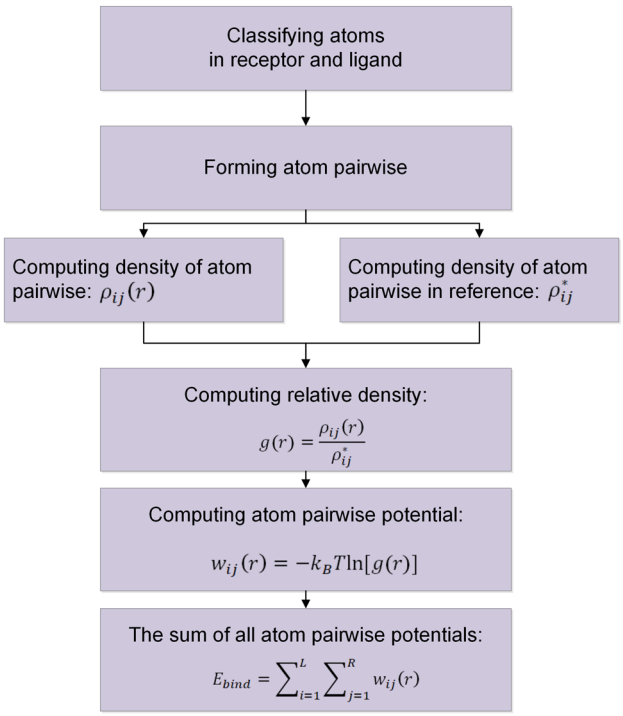


Fig. 3 The computational processing of knowledge-based SFs. Here, $\rho_{ij}(r)$ is the number density of the protein–ligand atom pair $i-j$ at distance r. ρ^*_{ij} is the pair density in a reference state and $g(r)$ is the relative number density of atom pairwise $i-j$ at distance r. k_B is the Boltzmann constant and T is the absolute temperature

item based on the first strategy to improve the accuracy of the knowledge-based SF, such as the volume factor correction method [48, 54, 55], physics-based iterative method [56–58], and so on. Currently, most researchers focus on extending the pairwise potentials to many-body potentials by introducing several new parameters [59–62], which will significantly increase the predictive accuracy while we have difficulty to locate these newly introduced parameters.

Since the training sets for these potentials only consist of structural information and are independent of the experimental binding affinity data, they can avoid possible binding affinity ambiguities caused by experimental conditions, indicating that knowledge-based SFs are suitable for binding poses prediction rather than the binding affinities [63].

5 Machine-Learning-Based Scoring Functions

Unlike the classical SFs (Fig. 1) with assumed mathematical functional form, machine-learning-based SFs employ a variety of machine-learning algorithms, such as support vector machine, random forest, neural network, deep-learning, etc. Figure 4 shows the common workflow to train a machine-learning-based SF. Although machine-learning-based SFs have outperformed classical SFs [12, 64, 65], they are seldom directly incorporated into docking software but are usually used for rescoring [66]. The reason is machine-learning-based scoring functions rely on the training dataset [67, 68]. If the protein and ligand are docked by classical docking software, and then the docked structure is rescored by machine-learning SFs, the accuracy will be improved.

5.1 support vector machine (SVM)

SVM is often employed in structure-based virtual screening to discriminate active poses of ligand from non-active poses, whereas the regression model of support vector machine (another type of SVM) is for binding affinity prediction [69].

Kinnings et al. and Brylinski et al. [70, 71] increased the predictive accuracy for virtual screening and the correlation between the predicted and experimental binding affinities by SVM while using the energy terms of the empirical function.

Additionally, the scoring functions proposed by Ding et al. [72–76] are based on a different set of descriptors related to protein–ligand interactions, training by a modified SVM to fit the experimental binding affinities data. This type of SF performs well when we have a large independent training set.

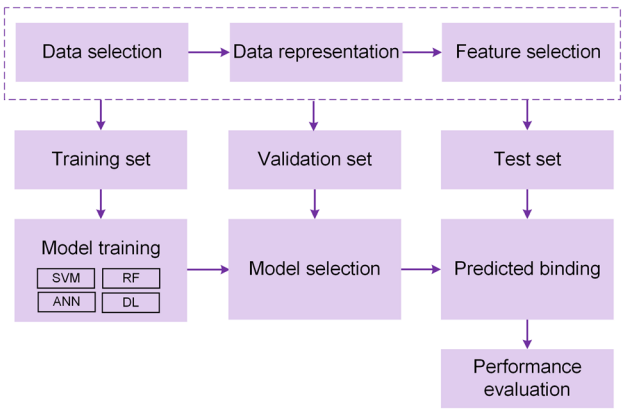


Fig. 4 Workflow of training a machine-learning-based SFs

5.2 Random Forest (RF)

Random forest (RF) is an ensemble learning method based on the aggregation of numerous decision trees, which can comprehensively capture the nonlinear nature of the big data without overfitting. It has been proven that the performance of conventional SFs can be improved when replacing the multiple linear regression method with RF [77–81]. While using simple geometric nature (protein–ligand atom pairwise counts in some distance) as descriptors in RF-Score [82], RF still has greater accuracy in binding affinity prediction than the best conventional scoring function, such as X-Score [78]. Moreover, the latest RF-Score [83, 84] significantly increased predictive accuracy by changing the atom pairwise distance threshold in descriptors. However, Gabel et al. [85] considered that, since the descriptors do not describe true protein–ligand interactions, and RF-Score only consider the relation between protein–ligand atom pairwise counts and the binding affinities, RF-Score does not work well for binding pose prediction and virtual screening. Recently, Wang et al. [79] employed different descriptors to describe the protein–ligand interactions and incorporating decoys to increase the predictive accuracy.

5.3 Artificial Neural Network (ANN)

ANN, a computational model for brain function simulation, has been widely used in recent drug discovery research. However, it is commonly applied for quantitative structure–activity relationship (QSAR) modeling problems [86–92], but is seldom used for binding affinity prediction. The classical neural-network scoring function (NNScore) [93] is used to screen candidate binders by a binary classification output layer. The NNScore version 2.0 [94] considers more binding characteristics, and the output can predict binding affinity. Additionally, Durrant et al. [95] report that, using NNScore rescoring, the results can significantly improve the scoring performance. Moreover, integrating boosting or bagging techniques into a classical ANN-based SF [96, 97] could greatly improve the predictive accuracy.

Although ANN-based SFs have considerably high predictive accuracy for binding affinity, they cannot perform well for high dimension data, which greatly limits their use for commercial docking applications.

5.4 Deep Learning (DL)

Compared with the scoring functions discussed above, the DL scoring function can extract features along with the process of fitting the model's parameters to the available data.

The convolutional neural network model is one of the common deep learning models for scoring functions. This model can be used for drug binding classification [98, 99] and direct binding affinity prediction [100, 101]. Currently, the convolutional neural network model is demonstrated to be better than classical machine-learning methods [102], while it becomes time-consuming as the nodes and depth of the network increase. Moreover, since deep neural networks (DNN) are good at developing multitask classifiers or regressors, we consider that it can predict protein–ligand binding affinities and explore the active ligands from non-actives.

6 Conclusion

We review the scoring functions used for protein–ligand interactions of molecular docking by classifying them into physics-based, empirical, knowledge-based, machine-learning-based scoring functions.

Physics-based, empirical and knowledge-based SFs are commonly used in previous studies [14, 35, 48, 49]. These SFs have been successfully incorporated into popular docking software, such as DOCK [103], SYBYL [104], Discovery Studio [105], Schrodinger [106], Autodock [107], Autodock vina [45], and so on. However, since each SF has obvious shortcomings, the docking software cannot perform well in every aspect. Physics-based SFs (Fig. 2) are good at binding free energy prediction [17, 18]. Nevertheless, since it is time-consuming, how to speed up the physics-based SFs has already become a major research field [108, 109]. Empirical SFs can obtain fast computing capacity, but their simple function form and the linear regression method (Eq. 2 in Fig. 2) may mask the relationship between binding affinity and crystal structure [35]. Knowledge-based SFs can compromise the demand between speed and accuracy; theoretically extending the pairwise potentials to many-body potentials will help with improving the accuracy, but it is difficult for us to locate too many introduced parameters [59–62].

Given that classical scoring function has its advantages and limitations, combining different scoring functions (hybrid SFs) is generally used to improve the accuracy of SFs. Examples of the hybrid SFs are MultiScore [110], GFscore [111], SeleX-CS [112], VoteDock [113] and so on. The hybrid SFs combine different individual scores by consensus strategies, such as vote-by-number, number-by-number, rank-by-number, average rank, linear combination, etc [3]. Hybrid SFs are prone to perform better than standalone scoring functions. The reason may be that combining dissimilar types of scoring functions can compensate for the weaknesses of each. However, such scoring functions are obviously more time-consuming.

For this reason, machine learning-based scoring function research [12, 64, 65] (Fig. 4) has become a current hot research area. Although the descriptors for this type of SF are difficult to explain, machine learning-based SFs can outperform conventional SFs in practice. More importantly, we can increase the predictive accuracy [114, 115] for machine-learning-based SFs by enlarging the training data set.

Since increasing structural and interaction data will be accumulated from academic and industrial fields, we consider that machine-learning SFs have the potential to dominate future SFs. However, there is no such universal SF that can work well for every molecular docking computation; thus, we have to develop the scoring functions for a specific aim [66, 116, 117] by integrating different type of SFs in the distant future [118–120]. Moreover, new types of features (intermolecular features, ligand-only and protein-only features) can be employed to improve the performance of the SFs, and it is also very important to generate these SFs as open software for more researchers in this field to use.

Acknowledgements This study is supported by the National Natural Science Foundation of China (No. 61372138), and National Science and Technology Major Project of China (No. 2018ZX10201002).

Author contributions Conception and design: LZ; Writing and revision of the manuscript: JL; ALF.

Funding This study is supported by the National Natural Science Foundation of China (No. 61372138), and National Science and Technology Major Project of China (No. 2018ZX10201002).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Irwin JJ, Lorber DM, Mcgovern SL, Wei B, Shoichet BK (2002) Molecular docking and drug design. Comput Nanosci Nanotechnol 2:50–51
2. Fenu LALR, Good AC, Bodkin M, Essex JW (2007) Structure-based drug discovery. Springer, Dordrecht, p 24
3. Huang SY, Grinter SZ, Zou X (2010) Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. Phys Chem Chem Phys PCCP 12(40):12899–12908. https://doi.org/10.1039/c0cp00151a

4. Brooijmans N, Kuntz ID (2003) Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 32:335–373. https://doi.org/10.1146/annurev.biophys.32.110601.142532
5. Wang JC, Lin JH (2013) Scoring functions for prediction of protein–ligand interactions. Curr Pharm Design 19(12):2174–2182
6. Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, Raushel FM (2007) Structure-based activity prediction for an enzyme of unknown function. Nature 448(7155):775–779
7. Joseph-Mccarthy D, Baber JC, Feyfant E, Thompson DC, Humblet C (2007) Lead optimization via high-throughput molecular docking. Curr Opin Drug Discov Devel 10(3):264–274
8. Jorgensen WL (2009) Efficient drug lead discovery and optimization. Acc Chem Res 42(6):724–733
9. Seifert MH, Kraus J, Kramer B (2007) Virtual high-throughput screening of molecular databases. Curr Opin Drug Discov Devel 10(3):298–307
10. Schneider G (2010) Virtual screening: an endless staircase? Nat Rev Drug Discov 9(4):273–276. https://doi.org/10.1038/nrd3139
11. Ain QU, Aleksandrova A, Roessler FD, Ballester PJ (2015) Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. Wiley Interdiscip Rev Comput Mol Sci 5(6):405–424. https://doi.org/10.1002/wcms.1225
12. Khamis MA, Gomaa W, Ahmed WF (2015) Machine learning in computational docking. Artif Intell Med 63(3):135–152. https://doi.org/10.1016/j.artmed.2015.02.002
13. Liu J, Wang R (2015) Classification of current scoring functions. J Chem Inf Model 55(3):475
14. Meng EC, Shoichet BK, Kuntz ID (1992) Automated docking with grid-based energy evaluation, vol 13. Wiley, New York
15. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79(2):926–935
16. Pullman B (1981) Intermolecular forces. D. Reidel Publishing Company, Dordrecht
17. Raha K, Peters MB, Wang B, Yu N, Wollacott AM, Westerhoff LM, Merz KM Jr (2007) The role of quantum mechanics in structure-based drug design. Drug Discov Today 12(17–18):725–731. https://doi.org/10.1016/j.drudis.2007.07.006
18. Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. Angew Chem (Int Edn Engl) 48(7):1198–1229. https://doi.org/10.1002/anie.200802019
19. Kramer B, Rarey M, Lengauer T (1999) Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. Proteins Struct Funct Bioinf 37(2):228–241
20. Yang Y, Lightstone FC, Wong SE (2013) Approaches to efficiently estimate solvation and explicit water energetics in ligand binding: the use of WaterMap. Expert Opin Drug Discov 8(3):277–287. https://doi.org/10.1517/17460441.2013.749853
21. Michel J, Tirado-Rives J, Jorgensen WL (2009) Prediction of the water content in protein binding sites. J Phys Chem B 113(40):13337–13346. https://doi.org/10.1021/jp9047456
22. Ross GA, Morris GM, Biggin PC (2012) Rapid and accurate prediction and scoring of water molecules in protein binding sites. PLoS One 7(3):e32036. https://doi.org/10.1371/journal.pone.0032036
23. Uehara S, Tanaka S (2016) AutoDock-GIST: incorporating thermodynamics of active-site water into scoring function for accurate protein–ligand docking. Molecules. https://doi.org/10.3390/molecules21111604
24. Kumar A, Zhang KY (2013) Investigation on the effect of key water molecules on docking performance in CSARdock exercise. J Chem Inf Model 53(8):1880–1892
25. Sun H, Li Y, Li D, Hou T (2013) Insight into crizotinib resistance mechanisms caused by three mutations in ALK tyrosine kinase using free energy calculation approaches. J Chem Inf Model 53(9):2376–2389. https://doi.org/10.1021/ci400188q
26. Sun HY, Hou TJ, Zhang HY (2014) Finding chemical drugs for genetic diseases. Drug Discov Today 19(12):1836–1840. https://doi.org/10.1016/j.drudis.2014.09.013
27. Chen F, Liu H, Sun H, Pan P, Li Y, Li D, Hou T (2016) Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking. Phys Chem Chem Phys PCCP 18(32):22129–22139. https://doi.org/10.1039/c6cp03670h
28. Kulik HJ (2018) Large-scale QM/MM free energy simulations of enzyme catalysis reveal the influence of charge transfer. Phys Chem Chem Phys PCCP 20(31):20650–20660. https://doi.org/10.1039/c8cp03871f
29. Orozco-Gonzalez Y, Manathunga M, Marin MDC, Agathangelou D, Jung KH, Melaccio F, Ferre N, Haacke S, Coutinho K, Canuto S, Olivucci M (2017) An average solvent electrostatic configuration protocol for QM/MM free energy optimization: implementation and application to rhodopsin systems. J Chem Theory Comput 13(12):6391–6404. https://doi.org/10.1021/acs.jctc.7b00860
30. Chaskar P, Zoete V, Röhrig UF (2017) On-the-Fly QM/MM docking with attracting cavities. J Chem Inf Model 57(1):73–84. https://doi.org/10.1021/acs.jcim.6b00406
31. Natesan S, Subramaniam R, Bergeron C, Balaz S (2012) Binding affinity prediction for ligands and receptors forming tautomers and ionization species: inhibition of mitogen-activated protein kinase-activated protein kinase 2 (MK2). J Med Chem 55(5):2035–2047. https://doi.org/10.1021/jm201217q
32. Chaskar P, Zoete V, Rohrig UF (2014) Toward on-the-fly quantum mechanical/molecular mechanical (QM/MM) docking: development and benchmark of a scoring function. J Chem Inf Model 54(11):3137–3152. https://doi.org/10.1021/ci5004152
33. Steinmann C, Olsson MA, Ryde U (2018) Relative ligand-binding free energies calculated from multiple short QM/MM MD simulations. Acs Nano 14(6):3228–3237
34. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J Comput Aided Mol Design 11(5):425–445
35. Murray CW, Auton TR, Eldridge MD (1998) Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand-receptor binding affinities and the use of Bayesian regression to improve the quality of the model. J Comput Aided Mol Design 12(5):503–519
36. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. J Med Chem 49(21):6177–6196. https://doi.org/10.1021/jm051256o
37. Zheng Z, Merz KM (2011) Ligand identification scoring algorithm (LISA). J Chem Inf Model 51(6):1296–1306. https://doi.org/10.1021/ci2000665
38. Kadukova M, Grudinin S (2017) Convex-PL: a novel knowledge-based potential for protein–ligand interactions deduced from structural databases using convex optimization. J Comput Aided Mol Design 31(10):943–958. https://doi.org/10.1007/s10822-017-0068-8
39. Fornabaio M, Spyrakis F, Mozzarelli A, Cozzini P, Abraham DJ, Kellogg GE (2004) Simple, intuitive calculations of free energy of binding for protein–ligand complexes. 3. The free energy contribution of structural water molecules in HIV-1 protease complexes. J Med Chem 47(18):4507–4516. https://doi.org/10.1021/jm030596b

40. Kerzmann A, Neumann D, Kohlbacher O (2006) SLICK—scoring and energy functions for protein–carbohydrate interactions. J Chem Inf Model 46(4):1635–1642. https://doi.org/10.1021/ci050422y
41. Catana CS, Novel PFW (2007) Customizable scoring functions, parameterized using N-PLS, for structure-based drug discovery. J Chem Inf Model 47(1):85–91
42. Sotriffer CA, Sanschagrin P, Matter H, Klebe G (2008) SFCscore: scoring functions for affinity prediction of protein–ligand complexes. Proteins 73(2):395–419. https://doi.org/10.1002/prot.22058
43. Bohm HJ (1994) 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 Design 8(3):243–256
44. Jain AN (2003) Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. J Med Chem 46(4):499–511. https://doi.org/10.1021/jm020406h
45. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31(2):455–461. https://doi.org/10.1002/jcc.21334
46. Li Y, Liu Z, Li J, Han L, Liu J, Zhao Z, Wang R (2014) Comparative assessment of scoring functions on an updated benchmark: 1. Compilation of the test set. J Chem Inf Model 54(6):1700–1716. https://doi.org/10.1021/ci500080q
47. Thornton BF, Wik M, Crill PM (2017) Double-counting challenges the accuracy of high-latitude methane inventories. Geophys Res Lett 43:(24)
48. Muegge I, Martin YC (1999) A general and fast scoring function for protein–ligand interactions: a simplified potential approach. J Med Chem 42(5):791–804. https://doi.org/10.1021/jm980536j
49. Gohlke H, Hendlich M, Klebe G (2000) Knowledge-based scoring function to predict protein–ligand interactions. J Mol Biol 295(2):337–356. https://doi.org/10.1006/jmbi.1999.3371
50. Velec HF, Gohlke H, Klebe G (2005) DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. J Med Chem 48(20):6296–6303. https://doi.org/10.1021/jm050436v
51. Neudert G, Klebe G (2011) DSX: a knowledge-based scoring function for the assessment of protein–ligand complexes. J Chem Inf Model 51(10):2731–2745. https://doi.org/10.1021/ci200274q
52. Mooij WT, Verdonk ML (2005) General and targeted statistical potentials for protein–ligand interactions. Proteins 61(2):272–287. https://doi.org/10.1002/prot.20588
53. Yang CY, Wang R, Wang S (2006) M-score: a knowledge-based potential scoring function accounting for protein atom mobility. J Med Chem 49(20):5903–5911. https://doi.org/10.1021/jm050043w
54. Huang SY, Zou X (2006) An iterative knowledge-based scoring function to predict protein–ligand interactions: I. Derivation of interaction potentials. J Comput Chem 27(15):1866–1875. https://doi.org/10.1002/jcc.20504
55. Huang SY, Zou X (2006) An iterative knowledge-based scoring function to predict protein–ligand interactions: II. Validation of the scoring function. J Comput Chem 27(15):1876–1882. https://doi.org/10.1002/jcc.20505
56. Huang SY, Zou X (2014) A knowledge-based scoring function for protein–RNA interactions derived from a statistical mechanics-based iterative method. Nucleic Acids Res 42(7):e55. https://doi.org/10.1093/nar/gku077
57. Forli S, Olson AJ (2012) A force field with discrete displaceable waters and desolvation entropy for hydrated ligand docking. J Med Chem 55(2):623–638. https://doi.org/10.1021/jm2005145
58. Huang SY, Zou X (2010) Inclusion of solvation and entropy in the knowledge-based scoring function for protein–ligand interactions. J Chem Inf Model 50(2):262–273. https://doi.org/10.1021/ci9002987
59. Lu M, Dousis AD, Ma J (2008) OPUS-PSP: an orientation-dependent statistical all-atom potential derived from side-chain packing. J Mol Biol 376(1):288–301. https://doi.org/10.1016/j.jmb.2007.11.033
60. Xu G, Ma T, Zang T, Sun W, Wang Q, Ma J (2017) OPUS-DOSP: a distance- and orientation-dependent all-atom potential derived from side-chain packing. J Mol Biol 429(20):3113–3120. https://doi.org/10.1016/j.jmb.2017.08.013
61. Li Y, Han L, Liu Z, Wang R (2014) Comparative assessment of scoring functions on an updated benchmark: 2. Evaluation methods and general results. J Chem Inf Model 54(6):1717–1736. https://doi.org/10.1021/ci500081m
62. Park J, Saitou K (2014) ROTAS: a rotamer-dependent, atomic statistical potential for assessment and prediction of protein structures. BMC Bioinform 15:307. https://doi.org/10.1186/1471-2105-15-307
63. Zheng Z, Merz KM Jr (2013) Development of the knowledge-based and empirical combined scoring algorithm (KECSA) to score protein–ligand interactions. J Chem Inf Model 53(5):1073–1083. https://doi.org/10.1021/ci300619x
64. Ma DL, Chan DS, Leung CH (2013) Drug repositioning by structure-based virtual screening. Chem Society Rev 42(5):2130–2141. https://doi.org/10.1039/c2cs35357a
65. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH (2012) Structure-based virtual screening for drug discovery: a problem-centric review. AAPS J 14(1):133–141. https://doi.org/10.1208/s12248-012-9322-0
66. Zhang L, Ai HX, Li SM, Qi MY, Zhao J, Zhao Q, Liu HS (2017) Virtual screening approach to identifying influenza virus neuraminidase inhibitors using molecular docking combined with machine-learning-based scoring function. Oncotarget 8(47):83142–83154. https://doi.org/10.18632/oncotarget.20915
67. Zhang L, Qiao M, Gao H, Hu B, Tan H, Zhou X, Li CM (2016) Investigation of mechanism of bone regeneration in a porous biodegradable calcium phosphate (CaP) scaffold by a combination of a multi-scale agent-based model and experimental optimization/validation. Nanoscale 8(31):14877–14887. https://doi.org/10.1039/c6nr01637e
68. Zhang L, Zhang S (2017) Using game theory to investigate the epigenetic control mechanisms of embryo development: Comment on: "Epigenetic game theory: How to compute the epigenetic control of maternal-to-zygotic transition" by Qian Wang et al. Phys Life Rev 20:140–142. https://doi.org/10.1016/j.plrev.2017.01.007
69. Zhang L, Zheng CQ, Li T, Xing L, Zeng H, Li TT, Yang H, Cao J, Chen BD, Zhou ZY (2017) Building up a robust risk mathematical platform to predict colorectal cancer. Complexity 2017:14. https://doi.org/10.1155/2017/8917258
70. Kinnings SL, Liu N, Tonge PJ, Jackson RM, Xie L, Bourne PE (2011) A machine learning-based method to improve docking scoring functions and its application to drug repurposing. J Chem Inf Model 51(2):408–419. https://doi.org/10.1021/ci100369f
71. Brylinski M (2013) Nonlinear scoring functions for similarity-based ligand docking and binding affinity prediction. J Chem Inf Model 53(11):3097–3112. https://doi.org/10.1021/ci400510e
72. Li GB, Yang LL, Wang WJ, Li LL, Yang SY (2013) ID-Score: a new empirical scoring function based on a comprehensive set of descriptors related to protein–ligand interactions. J Chem Inf Model 53(3):592–600. https://doi.org/10.1021/ci300493w
73. Koppisetty CA, Frank M, Kemp GJ, Nyholm PG (2013) Computation of binding energies including their enthalpy and entropy components for protein–ligand complexes using

support vector machines. J Chem Inf Model 53(10):2559–2570. https://doi.org/10.1021/ci400321r
74. Ding B, Li N, Wang W (2013) Characterizing binding of small molecules. II. Evaluating the potency of small molecules to combat resistance based on docking structures. J Chem Inf Model 53(5):1213–1222. https://doi.org/10.1021/ci400011c
75. Ding B, Wang J, Li N, Wang W (2013) Characterization of small molecule binding. I. Accurate identification of strong inhibitors in virtual screening. J Chem Inf Model 53(1):114–122. https://doi.org/10.1021/ci300508m
76. Yan Y, Wang W, Sun Z, Zhang JZH, Ji C (2017) Protein–ligand empirical interaction components for virtual screening. J Chem Inf Model 57(8):1793–1806. https://doi.org/10.1021/acs.jcim.7b00017
77. Li H, Leung KS, Wong MH, Ballester PJ (2014) Substituting random forest for multiple linear regression improves binding affinity prediction of scoring functions: cyscore as a case study. BMC Bioinf 15:291. https://doi.org/10.1186/1471-2105-15-291
78. Afifi K, Al-Sadek AF (2018) Improving classical scoring functions using random forest: the non-additivity of free energy terms' contributions in binding. Chem Biol Drug Design. https://doi.org/10.1111/cbdd.13206
79. Wang C, Zhang Y (2017) Improving scoring-docking-screening powers of protein–ligand scoring functions using random forest. J Comput Chem 38(3):169–177. https://doi.org/10.1002/jcc.24667
80. Zilian D, Sotriffer CA (2013) SFCscore(RF): a random forest-based scoring function for improved affinity prediction of protein–ligand complexes. J Chem Inf Model 53(8):1923–1933. https://doi.org/10.1021/ci400120b
81. Liu Q, Kwoh CK, Li J (2013) Binding affinity prediction for protein–ligand complexes based on beta contacts and B factor. J Chem Inf Model 53(11):3076–3085. https://doi.org/10.1021/ci400450h
82. Ballester PJ, Mitchell JBO (2010) A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. Oxford University Press, Oxford
83. Ballester PJ, Schreyer A, Blundell TL (2014) Does a more precise chemical description of protein–ligand complexes lead to more accurate prediction of binding affinity? J Chem Inf Model 54(3):944–955. https://doi.org/10.1021/ci500091r
84. Li H, Leung KS, Wong MH, Ballester PJ (2015) Improving AutoDock Vina using random forest: the growing accuracy of binding affinity prediction by the effective exploitation of larger data sets. Mol Inf 34(2–3):115–126. https://doi.org/10.1002/minf.201400132
85. Gabel J, Desaphy J, Rognan D (2014) Beware of machine learning-based scoring functions—on the danger of developing black boxes. J Chem Inf Model 54(10):2807–2815. https://doi.org/10.1021/ci500406k
86. Cang Z, Mu L, Wei GW (2018) Representability of algebraic topology for biomolecules in machine learning based scoring and virtual screening. PLoS Comput Biol 14(1):e1005929. https://doi.org/10.1371/journal.pcbi.1005929
87. Buiu C, Putz MV, Avram S (2016) Learning the relationship between the primary structure of HIV envelope glycoproteins and neutralization activity of particular antibodies by using artificial neural networks. Int J Mol Sci. https://doi.org/10.3390/ijms17101710
88. Winkler DA, Burden FR (2007) Nonlinear predictive modeling of MHC class II-peptide binding using Bayesian neural networks. Methods Mol Biol (Clifton NJ) 409:365–377. https://doi.org/10.1007/978-1-60327-118-9_27
89. Fabry-Asztalos L, Andonie R, Collar CJ, Abdul-Wahid S, Salim N (2008) A genetic algorithm optimized fuzzy neural network analysis of the affinity of inhibitors for HIV-1 protease. Bioorg Med Chem 16(6):2903–2911. https://doi.org/10.1016/j.bmc.2007.12.055
90. Shen J, Cui Y, Gu J, Li Y, Li L (2014) A genetic algorithm-back propagation artificial neural network model to quantify the affinity of flavonoids toward P-glycoprotein. Combinatorial Chem High Throughput Screen 17(2):162–172
91. O'Donnell TJ, Rubinsteyn A, Bonsack M, Riemer AB, Laserson U, Hammerbacher J (2018) MHCflurry: open-source class I MHC binding affinity prediction. Cell Syst 7(1):129–132.e124. https://doi.org/10.1016/j.cels.2018.05.014
92. Chupakhin V, Marcou G, Baskin I, Varnek A, Rognan D (2013) Predicting ligand binding modes from neural networks trained on protein–ligand interaction fingerprints. J Chem Inf Model 53(4):763–772. https://doi.org/10.1021/ci300200r
93. Durrant JD, Mccammon JA (2010) NNScore: a neural-network-based scoring function for the characterization of protein–ligand complexes. J Chem Inf Model 50(10):1865–1871
94. Durrant JD, McCammon JA (2011) NNScore 2.0: a neural-network receptor-ligand scoring function. J Chem Inf Model 51(11):2897–2903. https://doi.org/10.1021/ci2003889
95. Durrant JD, Friedman AJ, Rogers KE, McCammon JA (2013) Comparing neural-network scoring functions and the state of the art: applications to common library screening. J Chem Inf Model 53(7):1726–1735. https://doi.org/10.1021/ci400042y
96. Ashtawy HM, Mahapatra NR (2018) Boosted neural networks scoring functions for accurate ligand docking and ranking. J Bioinf Comput Biol 16(2):1850004. https://doi.org/10.1142/s021972001850004x
97. Ashtawy HM, Mahapatra NR (2015) BgN-Score and BsN-Score: bagging and boosting based ensemble neural networks scoring functions for accurate binding affinity prediction of protein–ligand complexes. BMC Bioinf 16(Suppl 4):S8. https://doi.org/10.1186/1471-2105-16-s4-s8
98. Wallach I, Dzamba M, Heifets A (2015) AtomNet: a deep convolutional neural network for bioactivity prediction in structure-based drug discovery. Math Z 47(1):34–46
99. Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR (2017) Protein–ligand scoring with convolutional neural networks. J Chem Inf Model 57(4):942–957. https://doi.org/10.1021/acs.jcim.6b00740
100. Gomes J, Ramsundar B, Feinberg EN, Pande VS (2017) Atomic convolutional networks for predicting protein–ligand binding affinity. arXiv preprint arXiv: 170310603
101. Stepniewska-Dziubinska MM, Zielenkiewicz P, Siedlecki P (2018) Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. Bioinformatics. https://doi.org/10.1093/bioinformatics/bty374
102. Bengio Y, Vincent AC,P (2013) Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35(8):31
103. Meng ECSB, Kuntz ID (1992) Automated docking with grid-based energy evaluation. J Comput Chem 13:20
104. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267(3):727–748. https://doi.org/10.1006/jmbi.1996.0897
105. Krammer A, Kirchhoff PD, Jiang X, Venkatachalam CM, Waldman M (2005) LigScore: a novel scoring function for predicting binding affinities. J Mol Graph Model 23(5):395–407. https://doi.org/10.1016/j.jmgm.2004.11.007
106. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47(7):1739–1749. https://doi.org/10.1021/jm0306430

107. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (2015) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19(14):1639–1662
108. Steinmann C, Olsson MA, Ryde U (2018) Relative ligand-binding free energies calculated from multiple short QM/MM MD simulations. J Chem Theory Comput 14(6):3228–3237. https://doi.org/10.1021/acs.jctc.8b00081
109. Chaskar P, Zoete V, Röhrig UF (2014) Toward on-the-fly quantum mechanical/molecular mechanical (QM/MM) docking: development and benchmark of a scoring function. J Chem Inf Model 54(11):3137–3152. https://doi.org/10.1021/ci5004152
110. Terp GE, Johansen BN, Christensen IT, Jorgensen FS (2001) A new concept for multidimensional selection of ligand conformations (MultiSelect) and multidimensional scoring (MultiScore) of protein–ligand binding affinities. J Med Chem 44(14):2333–2343
111. Betzi S, Suhre K, Chetrit B, Guerlesquin F, Morelli X (2006) GFscore: a general nonlinear consensus scoring function for high-throughput docking. J Chem Inf Model 46(4):1704–1712. https://doi.org/10.1021/ci0600758
112. Bar-Haim S, Aharon A, Ben-Moshe T, Marantz Y, Senderowitz H (2009) SeleX-CS: a new consensus scoring algorithm for hit discovery and lead optimization. J Chem Inf Model 49(3):623–633. https://doi.org/10.1021/ci800335j
113. Plewczynski D, Lazniewski M, von Grotthuss M, Rychlewski L, Ginalski K (2011) VoteDock: consensus docking method for prediction of protein–ligand interactions. J Comput Chem 32(4):568–581. https://doi.org/10.1002/jcc.21642
114. Zhang L, Liu Y, Wang M, Wu Z, Li N, Zhang J, Yang C (2017) EZH2-, CHD4-, and IDH-linked epigenetic perturbation and its association with survival in glioma patients. J Mol Cell Biol 9(6):477–488. https://doi.org/10.1093/jmcb/mjx056
115. Zhang L, Xiao M, Zhou J, Yu J (2018) Lineage-associated underrepresented permutations (LAUPs) of mammalian genomic sequences based on a Jellyfish-based LAUPs analysis application (JBLA). Bioinformatics 34(21):3624–3630. https://doi.org/10.1093/bioinformatics/bty392
116. Santos-Martins D, Forli S, Ramos MJ, Olson AJ (2014) AutoDock4(Zn): an improved AutoDock force field for small-molecule docking to zinc metalloproteins. J Chem Inf Model 54(8):2371–2379. https://doi.org/10.1021/ci500209e
117. Poli G, Jha V, Martinelli A, Supuran CT, Tuccinardi T (2018) Development of a fingerprint-based scoring function for the prediction of the binding mode of carbonic anhydrase II inhibitors. Int J Mol Sci. https://doi.org/10.3390/ijms19071851
118. Baek M, Shin WH, Chung HW, Seok C (2017) GalaxyDock BP2 score: a hybrid scoring function for accurate protein–ligand docking. J Comput Aided Mol Design 31(7):653–666. https://doi.org/10.1007/s10822-017-0030-9
119. Shin WH, Kim JK, Kim DS, Seok C (2013) GalaxyDock2: protein–ligand docking using beta-complex and global optimization. J Comput Chem 34(30):2647–2656. https://doi.org/10.1002/jcc.23438
120. Debroise T, Shakhnovich EI, Cheron N (2017) A hybrid knowledge-based and empirical scoring function for protein–ligand interaction: SMoG2016. J Chem Inf Model 57(3):584–593. https://doi.org/10.1021/acs.jcim.6b00610