Evaluating reverse docking on general and selective inhibitors: a case study about glide

Mingna Li[1], Xing Wu[2], Liangren Zhang[1*], Zhenming Liu[1*]

1. *State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University Health Science Center, Beijing 100191, China*
2. *Hengdian Group Holdings Limited, Hangzhou 310007, China*

Abstract: As a powerful tool for target prediction, reverse docking remains largely unexplored. The objective evaluation of reverse docking software can help us know better about the strength and weakness of these tools, hence guiding us in target prediction. In the present study, we evaluated the target prediction power of Glide (SP) against general inhibitors and selective inhibitors. The results showed that the scoring tendency could be different for each ligand, and overall scoring sampling was necessary for a better understanding of the docking score for a certain protein-ligand pair. Besides, the input conformation of the binding pocket could affect the docking result. Glide (SP) showed a preferable performance on the target prediction of the general inhibitors. However, the accuracy of the target prediction of the selective inhibitors was relatively low, indicating that Glide (SP) might not be capable for this task. The case study about COVID-19 proved that coagulation factor Xa might be a potential target of chloroquine. Therefore, we recommend the further development of reverse docking tools and rectification of inter-target scoring bias.

Keywords: Reverse docking; Target prediction; Software evaluation

CLC number: R916 **Document code:** A **Article ID:** 1003–1057(2020)10–679–10

1. Introduction

Reverse docking, also known as inverse docking, is a vital aspect in computational target prediction. By using 3D structural information, reverse docking can predict targets with binding mode and approximate binding free energy, providing a convenient way to explore potential mechanism of action of a small molecule[1,2]. Many reverse docking tools have been successfully used in target prediction. For example, Lim et al.[3] have found cyclin-dependent kinase 2 as a new target of curcumin by using DOCK6, which is then confirmed by *in vitro* and *ex vivo* kinase assay. Similarly, Wang et al.[4] have found several potential protein targets of glabridin using reverse docking, and the selective binding is then verified by *in vitro* pull-down assay.

With the rapid expansion of reverse docking tools, there is an urgent need to evaluate the target prediction power of docking software. Lapillo et al.[5] have reported the first extensive reliability evaluation work of reverse docking software with a new dataset consist of 60 targets and 600 ligands. By analyzing the docking score given by different tools, the authors have summarized docking software with better target prediction power and factors that may make an impact on reverse docking. However, except for the global success rate of target prediction, whether reverse docking can distinguish the true target for the selective inhibitors remains a major concern. For example, if a ligand is a selective inhibitor of

Received: 2020-04-16; Revised: 2020-07-15; Accepted: 2020-08-20.
Foundation items: National Key Research and Development Project (Grant No. 2019YFC1708900), the National Natural Science Foundation of China (Grant No. 81872730, 81673279, 21772005), National Major Scientific and Technological Special Project for Significant New Drugs Development (Grant No. 2018ZX09735001-003, 2019ZX09201005-001, 2019ZX09204-001) and Beijing Natural Science Foundation (Grant No. 7202088, 7172118).
*Corresponding author. Tel.: +86-10-82805281
E-mail: zmliu@bjmu.edu.cn

http://dx.doi.org/10.5246/jcps.2020.10.063

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

protein A, can a docking software discriminate protein B during target prediction while protein A and B are of the same protein class?

To the best of our knowledge, there is still no existing work which has reported the target prediction power of Glide[6,7], a classical docking software, on subtypes of protein. We used a well-established benchmark dataset for evaluation and PDBbind refined set for overall scoring sampling. Moreover, we analyzed the factors that might be relevant to the target prediction power.

In the present work, we aimed to answer the questions as follows: 1) How Glide performed for the target prediction of general and selective inhibitors; 2) Was the feature of ligand relevant to the docking score; 3) Would conformations of the protein influence the docking score?

2. Materials and methods

2.1. Evaluation of datasets

The selectivity data set proposed by Karen et al.[8] was used as the benchmark dataset for the evaluation of the target prediction power among protein subtypes. This dataset was composed of three target classes, namely phosphodiesterases, histone deacetylases and serine proteases. Proteins in the benchmark dataset were shown in Table 1. For each protein, two or three structures were used as input file to account for protein conformation changes. As query ligands for the proteins, 17 selective or general inhibitors were used, which were shown in Table 2.

Table 1. Proteins in the benchmark dataset and the PDB codes of the protein input file.

Protein	PDB codes of protein structures
Phosphodiesterases	
PDE4	2QYN, 2QYK, 3TVX
PDE5	3HC8, 3SHY, 3TGE
Histone Deacetylases	
HDAC2 (class I)	3MAX, 4LY1
HDAC8 (class I)	1T67, 1W22, 3SFF
HDAC4 (class II)	2VGJ, 4CBY
Serine Proteases	
Trypsin	2G5N, 2G8T, 3GY2
Thrombin	2BVR, 3RM2, 3SI4
Factor Xa	2JKH, 2Y5F, 3KL6

Table 2. General and selective inhibitors in the benchmark dataset.

Protein	General inhibitors (DrugBank ID or PubChem ID)	Selective inhibitors (DrugBank ID or PubChem ID)
Phosphodiesterases		
PDE4	Caffeine (DB00201) Paraxanthine (CID 4687) Theophylline (DB00277)	Piclamilast (DB01791) Drotaverine (DB06751) Roflumilast (DB01656)
PDE5		Avanafil (DB06237) Sildenafil (DB00203) Tadalafil (DB00820)
Histone Deacetylases		
HDAC2 (class I)	Vorinostat (DB02546) Trichostatin A (CID 444732)	Mocetinostat (CID 9865515)
HDAC8 (class I)		
HDAC4 (class II)		
Serine Proteases		
Trypsin	Benzamidine (DB03127) Pefabloc (DB07347)	
Thrombin		Melagatran (CID 183797)
Factor Xa		Apixaban (DB06605) Rivaroxaban (DB06228)

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

For the investigation of the general scoring tendency of each ligand, proteins in the PDBbind refined set[9] were used for the overall scoring sampling, which included 4854 structures.

2.2. Structure preparation

The protein structures of the targets in the benchmark dataset were downloaded from PDB database[10], while the ligand structures were downloaded from Drug Bank[11] or Pubchem[12]. The structure files in the PDBbind refined set were downloaded from the official website.

The ligand files were prepared by LigPrep, and the protein files were prepared using the Protein Preparation Wizard included in the Schrödinger Suite by adding hydrogens, deleting waters, assigning bond orders and determining protonation states[13]. The coordinate of the pocket was determined by the co-crystallized ligand of each protein.

2.3. Molecular descriptors

To characterize the feature of the ligands, five physicochemical properties were calculated using BIOVIA pipeline pilot components including molecular weight, AlogP, number of H-bond acceptors, number of H-bond donors and number of rotatable bonds.

2.4. Molecular docking

The standard precision (SP) docking modes was employed in our evaluation. The OPLS3 force field[14] and the default values of all docking parameters were used for docking. The docking score, which included Epik state penalties in the scoring, was used for data analysis and only the best score of each pair of protein and ligand was kept for data analysis.

3. Results and discussion

3.1. Physicochemical property analysis

In this study, 17 inhibitors were used for the evaluation of the reverse docking tools. We computed ALogP, molecular weight, number of H-bond acceptors, number of H-bond donors and number of rotatable bonds of the inhibitors (Table 3). It could be found that all these ligands complied with Lipinski's rules of five[15]

Table 3. Physicochemical properties of the inhibitors in the benchmark dataset.

Inhibitors (DrugBank ID or PubChem ID)	ALogP	Molecular weight	Number of H-bond acceptors	Number of H-bond donors	Number of rotatable bonds
DB00201	–0.100	194.2	3	0	0
CID 4687	–0.306	180.2	3	1	0
DB00277	–0.306	180.2	3	1	0
DB01791	4.162	381.3	4	1	5
DB06751	4.455	397.5	5	1	9
DB01656	4.402	403.2	4	1	7
DB06237	2.165	484.0	9	3	9
DB00203	2.247	474.6	7	1	7
DB00820	2.183	389.4	4	1	1
DB02546	2.005	264.3	3	3	8
CID 444732	2.772	302.4	4	2	6
DB03127	0.783	121.2	0	2	1
DB07347	0.879	203.2	3	1	3
CID 9865515	2.798	396.4	6	3	6
CID 183797	–2.133	429.5	6	5	9
DB06605	2.865	459.5	5	1	5
DB06228	1.797	435.9	5	1	5

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

(ALogP ≤ 5, Molecular Weight ≤ 500, Number of H-bond Acceptors ≤ 10, Number of H-bond Donors ≤ 5, Number of Rotatable Bonds ≤ 10), indicating that their physico-chemical properties were proper for the evaluation since reverse docking was always applied to drug-like molecules.

3.2. Overall scoring sampling

As mentioned before, we used PDBbind refined set as the protein structure database for the overall scoring sampling. Therefore, we could estimate whether a docking score was good or bad for a given protein and ligand afterwards. The distribution of docking scoring for each inhibitor was shown in Figure 1. It could be found that the median docking score was around –6 for the most of the inhibitors, and a docking score below –9 could be seen as a high scoring. However, there were certain inhibitors which had a special scoring tendency. For example, the docking score for the Vorinostat (DB02546) was remarkably worse than other ligands and the median was around –3. It proved that the overall scoring sampling was necessary for each ligand respectively since the scoring tendency could be different.

Another interesting phenomenon was that with the Tadalafil (DB00820), none of the protein could be successfully docked. It might be attributed to its less rotatable bonds and larger molecular weight, indicating that it was less likely for this type of molecule to find a rational conformation in the binding pocket. Due to this docking failure, Tadalafil was not used as a benchmark inhibitor in this study. This phenomenon also proved that though reverse docking was a useful tool for target prediction, it might be not applicable to certain molecules.

3.3. Influence of protein conformation changes

To investigate the influence of protein conformation changes, two or three structures of the protein were used as the docking input file. We found that the docking score could fluctuate sometimes due to the changes in protein conformation though the protein type was not changed. As a demonstration, the docking scores of PDE4 (PDB ID of the proteins: 2QYN, 2QYK, 3TVX) and general inhibitors caffeine (DB00201), paraxanthine (CID4687) and theophylline (DB00277) were shown in Table 4. It could be found that for DB00201, the docking score just showed a little change, while for CID4687 and DB00277 the docking score was changed significantly. It demonstrated that the input conformation of the binding pocket could sometimes affect the docking score between the ligand and the protein and even change the conclusion of the reverse docking (for DB00277, –8.40 was a good docking score, suggesting that this protein could be a potential target, while –6.07 was a bad score, which could make this protein neglected). Figure 2 shows that binding site alignment was also applied to these protein structures, and a little conformational difference could be found, which might be the reason why different docking scores were produced.

Since the aim of molecular docking was to find a proper docking result of the protein and the ligand, we only kept the highest score among these structures for data analysis for proteins with multiple conformations in the benchmark dataset. For example, for docking the PDE4 and CID4687, only the score –8.59 was used during analysis.

From the scoring fluctuation due to the protein input conformation, it could be inferred that for reverse docking, abundant conformation of each protein would make the prediction more accurate. The more input conformations of a binding pocket were used, the more likely a proper docking score could be given. However, the limited number of experimentally determined structures was still a big challenge.

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

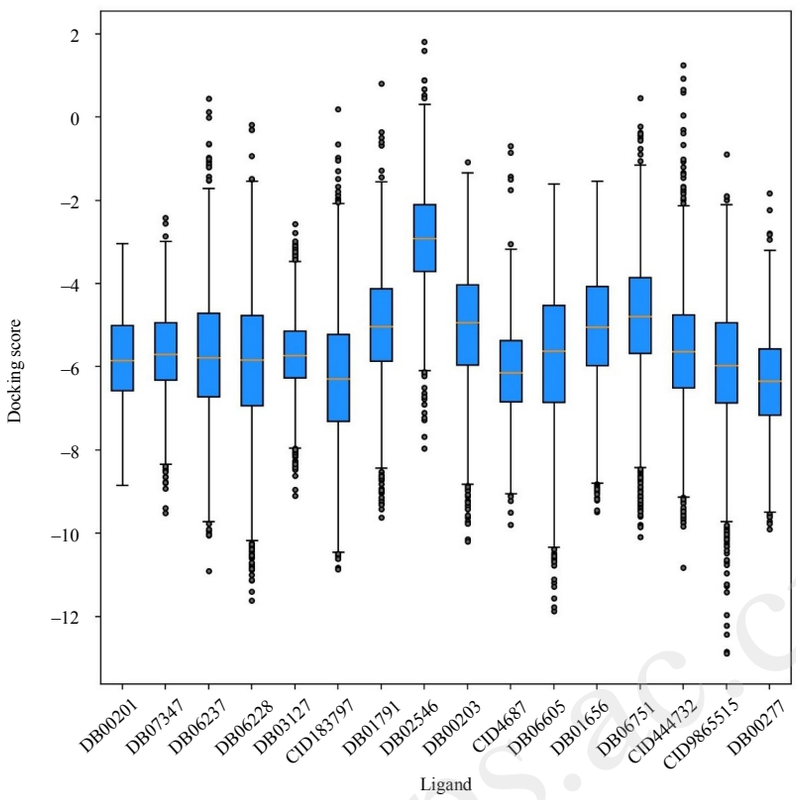

Figure 1. Distribution of docking score for each inhibitor in the benchmark dataset while using PDBbind refined set as the protein structure database.

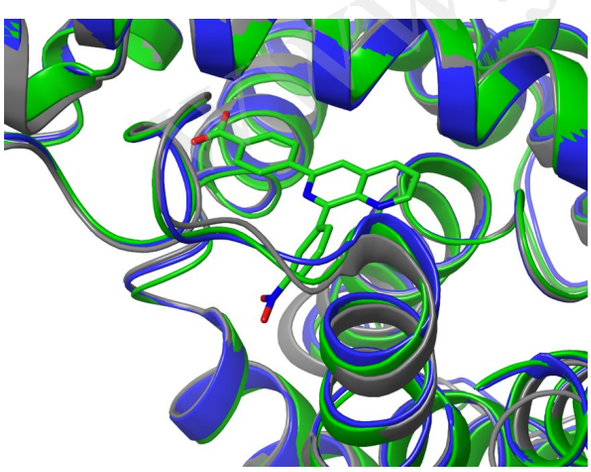

Figure 2. Binding site alignment (2QYN, 2QYK and 3TVX).

Table 4. Docking scores of PDE4 and general inhibitors.

	2QYN	2QYK	3TVX
DB00201	−7.79	−7.42	−7.89
CID4687	−6.94	−5.28	−8.59
DB00277	−8.06	−6.07	−8.40

3.4. Prediction precision of the general inhibitors

In this study, seven general inhibitors were used for target prediction (Table 2). Table 5 shows the docking score between proteins and their general inhibitors. Referring the overall scoring distribution in Figure 1, it could be found that the docking score was better than median or even better than lower quartile for most of the general inhibitors, indicating that the target prediction was quite successful. However, for certain inhibitors, such as DB02546, the prediction result was not very satisfactory. As mentioned before, the overall scoring distribution of DB02546 was also quite special, and reverse docking might not be applicable to certain molecules.

3.5. Prediction precision of the selective inhibitors

Except for Tadalafil (DB00820), nine selective inhibitors were used in this study. Theoretically, if protein A and

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

protein B are of the same target class, the selective inhibitor of protein A should only be highly scored against protein A, and a low docking score against protein B should be gained though they are similar in protein sequence. To verify whether Glide (SP) could distinguish the selective inhibitors correctly, we docked the selective inhibitors against the proteins in the target class of their known targets. The docking score was shown in Table 6.

From Table 6, it could be inferred that the correct recognition for the true target of the selective inhibitor was extremely challenging. As a demonstration, for the target class phosphodiesterases (including subclass PDE4 and PDE5), of the five selective inhibitors, only three inhibitors could get a higher score against their true targets (DB06751, DB06237 and DB00203). Similarly, for the target class histone deacetylases (including subclass HDAC2, HDAC8 and HDAC4), the inhibitor CID9865515 got a bad scoring ranking: HDAC8 > HDAC4 > HDAC2, while its true targets were HDAC8 and HDAC2. It proved that as a reverse docking tool, Glide (SP) was not able to recognize the correct target of a selective inhibitor against several proteins of the same protein class.

As reported by Wang et al.[16] and Luo et al.[17], the inter-target scoring noises, which are related to low-dimensional properties of binding pockets, have a great influence on target prediction. It means that the docking score of protein A and protein B could be not comparable due to the difference of the binding pocket properties. As mentioned above, we found that Glide (SP) had a poor performance in finding the true target for a selective inhibitor, and this might be due to the inter-target scoring noise. Though the protein sequence might be similar for the protein of the same target class, it was possible that the physicochemical properties of the binding sites were quite different.

Table 5. The docking score of proteins and their general inhibitors in the benchmark dataset. (Light gray bottom color for the correct protein-ligand pairs.)

	DB00201	CID4687	DB00277	DB02546	CID444732	DB03127	DB07347
PDE4	–7.89	–8.59	–8.40	–5.54	–8.62	–6.73	–6.22
PDE5	–7.10	–8.19	–8.29	–5.17	–6.54	–6.40	–6.57
HDAC2	–4.94	–4.83	–5.83	–2.01	–3.99	–5.61	–4.91
HDAC8	–6.64	–7.39	–6.57	–4.29	–6.97	–6.27	–5.54
HDAC4	–6.48	–7.36	–7.11	–3.49	–6.59	–6.33	–6.48
Trypsin	–6.38	–6.69	–7.04	–3.25	–5.34	–7.84	–6.52
Thrombin	–6.38	–6.53	–6.99	–2.92	–5.67	–7.43	–7.16
Factor Xa	–7.17	–7.47	–7.64	–3.63	–5.92	–6.38	–6.34

Table 6. The docking score of selective inhibitors and the proteins in the target class of their known target. (Light gray bottom color for the correct protein-ligand pairs.)

	DB01791	DB06751	DB01656	DB06237	DB00203	CID9865515	CID183797	DB06605	DB06228
PDE4	–6.76	–8.31	–6.89	–8.18	–7.38	–	–	–	–
PDE5	–8.03	–6.39	–8.50	–9.60	–7.42	–	–	–	–
HDAC2	–	–	–	–	–	–4.37	–	–	–
HDAC8	–	–	–	–	–	–7.72	–	–	–
HDAC4	–	–	–	–	–	–6.37	–	–	–
Trypsin	–	–	–	–	–	–	–8.53	–5.27	–5.09
Thrombin	–	–	–	–	–	–	–8.47	–8.35	–5.39
Factor Xa	–	–	–	–	–	–	–9.16	–9.95	–8.73

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

Table 7. Top 40 scored proteins while docking with chloroquine.

PDB ID	Protein name	Docking score	PDB ID	Protein name	Docking score
2pgz	ACh-Binding protein	–13.36	3cyz	Pheromone-binding protein ASP1	–10.05
5ioz	Transcriptional regulatory repressor protein	–11.07	6el5	Heat shock protein 90-alpha	–10.04
3v78	Transcriptional regulatory protein	–10.97	2wn9	ACh-Binding protein	–9.92
5nyh	Heat shock protein 90-alpha	–10.77	6ey8	Heat shock protein 90-alpha	–9.92
4b5d	ACh-Binding protein	–10.71	5j27	Heat shock protein 90-alpha	–9.83
5j9x	Heat shock protein 90-alpha	–10.70	4alx	ACh-Binding protein	–9.81
2wnc	ACh-Binding protein	–10.69	5fnc	Heat shock protein 90-alpha	–9.78
3t0x	Immunoglobulin variable lambda domain M8VLA4 (S55P)	–10.65	3ipu	Oxysterols receptor LXR-alpha	–9.76
5j86	Heat shock protein 90-alpha	–10.64	3t1a	Reverse transcriptase	–9.72
3u8l	ACh-Binding protein	–10.63	2yki	Heat shock protein 90-alpha	–9.67
4cwf	Heat shock protein 90-alpha	–10.45	4cwr	Heat shock protein 90-alpha	–9.62
4o09	Heat shock protein 90-alpha	–10.38	3c4h	Poly(ADP-ribose) polymerase 3	–9.53
4xir	Heat shock protein 90-alpha	–10.35	2wnj	ACh-Binding protein	–9.53
1e66	Acetylcholinesterase	–10.33	4xit	Heat shock protein 90-alpha	–9.51
5j20	Heat shock protein 90-alpha	–10.31	2yme	ACh-Binding protein	–9.46
3qdd	Heat shock protein 90-alpha	–10.31	4bny	3-Oxoacyl-(acyl-carrier-protein) reductase	–9.44
4qac	ACh-Binding protein	–10.24	3b2q	V-type ATP synthase beta chain	–9.44
5j82	Heat shock protein 90-alpha	–10.22	2ha2	Acetylcholinesterase	–9.43
3d0b	Heat shock protein 90-alpha	–10.09	4xip	Heat shock protein 90-alpha	–9.43
4o04	Heat shock protein 90-alpha	–10.07	5j6l	Heat shock protein 90-alpha	–9.40

3.6. A case study about COVID-19

Since Glide (SP) was proved to be a powerful tool on target prediction, we decided to conduct a case study about Coronavirus Disease 2019 (COVID-19). As reported, chloroquine is found to be capable of inhibiting the SARS-CoV-2 infection[18]. To search potential targets of chloroquine, we docked it against the proteins in the PDBbind refined set, and the top 40 scored (corresponding to top 1%) proteins were shown in Table 7.

From Table 7, it could be found that some proteins appeared more than once among top 40 scored proteins. One reason was that the frequency of occurrence of some popular proteins was relatively high in PDBbind refined set. On the other hand, this result also proved that the input conformation of protein would affect the docking score, since the docking scores of same protein name showed differences.

Among the top 40 scored proteins, no proteins were related to the biological activity of chloroquine. Therefore, we investigated other proteins with high docking score and coagulation factor Xa (docking score = –8.57) was found as a possible target. Coagulation factor Xa was the activated form of zymogen factor X, which was related to the formation of prothrombinase complex and clot. Inhibitors of coagulation factor Xa could control thrombin levels, therefore a candidate for abnormal coagulation[19]. As reported, critical patients of COVID-19 always show coagulation dysfunction, and the formation of hyaline thrombus is found in pulmonary capillary. The application of anticoagulant is also suggested, showing a good effect on correcting local clotting disorders in the lungs[20]. We suspected that coagulation factor Xa was one target of chloroquine and therefore made chloroquine effective to some extent.

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

4. Conclusions

As an important tool for target prediction, reverse docking remains essential to drug development. The comprehensive evaluation of reverse docking tools is quite a significant task both for the users and developers of this research field. For different ligands, the overall scoring tendency could have their own characteristics, while the conformation of the binding site could also affect the docking score. Glide (SP) was capable for finding potential general targets among thousands of proteins. However, it was not able to recognize the true target for a selective inhibitor correctly among proteins of the same protein class. Therefore, we recommend the further development of reverse docking tools and rectification of inter-target scoring bias.

Acknowledgements

This research was supported by the National Key Research and Development Project (Grant No. 2019YFC1708900), the National Natural Science Foundation of China (Grant No. 81872730, 81673279, 21772005), National Major Scientific and Technological Special Project for Significant New Drugs Development (Grant No. 2018ZX09735001-003, 2019ZX09201005-001, 2019ZX09204-001) and Beijing Natural Science Foundation (Grant No. 7202088, 7172118).

References

[1] Sydow, D.; Burggraaff, L.; Szengel, A.; van Vlijmen, H.W.T.; IJzerman, A.P.; van Westen, G.J.P.; Volkamer, A. Advances and challenges in computational target prediction. *J. Chem. Inf. Model.* **2019**, *59*, 1728–1742.

[2] da Matta, C.B.B.; de Queiroz, A.C.; Santos, M.S.; Alexandre-Moreira, M.S.; Gonçalves, V.T.; de Nigris del Cistia, C.; Sant'Anna, C.M.R.; DaCosta, J.B.N. Novel dialkylphosphorylhydrazones: Synthesis, leishmanicidal evaluation and theoretical investigation of the proposed mechanism of action. *Eur. J. Med. Chem.* **2015**, *101*, 1–12.

[3] Lim, T.G.; Lee, S.Y.; Huang, Z.N.; Lim, D.Y.; Chen, H.Y.; Jung, S.K.; Bode, A.M.; Lee, K.W.; Dong, Z.G. Curcumin suppresses proliferation of colon cancer cells by targeting CDK2. *Cancer Prev. Res.(Phila).* **2014**, *7*, 466–474.

[4] Wang, Z.Y.; Luo, S.Q.; Wan, Z.; Chen, C.Y.; Zhang, X.N.; Li, B.B.; Huang, G.L.; Chen, L.Y.; He, Z.W.; Huang, Z.N. Glabridin arrests cell cycle and inhibits proliferation of hepatocellular carcinoma by suppressing braf/MEK signaling pathway. *Tumor Biol.* **2016**, *37*, 5837–5846.

[5] Lapillo, M.; Tuccinardi, T.; Martinelli, A.; Macchia, M.; Giordano, A.; Poli, G. Extensive reliability evaluation of docking-based target-fishing strategies. *Int. J. Mol. Sci.* **2019**, *20*, 1023.

[6] Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; *Perry*, J.K.; Shaw, D.E.; Francis, P.; Shenkin, P.S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.

[7] Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.

[8] Schomburg, K.T.; Rarey, M. Benchmark data sets for structure-based computational target prediction. *J. Chem. Inf. Model.* **2014**, *54*, 2261–2274.

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

[9] Liu, Z.H.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.X.; Nie, W.; Liu, Y.C.; Wang, R.X. PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics*. **2015**, *31*, 405–412.

[10] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

[11] Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.F.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082.

[12] Kim, S.; Chen, J.; Cheng, T.J.; Gindulyte, A.; He, J.; He, S.Q.; Li, Q.L.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E.E. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **2019**, *47*, D1102–D1109.

[13] Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27*, 221–234.

[14] Harder, E.; Damm, W.; Maple, J.; Wu, C.J.; Reboul, M.; Xiang, J.Y.; Wang, L.L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; Kaus, J.W.; Cerutti, D.S.; Krilov, G.; Jorgensen, W.L.; Abel, R.; Friesner, R.A. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281–296.

[15] Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.

[16] Wang, W.; Zhou, X.; He, W.L.; Fan, Y.; Chen, Y.Z.; Chen, X. The interprotein scoring noises in glide docking scores. *Proteins: Struct. Funct. Bioinform.* **2012**, *80*, 169–183

[17] Luo, Q.Y.; Zhao, L.; Hu, J.X.; Jin, H.W.; Liu, Z.M.; Zhang, L.R. The scoring bias in reverse docking and the score normalization strategy to improve success rate of target fishing. *PLoS One.* **2017**, *12*, e0171433.

[18] Kakodkar, P.; Kaka, N.; Baig, M.N. A comprehensive literature review on the clinical presentation, and management of the pandemic coronavirus disease 2019 (COVID-19). *Cureus.* **2020**, *12*, e7560.

[19] Ewing, W.R.; Becker, M.R.; Manetta, V.E.; Davis, R.S.; Pauls, H.W.; Mason, H.; Choi-Sledeski, Y.M.; Green, D.; Cha, D.; Spada, A.P.; Cheney, D.L.; Mason, J.S.; Maignan, S.; Guilloteau, J.P.; Brown, K.; Colussi, D.; Bentley, R.; Bostwick, J.; Kasiewski, C.J.; Morgan, S.R.; Leadley, R.J.; Dunwiddie, C.T.; Perrone, M.H.; Chu, V. Design and structure-activity relationships of potent and selective inhibitors of blood coagulation factor Xa. *J. Med. Chem.* **1999**, *42*, 3557–3571.

[20] Li, H.C.; Ma, J.; Zhang, H.; Cheng, Y.; Wang, X.; Hu, Z.W.; Li, N.; Deng, X.R.; Zhang, Y.; Zheng, X.Z.; Yang, F.; Weng, H.Y.; Dong, J.P.; Liu, J.W.; Wang, Y.Y.; Liu, X.M. Thoughts and practice on the treatment of severe and critical new coronavirus pneumonia. *Chin. J.Tuberculosis Res. Dis.* **2020**, *43*, E038.

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn

反向对接的选择性与非选择性抑制剂找靶评价: Glide个案研究

李明娜[1], 吴兴[2], 张亮仁[1*], 刘振明[1*]

1. 北京大学医学部 药学院 天然药物及仿生药物国家重点实验室, 北京 100191

2. 横店集团控股有限公司, 浙江 杭州 310007

摘要: 反向对接作为一种靶标预测的有效工具, 有许多方面仍待探究。对反向对接软件的客观评价可以帮助我们更好地了解这些工具的长处与短处, 并在靶标预测的过程中起到指导作用。本研究中, 我们评估了Glide (SP)针对选择性抑制剂与非选择性抑制剂的靶标预测能力。结果说明针对不同配体, 对接打分的倾向可能存在差异, 因此总体打分抽样对帮助我们更好地理解某对配体受体间的对接打分具有重要的意义。另外, 对接时结合口袋的输入构象对对接结果存在一定的影响。Glide (SP)显示出较好的对非选择性抑制剂的靶标预测能力。然而, 其对于选择性抑制剂的靶标预测准确度相对较低, 说明该软件不适用于这方面的工作。针对COVID-19的案例研究表面凝血因子Xa可能是氯喹的潜在靶点。因此, 我们认为对反向对接软件的进一步开发与靶点间打分差异的修正十分必要。

关键词: 反向对接; 靶标预测; 软件评估

www.jcps.ac.cn

Copyright © 2020 Journal of Chinese Pharmaceutical Sciences, School of Pharmaceutical Sciences, Peking University http://www.jcps.ac.cn