The Study of Box Size Effect on Shape-based Docking of Anti-tyrosinase Compounds with ArgusLab 4.0.1 Program

Prasan Tangyuenyongwatana[1, *], Pavinee Sengsunt[1], Sanhajutha Puangmala[1], Napa Boonma[1] and Wandee Gritsanapan[2].

[1]College of Oriental Medicine, Rangsit university, Phathum Thani, Thailand
[2]Phytomedical research, Rimklongprapra Road, Bangsue, Bangkok, Thailand
[*]Corresponding author, E-mail: prasan.t@rsu.ac.th

Abstract

ArgusLab 4.0.1 program is a free molecular docking program in Windows platform that composes of shape-based (ArgusDock) and genetic algorithm (GA Dock) modes. The easy docking procedure and promising results of this program make it a popular one. For shape-based mode, some compounds were rejected or undocked from the process with the default calculated box size. This research focused on the box size effect involved in docking results. The study started with calculating the radius of gyration of anti-tyrosinase compounds in a database which obtained 4.99 Å and made the smallest box size equaling to 14.4885 Å on each size. Next, the cubic box sizes from 15 × 15 × 15 to 22 × 22 × 22 Å by a small incremental size of 1 Å were used to dock the database set to find the good correlation coefficient between anti-tyrosinase activity and binding energy. From the experimental result, the box size 15 × 15 × 15 to 17 × 17 ×17 Å showed some rejection or undocked compounds in the database set. For 18 × 18 × 18 to 22 × 22 × 22 Å, all compounds were docked well, and the correlation coefficient ranged from 0.7415 – 0.8303 and the rational one was 18 × 18 × 18 Å which had a good correlation coefficient of 0.7879 and agreed with the good anti-tyrosinase activity than other box sizes. In conclusion, a good and practical box size for any set of compounds for shape-based docking should be tested before proceeding with the docked screening.

Keywords**:** *ArgusLab, ArgusDock, Shape-based docking, Radius of gyration, Box size, Tyrosinase*

1. Introduction

Computational methods have grown in importance as a result of developments in information technology and are now a crucial part of contemporary biological research. In these few decades, bio-algorithms have rapidly developed and are widely applied to the molecular modeling arena. For unknown proteins, amino acid sequences can be used to accurately simulate protein tertiary structures to help infer their molecular functions (Meng et al., 2011). In addition, these computer-generated protein models can be used to identify possible ligand-binding pockets, which can then be utilized to find new drug targets (Halperin et al., 2002). Among the many technologies created to date, molecular docking has many uses in the field of drug design, such as supporting the identification of novel lead compounds and drug repositioning (Naqvi et al., 2019). Because of this, molecular docking is commonly used as a crucial component of many ongoing drug research efforts. It holds enormous promise for accelerating drug discovery.

Predicting non-covalent interactions between a ligand and its receptor protein is the aim of molecular docking. The binding pose prediction and the estimation of binding affinity are two key elements of a conventional docking procedure (Tang et al., 2017). It is crucial to remember that both molecules may alter conformation when ligands bind to their receptor proteins (Du et al., 2016). However, docking is computationally complicated because of the high number of rotatable bonds, or degrees of freedom, in molecules (Feyza et al., 2022). The docking problem can often be solved using one of two methods. The first strategy takes a geometrical or shape-based approach, and the second is the structural-based approach (Ferreira et al., 2015). For structural-based molecular docking approach programs, such as AutoDock or Autodock vina, are widely used for identification of new lead compounds for specific targets (Trott &

Published online: Copyright © 2016-2023 Rangsit University

Olson, 2010). In contrast, the shape-based method is based on the assumption that the molecular surfaces of the receptor and the ligand must match in order for the molecules to bind to each other with high affinity (Kumar et al., 2015).

The shape of molecules has been demonstrated to be an effective virtual screening performance indicator because it offers greater specificity than docking scores (Bhutoria et al., 2016). In 2004, Mark Thompson introduced ArgusLab docking program to the community. ArgusLab is free distributed molecular docking software in Windows platform. This program can perform a shaped-base and genetic algorithm (GA) docking which is very convenient to carry out the docking experiment. Even those new to molecular docking can use ArgusLab's user-friendly interface because it is simple to use (Bitencourt-Ferreira & de Azevedo, 2019). In contrast to AutoDock 4.0 or Autodock vina which requires the user to know and specify the coordinates of the binding site in order to perform docking, ArgusLab offers new researchers in molecular docking a quick and reliable method of binding site optimization that allows the program to locate binding sites automatically and speed up the docking process (Tangyuenyongwatana & Jongkon, 2016). Oda and Takahashi (2009) used ArgusLab to determine the binding free energy between proteins and ligands. The findings demonstrated that ArgusLab was useful for virtual screening in addition to posture building and pose selection (Oda & Takahashi, 2009).

Typically, molecular docking involves searching a user-defined docking search space for potential ligand binding conformations. The choice of a suitable search space, or docking box, is a difficult issue. While a sufficiently broad docking area might provide an excessive number of irrelevant binding poses, a small search space might result in an insufficient number of conformations. Thus, an optimally confined search space is critical for the success of molecular docking. Many current docking protocols offer a default method for estimating the box size. Currently, available docking packages often come with a default protocol for calculating the box size, however, many of these procedures have not been systematically evaluated (Feinstein & Brylinski, 2015).

For ArgusLab suite with ArgusDock, there is no reported study about the box size estimation, and we encounter that some molecules in ArgusDock docking experiment using default box size were rejected from the experiment. Which means the docking process was stopped or halted. The objective of this research was to demonstrate the effect of box size on docking results using ArgusDock, a shape-based algorithm, in a variety of box sizes to perform docking and also showed the efficiency of shape-based approach to show the correlation of binding energy related to the enzyme inhibition data.

2. Objective

The objective of this research was to demonstrate the effect of box size on docking results using ArgusDock, a shape-based approach, in a variety of box sizes to perform molecular docking.

3. Materials and Methods

The tyrosinase crystal structure (PDB ID: 4P6T) was downloaded from Protein Data Bang to ArgusLab program, and the ligand 4-(2-hydroxyethyl)phenol was chosen, “centered,” and “hydrogen atoms added.” After that, the binding site was assigned by clicking a “create the binding site for the group” button. The 2D structures of anti-tyrosinase compounds, oxyresveratrol, resveratrol, kojic acid, hydroquinone, HS1713, HS1784, HS1791, HS1792, and HS1793 (Figure 1) were prepared by ChemSketch program (https://www.acdlabs.com/resources/free-chemistry-software-apps/chemsketch-freeware/) and then converted to 3D structure with energy minimization with MMFF99 by Avogrado program (https://avogadro.cc). All structures were collected in an SDF file which was prepared by VEGAZZ software (https://www.ddl.unimi.it/cms/index.php?Software_projects:VEGA_ZZ) (Tangyuenyongwatana & Jongkon, 2016). The database was selected by clicking a database docking button on the ArgusLab user interface.

The docking engine employed the exhaustive search method for the ArgusDock settings. Grids were built to cover the binding location and a torsion tree was used to represent the flexible ligand. On a search point in the enzyme pocket, a group of bound atoms without rotatable bonds or ligand root nodes

Published online: Copyright © 2016-2023 Rangsit University

was inserted, and a collection of interesting and energetically advantageous rotations was produced. Torsion searches were conducted for each search and the poses that were found after the torsion search were graded. The final set of postures was subjected to coarse reduction, re-clustering, and ranking (Tangyuenyongwatana & Gritsanapan, 2017), with the N-lowest energy, poses being reserved. The default unsymmetrical binding site box size of 4-(2-hydroxyethyl)phenol was 15.293 × 18.263 × 13.633 Å and the grid resolution was 0.4 Å. For the test box size option, the radius of gyration of each compound was calculated to guide the suitable box size for a promising docking experiment by using an online calculator from http://www.scfbio-iitd.res.in/software/proteomics/rgnew1.jsp website. Next, the box sizes were varied and selected from 15 × 15 × 15 to 22 × 22 × 22 Å by a small incremental size of 1 Å to cover the binding site for docking. ArgusDock ran with the AScore scoring function enabled.

The "Dock" and "Flexible" ligand docking modes were used as the docking calculation type. The screening output file was obtained at the conclusion of the docking. The binding values' free energies were sorted from best binding energy to worst binding energy. The docking pose was exported by saving the result as pdb file. The interaction of the ligand with the binding site in relation to the orientation of the 4-(2-hydroxyethyl)phenol was demonstrated using Discovery Studio Visualizer 4.0.

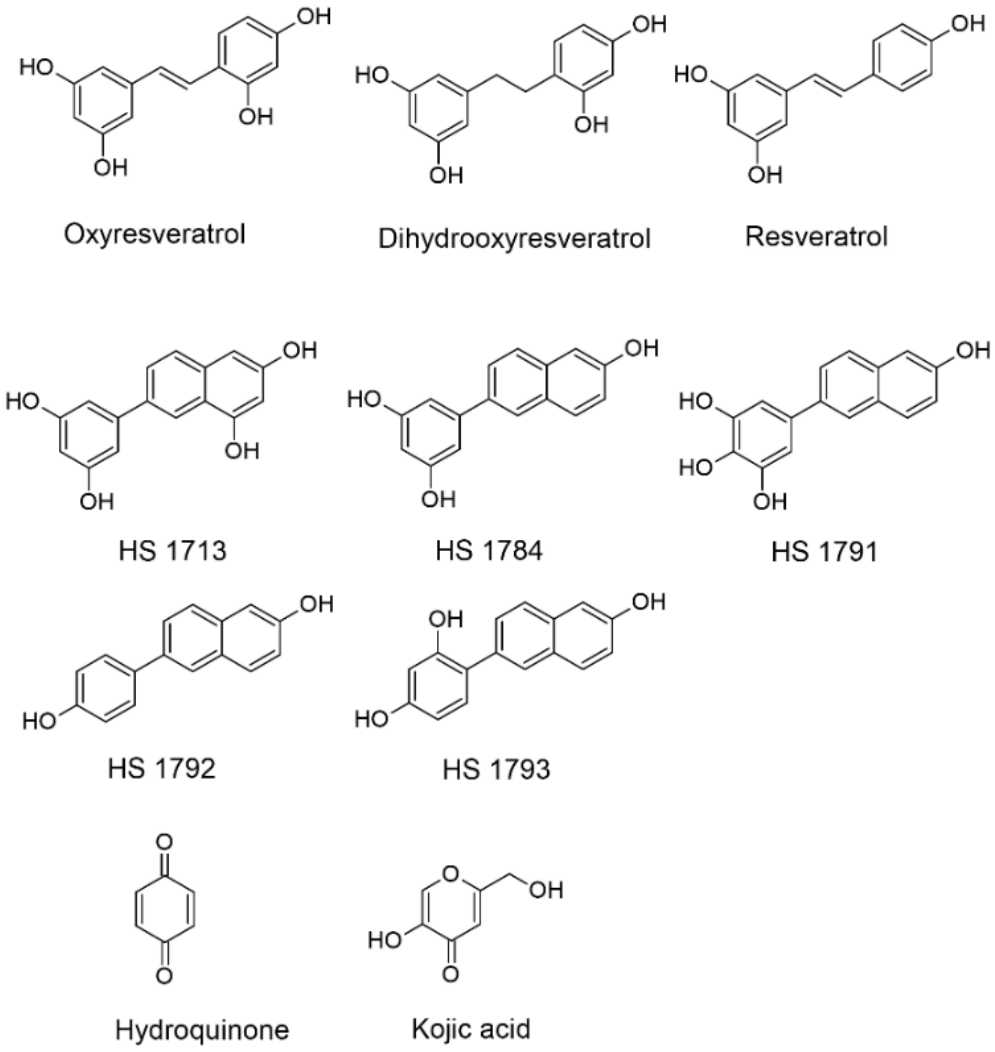

Figure 1 Structures of anti-tyrosinase compounds

[3]

Published online: Copyright © 2016-2023 Rangsit University

4. Results and Discussion

First, to support the box size selection, the radius of gyration of each docking compound was calculated from the online calculator (http://www.scfbio-iitd.res.in/software/proteomics/rgnew1.jsp) and showed the results in Table 1.

Table 1 Radius of gyration of anti-tyrosinase compounds

Compound	Radius of gyration (R_g, Å)
Oxyresveratrol	3.6239
Resveratrol	4.9961
HS 1713	3.7512
HS 1784	4.9029
HS 1791	3.4444
HS 1792	3.4136
HS 1793	2.8854
Hydroquinone	1.5397
Kojic acid	2.2293

Feinstein and Brylinski (2015) operate a systematic study of ligand binding poses generated by AutoDock Vina exhibiting the highest accuracy of docking when the dimensions of the search space are 2.9 times larger than the radius of gyration of the docking compound (Feinstein & Brylinski, 2015). From the test, compounds in this experiment showed the radius of gyration ranging from 1.5397 to 4.9961 Å, the box size of this set of compounds should be larger than the largest compound which was 4.9961 × 2.9 equal to 14.4885 Å. That means the box size should be larger than 14.4885 Å in each size.

From the study, the docking results with ArgusDock in ArgusLab 4.0.1 program were shown in Table 2-5. The linear correlations between tyrosinase inhibitory concentrations (IC_{50}) in different box sizes were reported in Table 6 and an example of a graph plot was depicted in Figure 2.

Table 2 Docking energy of anti-tyrosinase compounds box size default, 15 × 15 × 15 and 16 × 16 × 16 Å

Compound	IC_{50} (µM)	Docking energy (kcal/mol) 15.29 × 18.26 × 13.63 Å	Docking energy (kcal/mol) 15 × 15 × 15 Å	Docking energy (kcal/mol) 16 × 16 × 16 Å
Oxyresveratrol	12.7	-6.81	N/A	N/A
Resveratrol	26.63	-6.18	N/A	N/A
HS 1713	0.49	N/A	N/A	N/A
HS 1784	16.52	N/A	N/A	N/A
HS 1791	2.95	N/A	N/A	N/A
HS 1792	6.4	N/A	N/A	N/A
HS 1793	0.034	N/A	N/A	N/A
Hydroquinone	33.48	-5.51	-5.59	-5.74
Kojic acid	38.24	-5.77	-5.93	-6.01

N/A = not available

Table 3 Docking energy of anti-tyrosinase compounds at box size 17 × 17 × 17 and 18 × 18 × 18 Å

Compound	IC_{50} (µM)	Docking energy (kcal/mol) 17 × 17 × 17 Å	Docking energy (kcal/mol) 18 × 18 × 18 Å
Oxyresveratrol	12.7	-7.16	-7.35
Resveratrol	26.63	-7.50	-7.19
HS 1713	0.49	N/A	-8.20
HS 1784	16.52	N/A	-7.70
HS 1791	2.95	-8.04	-8.30

Published online: Copyright © 2016-2023 Rangsit University

HS 1792	6.4	-8.99	-7.69
HS 1793	0.034	-8.67	-10.13
Hydroquinone	33.48	-5.74	-5.68
Kojic acid	38.24	-5.89	-5.96

N/A = not available

Table 4 Docking energy of anti-tyrosinase compounds at box size 19 × 19 × 19 and 20 × 20 × 20 Å

Compound	**IC_{50} (μM)**	**Docking energy (kcal/mol) 19 × 19 × 19 Å**	**Docking energy (kcal/mol) 20 × 20 × 20 Å**
Oxyresveratrol	12.7	-7.68	-7.63
Resveratrol	26.63	-7.87	-7.77
HS 1713	0.49	-9.24	-9.20
HS 1784	16.52	-9.54	-9.59
HS 1791	2.95	-10.07	-9.50
HS 1792	6.4	-9.64	-9.13
HS 1793	0.034	-9.31	-8.95
Hydroquinone	33.48	-5.81	-5.87
Kojic acid	38.24	-5.93	-6.10

Table 5 Docking energy of anti-tyrosinase compounds at box size 21 × 21 × 21 and 22 × 22 × 22 Å

Compound	IC_{50} (μM)	Docking energy (kcal/mol) **21 × 21 × 21** Å	Docking energy (kcal/mol) **22 × 22 × 22** Å
Oxyresveratrol	12.7	-7.57	-8.00
Resveratrol	26.63	-7.74	-8.01
HS 1713	0.49	-8.91	-8.78
HS 1784	16.52	-8.93	-9.19
HS 1791	2.95	-9.14	-8.91
HS 1792	6.4	-9.02	-9.50
HS 1793	0.034	-9.32	-9.25
Hydroquinone	33.48	-5.85	-5.80
Kojic acid	38.24	-5.89	-6.02

N/A = not available

Table 6 Linear regression equation and correlation coefficient (*r*) of docking result in each box size

Box size (Å)	**Linear regression equation**	**Correlation coefficient (*r*)**
15.29 × 18.26 × 13.63	N/A	N/A
15 × 15 × 15	N/A	N/A
16 × 16 × 16	N/A	N/A
17 × 17 × 17	N/A	N/A
18 × 18 × 18	Y = 0.0859x – 8.9485	0.7879
19 × 19 × 19	Y = 0.0981x - 9.8409	0.7778
20 × 20 × 20	Y = 0.0851x – 9.4935	0.7415
21 × 21 × 21	Y = 0.0864x – 9.3608	0.8303
22 × 22 × 22	Y = 0.0829x – 9.4280	0.7587

N/A = not available

Published online: Copyright © 2016-2023 Rangsit University

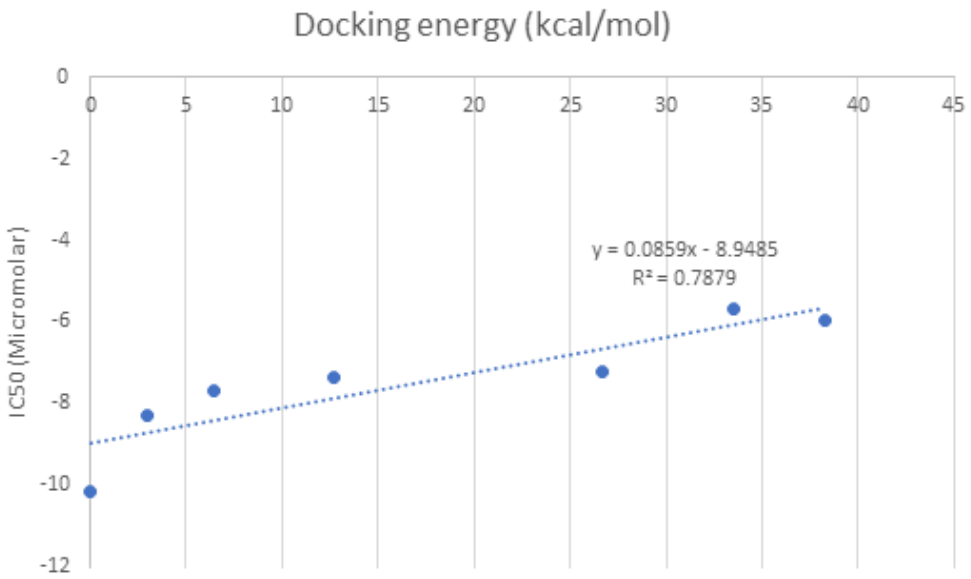

Figure 2 Graph plot between IC_{50} of anti-tyrosinase activity and docking energy of anti-tyrosinase compounds

Normally, ArgusDock needs a receptor that is bound with a ligand or inhibitor to calculate the box size. In the default mode, when clicking calculate box size button, ArgusDock will automatically calculate the fit and unsymmetrical box size. For the ligand of 4P6T enzyme, ArgusDock calculates a box size as 15.29 × 18.26 × 13.63 Å. The anti-tyrosinase ligands were docked in this fit box size and some compounds were rejected. The cause of this phenomenon involved the size of the molecule which was too big. In our experiment, the protocol started with symmetrical (x, y, z) 15 × 15 × 15 to 22 × 22 × 22 Å box sizes to explore the limitation of box sizes on molecular docking in the shape-based algorithm which has never been reported before. From Table 1, the 15 × 15 × 15 and 16 × 16 × 16 Å box sizes showed the almost same result that most compounds were rejected from the docking process. Hydroquinone and kojic acid which are small molecules can be docked into the receptor site of tyrosinase enzyme with similar binding energy.

When the box size was increased to 17 × 17 × 17 and 18 × 18 × 18 Å, most of the compounds could be docked and gave the binding energy except for HS 1713 and HS 1784 which were rejected in 17 × 17 × 17 Å box size. For 19 × 19 × 19 to 22 × 22 × 22 Å box sizes, all compounds were docked completely without any rejection. From the result, it can be concluded that the shape-based approach needs the primary docked screening which is an important task to process with ArgusDock for better docking results.

For the best box size to be used in ArgusDock experiment, the correlation coefficient (r) between IC_{50} and docked energy of the compounds (Table 5) was the key parameter to consider. However, when considering the correlation coefficient value of 18 × 18 × 18 to 22 × 22 × 22 Å, the correlation coefficient values look very close to each other (0.7415-0.7879) except for 21 × 21 × 21 Å box size which the value was higher than others. The 18 × 18 × 18 Å box size was chosen because the binding energy of HS1792 and HS1793 (-7.69 and -10.13 kcal/mol) correlated well with IC_{50} (6.4 and 0.034 µM) and the analysis of the best pose and bonding interaction with the receptor was accepted with this box size as shown in Figure 3-5. The 4P6T ligand exhibited interaction with the receptor amino acids (HIS208, VAL218, and ALA221). HS1792 had one crucial amino acid (VAL218) interaction with the receptor while HS1763 showed two crucial amino acids (HIS208 and VAL218) interaction with the receptor.

Published online: Copyright © 2016-2023 Rangsit University

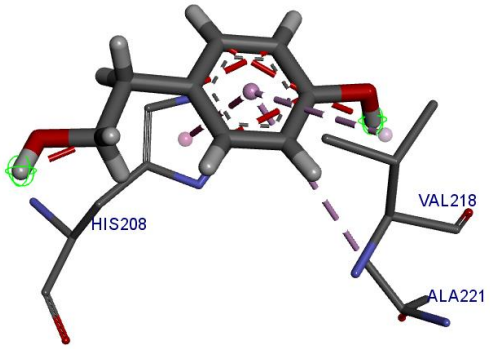

Figure 3 The 4p6t ligand exhibited interaction with HIS208, VAL218, and ALA221

5. Conclusion

Shape-based docking algorithm in ArgusLab 4.0.1 program is another tool for finding active molecules by molecular docking. The easy docking procedure and promising results of this program make it a popular one. For shape-based mode, some compounds were rejected or undocked from the process with the default box size. This research focused on calculating the radius of gyration and used symmetrical box sizes to find the practical one which was 18 × 18 × 18 Å. A good and practical box size for any set of compounds should be tested before proceeding with the docked screening. Further study of this project is to test the bigger data set with various sizes of molecules.

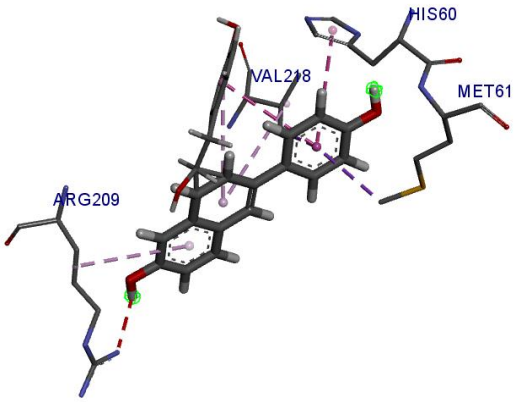

Figure 4 HS1792 showed one crucial amino acid (VAL218) interaction

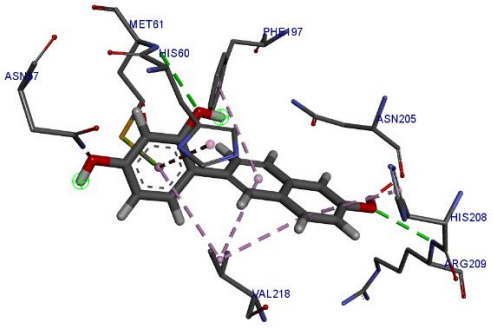

Figure 5 HS1763 showed two crucial amino acids (HIS208 and VAL218) interaction with the receptor.

Published online: Copyright © 2016-2023 Rangsit University

6. Acknowledgements

This research was supported by the College of Oriental Medicine research fund for the year 2023.

7. References

Bhutoria, S., Das, B., & Ghoshal, N. (2016). Molecular Shape Analysis-Guided Virtual Screening Platform for Adenosine Kinase Inhibitors. *Bioinformatics and Biology Insights*, *10*, 97–103. https://doi.org/10.4137/BBI.S38430

Bitencourt-Ferreira, G., & de Azevedo, W. F. (2019). Molecular docking simulations with ArgusLab. *Methods in Molecular Biology*, *2053*, 203–220. https://doi.org/10.1007/978-1-4939-9752-7_13

Du, X., Li, Y., Xia, Y. L., Ai, S. M., Liang, J., Sang, P., Ji, X. L., & Liu, S. Q. (2016). Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences*, *17*(2). https://doi.org/10.3390/IJMS17020144

Feinstein, W. P., & Brylinski, M. (2015). Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *Journal of Cheminformatics*, *7*(1), 1–10. https://doi.org/10.1186/S13321-015-0067-5/FIGURES/5

Ferreira, L. G., dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. *Molecules (Basel, Switzerland)*, *20*(7), 13384–13421. https://doi.org/10.3390/MOLECULES200713384

Feyza, M. S., Selin, S., Ece, A. S., Feyza, M. S., Selin, S., & Ece, A. S. (2022). *Fundamentals of Molecular Docking and Comparative Analysis of Protein–Small-Molecule Docking Approaches*. https://doi.org/10.5772/INTECHOPEN.105815

Halperin, I., Ma, B., Wolfson, H., & Nussinov, R. (2002). Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins*, *47*(4), 409–443. https://doi.org/10.1002/PROT.10115

Kumar, R., Garg, P., & Bharatam, P. v. (2015). Shape-based virtual screening, docking, and molecular dynamics simulations to identify Mtb-ASADH inhibitors. *Journal of Biomolecular Structure & Dynamics*, *33*(5), 1082–1093. https://doi.org/10.1080/07391102.2014.929535

Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, *7*(2), 146–157. https://doi.org/10.2174/157340911795677602

Naqvi, A. A. T., Mohammad, T., Hasan, G. M., & Hassan, Md. I. (2019). Advancements in Docking and Molecular Dynamics Simulations Towards Ligand-receptor Interactions and Structure-function Relationships. *Current Topics in Medicinal Chemistry*, *18*(20), 1755–1768. https://doi.org/10.2174/1568026618666181025114157

Oda, A., & Takahashi, O. (2009). Validation of ArgusLab Efficiencies for Binding Free Energy Calculations. *Chem-Bio Informatics Journal*, *9*, 52–61. http://www.cbi.or.jp

Tang, Z., Roberts, C. C., & Chang, C. E. A. (2017). Understanding ligand-receptor non-covalent binding kinetics using molecular modeling. *Frontiers in Bioscience (Landmark Edition)*, *22*(6), 960. https://doi.org/10.2741/4527

Tangyuenyongwatana, P., & Gritsanapan, W. (2017). Virtual screening for novel 1-deoxy-D-xylulose-5-phosphate reductoisomerase inhibitors: A shape-based search approach. *Thai Journal of Pharmaceutical Sciences*, *41*(1), 1-5.

Tangyuenyongwatana, P., & Jongkon, N. (2016). Molecular docking study of tyrosinase inhibitors using ArgusLab 4.0.1. *Thai Journal of Pharmaceutical Sciences*, *40*(1), 21-25.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, *31*(2), 455. https://doi.org/10.1002/JCC.21334

Published online: Copyright © 2016-2023 Rangsit University