CASE REPORT

ISSN 2307-9266 e-ISSN 2413-2241

Scientific and Practical Journal
PHARMACY & PHARMACOLOGY

NUMBER OF RUNS VARIATIONS ON AUTODOCK 4 DO NOT HAVE A SIGNIFICANT EFFECT ON RMSD FROM DOCKING RESULTS

M.R.F. Pratama[1,2], Siswandono[1]

[1] Airlangga University
Dr. Ir. Soekarno St. Campus C, Mulyorejo, Surabaya, East Java, Indonesia 60115
[2] Muhammadiyah Palangkaraya University
RTA Milono St. Km. 1.5, Pahandut, Palangka Raya, Central Kalimantan, Indonesia 73111

E-mail: prof.sis@ff.unair.ac.id

Received 15 May 2019 Accepted 11 Nov 2020

The aim. The number of runs in the docking process with AutoDock 4 is known to play an important role in the validity of the results obtained. The greater the number of runs it is often associated with the more valid docking results. However, it is not known exactly how the most ideal runs in the docking process with AutoDock 4. This study aims to determine the effect of the number of runs docking processes with AutoDock 4 on the validity of the docking results.
Materials and methods. The method used is the redocking process with AutoDock 4.2.6. The receptor used is an estrogen receptor with ligand reference estradiol (PDB ID 1GWR). Variations were made on the number of runs from 10 to 100 in multiples of 10. The parameters observed were RMSD, free energy of binding, inhibition constants, amino acid residues, and the number of hydrogen bonds.
Results. All experiments produce identical bond free energy, where the maximum difference in inhibition constant is only 0.06 nM. The lowest RMSD is indicated by the number of runs of 60, with a RMSD value of 0.942. There is no linear relationship between the number of runs and RMSD, with R in the linear equation of 0.4607.
Conclusion. Overall, the number of runs does not show a significant contribution to the validity of the results of docking with AutoDock 4. However, these results have only been proven with the receptors used.
Keywords: AutoDock 4; docking; number of runs; RMSD, variation

ИЗУЧЕНИЕ ВЛИЯНИЯ КОЛИЧЕСТВА ЗАПУСКОВ AUTODOCK 4 НА СРЕДНЕКВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ РЕЗУЛЬТАТОВ ДОКИНГА

М.Р.Ф. Пратама[1,2], Сисвандоно[1]

[1] Университет Айрланга
60115, Индонезия, Восточная Ява, г. Сурабая, Доктор Ир. Сукарно, ул. Кампус С, Мулйореё
[2] Университет Мухаммадии Палангкарая
73111, Индонезия, Центральный Калимантан, г. Палангка Рая, Пахандут РТА Милоно, ул. 1.5 Км

E-mail: prof.sis@ff.unair.ac.id

Получена 15.05.2019 Принята к печати 11.11.2020

Цель. Известно, что количество запусков докинга с AutoDock 4 играет важную роль в достоверности получаемых результатов. Чем больше количество запусков, тем больше достоверных результатов докинга. Однако точно не известно, как наиболее оптимально работает докинг с AutoDock 4. Это исследование направлено на определение влияния количества запусков процесса докинга в AutoDock 4 на достоверность результатов.
Материалы и методы. В качестве метода исследования использовали процесс редокинга с AutoDock 4.2.6. Исполь-

For citation: M.R.F. Pratama, Siswandono. Number of runs variations on AutoDock 4 do not have a significant effect on RMSD from docking results. *Pharmacy & Pharmacology*. 2020;8(6):476-480. **DOI:** 10.19163/2307-9266-2020-8-6-476-480

Для цитирования: М.Р.Ф. Пратама, Сисвандоно. Изучение влияния количества запусков AutoDock 4 на среднеквадратическое отклонение результатов докинга. *Фармация и фармакология.* 2020;8(6):476-480. **DOI:** 10.19163/2307-9266-2020-8-6-476-480

КРАТКОЕ СООБЩЕНИЕ
DOI: 10.19163/2307-9266-2020-8-6-476-480

зуемый рецептор представляет собой рецептор эстрогена с эталонным лигандом эстрадиола (PDB ID 1GWR). Варьировали количество прогонов от 10 до 100, кратные 10. Наблюдаемыми параметрами были RMSD, свободная энергия связывания, константы ингибирования, аминокислотные остатки и количество водородных связей.

Результаты. Все эксперименты вырабатывают идентичную свободную энергию связи, где максимальная разница в константе ингибирования составляет всего 0,06 нМ. Наименьшее среднеквадратичное отклонение (RMSD) определяется количеством прогонов, равным 60, при значении среднеквадратичного отклонения, равного 0,942. Установлено, что между количеством прогонов и пространственным выравниванием (RMSD) нет линейной зависимости, так как коэффициент корреляции (R) равен 0,4607.

Заключение. В целом, количество запусков не оказывает значительного вклада в достоверность результатов процесса докинга с AutoDock 4. Однако эти результаты справедливы только с рецепторами, использованными в данном исследовании.

Ключевые слова: AutoDock 4; докинг; количество запусков; RMSD; вариация

INTRODUCTION

Molecular docking (or simply called docking) is one of the most common *in silico* experimental methods. The ease in the process of preparation, implementation, and analysis of results is one of the attractions of the method [1]. However, docking is one simple, yet complex method. As one of the most popular *in silico* methods, docking has several advantages and disadvantages. In addition to the speed and practicality, docking also offers a fairly complete observation to analyze complexes between drug-receptors, which are even very difficult to do in vitro or in vivo. However, the results obtained generally have a large degree of variation, where reproducibility is one of the important issues in the level of confidence in the results of docking [2]. For this reason, the validation of docking results is absolutely necessary.

The results of the validation docking are influenced by various factors, both related to the conditions and settings of the software-hardware and protocol used. The different settings and variations of various factors can cause differences in the value of Root-mean-square deviation (RMSD), which shows the degree of difference in the position of each atom before and after the docking process is carried out [3]. The RMSD value itself is the main indicator of the docking validation process, where the smaller the RMSD shows the position of the atom which is getting closer to the original position before the docking process is carried out. Ideally, to be declared valid, the RMSD value docking from the crystallographic ligand should not be greater than 2 Å [4]. This provision also applies to AutoDock 4, one of the most popular docking software.

In AutoDock 4, there are several parameters that can affect the RMSD value of docking results, such as position and size of grid box, regulation of rigid/flexible amino acid residues, presence of water molecules, and number of runs of docking processes [5]. Of these factors, the number of runs is interesting for further investigation because different from other factors, the number of runs does not depend on the type of ligand or the receptor used [6]. In general, number of runs shows the number of repetitions of the Lamarckian Genetic Algorithm performed to obtain variations in the ligand pose on the grid box that has been determined [7]. Logically, the larger number of runs will provide an opportunity for the ligand to reach the most ideal pose and approach the crystallographic pose, although this will certainly increase the time needed [8]. However, there is currently no clear limit on how many number of runs are ideal for obtaining the smallest RMSD with the shortest docking time.

This study aims to determine the effect of the number of runs docking processes with AutoDock 4 on the validity of the docking results as indicated by the RMSD value. We will try to prove whether the larger number of runs will also give smaller RMSD or vice versa. In addition, we also determined the linearity of the relationship between number of runs to the RMSD value seen from the correlation coefficient (R).

MATERIALS AND METHODS

The hardware used is the ASUS A46CB series Ultrabook with an Intel™ Core i5-3337U@1.8 GHz and Windows 7 Ultimate 64-bit SP-1 operating system. The software used is AutoDockTools 1.5.6 and Autodock 4.2.6 software from The Scripps Research Institute, Inc (USA). Information on 3D dimensional structures of receptor proteins obtained from the website of Protein Data Banks (http://www.rcsb.org).

The receptor used is an estrogen receptor with ligand reference estradiol (PDB ID 1GWR). Receptors are downloaded in the format. pdb then the unused part including water molecules is removed, added non-polar hydrogen, given charged, and arranged size and coordinate grid box using AutoDockTools 1.5.6 [9]. The reference ligand, estradiol, is extracted and then reused for the redocking process after added non-polar hydrogen, given charged, and set torque. The size and coordinates of the grid box are adjusted automatically with the ligand co-crystal position of each receptor by making the ligand position the center of the grid box [10]. The grid

coordinate used for x, y, and z dimensions is –7.351, –4.233 and 12.804, respectively. While the grid box size is 40 x 40 x 40 Å.

Docking search parameter used are Lamarckian Genetic Algorithm with the number of genetic algorithm varies from 10 to 100 runs with multiples of 10, population size 150, the maximum number of energy evaluation is medium with 2500000, the maximum number of generations 27000, with the default docking parameter used for run options.

The main parameter observed is the RMSD value to determine the validity of the docking results, where the linearity of the relationship between number of runs as a control variable and RMSD as the dependent variable is determined in the form of a correlation coefficient (R). In addition, observations of other parameters such as the free energy of binding (ΔG), the inhibition constant (K_i), amino acid residues, and the number of hydrogen bonds [9]. The selection of poses for each run is done by looking at the re-rank score of the ΔG and K_i value, where the conformation with the most negative ΔG and the smallest K_i value is representative for the conformation of the number of runs. Observation of these parameters is done to observe differences in the results of docking that occur due to differences in number of runs.

RESULTS AND DISCUSSION

The docking results show that the difference in number of runs does not give a significant difference to all observed parameters, even parameters such as ΔG and the number of hydrogen bonds show identical results for all number of runs. The RMSD value of each number of runs is in the range 0.942 to 0.97, with the smallest RMSD produced at number of runs 60. This result is interesting, because it turns out that the smallest value of RMSD is not indicated by the largest number of runs. These results become information updates, where researchers generally use a standard number of runs of at least 100 in the docking process which requires a longer processing time [9, 11], where the number of runs that are lower can actually provide a better RMSD value. This result provides a new dimension for the docking process with AutoDock 4, where in addition to conducting orientation to determine the ideal coordinates and grid box size [12], the number of runs that provide the smallest RMSD must also be determined in advance.

On the other hand, other parameters such as K_i and amino acid residues also did not show a significant difference. The K_i value shown varies in the range 15.57 to 15.63 nM with a difference of only 0.06 nM and the lowest value produced on number of runs 90, indicating there is no difference in the K_i value of the difference in number of runs used. Amino acid residues show almost the same results for each number of runs with interactions shown in the amino acids 353-Glu, 384-Leu, 388-Met, 391-Leu, 404-Phe, 424-Ile, 521-Gly, 524-His, and 525-Leu. The difference in amino acid residues is shown on number of runs 20, 30, 40, and 90, with variations in interactions occurring in amino acids 384-Leu and 391-Leu. These results indicate that variations in the number of runs can lead to small variations in the K_i value and the residual amino acids obtained. Again, these results emphasize the importance of carrying out an orientation to determine the most ideal number of runs before carrying out the docking process [11]. The complete results of the docking process include the values of RMSD, ΔG, K_i, amino acid residues, and the number of hydrogen bonds from each number of runs can be seen in Table 1.

Table 1 – The results of redocking 1GWR receptors with variations in number of runs

Parameter	Number of runs									
	100	90	80	70	60	50	40	30	20	10
RMSD (Å)	0.950	0.948	0.951	0.953	0.942	0.948	0.956	0.970	0.958	0.951
ΔG (kcal/mol)	–10.65	–10.65	–10.65	–10.65	–10.65	–10.65	–10.65	–10.65	–10.65	–10.65
Amino acid residues	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu	353-Glu
	384-Leu	-	384-Leu	384-Leu	384-Leu	384-Leu	–	–	384-Leu	384-Leu
	388-Met	388-Met	388-Met	388-Met	388-Met	388-Met	388-Met	388-Met	388-Met	388-Met
	391-Leu	391-Leu	391-Leu	391-Leu	391-Leu	391-Leu	391-Leu	391-Leu	–	391-Leu
	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe	404-Phe
	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile	424-Ile
	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly	521-Gly
	524-His	524-His	524-His	524-His	524-His	524-His	524-His	524-His	524-His	524-His
	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu	525-Leu
Number of hydrogen bonds	3	3	3	3	3	3	3	3	3	3

КРАТКОЕ СООБЩЕНИЕ
DOI: 10.19163/2307-9266-2020-8-6-476-480

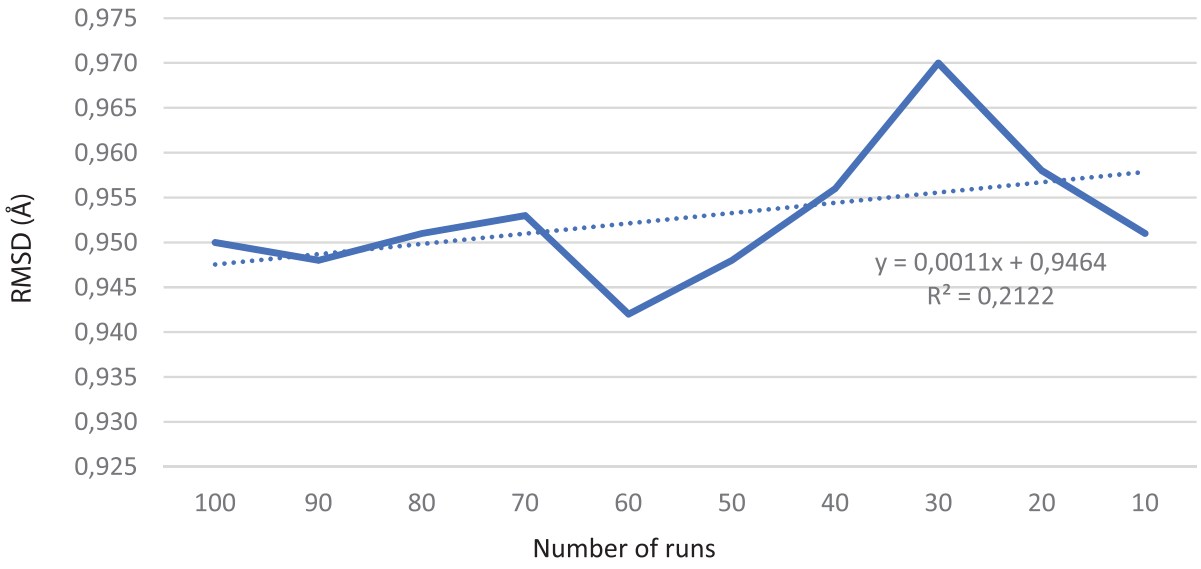

Figure 1 – The relationship between number of runs to the RMSD value at the 1GWR receptor

To observe the linearity of the relationship between the number of runs and the RMSD obtained, a comparison was made with the chart to obtain the R. The results of the observation show that the equation of linearity between the number of runs and RMSD is y = 0.0011x + 0.9464 with the value R = 0.4607, as shown in Figure 1. These results confirm that the relationship between the number of runs and RMSD is indeed not linear, where increasing the number of runs does not necessarily increase the validity of the docking results. In other words, an increase in number of runs does not guarantee the better validity of docking results. This result also supports opinions which state that the most ideal number if runs are in the range of around 50 to 60, where RMSD in that range is at the smallest value [9].

CONCLUSION

This simple study successfully proved that even though number of runs has an influence on the RMSD value of the docking results, the influence shown does not have a linear relationship. The most ideal number of runs in the range 50 to 60, with larger numbers of runs not showing a smaller RMSD value. However, this study has limitations, including the type of receptor and the docking software used. Further research using other types of receptors in large quantities and other docking software such as AutoDock Vina can be done to increase confidence in these findings. Still, this discovery can be a basic guideline for research using the docking method with AutoDock 4.

FUNDING

This study did not have any financial support from outside organizations

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS' CONTRIBUTION

Pratama M.R.F. – conceptualization, data curation, formal analysis, investigation, validation, visualization, writing original draft; Siswandono – conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, writing review & editing

REFERENCES

1. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;20(7):13384–421. DOI: 10.3390/molecules200713384.
2. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational Methods in Drug Discovery. Pharmacological Reviews. 2014;66(1):334–95. DOI: 10.1124/pr.112.007336.
3. Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. Biophysical Reviews. 2017;9(2):91–102. DOI: 10.1007/s12551-016-0247-1.
4. Kontoyianni M, McClellan LM, Sokol GS. Evaluation of docking performance: comparative data on docking algorithms. Journal of Medicinal Chemistry. 2004;47(3):558–65. DOI: 10.1021/jm0302997
5. Kufareva I, Abagyan R. Methods of protein structure comparison. Methods in Molecular Biology. 2012;857:231–57. DOI: 10.1007/978-1-61779-588-6_10

6. Guedes IA, de Magalhaes CS, Dardenne LE. Receptor–ligand molecular docking. Biophysical Reviews. 2014;6(1):75–87. DOI: 10.1007/s12551-013-0130-2
7. Lape M, Elam C, Paula S. Comparison of current docking tools for the simulation of inhibitor binding by the transmembrane domain of the sarco/endoplasmic reticulum calcium ATPase. Biophysical Chemistry. 2010;150(1–3):88–97. DOI: 10.1016/j.bpc.2010.01.011
8. Ramirez D, Caballero J. Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? Molecules. 2018;23(5):1038. DOI: 10.3390/molecules23051038
9. Forli S, Huey R, Pique ME, Sanner M, Goodsell DS, Olson AJ. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nature Protocols. 2016;11(5):905–19. DOI: 10.3390/molecules23051038
10. Arba M, Yamin, Ihsan S, Tjahjono DH. Computational approach toward targeting the interaction of porphyrin derivatives with Bcl-2. Journal of Applied Pharmaceutical Science. 2018;8(12):60–6. DOI: 10.7324/JAPS.2018.81208
11. Atkovska K, Samsonov SA, Pszkowski-Rogacz M, Pisabarro MT. Multipose Binding in Molecular Docking. International Journal of Molecular Sciences. 2014;15(2):2622–45. DOI: https://doi.org/10.3390/ijms15022622
12. Feinstein WP, Brylinski M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. Journal of Cheminformatics. 2015;7:18. DOI: https://dx.doi.org/10.1186/s13321-015-0067-5

AUTHORS

Mohammad Rizki Fadhil Pratama – Ph.D. student of Pharmaceutical Chemistry from Doctoral Program of Pharmaceutical Sciences, Universitas Airlangga, Indonesia; and Assistant Professor of Medicinal Chemistry from Department of Pharmacy, Universitas Muhammadiyah Palangkaraya, Indonesia. ORCID ID: 0000-0002-0727-4392. E-mail: m.rizkifadhil@umpalangkaraya.ac.id

Siswandono – Professor of Medicinal Chemistry from Department of Pharmaceutical Chemistry, Universitas Airlangga, Indonesia. ORCID ID: 0000-0002-9579-8929. E-mail: E-mail: prof.sis@ff.unair.ac.id