УДК 577.15

ИЗУЧЕНИЕ ВЛИЯНИЯ КОЛИЧЕСТВА ЗАПУСКОВ AUTODOCK 4 НА СРЕДНЕКВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ РЕЗУЛЬТАТОВ ДОКИНГА

М.Р.Ф. Пратама^{1,2}, Сисвандоно¹¹ Университет Айрланга

60115, Индонезия, Восточная Ява, г. Сурабая, Доктор Ир. Сукарно, ул. Кампус С, Мульореё

² Университет Мухаммадии Палангкая

73111, Индонезия, Центральный Калимантан, г. Паланга 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. Используемый 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

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

M.R.F. Pratama^{1,2}, Siswandono¹¹ 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

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

© М.Р.Ф. Пратама, Сисвандоно, 2020

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

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

ВВЕДЕНИЕ

Молекулярный докинг (или просто докинг) – один из наиболее распространенных экспериментальных методов *in silico*. Простота в процессе подготовки, внедрения и анализа результатов – одно из достоинств метода [1]. Однако докинг – это простой, но сложный метод. Как один из самых популярных методов *in silico*, докинг имеет ряд преимуществ и недостатков. Помимо скорости и практичности, докинг также предлагает достаточно полное наблюдение анализа комплексов между лекарственными средствами и рецепторами, что даже очень сложно сделать *in vitro* или *in vivo*. Однако полученные результаты обычно имеют большую степень вариации, причем воспроизводимость является одним из важных вопросов, влияющих на уровень достоверности результатов докинга [2]. По этой причине абсолютно необходима проверка результатов докинга.

На результаты валидации докинга влияют различные факторы, связанные как с условиями, так и с настройками программно-аппаратного обеспечения и используемого протокола. Различные настройки и вариации различных факторов могут вызывать различия в значении среднеквадратичного отклонения (RMSD), которое показывает степень различия в положении каждого атома до и после выполнения процесса докинга [3]. Само значение RMSD является основным индикатором процесса проверки докинга, где меньшее значение RMSD показывает положение атома, который приближается к исходному положению до выполнения процесса докинга. В идеале, чтобы быть достоверным, значение RMSD докинга с кристаллографическим лигандом не должно превышать 2 Å [4]. Это положение также относится к AutoDock 4, одному из самых популярных программ для осуществления докинга.

В AutoDock 4 есть несколько параметров, которые могут влиять на значение RMSD результатов докинга, таких как: положение и размер ячейки сетки; регулирование жестких/гибких аминокислотных остатков; присутствие молекул воды; количество запусков процессов докинга [5]. Из-за этих факторов количество прогонов представляет интерес для дальнейшего исследования, поскольку в отличие от

других факторов, количество прогонов не зависит от типа лиганда или используемого рецептора [6]. В общем, количество прогонов показывает количество повторов генетического алгоритма Ламарка, выполненных для получения вариаций установленной позиции лиганда на сетке [7]. Логично, что большее количество прогонов даст лиганду возможность достичь наиболее идеальной позиции и приблизиться к кристаллографической, хотя это, безусловно, увеличит требуемое время [8]. Однако в настоящее время нет четкого ограничения на количество прогонов, которое идеально подходит для получения наименьшего RMSD с наименьшим временем докинга.

Это исследование направлено на определение влияния количества запусков процесса докинга с AutoDock 4 на достоверность результатов докинга, определяемого значением RMSD. Мы попытаемся доказать, дает ли большее количество прогонов меньшее среднеквадратичное отклонение или наоборот. Кроме того, мы также определили линейность отношения между количеством запусков и значением RMSD, исходя из коэффициента корреляции (R).

МАТЕРИАЛЫ И МЕТОДЫ

Используемое оборудование: ультрабук ASUS серии A46CB с процессором Intel™ Core i5-3337U @ 1,8 ГГц и 64-битной операционной системой Windows 7 Ultimate SP-1. Используемое программное обеспечение – AutoDockTools 1.5.6 и Autodock 4.2.6 от Scripps Research Institute, Inc (USA). Информация о трехмерных структурах рецепторных белков взята с сайта Protein Data Banks (<http://www.rcsb.org>).

Используемый рецептор представляет собой рецептор эстрогена с эталонным лигандом эстрадиола (PDB ID 1GWR). Рецепторы скачиваются в формате .pdb, затем неиспользуемая часть, включая молекулы воды, удаляется, а вместо них добавляется неполярный водород с заданным зарядом и с помощью AutoDockTools 1.5.6 упорядочивается размер и блок координатной сетки [9]. Эталонный лиганд, эстрадиол, извлекается и затем повторно используется для процесса повторной фиксации после добавления неполярного водорода, заданного заряда и заданного

крутящего момента. Размер и координаты ячейки сетки автоматически регулируются в соответствии с положением сокристалла лиганда каждого рецептора, устанавливая лиганд в центре ячейки сетки [10]. Координата сетки, используемая для размеров x , y и z , равна -7.351 , -4.233 и 12.804 соответственно. При этом размер ячейки сетки составляет $40 \times 40 \times 40$ Å.

Используемый параметр поиска докинга: генетический алгоритм Ламарка с количеством генетических алгоритмов от 10 до 100, количество запусков которых кратно 10, размер популяции 150, максимальное количество оценок энергии – среднее от 2500000, максимальное количество поколений – 27000, по умолчанию параметр стыковки, используемый для параметров запуска.

Основным наблюдаемым параметром является значение RMSD для определения достоверности результатов докинга, где линейность зависимости между количеством прогонов в качестве контрольной переменной и RMSD в качестве зависимой переменной определяется в форме коэффициента корреляции (R). Кроме того, наблюдаются другие параметры, такие как свободная энергия связывания (ΔG), константа ингибирования (K_i), аминокислотные остатки и количество водородных связей [9]. Выбор позиции для каждого прогона осуществляется путем просмотра оценки повторного ранжирования значений ΔG и K_i , где конформация с наиболее отрицательным ΔG и наименьшим значением K_i является репрезентативной для соответствия количеству прогонов. Наблюдение за этими параметрами проводится для выявления различий в результатах докинга, возникающих из-за различий в количестве запусков.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Результаты докинга показывают, что разница в количестве прогонов не дает существенной разницы для всех наблюдаемых параметров, даже таких, как ΔG и количество водородных связей показывают идентичные результаты для всего количества прогонов. Значение RMSD для каждого количества прогонов находится в диапазоне от 0,942 до 0,97, причем наименьшее RMSD получается при количестве прогонов равном 60. Этот результат интересен, потому что оказывается, что наименьшее значение RMSD не указывает на наибольшее количество прогонов. Эти результаты становятся обновлениями информации, так как исследователи обычно используют стандартное количество прогонов, по крайней мере, равное 100 в процессе докинга, который требует более длительного времени обработки [9, 11], где меньшее количество прогонов может фактически обеспечить лучшее значение RMSD. Этот результат обеспечивает новое измерение для процесса докинга с AutoDock 4, где, помимо проведения ориентации для определения идеальных координат и размера ячейки сетки [12], необходимо также заранее определить количество

прогонов, обеспечивающих наименьшее среднеквадратичное отклонение.

С другой стороны, другие параметры, такие как K_i и количество аминокислотных остатков, также не показали значительной разницы. Показанное значение K_i показывает варьирование в диапазоне от 15,57 до 15,63 нМ с разницей всего в 0,06 нМ и наименьшим значением, полученным при количестве запусков равном 90, что указывает на отсутствие разницы в значении K_i при разнице в количестве использованных запусков. Количество аминокислотных остатков показывает почти те же результаты для каждого количества запусков для таких аминокислот, как: 353-Глу, 384-Лей, 388-Мет, 391-Лей, 404-Фен, 424-Иле, 521-Глу, 524-Гис и 525-Лей. Разница в аминокислотных остатках показана на количестве запусков равных 20, 30, 40 и 90 с вариациями взаимодействий, происходящих в аминокислотах 384-Лей и 391-Лей. Эти результаты показывают, что вариации в количестве прогонов могут привести к небольшим вариациям значения K_i и полученных аминокислотных остатках. Опять же, эти результаты подчеркивают важность проведения ориентации для определения наиболее идеального количества прогонов перед выполнением процесса докинга [11]. Полные результаты процесса докинга включают значения RMSD, ΔG , K_i , аминокислотных остатков, а количество водородных связей для каждого количества запусков представлены в таблице 1.

Для определения линейной зависимости между количеством прогонов и полученным RMSD, было проведено сравнение с диаграммой для получения R . Результаты наблюдения показывают, что уравнение линейности между количеством прогонов и RMSD имеет вид $y = 0,0011x + 0,9464$ со значением $R = 0,4607$, как показано на рисунке 1. Эти результаты подтверждают, что соотношение между количеством прогонов и RMSD действительно не линейное, что говорит о том, что увеличение количества запусков не обязательно увеличивает достоверность полученных результатов докинга. Другими словами, увеличение количества запусков не гарантирует лучшей достоверности результатов докинга. Этот результат также подтверждает утверждения, согласно которым наиболее идеальное число запусков находится в диапазоне от 50 до 60, так как RMSD в этом диапазоне является наименьшим [9].

ЗАКЛЮЧЕНИЕ

Это простое исследование успешно доказало, что даже несмотря на то, что количество запусков влияет на значение RMSD результатов докинга, оно не имеет линейной зависимости. Наиболее идеальное количество запусков находится в диапазоне от 50 до 60, при этом большее количество запусков не показывает меньшее значение RMSD. Однако это исследование имеет ограничения, в том числе тип рецептора и используемое программное обеспечение для докинга.

Таблица 1 – Результаты повторного докинга рецепторов 1GWR с вариацией в количестве запусков

Параметры	Число запусков									
	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 (ккал / моль)	-10.65	-10.65	-10.65	-10.65	-10.65	-10.65	-10.65	-10.65	-10.65	-10.65
Аминокислотные остатки	353-Глу	353-Глу	353-Глу	353-Glu	353-Глу	353-Глу	353-Глу	353-Глу	353-Глу	353-Глу
	384-Лей	–	384-Лей	384-Лей	384-Лей	384-Лей	–	–	384-Лей	384-Лей
	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет	388-Мет
	391-Лей	391-Лей	391-Лей	391-Лей	391-Лей	391-Лей	391-Лей	391-Лей	–	391-Лей
	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен	404-Фен
	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле	424-Иле
	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу	521-Глу
	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис	524-Гис
	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей	525-Лей
Количество водородных связей	3	3	3	3	3	3	3	3	3	3

Рисунок 1 – Связь между количеством запусков и значением RMSD на рецепторе 1GWR

Дальнейшие исследования с использованием других типов рецепторов в больших количествах и другого программного обеспечения для проведения докинга, такого как AutoDock Vina, могут быть

проведены для того, чтобы удостовериться в данных результатах. Тем не менее, это открытие может быть основным ориентиром для исследований методом докинга с использованием программы AutoDock 4.

ФИНАНСОВАЯ ПОДДЕРЖКА

Данное исследование не имело финансовой поддержки от сторонних организаций.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ВКЛАД АВТОРОВ

М.Р.Ф. Пратама – концептуализация, курирование данных, формальный анализ, исследование, проверка, визуализация, написание оригинального проекта; Сисвандоно – концептуализация, курирование данных, формальный анализ, исследование, методология, администрирование проекта, ресурсы, программное обеспечение, надзор, написание обзора и редактирование.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ferreira L.G., Dos Santos R.N., Oliva G., Andricopulo A.D. Molecular docking and structure-based drug design strategies // *Molecules*. – 2015. – Vol.20. – No.7. – P. 13384–13421. DOI: 10.3390/molecules200713384.
2. Sliwoski G., Kothiwale S., Meiler J., Lowe E.W. Computational Methods in Drug Discovery // *Pharmacological Reviews*. – 2014. – Vol. 66. – No.1. – P. 334–395. DOI: 10.1124/pr.112.007336.
3. Pagadala N.S., Syed K., Tuszynski J. Software for molecular docking: a review // *Biophysical Reviews*. – 2017. – Vol. 9. – No.2. – P. 91–102. DOI:10.1007/s12551-016-0247-1.
4. Kontoyianni M., McClellan L.M., Sokol G.S. Evaluation of docking performance: comparative data on docking algorithms // *Journal of Medicinal Chemistry*. – 2004. – Vol. 47. – No.3. – P. 558–65. DOI: 10.1021/jm0302997
5. Kufareva I., Abagyan R. Methods of protein structure comparison // *Methods in Molecular Biology*. – 2012. – No.857. – P. 231–257. DOI: 10.1007/978-1-61779-588-6_10.
6. Guedes I.A., de Magalhaes C.S., Dardenne L.E. Receptor – ligand molecular docking // *Biophysical Reviews*. – 2014. – Vol. 6. – No.1. – P. 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. – Vol. 150. – No.1–3. – P. 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. – Vol. 23. – No.5. – P. 1038. DOI: 10.3390/molecules23051038.
9. Forli S., Huey R., Pique M.E., Sanner M., Goodsell D.S., Olson A.J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite // *Nature Protocols*. – 2016. – Vol. 11. – No.5. – P. 905–919. DOI:10.3390/molecules23051038.
10. Arba M., Yamin, Ihsan S., Tjahjono D.H. Computational approach toward targeting the interaction of porphyrin derivatives with Bcl-2 // *Journal of Applied Pharmaceutical Science*. – 2018. – Vol. 8. – No.12. – P. 60–66. DOI: 10.7324/JAPS.2018.81208.
11. Atkovska K., Samsonov S.A., Pszkowski-Rogacz M., Pisabarro M.T. Multipose Binding in Molecular Docking // *International Journal of Molecular Sciences*. – 2014. – Vol. 15. – No.2. – P. 2622–2645. DOI: 10.3390/ijms15022622.
12. Feinstein W.P., Brylinski M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets // *Journal of Cheminformatics*. – 2015. – No.7. – P. 18. DOI: 10.1186/s13321-015-0067-5.

АВТОРЫ

Мохаммад Ризки Фадхил Пратама – кандидат фармацевтических наук, докторант факультета фармацевтической химии Университета Айрланга, Индонезия; доцент кафедры медицинской химии фармацевтического факультета Палангкарайского университета Мухаммадии, Индонезия. ORCID

ID: 0000-0002-0727-4392. E-mail: m.rizkifadhil@umpalangkaraya.ac.id

Сисвандоно – профессор медицинской химии факультета фармацевтической химии Университета Айрланга, Индонезия. ORCID ID: 0000-0002-9579-8929. E-mail: prof.sis@ff.unair.ac.id