Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 1030–1033
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Short communication

Prediction of protein–ligand complex structure by docking software guided by other complex structures

Yoshifumi Fukunishi [a,b,*], Haruki Nakamura [a,c]

[a] *Biological Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6 Aomi, Koto-ku, Tokyo 135-0064, Japan*
[b] *Pharmaceutical Innovation Value Chain, BioGrid Center Kansai, 1-4-2 Shinsenri-Higashimachi, Toyonaka, Osaka 560-0082, Japan*
[c] *Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan*

Received 5 May 2007; received in revised form 5 July 2007; accepted 6 July 2007
Available online 12 July 2007

Abstract

We developed a new scoring method that selects a protein–ligand complex structure with higher geometrical accuracy than the top-scoring complex structure, using the structural information of known protein–ligand complexes. To apply this method, one or more protein–ligand complex structures must be known for the target protein. A number of predicted structures were generated by the protein-compound docking program for a new ligand, and one of these structures, which showed the maximum overlap with the ligand coordinates of the known protein–ligand complex, was selected as the most probable complex structure.

Keywords: Protein–ligand docking; Docking score; Protein–ligand complex

One of the most important problems in drug development is the prediction of protein–ligand complex structures. In the lead optimization process, a number of compounds are designed based on the complex structure of a target protein with a known active ligand. It is important to be able to predict complex structures using these newly designed compounds in order to evaluate the compounds before actual synthesis. In the present study, we first applied a conventional docking method to generate 120 predicted structures for each ligand, and then used the geometric fit scoring function, $S(n)$, given in Eq. (1), as a means of ranking solutions based on their fit to a known protein–ligand complex. This method could be applied to improve the docking results obtained with docking programs other than the program used in this report.

Many docking programs have been developed [1–7], but the structural accuracy of existing docking software is relatively low [1,2]. The success rate of reproduction of the protein–ligand complex structure within a root mean square deviation (RMSD) of <2 Å is almost 50% [6,7]. In addition, the accuracy of the binding free energy estimation remains at approximately 2–3 kcal/mol [6,7]. The low accuracy of the binding free energy and docking scores causes low database enrichment from *in silico* screening [1,2]. The parameters of the protein-compound docking programs govern both computational speed and docking accuracy. Usually, the values of the parameters are selected to allow docking to take place within a few seconds to a few minutes for one protein-compound docking for large-scale *in silico* drug screening. Alternatively, allowing more computational time (several tens of minutes to several hours) is permissible to achieve greater accuracy for one protein-compound docking, where high-throughput docking is not required.

A more radical approach is modification of the docking scheme itself. To improve docking accuracy and database enrichment, one approach is to improve the docking score itself [8,9]. However, the limitations of this improvement are obvious. Free energy is calculated from the partition function, which is based on a structural ensemble of numerous structures at a particular temperature; in contrast, the docking score is calculated from a single protein–ligand complex structure. Another method of improving database enrichment is the

* Corresponding author at: Biological Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6 Aomi, Koto-ku, Tokyo 135-0064, Japan. Tel.: +81 3 3599 8290; fax: +81 3 3599 8099.
E-mail address: y-fukunishi@jbirc.aist.go.jp (Y. Fukunishi).

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2007.07.001

application of the protein-compound affinity matrix. The multiple active site correction (MASC) scoring method uses the deviation of the docking score instead of the raw docking score [10], and the multiple target screening (MTS) method compares the docking scores of many proteins for one compound instead of comparing the docking scores of many compounds for one target protein [11,12]. These methods indirectly improve the estimation of the binding free energy, but do not improve docking accuracy.

In ligand-based docking, the pharmacophore is determined by superimposing several active compound structures. In protein structure-based docking, the pharmacophore is not necessary because the protein–ligand interaction is estimated based on the protein 3D structure. However, the idea of the pharmacophore remains important in structure-based docking. It is usually difficult to evaluate the entropy and structure changes of a protein upon ligand binding. The protein–ligand interaction in a rigid region of the protein may be stronger than that in a flexible region of the protein because the entropy change in the former case is smaller than that in the latter case. Thus, it is difficult to take into account the flexibility of a protein in structure-based docking. In the present study, we therefore tried to take the pharmacophore into account in protein–ligand binding by using the known protein–ligand complex structure and evaluating the overlap between the predicted ligand coordinates and the reference ligand coordinates of the known complex structure.

The complex structure predicted by the best docking score is not necessarily the correct structure. In the present study, a correct structure means a complex structure with an RMSD value of <2 Å from the crystallographically observed structure. The probability of finding a correct structure among the top 100 complex structures predicted by the docking program, Sievgene [7], is approximately 90%. In contrast, in our previous study [7], only 50.8% of the top-scoring structures were correct. Thus, in many cases, the correct complex structure can be found among the many predicted structures. In the present study, we selected a predicted structure, which gives the maximum structural overlap with the ligand structure(s) of the other known complex structure(s).

Let $\{x_n^i, y_n^i, z_n^i\}$ be the $\{x, y, z\}$ coordinates of the ith atom of the nth predicted binding pose, and let $\{x_r^j, y_r^j, z_r^j\}$ be the $\{x, y, z\}$ coordinates of the jth atom of the rth reference ligand structure of the known protein–ligand complex structure. Let q_i and q_j^r be the atomic charge of the ith atom of the ligand and the atomic charge of the jth atom of the rth reference ligand, respectively. $S(n)$ is the measure used to evaluate the overlap between the nth predicted binding pose and the reference ligand structure(s) of the known protein–ligand complex structure(s). Then,

$$S(n) = \sum_{r=1}^{N_{\mathrm{ref}}} \sum_{j=1}^{N(r)} \sum_{i=1}^{M_{\mathrm{atom}}} w(i,j,r) \exp(-c((x_n^i - x_r^j)^2 + (y_n^i - y_r^j)^2 + (z_n^i - z_r^j)^2)) \quad (1)$$

and

$$w(i,j,r) = \begin{cases} 1; & |q_i - q_j^r| < q_{\mathrm{thr}} \\ 0; & |q_i - q_i^r| \geq q_{\mathrm{thr}} \end{cases} \quad (2)$$

where N_{ref}, $N(r)$, M_{atom}, c, w and q_{thr} are the number of known protein–ligand complex structures, the number of atoms of the rth reference ligand, the number of atoms of the ligand to be docked, a coefficient, a switching function and a threshold value, respectively. We select the binding pose, which gives the maximum $S(n)$ value, as the most probable prediction result. In the present study, c and q_{thr} were set as 1.0 Å^{-2} and 0.2 a.u., respectively. This method is called the maximum volume overlap (MVO) method.

To evaluate our method, a cross-docking test was performed. For this test, 18, 5 and 5 complex structures were prepared for thermolysin (THR), cyclooxygenase-2 (COX2) and tumor necrosis factor (TNF)-alpha converting enzyme (TACE), respectively. The ligands of these complexes are rich in variety; their structures are summarized in the supplementary figure. The names of these protein–ligand complex structures are listed in Appendix A. In each case, a single one of these complex structures was selected as a reference structure: 1hyt for THR, 1cx2 for COX2 and 1zxc for TACE. The other structures were superimposed onto the reference structure by minimizing the RMSD of the coordinates of the main chain Cα carbons atoms by Kearsley's method [13] using our in-house program. When the lengths of the two proteins were different, only the Cα carbons of the overlapped amino acids were used. The first chain in the Protein Data Bank (PDB) file was selected when more than one complex structure was presented in the PDB file.

The 3D coordinates of the chemical compounds were generated by the Chem3D program (Cambridge Software, Cambridge, MA, USA). The atomic charges of each ligand were determined by the Gasteiger method [14,15]. The atomic charges of the proteins were the same as the atomic charges of AMBER parm99 [16]. The protein–ligand dockings were performed by the docking program Sievgene/myPresto (http://www.jbic.or.jp/activity/st_pr_pj/mypresto/index_mypr.html) [7]. Almost all parameters used in Sievgene were identical to the values described in our previous paper [7]. In the present study, up to 200 conformers were generated for each ligand. The global search was performed once.

Suppose that complex A is a complex of protein PA and its ligand LA, that complex B is a complex of protein PB and its ligand LB, and that proteins A and B are the same. To use complex A as the reference complex structure, complex B is superimposed onto complex A in order to minimize the RMSD value of the Cα carbons of PA and PB. LA then gives the reference coordinates $\{x_r^j, y_r^j, z_r^j\}$. The Sievgene program utilizes the grid potential to calculate each energy term except for the intramolecular interaction. In the present study, the adopted mesh size of $100 \times 100 \times 100$ generated, 120 predicted structures for each ligand. The search space was around the reference ligand coordinates and allowed margins of 6.5 Å; thus, the cell length of the search space was

Table 1
Cross-docking results obtained by the original docking method, the single-reference docking method, the two-reference docking method and the three-reference docking method

Method	Target protein	Frequency		
		RMSD < 1 Å	RMSD < 2 Å	RMSD < 3 Å
Original	THR	2.3% (7)	25.5% (78)	38.6% (118)
	COX2	15.0% (3)	65.0% (13)	65.0% (13)
	TACE	0% (0)	25.0% (5)	30.0% (6)
Average		2.9%	27.7%	39.6%
Single-reference	THR	37.3% (114)	57.8% (177)	64.7% (198)
	COX2	25.0% (5)	65.0% (13)	70.0% (14)
	TACE	30.0% (6)	40.0% (8)	55.0% (11)
Average		36.1%	57.2%	64.5%
Two-reference	THR	42.8% (131)	69.3% (212)	80.7% (247)
	COX2	27.5% (5.5)	65.0% (13)	75.0% (15)
	TACE	34.0% (6.8)	56.5% (11.3)	66.5% (13.3)
Average		41.4%	68.3%	79.6%
Three-reference	THR	44.9% (137.5)	78.9% (241.5)	83.5% (255.5)
	COX2	27.5% (5.5)	65.0% (13)	75.0% (15)
	TACE	34.0% (6.8)	57.5% (11.5)	66.5% (13.3)
Average		43.3%	76.9%	82.0%

For the two-reference docking method, two-reference ligand structures were randomly selected and the results are the average of four trial dockings. The numbers of dockings were 306 (=18 × 17) for THR and 20 (=5 × 4) for COX2 and TACE. The numbers in parentheses are the numbers of observed structures. The RMSD value was calculated for the coordinates of the heavy atoms of the ligand.

approximately 20–30 Å. We performed a cross-docking test using 18 vs. 17 dockings for THR, 5 vs. 4 dockings for COX2 and 5 vs. 4 dockings for TACE. This test therefore served as an N-fold cross-validation test. With a single-reference, the reference structure coordinates were the ligand coordinates of the protein–ligand complex itself, and the other $N-1$ ligands were docked to the protein. The docking results are summarized in Table 1. The RMSD was calculated between all atom positions, with the exception of the H atoms of each docked compound and the corresponding atom positions in the complex crystal structure. Two- and three-references were randomly selected, and each trial was performed four times for each structure. The values in Table 1 are the averaged values of the 1224 (=18 × 17 × 4), 80 (=5 × 4 × 4) and 80 (=5 × 4 × 4) dockings for THR, COX2 and TACE, respectively.

Docking accuracy calculated by the original docking method, which utilizes only a docking score to select the optimal protein–ligand complex structure, is much lower than the self-docking results previously reported [17]. In fact, the Sievgene program reconstructed 18.9%, 50.8% and 59.8% of a total of 132 complexes with RMSD values of <1 Å, 2 Å and 3 Å, respectively, for the self-docking result [7]. Using almost the same dataset, the DOCK [3], FlexX [4] and GOLD [5] programs have been reported to reconstruct complexes of 39%, 51% and 56%, respectively, with RMSDs of < 2 Å [18,19] for the self-docking results. The cross-docking test, whose results are summarized in Table 1, is a much more rigorous test than the self-docking test.

The number of precisely predicted structures (RMSD < 1 Å) was drastically improved by the MVO method; specifically, 10- and 15-fold improvements were achieved by the single-reference and two-reference docking methods, respectively, compared to the conventional docking method that uses only the docking score to select the complex structure. The number of correctly predicted structures (RMSD < 2 Å) was also drastically improved, with 2- and 2.5-fold improvements obtained by the single-reference and two-reference docking methods compared to the conventional docking method. The three-reference docking method improved the docking result by only 2–3% compared to the two-reference docking method. The average results were close to the results of the THR docking results, since the number of THR complexes is much bigger than those of COX2 and TACE. All targets (THR, COX2 and TACE) showed the same trend: 70% of complex structures were correctly predicted (RMSD < 2 Å) by the two- or three-references docking method, which takes only an average of 1–10 min to yield results. The conventional docking method requires several tens of minutes or even hours to obtain the same level of docking accuracy [19].

Similarity between the ligand in question and the reference ligand should be critical in the MVO method. To define the similarity between two ligands, a definition of distance (D_g) is introduced. Substructures of molecule A are overlapped onto molecule B by topological graph matching, and D_g is the distance defined by the maximum common subgraph method [20].

$$D_g = 1 - \frac{V(G_S)}{\max(V(G_A), V(G_B))}. \quad (3)$$

Here, $V(G_S)$, $V(G_A)$ and $V(G_B)$ are the number of atoms of the maximum common subgraph, that of molecule A and that of

molecule B, respectively. In the maximum common subgraph, mismatched atoms are ignored, as are differences in atom types (C, N, O, etc.). Furthermore, hydrogen atoms are ignored to calculate $V(G_S)$, $V(G_A)$ and $V(G_B)$. For the single-reference case, the correlation between D_g and RMSD was examined. There was almost no correlation: RMSD = 1.53 D_g + 2.12, and the correlation coefficient was 0.12. Thus, the MVO method can be applied even if the ligand to be docked is not similar to the ligand in the reference crystal structure.

The MVO method can improve docking accuracy, and can be applied to the predicted results by any kind of docking software. It is based on the experimentally determined protein–ligand complex structure, and selects the structure from many predicted structures generated by docking software according to overlap between the predicted ligand structure and the experimentally determined structure. A set of two-reference complex structures is better than a single-reference complex structure. Instead of $S(n)$, a linear combination of $S(n)$ and the docking score was also applied to the docking test, however, our results showed that $S(n)$ without modification is sufficient to achieve precise docking results.

Acknowledgements

This work was supported by grants from the New Energy and Industrial Technology Development Organization of Japan (NEDO) and the Ministry of Economy, Trade and Industry (METI) of Japan.

Appendix A

The following PDB identifier list of complexes was used as the THR complex data: 1hyt, 1lna, 1os0, 1pe7, 1pe8, 1qf0, 1qf1, 1qf2, 1tlp, 1tmn, 1thl, 1z9g, 1zdp, 2tmn, 4tln, 4tmn, 5tln and 5tmn.

The following PDB identifier list of complexes was used as the COX2 complex data: 1cx2, 1pxx, 3pgh, 4cox and 6cox.

The following PDB identifier list of complexes was used as the TACE complex data: 1zxc, 2a8h, 2ddf, 2fv5 and 2fv9.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2007.07.001.

References

[1] G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff, M.S. Head, A critical assessment of docking programs and scoring functions, J. Med. Chem. 49 (2006) 5912–5931.
[2] M. Kontoyianni, G.S. Sokol, L.M. McClellan, Evaluation of library ranking efficacy in virtual screening, J. Comput. Chem. 26 (2005) 11–22.
[3] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, A geometric approach to macromolecule–ligand interactions, J. Mol. Biol. 161 (1982) 269–288.
[4] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A fast flexible docking method using an incremental construction algorithm, J. Mol. Biol. 261 (1996) 470–489.
[5] G. Jones, P. Willet, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[6] C.A. Baxter, C.W. Murray, D.E. Clark, D.R. Westhead, M.D. Eldridge, Flexible docking using tabu search and an empirical estimate of binding affinity, Proteins 33 (1998) 367–382.
[7] Y. Fukunishi, Y. Mikami, H. Nakamura, Similarities among receptor pockets and among compounds: analysis and application to in silico ligand screening, J. Mol. Graph. Model. 24 (2005) 34–45.
[8] C. Zhang, S. Liu, Q. Zhu, Y. Zhou, A knowledge-based energy function for protein–ligand, protein–protein, and protein–DNA complexes, J. Med. Chem. 48 (2005) 2325–2335.
[9] I. Muegge, Y.C. Martin, A general and fast scoring function for protein–ligand interactions: a simplified potential approach, J. Med. Chem. 42 (1999) 791–804.
[10] G.P.A. Vigers, J.P. Rizzi, Multiple active site corrections for docking and virtual screening, J. Med. Chem. 47 (2004) 80–89.
[11] Y. Fukunishi, Y. Mikami, S. Kubota, H. Nakamura, Multiple target screening method for robust and accurate in silico ligand screening, J. Mol. Graph. Model. 25 (2005) 61–70.
[12] Y. Fukunishi, S. Kubota, H. Nakamura, Noise reduction method for molecular interaction energy: application to in silico drug screening and in silico target protein screening, J. Chem. Inf. Model. 46 (2006) 2071–2084.
[13] S.K. Kearsley, On the orthogonal transformation used for structural comparisons, Acta Crystallogr. A45 (1989) 208–210.
[14] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, Tetrahedron 36 (1980) 3219–3228.
[15] J. Gasteiger, M. Marsili, A new model for calculating atomic charges in molecules, Tetrahedron Lett. (1978) 3181–3184.
[16] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, B. Wang, D.A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J.W. Caldwell, W.S. Ross, P.A. Kollman, Amber 8, UCSF, 2004.
[17] S.L. McGovern, B.K. Shoichet, Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes, J. Med. Chem. 46 (2003) 2895–2907.
[18] N. Paul, D. Rognan, ConsDock: a new program for the consensus analysis of protein–ligand interactions, Protein Struct. Funct. Genet. 47 (2002) 521–533.
[19] J.W.M. Nissink, C. Murray, M. Hartshorn, M.L. Verdonk, J.C. Cole, R. Taylor, A new test set for validating predictions of protein–ligand interaction, Protein Struct. Funct. Genet. 49 (2002) 457–471.
[20] H. Bunke, K. Shearer, A graph distance metric based on the maximal common subgraph, Pattern Recognit. Lett. 19 (1998) 255–259.