Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

An empirical molecular docking study of a di-iron binding protein with iron ions*

Huan WANG[†1,2], Ping LIU[3], Hao XIE[†‡1,2]

([1]*School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China*)

([2]*Institute of Sciences, Wuhan University of Technology, Wuhan 430070, China*)

([3]*School of Information Management, Wuhan University, Wuhan 430072, China*)

[†]E-mail: shrekwang1@gmail.com; drxiehao@126.com

Received Mar. 15, 2012; Revision accepted Sept. 28, 2012; Crosschecked Jan. 23, 2013

Abstract: Various molecular docking software packages are available for modeling interactions between small molecules and proteins. However, there have been few reports of modeling the interactions between metal ions and metalloproteins. In this study, the AutoDock package was employed to example docking into a di-iron binding protein, bacterioferritin. Each binding site of this protein was tested for docking with iron ions. Blind docking experiments showed that all docking conformations converged into two clusters, one for internal iron binding in sites within the metalloprotein and the other for external iron binding on the protein surface. Local docking experiments showed that there were significant differences between two internal iron binding sites. Docking at one site gave a reasonable root-mean-square deviation (RMSD) distribution with relatively low binding energy. Analysis of the binding mode quality for this site revealed that more than half of the docking conformations were categorized as having good binding geometry, while no good conformations were found for the other site. Further investigations indicated that coordinating water molecules contributed to the stability of binding geometries. This study provides an empirical approach towards the study of molecular docking in metalloproteins.

Key words: AutoDock, Docking, Iron ions, Metalloproteins, Binding modes
doi:10.1631/jzus.C1200072 **Document code:** A **CLC number:** R857.3

1 Introduction

In silico analysis of possible interactions between small molecules and proteins can be achieved by various software packages including AutoDock (Morris *et al.*, 1996; 1998), LigandFit/Cerius (Morris *et al.*, 1998; Venkatachalam *et al.*, 2003), FlexX (Böhm, 1998; Kramer *et al.*, 1999), GOLD (Jones *et al.*, 1997), Glide (Friesner *et al.*, 2004; Halgren *et al.*, 2004), and DOCK (Bissantz *et al.*, 2000), which are widely used in screening large compound libraries, especially when seeking inhibitors of enzymes where assays are not suitable for high throughput screening. Many enzymes require metal ions, such as Fe^{3+}, Mg^{2+}, Ca^{2+}, and Zn^{2+}, in their active sites for successful substrate binding and catalysis (Ferrara *et al.*, 2004; Ananthanarayanan and Kerman, 2006). Although metal ions can be accounted for in available crystal structures, models for docking metal ions into metal-binding sites are very challenging due to multiple coordination geometries and lack of sufficiently accurate force field parameters for protein-metal interactions (Khandelwal *et al.*, 2005). The enhancement of docking accuracy in these programs also requires optimizing metal ion parameters such as radius, partial charge, and well depth.

The general objective of molecular docking is to search for the energetically most favorable conformation of a protein-ligand complex and the scoring of resultant geometries with respect to binding energy (Halperin *et al.*, 2002; Warren *et al.*, 2006). Many

‡ Corresponding author

* Project (No. 2011-II-010) supported by the Fundamental Research Funds for the Central Universities, China

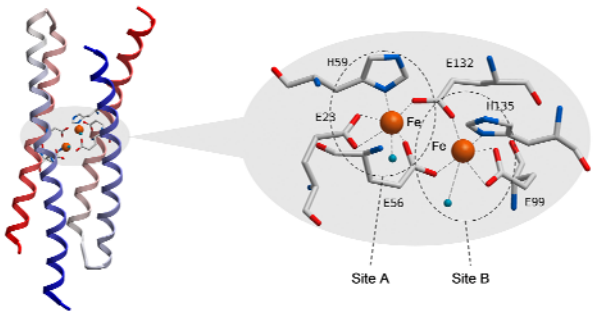
docking programs consider the scoring and geometry prediction as one issue and use the scoring function to fit the complex with the lowest energy (Spyrakis *et al.*, 2007). Therefore, minimizing the root-mean-square deviation (RMSD) between the predicted and experimentally determined complex geometries could result in accurate predictions with reasonable binding energies.

In the present study, the AutoDock 4.0 package was employed for docking metal ions into a di-iron protein, bacterioferritin (PDB ID: 1NF4). This protein belongs to a ferritin-like superfamily of enzymes that catalyze a wide range of important dioxygen-dependent biological reactions. This protein has a four-helix bundle that houses a di-iron binding center, wherein two iron ions are coordinated by two histidine residues and four aspartate residues (Fig. 1) (Kurtz, 1997). There are 16 recorded chain conformations of this protein in PDB. In this study, one of these conformations was tested for blind docking and local docking with iron ions in the force-field of AutoDock. Docking geometries were analyzed by comparing the RMSD with experimental X-ray structures and binding energies as calculated using the scoring function. Appropriate docking accuracy was defined when the RMSD of the first rank docking result complex was within 0.2 nm of experimentally determined X-ray structure. Effects of coordinating water molecules on binding energy and the RMSD were also explored.

2 Materials and methods

2.1 Di-iron protein model

In the present study, a di-iron protein, bacterioferritin (PDB ID: 1NF4), was tested for docking with iron ions. Among the 16 recorded chain conformations of this protein (from Chain A to Chain P) in PDB, the Chain J conformation was selected and used as the di-iron protein model (Fig. 1). Two iron ions were coordinated in the binding pocket of this metalloprotein by six residues (Glu23, Glu56, His59, Glu99, Glu132, His135) and two water molecules (Fig. 1). One iron binding site was defined as 'Site A' where the water molecule was closer to the iron (Fe-O distance: 0.2401 nm), and the other site was defined as 'Site B' (Fe-O distance: 0.3859 nm).



Coordination residue	Distance from residues to iron ions	
	Site A	Site B
Glu23	Fe-O: 0.2258 nm, 0.2244 nm	–
Glu56	Fe-O: 0.2097 nm	Fe-O: 0.2145 nm
His59	Fe-N: 0.2228 nm	–
Water	Fe-O: 0.2401 nm	–
Glu99	–	Fe-O: 0.2464 nm, 0.2194 nm
Glu132	Fe-O: 0.2047 nm	Fe-O: 0.2044 nm
His135	–	Fe-N: 0.2274 nm
Water	–	Fe-O: 0.3859 nm

Fig. 1 Structure of the di-iron metalloprotein bacterioferritin (PDB ID: 1NF4)

Residues involved in metal binding are depicted as sticks. Brown spheres represent iron ions. Small blue spheres represent water molecules

References to color refer to the online version of this figure

2.2 Docking

Input protein structures were prepared by adding hydrogen atoms and removing non-functional water molecules. Partial charges were calculated using Gasteiger's method. Iron ions were docked into Site A, or Site B, or both sites, respectively. When iron ions were docked into Site A, four residues (His59, Glu23, Glu56, Glu132) were set as flexible residues, while the other two residues (His135, Glu99) were kept as rigid residues. When iron ions were docked into Site B, four residues (His135, Glu56, Glu99, Glu132) were set as flexible residues and the other two (His59, Glu23) were kept rigid. When iron ions were docked into both sites, all six (2 His, 4 Glu) were set as flexible residues. The root of the torsion was defined as the iron itself. Partial charges of irons were set as +1.000.

Grid parameter files and atom-specific affinity maps were built using AutoGrid 4 (Huey *et al.*, 2007). Grid points of map files were either 3.75 nm×4.725 nm×4.425 nm to cover the whole protein in blind docking experiments or 1.875 nm×1.875 nm×

1.35 nm to focus on the di-iron binding pocket in local docking experiments. Both were with grid spacing of 0.0375 nm and were centered at the experimentally determined binding sites. Docking simulations were carried out using the Lamarckian genetic algorithm. The number of iterations was set to 200 and the population size was 150. Other Autodock parameters were used with default values. The RMSD between the lowest energy docked pose and the crystal structure was evaluated for each docking simulation.

3 Docking experiment results

3.1 Blind docking

Blind docking is aimed to examine the possibility and potential sites in metalloproteins for docking with iron ions. In a blind docking experiment, the grid box volume covered the whole protein structure. When 'Site A' was tested for docking with iron ions, all docking conformations converged into two clusters (Fig. 2a). One cluster gathered conformations of lower RMSD (≤0.2 nm) but higher binding energy (≥−10 kJ/mol), indicating that the corresponding iron ions were inside the target binding pocket. The other cluster gathered conformations of lower binding energy (≤−15 kJ/mol) but higher RMSD (≥0.2 nm), indicating that the corresponding iron ions were outside the binding pocket, specifically on the protein surface. Similar docking results were obtained when 'Site B' (Fig. 2b) or 'Site A&B' (Fig. 2c) were tested. The percentages of conformations of lower RMSD were 34.33% in Site A, 22.50% in Site B, and 48.00% in Site A&B.

3.2 Local docking

In local docking experiments, the grid box volume was modified and contained two internal binding sites. Correlations were investigated between the RMSD, binding energy, and conformation distributions (Fig. 3). When Site A, or Site B, or Site A&B was tested for local docking, the binding energy of the most populated conformation was about −3.35 kJ/mol for Site A, 0.50 kJ/mol for Site B, and −7.12 kJ/mol for Site A&B. The RMSD varied between 0.08 nm and 0.12 nm for Site A or Site B, and between 0.12 nm and 0.16 nm for Site A&B. The lowest binding energy and RMSD distribution of Site A implies that the docking results of this site were more reasonable.

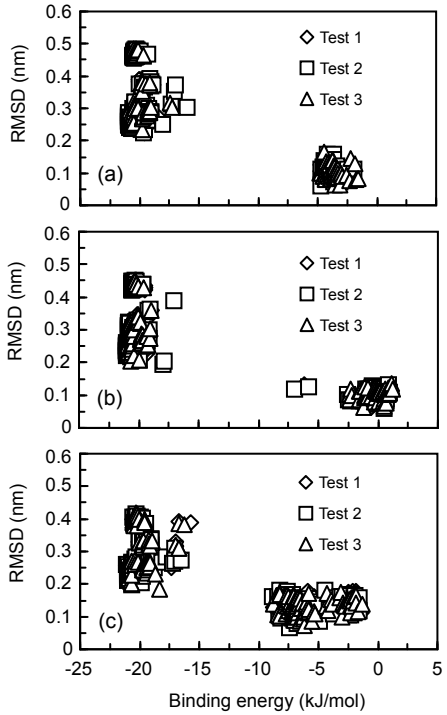


Fig. 2 Docking conformation distributions: RMSD versus binding energy
(a), (b), and (c) are the docking results of Site A, Site B, and Site A&B, respectively

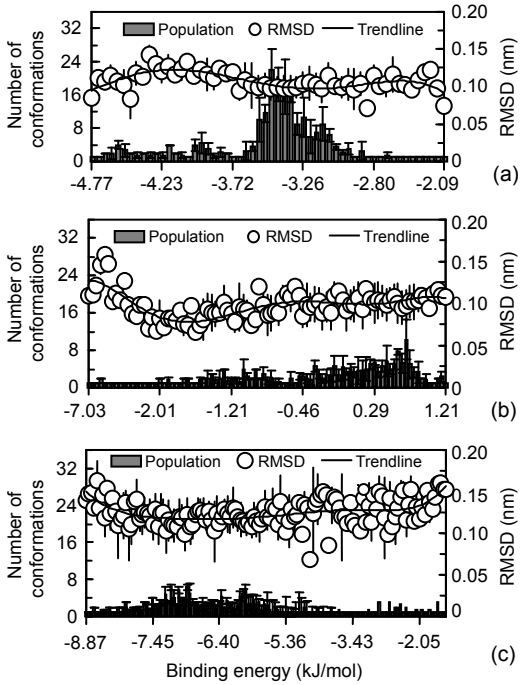


Fig. 3 Local docking results: correlations between docking conformations, RMSD, and binding energy
(a), (b), and (c) are the docking results of Site A, Site B, and Site A&B, respectively

Coordinating residues lock iron ions in a specific geometry and play a critical role in stabilizing binding conformations. To evaluate the quality of binding modes, the RMSD of the distance between iron ions and bound atoms was used as a primary measure of the bond between irons and binding residues. The overall geometry fit was calculated using the RMSD of the angles between docked and experimental geometries. The binding geometry was defined as 'good' when the RMSD of distance fit was within 0.10 nm and the angle was within 35°, or it was defined as 'fair' (distance fit RMSD <0.10 nm, angle fit RMSD >35°). When the conformation was categorized into neither 'good' nor 'fair', it was defined as 'poor' (Hu and Shelver, 2003). Results showed that when Site A was tested for docking, over half of conformations were good and 15% were fair. However, when Site B was tested for docking, none was good and 46% were fair (Fig. 4a). A comparison of coordinating geometries of irons in binding sites was made between the best ranked conformation and the experimental structure (Fig. 4b).

3.3 Effects of coordinating water molecules on docking experiments

Docking experiments for Site A and Site B provided different results, which might be due to the different Fe-O (H_2O) distances in the two binding sites. Therefore, we further investigated the effects of coordinating water molecules on iron docking experiments. When coordinating water molecules were removed in docking experiments, the binding energies were significantly increased (Fig. 5). Although the quality of binding did not change for Site A, the percentage of good conformations for Site B was increased from 0 to 30%.

When the Fe-O (H_2O) distance in Site B was set to 0.15, 0.25, 0.35, or 0.45 nm, distributions of docking conformations and RMSD in relation to the change of binding energy were explored (Fig. 6). The binding energy was increased when the Fe-O (H_2O) distance was increased. Quality analysis of binding modes showed that at the distance of 0.25 nm, the highest ratio of good conformations and the lowest ratio of poor conformations were obtained (Fig. 6e). Fig. 7 shows the resultant docking conformations when the distance from the water molecule to iron was set to 0.25 nm in Site B.

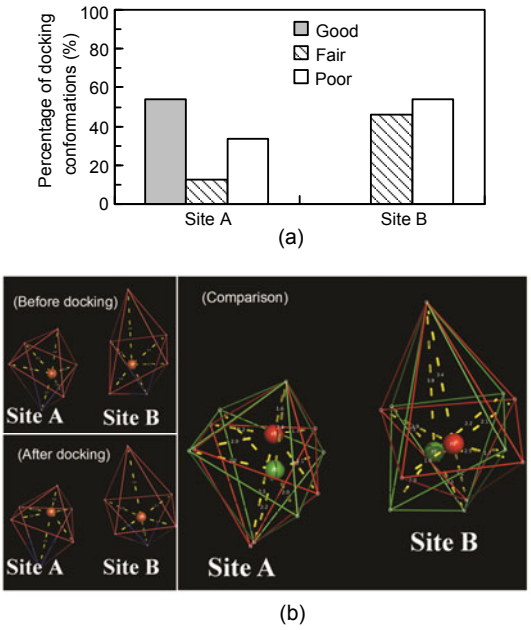


Fig. 4 Quality comparison of docking conformations between Site A and Site B (a) and relevant geometries before and after docking (b)

In (b), coordinating geometries of irons in binding sites are shown before (left up panel in (b)) and after (left bottom panel in (b)) docking experiments. Brown balls represent iron ions. The right panel of (b) shows a comparison between the best ranked conformation and the experimental structure. Green balls and red balls represent iron ions before and after docking, respectively

References to color refer to the online version of this figure

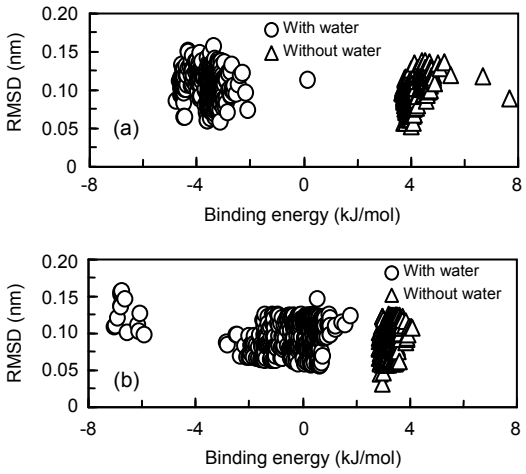


Fig. 5 Comparison of docking conformation distributions with or without coordinating water molecules

(a) and (b) are the docking results of Site A and Site B, respectively

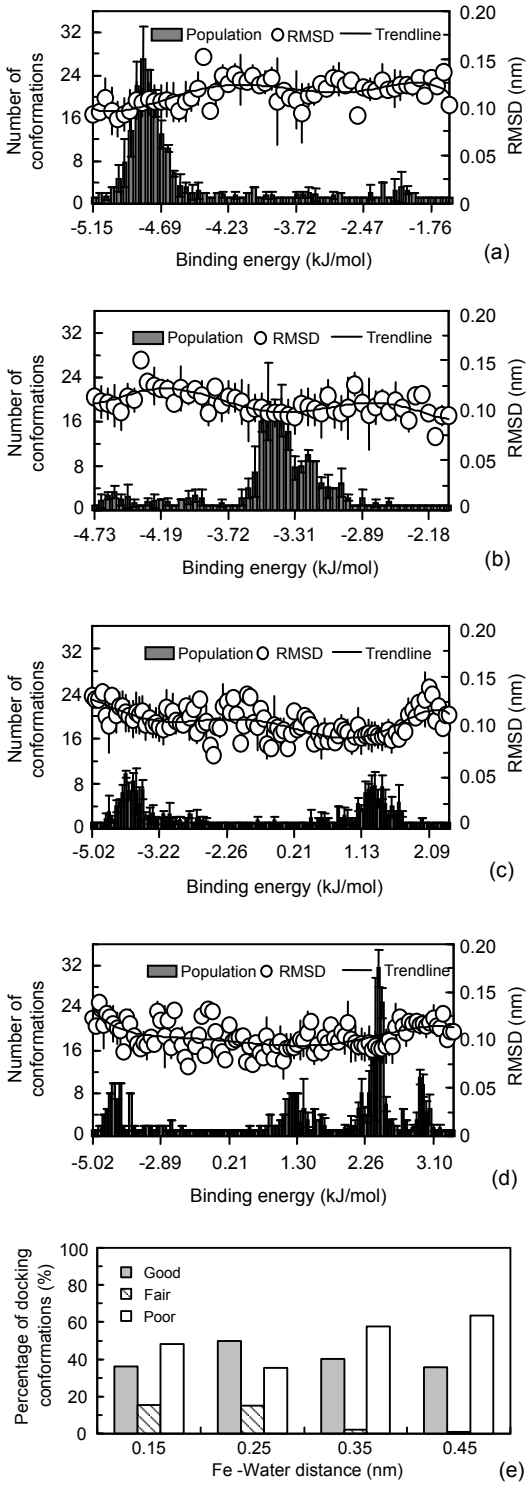


Fig. 6 Effects of Fe-O (H_2O) distance on local docking experiments

(a)–(d) are the docking results when the distance was set as 0.15, 0.25, 0.35, or 0.45 nm for the Fe-O (H_2O) bond, respectively; (e) shows the comparison of the quality of docking conformations of these different Fe-O (H_2O) distances in Site B

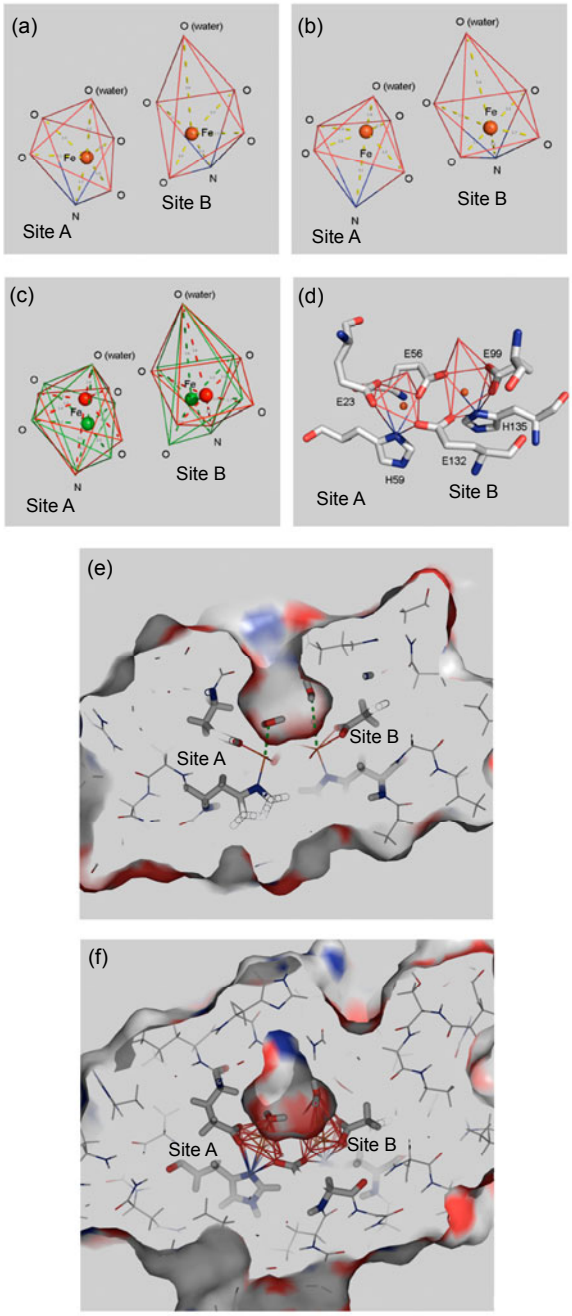


Fig. 7 Resultant docking conformations when the Fe-O (H_2O) distance was set to 0.25 nm in Site B

Coordinating geometries of irons and water molecules in binding sites are shown before (a) and after (b) docking experiments; (c) shows a comparison between the best ranked docking conformation and the experimental structure, where green balls and red balls represent iron ions before and after docking, respectively; (d) shows the positions of residues involved in docking geometry; (e) shows the positions of the coordinating water molecules involved in docking geometry, where water molecules are linked to irons by dashed green lines; (f) shows the docking geometries of irons and coordinating water molecules

References to color refer to the online version of this figure

4 Discussion

In blind docking experiments, lower RMSD of docking conformations implied that iron ions had been docked into targeted sites. However, the binding energies of these conformations were higher than those with higher RMSD (Fig. 2), suggesting that some iron ions were docked into sites on the protein surface rather than internal binding sites. These results indicate that iron ions can be docked into unknown sites in metalloproteins. Docking experiment results are well favorably ranked by the force field of Autodock.

Local docking experiments provided higher docking accuracy than blind docking experiments in this study. In local docking experiments, a lower binding energy distribution of reasonable RMSD (in the range of 0.07 to 0.16 nm) was obtained for the most populated conformations. The binding energies of the most populated conformations were almost the same when Site A and Site A&B were tested for docking. The best ranked poses obtained from local docking were close to the experimental position.

When coordinating water molecules were removed from binding sites, similar docking results were obtained for Site A and Site B (Fig. 5), implying that water molecules significantly affect docking experiments. Analysis of the quality of binding modes showed that the Fe-O (H_2O) distance of 0.25 nm in Site B gave the most stable geometry. Compared with the experimental Fe-O (H_2O) distance in Site A (0.2401 nm) and Site B (0.3859 nm), it is possible that water molecules contribute to the stability of binding geometries (Chen *et al.*, 2007). Therefore, coordinating water molecules in binding sites should not be removed when they are observed in binding sites of experimental X-ray geometries. Relative positions of water molecules can be optimized to enhance the docking accuracy and stability.

Due to multiple iron coordinating geometries and dynamic binding processes, it is not easy to evaluate the accuracy of docking results (Hu and Shelver, 2003). In this study, iron ions interact with relevant residues in an octahedral geometry. Results showed that the influence of binding sites and coordinating water molecules on docking experiments was easily measured with changes of binding modes (distance and angle fit of the RMSD). Therefore, the quality of binding modes can be used as a good measure for evaluating metal docking processes in metalloproteins.

References

Ananthanarayanan, V.S., Kerman, A., 2006. Role of metal ions in ligand receptor interaction: insights from structural studies. *Mol. Cell. Endocr.*, **246**(1-2):53-59. [doi:10.1016/j.mce.2005.11.023]

Bissantz, C., Folkers, G., Rognan, D., 2000. Protein-based virtual screening of chemical databases. I. Evaluation of different docking/scoring combinations. *J. Med. Chem.*, **43**(25):4759-4767. [doi:10.1021/jm001044l]

Böhm, H.J., 1998. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput. Aided Mol. Des.*, **12**(4):309-323. [doi:10.1023/A:1007999920146]

Chen, D., Menche, G., Power, T.D., Sower, L., Peterson, J.W., Schein, C.H., 2007. Accounting for ligand-bound metal ions in docking small molecules on adenylyl cyclase toxins. *Proteins*, **67**(3):593-605. [doi:10.1002/prot.21249]

Ferrara, P., Gohlke, H., Price, D.J., Klebe, G., Brooks, C.L.I., 2004. Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.*, **47**(12):3032-3047. [doi:10.1021/jm030489h]

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., *et al.*, 2004. Glide: a new approach for rapid, accurate docking and scoring. I. Method and assessment of docking accuracy. *J. Med. Chem.*, **47**(7):1739-1749. [doi:10.1021/jm0306430]

Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T., Banks, J.L., 2004. Glide: a new approach for rapid, accurate docking and scoring. Part 2: Enrichment factors in database screening. *J. Med. Chem.*, **47**(7):1750-1759. [doi:10.1021/jm030644s]

Halperin, I., Ma, B.Y., Wolfson, H., Nussinov, R., 2002. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins*, **47**(4):409-443. [doi:10.1002/prot.10115]

Hu, X., Shelver, W.H., 2003. Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. *J. Mol. Graph. Model.*, **22**(2):115-126. [doi:10.1016/S1093-3263(03)00153-0]

Huey, R., Morris, G.M., Olson, A.J., Goodsell, D.S., 2007. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.*, **28**(6):1145-1152. [doi:10.1002/jcc.20634]

Jones, G., Willett, P., Glen, R.C., Leach, A.R., Taylor, R., 1997. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.*, **267**(3):727-748. [doi:10.1006/jmbi.1996.0897]

Khandelwal, A., Lukacova, V., Comez, D., Kroll, D., Raha, S., Balaz, S., 2005. A combination of docking, QM/MM

methods, and MD simulation for binding affinity estimation of metalloprotein ligands. *J. Med. Chem.*, **48**(17): 5437-5447. [doi:10.1021/jm049050v]

Kramer, B., Rarey, M., Lengauer, T., 1999. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins*, **37**(2):228-241. [doi:10.1002/(SICI)1097-0134(19991101)37:2<228::AID-PROT8>3.0.CO;2-8]

Kurtz, D.M.Jr., 1997. Structural similarity and functional diversity in diiron-oxo proteins. *J. Biol. Inorg. Chem.*, **2**(2):159-167. [doi:10.1007/s007750050120]

Morris, G.M., Goodsell, D.S., Huey, R., Olson, A.J., 1996. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput. Aided Mol. Des.*, **10**(4):293-304. [doi:10.1007/BF00124499]

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J., 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, **19**(14): 1639-1662. [doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B]

Spyrakis, F., Amadasi, A., Fornabaio, M., Abraham, D.J., Mozzarelli, A., Kellogg, G.E., Cozzini, P., 2007. The consequences of scoring docked ligand conformations using free energy correlations. *Eur. J. Med. Chem.*, **42**(7):921-933. [doi:10.1016/j.ejmech.2006.12.037]

Venkatachalam, C.M., Jiang, X., Oldfield, T., Waldman, M., 2003. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model.*, **21**(4):289-307. [doi:10.1016/S1093-3263(02)00164-X]

Warren, G.L., Andrews, C.W., Capelli, A.M., Clarke, B., LaLonde, J., Lambert, M.H., Lindvall, M., Nevins, N., Semus, S.F., Senger, S., *et al.*, 2006. A critical assessment of docking programs and scoring functions. *J. Med. Chem.*, **49**(20):5912-5931. [doi:10.1021/jm050362n]