BIOCHIP JOURNAL, Vol. 1, No. 1, 70-75, March 2007

Effect of Quantum Mechanical Charges in Binding Sites of Metalloproteins

Art E. Cho[1]

[1]Department of Biotechnology and Bioinformatics, Korea University Jochiwon, Chungnam 339-700, Korea
Correspondence and requests for materials should be addressed to Art E. Cho (artcho@korea.ac.kr)

Accepted 15 February 2007

Abstract

Conventional docking methods assume fixed charge model from force field parameters. Combined quantum mechanics/molecular mechanics (QM/MM) method has been applied to docking as a variable charge model and shown to exhibit improvement on the docking accuracy over fixed-charge-based methods. However, there are a number of examples for which adoption of variable charge model fails to reproduce the native binding mode. In particular, the method fails more often for metal-ion-containing proteins, metalloproteins. This class of proteins has highly polarized binding sites at which high-coordinate-numbered metal ions reside. We examine the docking results of this group of proteins and analyze the detailed interactions involved. We deduce the mechanism for success and failure of variable charge model. It is argued that extension of QM/MM docking method would correct the over-fitted charges so as to lead to better docking accuracy for docking of metalloproteins.

Keywords: Docking, QM/MM, Charge model, Metalloproteins

Introduction

Protein docking is one of the most used molecular modeling methods in pharmaceutical industry[1-5]. It aims to find the right binding modes of drug-like molecules to target proteins in question. In the past decade, a number of docking approaches have been developed to improve the docking accuracy. Most of current docking methods utilize force-field-parametrized fixed electric charges for ligand atoms[2,6-9]. Employment of the combined quantum mechanics/molecular mechanics (QM/MM) methods to correct these charges on ligand atoms has been proven to lead to better docking accuracy[10]. However, in that work, there are some examples in which the new method fails to improve the results of conventional docking practice. In an attempt to understand why QM/MM docking failed in some examples and to find a way to improve on it, we analyze in detail a particular group of cases, namely, metalloproteins.

Metalloproteins play important roles in physiological processes (hemoglobin, cytochrome oxidase, catalase and superoxide dismutase), the receptor binding of potential drugs (carbonic anhydrases, matrix metalloproteinases, thermolysin, leucine aminopeptidase, phospolipase C, carboxypeptidases A and B, adenosine and cytidine deaminase) and drug metabolism (cytochromes P450 and methane monooxygenase), just to name a few[11]. Docking studies involving metalloproteins pose serious challenge since the ligand interactions with transition metals can be treated appropriately only at the quantum mechanical level[12-14].

It was originally postulated by the authors of aforementioned work that QM/MM docking would aid to improve docking accuracy in highly polarized binding sites, such as in metalloproteins. In fact, in some cases, the new method does help improve the results while retaining already good results by conventional method. However, the improvement is inconsistent over the same target proteins. Although the overall success rate on metalloprotein targets for QM/MM docking protocol is significantly higher than the conventional docking methods, it still remains to be explained why it fails on some of the examples. There have been other efforts to improve docking accuracy on metalloproteins, especially zinc complexes, by deviating from force-field-based atomic charges[15,16]. The authors of these works emphasize on practicality of the methods and therefore attempt to improve the docking accuracy and binding affinity prediction by reparametrization of the metal ion force field. In another work, Sternberg *et al.* used fluctuating-atomic-charge model of force field, which is parametrized by semi-empirical quantum chemical method to study zinc complexes[17]. We employ quantum mechanical calculations, as opposed to different model of force field charges or reparametrization, in an attempt to account for the full quantum effect in the binding sites of metalloproteins.

Results and Discussion

We chose a set of co-crystals, which are complexes of ligands and metalloproteins containing Fe, or Zn. The target proteins are carbonic anhydrase II, methylparaben insulin, and protocatechuate. Structure files were downloaded from Protein Data Bank (PDB) depository and prepared manually for docking. We ran Glide 4.0 standard precision (SP) mode[19] on examples chosen and selected only the ones for which Glide predicts the top scoring poses to be 2.0Å or worse for further analysis but verified if QM/MM docking reproduces the accurate binding mode (below 2.0Å), which regular Glide is able to generate based on its *Emodel* scoring function. The PDB id's and ligands for the complexes used for analysis are depicted in Figure 1. After running QSite[20] for QM/MM energy calculation regarding only ligand as quantum mechanically treated region on the native pose, we substitute atomic charge values on the ligand file with electrostatic potential (ESP) fitted charges. We run Glide using these charge values and compare the result.

Table 1 shows the RMSD values of top scoring pose after redocking of the native ligands for regular Glide and QM/MM docking protocol with QM/MM charges generated at native poses (QM Dock). Among the examples shown in Table 1, 1G46, 1G52, and 2BUR are the ones in which both Glide and QM Dock failed to produce poses under 2Å of RMSD as top scoring ones. For all others, QM Dock was able to either reproduce good results of Glide or improve poor results of Glide. We focus on examples in which Glide failed (1G46, 1G52, 2BUR, 3PCB, 3PCC, 3PCH, 3PCN).

Let us first take a look at 3PCC case. Figure 2 shows the native pose, top Glide SP pose, and top QM Dock pose. The blue sphere represents Fe^{3+} ion. Though in the native pose the ligand is positioned so that the ring part is closer to Fe^{3+}, top Glide SP pose has carboxyl group near the iron instead. This actually seems trivial since in force field parameter, Fe^{3+} is given the partial charge of positive 3, and oxygens in carboxylate group has strong negative charges. This strong Coulombic attraction seems unavoidable

Table 1. RMSD in Å of top scoring poses from native poses for regular Glide and QM/MM docking protocol.

	Glide SP	QM Dock
1G46	4.85	3.59
1G52	4.50	5.33
2BUR	4.12	4.01
3MTH	5.53	0.36
3PCB	3.88	0.11
3PCC	4.09	0.16
3PCE	0.10	0.36
3PCG	0.44	0.22
3PCH	4.05	0.34
3PCJ	0.33	0.12
3PCK	0.25	0.21
3PCN	4.54	0.14

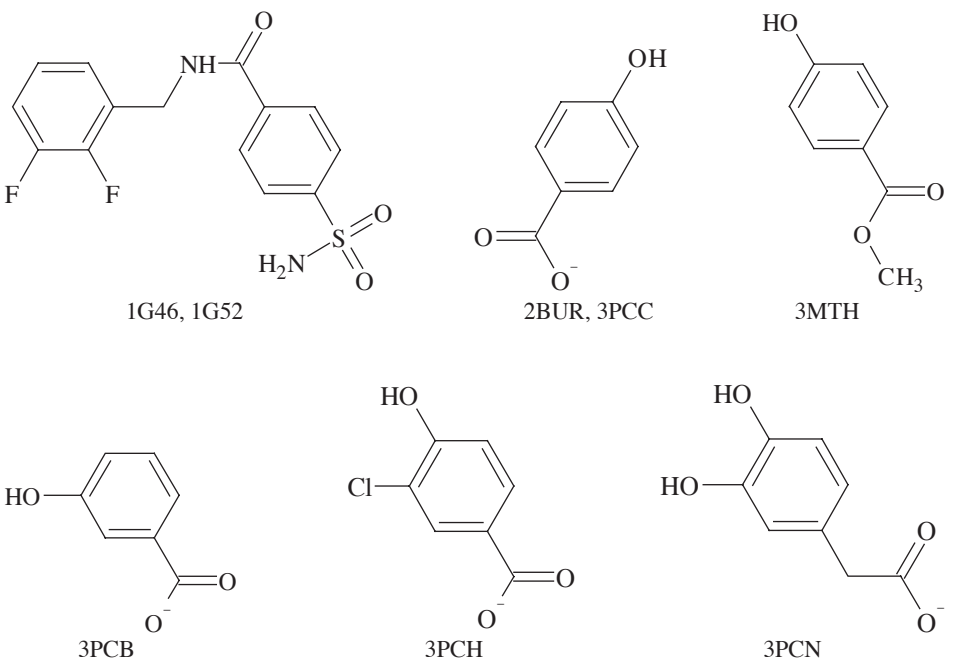

Figure 1. Structures of ligands used for analysis. From top, in PDB id: 1G46 and 1G52, 2BUR and 3PCC, 3MTH, 3PCB, 3PCH, 3PCN.

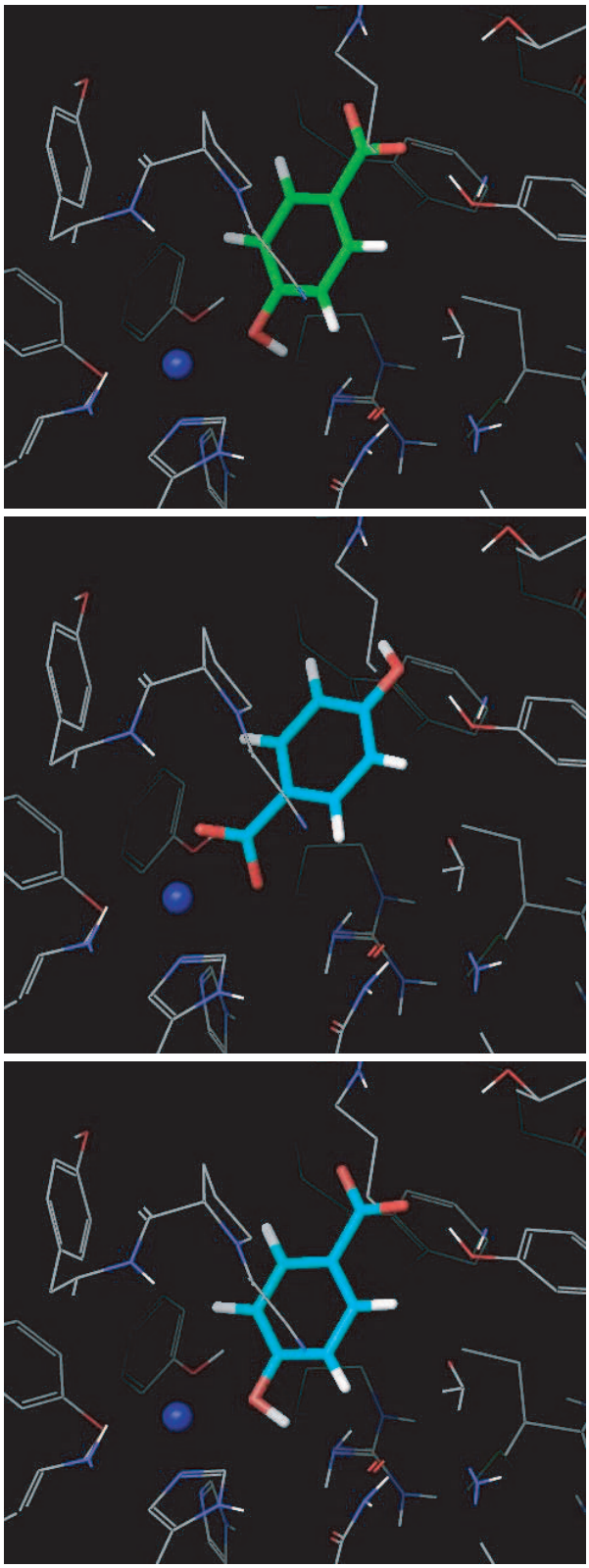

Figure 2. Docked poses for 3PCC. Native pose (top), top Glide SP pose (middle), and top QM Dock pose (bottom) are depicted.

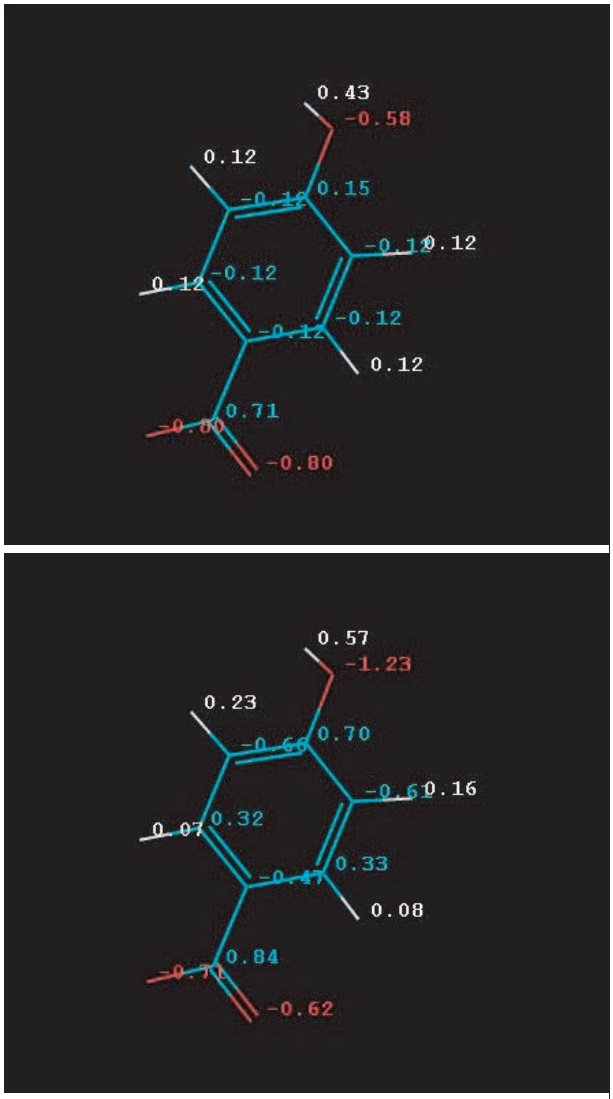

Figure 3. Partial charges on 3PCC ligand. Top picture shows partial charges assigned by force field parameters while bottom picture shows quantum mechanically calculated charges.

in the force field charge configuration. QM/MM calculation at native pose altered the charge configuration so that the native pose entails lower electrostatic energy as seen in Figure 3. The unusually high negative charge on the oxygen attached to the ring compensates the Coulombic attraction between carboxyl group and the iron. This quantum mechanically corrected charge configuration obviously enables QM Dock to find the correct pose to be the top scoring one. The same analysis can explain other cases in which QM Dock improves the docking accuracy dramatically over Glide SP (3MTH, 3PCC, 3PCH, 3PCN).

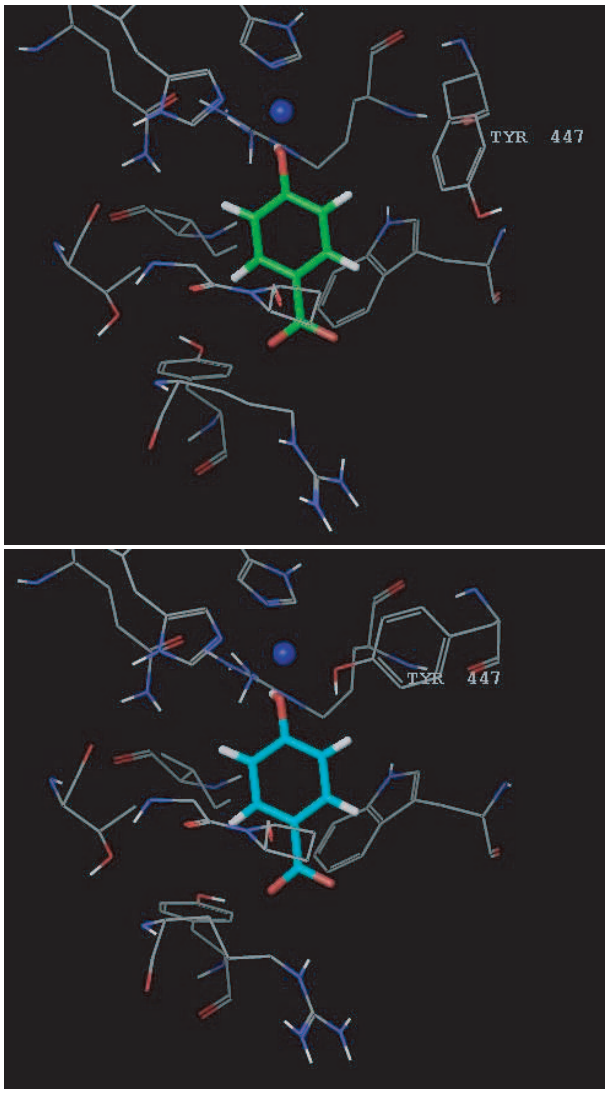

Figure 4. Binding sites of protocatechuate 3, 4-dioxygenase co-crystals. The difference between 2BUR (top picture) and 3PCC (bottom picture) is in conformation of Tyr 447.

Let us now turn our attention to 3 cases in which both Glide SP and QM Dock failed (1G46, 1G52, 2BUR). 2BUR has the same ligand as in 3PCC but mutated protein. The mutation causes conformational change within the binding site. Figure 4 shows the binding sites and their proximity of 2BUR and 3PCC complexes at their native poses. There is conformational change of the residues all around but most notably, Tyr 447 essentially flips between two conformations. In 3PCC, Tyr 447 is flipped towards the iron center, which makes the proximity around the center less positive, whereas in 2BUR, it points away leaving wide open vicinity. In case of 2BUR, the carboxylate group of the ligand would be attracted towards

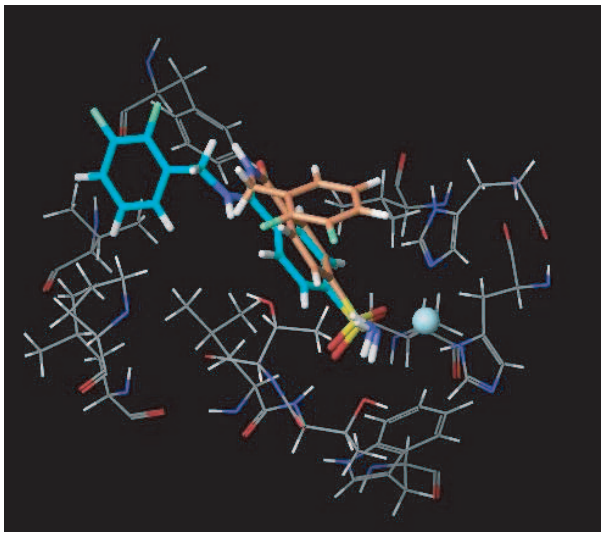

Figure 5. Binding site of 1G52. The native pose is on the left side and the top QM Dock pose is on the right. QM Dock pose has the chloride ring bending towards the zinc center.

the iron center, even though with quantum correction the oxygen attached to the ring has large negative charge. The correction is not enough to align the ligand so that the ring points toward the iron center.

In Figure 5 are the conformations of the native pose and top QM Dock pose of 1G52. Top-scoring QM Dock pose has rather high RMSD of 5.33Å because of the chloride ring bending towards zinc center. Obviously, the negative charges on chlorines would be attracted to positive charge of zinc yielding lower electrostatic energy. The difficulty of reproducing native pose of 1G46 can be explained in a similar manner.

Conclusions

The failure of QM Dock in the above examples of metalloproteins can be attributed to the excessive electrostatic energies stemming from large positive charges assigned to metal ions. In fact, it has been pointed out by other researchers that partial charge for zinc from force field parameter is excessive and they have tried to correct the problem by the use of optimized zinc parameters[15,21-25]. The metal centers located in binding sites surrounded by protein residues would inevitably absorb negative charge from the surroundings and thus alter the electrostatic potential surface of the binding pocket and the protein residues as well as their own atomic charges. Without addressing this effect in detail, docking practice to these target proteins would result in futility even

though occasional success can be achieved by compensating scoring function in other ways. We claim that in order to properly model the charge transfer within binding sites of metalloproteins, one must resort to quantum mechanical treatment although that would mean great increase in computational cost, which in most cases is impractical especially for industrial virtual ligand screening environment. Further work should be carried out in this regard.

Methods

In the earlier work of QM/MM docking, a docking protocol coined as "Survival of the Fittest[10]" was implemented. In that work, an initial docking with regular force field is performed to produce a set of poses that will feed into QM/MM one-point energy calculations regarding only ligands as quantum mechanical region, which in turn were fitted to generate a new set of atomic charges based on density functional theory (DFT)[18] quantum mechanical calculations. Using this new set of charges, a new generation of docking runs is performed and in the end the best scoring pose is selected. The key idea in this protocol is that at a pose that is close to the native structure, quantum mechanical calculation will produce atomic charges that give rise to lower Coulombic energy values at native pose. As a validation study, the authors performed QM/MM calculation at the native pose of each complex to adjust the atomic charges on ligand atoms. This calculation would generate a set of charges in a given binding site environment that is theoretically the best possible one for docking. We follow the same procedure for our examples and submit the result for further analysis.

Acknowledgements

The author would like to thank Schrödinger, Inc. for providing its software. This work was partially supported by the Second Stage of Brain Korea 21 project in 2006.

References

1. Halperin, I., Ma, B.Y., Wolfson, H. & Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins-Structure Function and Genetics* **47**, 409-443 (2002).
2. Gschwend, D.A., Good, A.C. & Kuntz, I.D. Molecular docking towards drug discovery. *Journal of Molecular Recognition* **9**, 175-186 (1996).
3. Abagyan, R. & Totrov, M. High-throughput docking for lead generation. *Current Opinion in Chemical Biology*. **5**, 375-382 (2001).
4. Kuntz, I.D. Structure-Based Strategies for Drug Design and Discovery. *Science* **257**, 1078-1082 (1992).
5. Drews, J. Drug discovery: A historical perspective. *Science* **287**, 1960-1964 (2000).
6. Cho, A.E. *et al*. The MPSim-Dock hierarchical docking algorithm: Application to the eight trypsin inhibitor cocrystals. *Journal of Computational Chemistry* **26**, 48-71 (2005).
7. Brooijmans, N. & Kuntz, I.D. Molecular recognition and docking algorithms. *Annual Review of Biophysics and Biomolecular Structure* **32**, 335-373 (2003).
8. Kramer, B., Metz, G., Rarey, M. & Lengauer, T. Ligand docking and screening with FlexX. *Medicinal Chemistry Research* **9**, 463-478 (1999).
9. Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. & Taylor, R.D. Improved protein-ligand docking using GOLD. *Proteins-Structure Function and Genetics* **52**, 609-623 (2003).
10. Cho, A.E., Guallar, V., Berne, B.J. & Friesner, R. Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *Journal of Computational Chemistry* **26**, 915-931 (2005).
11. Lipscomb, W.N. & Strater, N. Recent advances in zinc enzymology. *Chemical Reviews* **96**, 2375-2433 (1996).
12. Deerfield, D.W., Carter, C.W. & Pedersen, L.G. Models for protein-zinc ion binding sites. II. The catalytic sites. *International Journal of Quantum Chemistry* **83**, 150-165 (2001).
13. Dudev, T. & Lim, C. Metal binding in proteins: The effect of the dielectric medium. *Journal of Physical Chemistry B* **104**, 3692-3694 (2000).
14. Rulisek, L. & Havlas, Z. Theoretical studies of metal ion selectivity. 1. DFT calculations of interaction energies of amino acid side chains with selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). *Journal of the American Chemical Society* **122**, 10428-10439 (2000).
15. Hu, X. & Shelver, W.H. Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. *Journal of Molecular Graphics & Modelling* **22**, 115-126 (2003).
16. Khandelwal, A. *et al*. A combination of docking, QM /MM methods, and MD simulation for binding affinity estimation of metalloprotein ligands. *Journal of Medicinal Chemistry* **48**, 5437-5447 (2005).
17. Sternberg, U., Koch, F.T., Brauer, M., Kunert, M. & Anders, E. Molecular mechanics for zinc complexes with fluctuating atomic charges. *Journal of Molecular Modeling* **7**, 54-64 (2001).
18. Kohn, W. & Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **140**, 1133-1138 (1965).

19. Schrodinger, I., Portland, Oregon. *Glide* 4.0, (2006).
20. Schrodinger, I., Portland, Oregon. *QSite*, (2006).
21. Stote, R.H. & Karplus, M. Zinc-Binding in Proteins and Solution-a Simple but Accurate Nonbonded Representation. *Proteins-Structure Function and Genetics* **23**, 12-31 (1995).
22. Donini, O.A.T. & Kollman, P.A. Calculation and prediction of binding free energies for the matrix metalloproteinases. *Journal of Medicinal Chemistry* **43**, 4180-4188 (2000).
23. Hou, T.J., Guo, S.L. & Xu, X.J. Predictions of binding of a diverse set of ligands to gelatinase-A by a combination of molecular dynamics and continuum solvent models. *Journal of Physical Chemistry B* **106**, 5527-5535 (2002).
24. Pang, Y.P., Xu, K., El Yazal, J. & Prendergast, F.G. Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach (vol 9, pg 1857, 2000). *Protein Science* **9**, 2583-2583 (2000).
25. Pang, Y.P. Novel zinc protein molecular dynamics simulations: Steps toward antiangiogenesis for cancer treatment. *Journal of Molecular Modeling* **5**, 196-202 (1999).