Correlation between Performance of QM/MM Docking and Simple Classification of Binding Sites

Jae Yoon Chung,[†,‡] Jung-Mi Hah,*[,‡] and Art E. Cho*[,†]

Department of Bioinformatics and Biotechnology, Korea University, Jochiwon, Chungnam, Korea, and Life Sciences Research Division, Korea Institute of Science and Technology, Seoul, Korea

Received July 1, 2009

Use of SiteMap for binding site classification and its connection to QM/MM (quantum mechanics/ molecular mechanics) docking performance were investigated. Using the hydrophilic/hydrophobic character values along with balance between them which SiteMap calculates, we sorted 455 cocrystal complexes available from protein data bank into three groups and tested how Glide, a conventional docking program, and QPLD, a QM/MM docking program, perform on them. QPLD showed improvements on all three groups over Glide but with varying degrees. Analysis of the results was carried out, and establishment of correlations between the classification of binding sites and QM/MM docking performance was attempted. It was found that QM/MM docking delivers the most improvements for primarily hydrophobic binding sites with substantial hydrophilic interactions. Based on our findings, we make suggestions for use of QM/MM docking and directions for further developments.

INTRODUCTION

Being an integrated part of computer-aided drug design (CADD), protein docking has been a subject of intensive research for the last couple of decades.[1–3] Docking programs are usually used as a screening tool for lead compounds to predict docking poses of ligands and target proteins.[4] In doing so, a docking program generates a set of poses, which are relative conformations of ligand and receptor complex, and then scores them to finally produce pose predictions.[5–7] During this process, "scoring" is done with energy-like quantity that resembles more or less binding affinity.[8,9] Though there have been attempts to develop scoring functions in docking to be used as binding affinity prediction engine,[10] it is widely believed that today's docking algorithms are only good for pose predictions. It is also of general consensus that search algorithm in docking, which exploits the configuration space of ligand−receptor poses, is reliable enough, while scoring function still needs more development.[11] In an attempt to develop a docking algorithm which can predict poses accurately for cases which conventional docking programs fail to describe, we had combined quantum mechanical/molecular mechanical (QM/MM) calculations with scoring in docking.[12] The result was substantial improvement over conventional docking practice across various targets in terms of docking pose prediction. In essence, what we did was to distort the scoring function space by accentuating electrostatic interactions so that optimal pose with better electrostatic energy can be selected. This modification might complicate the predictive power of scoring function for binding affinity, but searching for the right binding mode would be improved provided that electrostatic interaction plays a substantial role in the given protein−ligand binding. For this reason, we initially conjectured that targets with polar binding sites can be described better by the new algorithm. Instead, however, we found that for metalloproteins, which are extreme cases of polar binding sites, our new method fails more than it does for other cases.[13,14] This fact led to an extension of the QM/MM docking method, which was successfully applied to the description of proteins with metal ions in the binding sites.[15] Still, we have not proven the better applicability of QM/MM docking to polar binding sites other than metalloproteins. To this end, we carried out a project in which we used SiteMap of the Schrödinger suite to classify binding sites of 455 cocrystals downloaded from PDB (protein data bank) in a simple way and checked the correlation between this classification and performance of QM/MM docking methods on them. The idea is that with a simple classification of binding sites and by correlating it to the QM/MM docking performance, one can determine which method should be used during preparation for docking. This is necessary since QM/MM docking is significantly more time-consuming and should be used only if one can gain significantly by using it. In the following sections, we describe SiteMap that was used for classification of binding sites, and then we briefly review our QM/MM docking protocol. Statistical analysis of the results correlated with binding site classification will follow. We discuss our findings and conclude with suggestions of QM/MM docking usage.

METHODS

Preparation of the Test Set. We used the list of 455 complexes of the Schrödinger test set, which is a subset of the PDBbind database.[16,17] This list represents a variety of targets that are relevant for drug development. We downloaded each of the structure files from the protein data bank (PDB) Web site and prepared them for docking using Protein Preparation Wizard, which is available as part of the

* Corresponding author e-mail: artcho@korea.ac.kr, jhah@kist.re.kr.
† Korea University.
‡ KIST.

10.1021/ci900231p CCC: $40.75
Published on Web 10/02/2009

Schrödinger suite. Hydrogen atoms were added according to Epik[18] calculation for pK_a values, and minimization up to 0.3 Å rmsd was performed for each complex.

Binding Site Classification. Based on our previous research,[14,15] we set a rule for classification of the test set into three groups: predominantly hydrophobic binding sites, predominantly hydrophilic binding sites, and metal-ion contained binding sites. SiteMap[19] scans the surface of protein and finds possible binding sites. At the same time, it analyzes those binding sites. During this analysis, a number of property values are calculated including hydrophobic and hydrophilic character. SiteMap constructs a measure of hydrophilicity by adding an "electric-field reward" term to the van der Waals (vdW) interaction energy given by

$$\text{Grid\_philic} = \text{vdW\_energy} + \text{oriented-dipole\_energy}$$

where the oriented-dipole energy is calculated as electrostatic energy of a point dipole simulating a water molecule oriented along the electric field at the grid point and thus necessarily negative. Hydrophilic regions are those within which the sum of the two terms is sufficiently negative. Conversely, the quantity representing hydrophobicity is constructed by adding an oppositely signed (positive) "electric-field penalty" term to the vdW term

$$\text{Grid\_phobic} = \text{vdW\_energy} - 0.30 * \text{oriented-dipole\_energy}$$

The hydrophobic and hydrophilic character of the site is computed by averaging the Grid_phobic or Grid_philic potential over the original site points and the extension points. The balance expresses the ratio of the two. The hydrophobic and hydrophilic scores are calibrated so that the average score for a submicromolar site is 1.0. The average balance score, on the other hand is about 1.6, not 1.0, because the ratios computed for sites that have high hydrophobic but low hydrophilic scores make large contributions to the average. We use the balance as an indicator if the binding site is more hydrophobic than hydrophilic or vice versa. Since the balance is defined as hydrophobicity divided by hydrophilicity, if it is bigger than 1, the binding site is more hydrophobic, whereas if it is smaller than 1, more hydrophilic. This classification is probably too crude, but it will nevertheless give a measure of the polarization within the binding site. Metalloproteins are a peculiar group in that not only they have highly polarized binding sites but also there is a high probability that electron transfer can happen within the binding sites due to the presence of the metal ions. The original QM/MM docking protocol cannot describe the electron transfer occurring between ligand and protein atoms (and/or metal ions) since the definition of QM region is limited to ligands. For this group of proteins, we have proposed extended QM/MM docking,[15] which we will not discuss in this paper because implementation of it requires a wholly different protocol. A thorough test of the protocol, which would focus on metalloproteins, should be left for another project. We nonetheless identify metalloproteins, with the definition that those proteins in the binding site of which exist(s) metal ion(s) within 3 Å, and group them separately, resulting in a total of three separate groups. It should be noted that the balance (of hydrophobic/hydrophilic) can be either bigger or smaller than 1 for metalloprotein group members as we have classified, since the hydrophobic character calculated often exceeds that of hydrophilic ones even if a metal ion exists in the binding site.

Docking. To test the improvement QM/MM docking might bring, we compared the results with those of regular docking method. To make a fair comparison, we used Glide 4.0[20] for this part. The docking algorithm in Glide utilizes a hierarchical search protocol, in which the final step is minimization of a flexible ligand in the field of the Coulomb and van der Waals potential of the protein, as represented by the OPLS-AA molecular mechanics potential energy function. Selection of the final ligand pose is primarily determined by the total Coulomb-van der Waals energy with the Coulomb energy screened by a distance-dependent dielectric constant. The scoring function, called GlideScore, for computing binding affinity is an extension of an empirically based Chem-Score function of Eldridge et al.[21]

QM/MM Docking. We used QPLD[22] (QM-Polarized Ligand Docking) workflow, available as part of Schrödinger suite, for QM/MM docking implementation. QPLD utilizes QSite and Glide to carry out our SOF (survival of the fittest) algorithm.[12] In the SOF algorithm, one performs conventional docking to produce a prescribed number of initial poses for a given target and a ligand. With these initial docked poses, one subsequently runs QM/MM single energy calculations on each of them with the ligand only as the QM region, producing new sets of atomic charges on the ligand by ESP (electrostatic potential) fitting. As previously noted, inclusion of atoms other than those of ligands (such as metal ions and/or surrounding protein atoms) in the QM region requires a protocol different from SOF because of the QM/MM border treatment problem and such a method is not featured in QPLD. Hence, we set only ligands as the QM region. Finally, redocking with these new atomic charges is done, and the best scoring pose is selected. Through QM/MM calculations, the pose which is close enough to the crystal structure would yield a new set of charges that will give rise to better redocking results. Thus, one can obtain an improved docking pose prediction. The QPLD version we used employs Glide version 4.0 and QSite[23] version 3.5. The initial number of poses was set to 5 and for the QM part, the 6-31G* basis set and the B3LYP functional were used in density functional theory (DFT) calculations.

RESULTS AND DISCUSSION

Overall Performance Comparison. The complete list of the test set is provided in Supporting Information. This test set is separated into three groups as previously noted. The three groups are hydrophobic, hydrophilic, and metalloprotein. The hydrophobic group had 147 members, hydrophilic group 201, and metalloprotein group 107.

The complete results of Glide and QPLD are listed in Table 1. Glide 4.0 overall predicts poses under 2.0 Å rmsd (root-mean-square deviation) to crystal structures about 77% of time. For under-1.0 Å, the success rate was 61%. The average rmsd of predicted poses by Glide for all the test cases is 1.485 Å. Glide 4.0 performed significantly better in the hydrophilic group than in the hydrophobic group. This reflects the fact that even regular force field based docking programs have electrostatic energy as a main component of their scoring function. It is generally more difficult to model

Table 1. Results of Glide and QPLD Runs

	rmsd < 1.0				rmsd < 2.0					
	Glide		QPLD		Glide		QPLD		average	
	#	%	#	%	#	%	#	%	Guide	QPLD
overall	278	61.10	319	70.11	351	77.14	397	87.25	1.485	1.078
hydrophilic	138	68.66	149	74.13	168	83.58	181	90.05	1.244	0.957
hydrophobic	81	55.10	98	66.67	111	75.51	127	86.39	1.595	1.148
metalloprotein	59	55.14	72	67.29	72	67.29	89	83.18	1.788	1.208

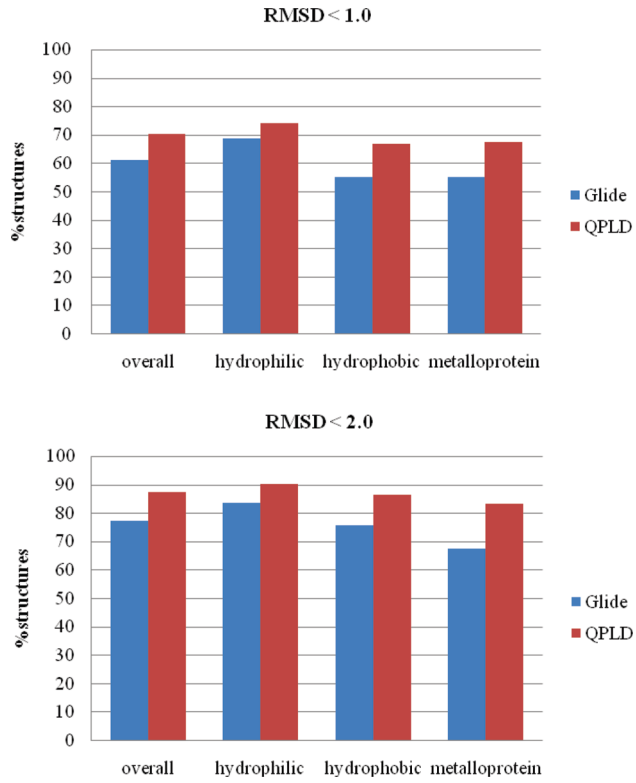

Figure 1. Comparison of Glide and QPLD performances.

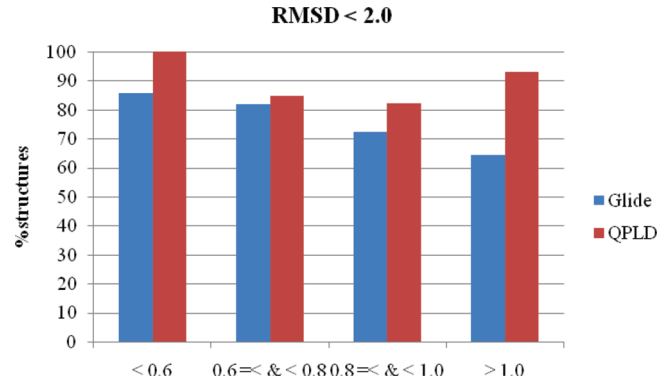

Figure 2. Performance comparison on hydrophobic group divided by hydrophilic character value (blue: Glide, red: QPLD).

hydrophobic interactions, which lack the traditional physics-based binary interaction description. For the metalloprotein group, it is well-known that most conventional docking programs do not work well.[24] This is due to a few reasons including inadequate force field parameters for metal ions. Glide 4.0 achieves 67% of the success rate for predicting poses that are under 2.0 Å for the metalloprotein group, which is 10% lower than overall and more than 16% lower than hydrophilic cases.

QPLD performed better than Glide in all groups as illustrated by comparative histograms in Figure 1. Overall, QPLD scored more than 10% better in a number of poses under 2.0 Å. This is also reflected in the average rmsd values with improvement of more than 0.4 Å as QPLD's average was 1.078 Å. QPLD's prediction of binding poses for hydrophilic cases was especially notable exceeding 90% for poses under 2.0 Å. However, it turned out that QPLD's predictive power is nearly as good for hydrophobic cases being more than 86% for the same measure. In fact, the improvement QPLD achieved over Glide on the hydrophobic group was more than on the hydrophilic group: 6.5% for hydrophilic vs almost 11% for hydrophobic. This fact is rather counterintuitive that we will elaborate on it in the next section. For metalloprotein, the improvement was the greatest yielding nearly 16% in the number of poses under 2.0 Å. Though the performance of QPLD for metalloproteins are not as good as in hydrophilic group and the poorest among the three groups, the advantage one gains over Glide was actually the biggest. Unlike the hydrophobic group, this result was certainly expected, nonetheless. More often than not, highly polarized characteristic in binding sites of metalloproteins will be grasped by QPLD better than Glide. The failure of QPLD for metalloproteins because of overestimation of the polarization in the binding sites without description of charge transfer between metal ions and protein atoms was addressed and dealt with in our earlier work,[15] which accounts for the fact that QPLD still fails more often for metalloprotein group than other two.

Hydrophobicity Dependence. Glide performed reasonably well for hydrophobic binding sites in our test. It has been reported that Glide, in contrast to other docking programs, performs rather well for hydrophobic binding sites. In fact, in the work of Perola et al.[24] it was shown that Glide performs better for hydrophobic sites than highly hydrophilic sites (their classification was based on the number of hydrogen bonds). Although our result indicates that, according to our classification, Glide still performs better for hydrophilic binding sites, its performance on hydrophobic sites is respectable yielding 75% success rate for prediction of binding poses under 2 Å rmsd. QPLD surprisingly improved this number by more than 10%. This is curiously in disaccord with our original idea that since QPLD rescales atomic charges using quantum mechanical calculations, it will give a better description for polarized binding sites rather than nonpolar ones. Noting that our definition of hydrophobic (hydrophilic) binding sites is based on the balance between the 2 character values, we subdivided the hydrophobic group according to their hydrophilic character values. Figure 2 shows the plot of the success rate against the hydrophilic value for the hydrophobic group. While the group with the lowest hydrophilic value exhibited a high success rate for both Glide and QPLD, for intermediate and high hydrophilic

values, Glide's success rate dropped gradually. As a result, the improvement QPLD brought was the biggest in high hydrophilic value group, reaching almost 30%. This result can be explained by the fact that indeed with QPLD large polarization can help the program to find correct binding modes. In Figure 3, this point is clearly illustrated. Panel a) is QPLD predicted pose and b) Glide predicted pose with SiteMap calculated hydrophilic/hydrophobic regions shown by colored surface representation for penicillin acylase-dihydroxyphenyl acetic acid complex (PDB id: 1ai4). The yellow colored region is designated hydrophobic, whereas the green colored region is hydrophilic. The hydrophobic region in this binding site is bulky and concentrated with a character value of about 2.0. The hydrophilic region is scattered with a character value of about 1.0. The balance thus indicates that this binding site is hydrophobic. Naturally, the ring of the ligand lies in the hydrophobic region for both poses. However, since the ligand has a carboxylate group, in order to find the correct binding mode, an accurate description of the interaction between the charged group and hydrophilic region is necessary. The atomic charges on two oxygen atoms of carboxylate group assigned by the OPLS force field are both −0.8e giving a symmetric charge configuration for the functional group. On the contrary, the QM/MM calculation on a near-native pose gives rise to an asymmetric charge configuration with one −0.95e and the other −0.76e. This difference in charge configuration leads to two distinctive hydrogen bonding patterns. While both oxygen atoms of carboxylate group form hydrogen bonds with protein atoms in the Glide-predicted pose, only the oxygen atom with charge −0.9e forms multiple hydrogen bonds in the QPLD-predicted pose. These hydrogen bond patterns explain the deviation in the overall geometry. In panel c), we see that the native pose shows the same hydrogen bond pattern as the QPLD-predicted pose does. As a consequence, Glide's top scoring pose has an rmsd of 3.75 Å, while QPLD's has an rmsd of 0.42 Å. QPLD's modification of charges on the carboxylate group has indeed made the ligand be oriented in the right way.

Hydrogen Bond Dependence. As in ref 24, we also adopted an approach in which counting hydrogen bonds between protein and ligand works as an indicator for hydrophilic character of binding sites. Figure 4 shows the plot of degree of hydrogen bonding (DHB), which is the number of hydrogen bodns (HB) divided by the number of heavy atoms (HA), against docking success rates. The trend is clearly that both Glide and QPLD perform gradually better as DHB increases, but the improvement of QPLD over Glide is the greatest for medium range of DHB. This is probably so because though more hydrogen bonds mean more interactions QPLD can describe accurately, in primarily hydrophilic environment Glide already does it well leaving little room for improvement. A large number of hydrogen bonds also means that the size of the ligand is big enough that conformational search becomes the limiting factor, which applies to both Glide and QPLD in the same way.

Summary of Findings and Suggestions for Use. QPLD was predicted to perform best under a hydrophilic environment because of its atomic charge rescaling protocol. Indeed, in our classification of hydrophilic/hydrophobic binding sites, QPLD's result substantially exceeded in the former than the latter. However, the same is true with Glide, which is the underlying docking program of QPLD. What was surprising was the fact that the improvement QPLD made over Glide for the hydrophobic group was significantly more than that for the hydrophilic group. This result was examined in more detail. Since the hydrophobic group was defined to be those proteins whose binding sites have more hydrophobic character than hydrophilic one, one can still divide this group by the degree of hydrophilicity. Within this group with such division, QPLD brought the most improvement for hydrophobic group members with a higher number of hydrophilic character. This characteristic of mainly hydrophobic with a number of hydrophilic interactions (such as hydrogen bonds) is possessed by many important drug targets.[25–27] Running QPLD is unquestionably more time-consuming than Glide, and hence it would be important to sort out cases in which it is most effective. In this regard, the hydrophobic group, according to our definition with SiteMap, with a high value of hydrophilic character should be the targets that can benefit from the use of QPLD most.

It is interesting to note that recent work by Friesner and co-workers about hydrophobic enclosure in protein−ligand binding focuses on a similar binding site - namely, a well-defined hydrophobic pocket with hydrophilic character.[28] In that work, the authors used molecular dynamics simulation to argue that the high affinity of the streptavidin−biotin complex can be attributed to the hydrophobic nature of the binding site. In addition, they showed that correlated hydrogen bonds can explain the geometry of the ligand in the binding site. SiteMap calculation shows that this binding site has a hydrophobic characteristic (balance ∼2.0) with strong hydrophilic character (∼0.9). QPLD apparently can help find the right geometry by exclusively elaborating on the latter.

For the metalloprotein group, QPLD also makes great improvement, but the resulting success rate is not as high as that for the hydrophobic/hydrophilic group. As we alluded to in other work, this problem could be remedied by rescaling the atomic charges on metal ions and then running QPLD.

CONCLUSIONS

On an extended set of cocrystals, we tested and compared the performance of Glide and QPLD to elucidate the effect of quantum mechanically modified atomic charges on docking accuracy. In doing so, we utilized SiteMap, which is a program that can identify binding sites on a target protein and characterize them. Within our definition of hydrophobic and hydrophilic groups, QPLD seemed to outperform Glide by a large degree for hydrophobic binding sites with hydrophilic interactions. This finding suggests the use of QPLD on specific targets, which can be isolated by SiteMap calculations. Further work should be in the following directions.

First, QPLD seems to give great improvement over Glide for the metalloprotein group, which confirms the conjecture that it will be useful for the description of highly polarized binding sites. Yet, it still fails in more cases than other groups. We have shown that the failure of QPLD for metalloproteins originates from the simple fact that force field parameters are not apt for representation of charges on metal ions that are shared by surrounding protein atoms. From the result presented in this paper, it can be readily guessed that

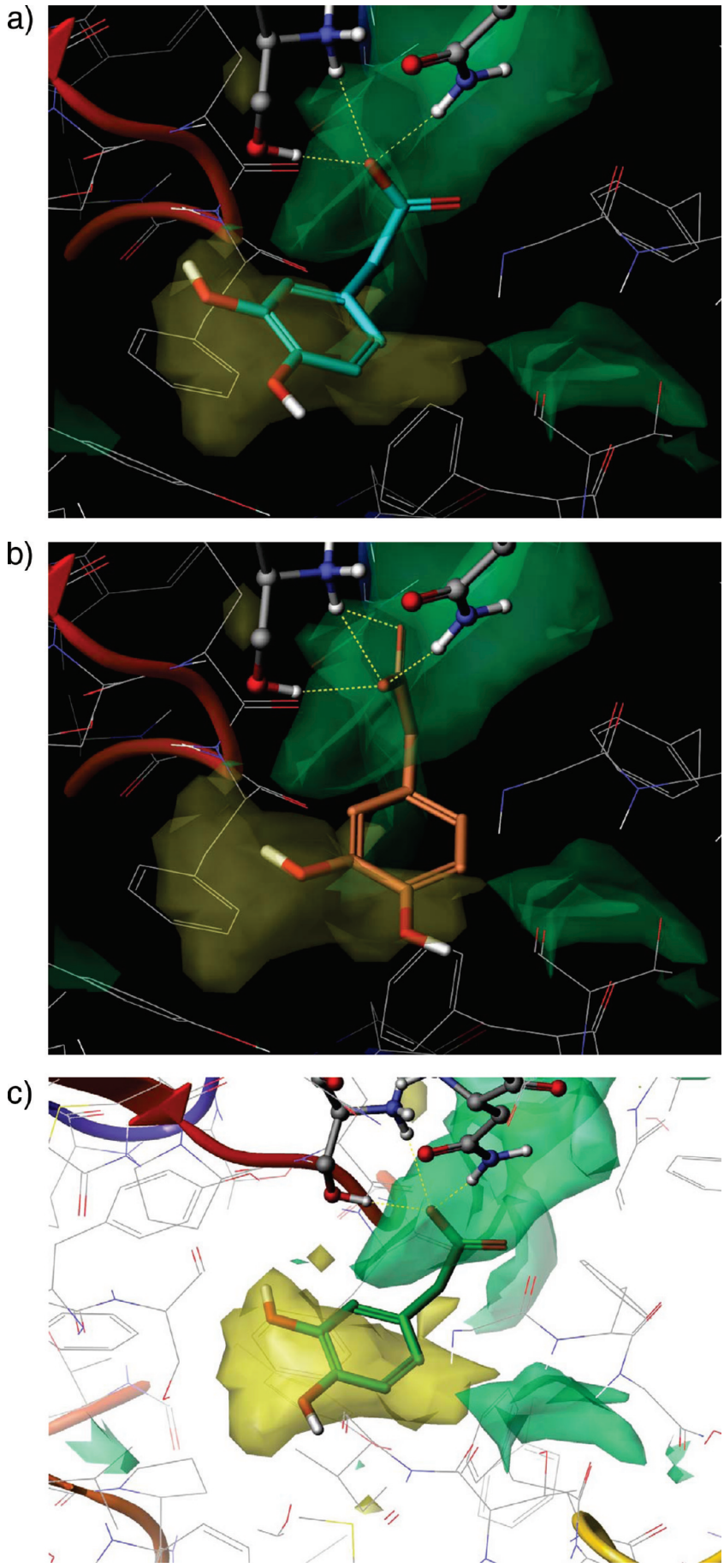

Figure 3. QPLD-predicted binding pose (a), Glide-predicted binding pose (b), and native pose (c) with colored region representation of SiteMap output for the binding site of 1ai4. Yellow color symbolizes hydrophobic and green color hydrophilic.

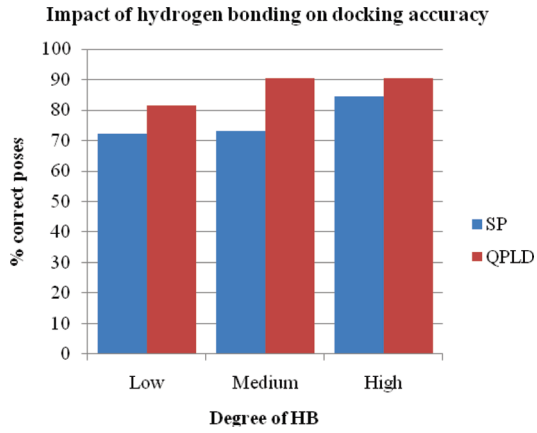

Figure 4. Performance comparison according to degrees of hydrogen bonding.

combining QPLD with prerescaling of metal atomic charges would produce superior results.

Second, SiteMap provides a wealth of information in the form of computed properties other than hydrophobic/hydrophilic character. Our simple classification can be enhanced by this array of values so as to characterize binding sites in a more refined fashion and make suggestions for appropriate programs/methods to be used for drug screening.

Third, the previously mentioned work of Young et al.[28] tried to explain the high binding affinity of a particular kind of binding sites using hydrophobic enclosure argument. At the moment, we do not know how QPLD calculations can be used to explain "super affinities" in a similar way; however, we predict that they will certainly add to the accuracy of the description of the protein−ligand binding with enhanced electrostatic energy term in the scoring function and hence better binding free energy calculations.

ACKNOWLEDGMENT

We thank Schrödinger LLC for the generous supply of the software used for this research. Part of the work was completed while one of the authors (A.E.C.) was at Schrödinger. We also thank Professor R. A. Friesner of Columbia University for helpful discussions and encouragement to initiate the work. This work was supported by KOSEF grant No. 2008-02807 (A.E.C., J.Y.C.). We are also grateful to KIST for financial support (J.M.H., J.Y.C.).

Supporting Information Available: Table of PDB structures used in the research. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Gschwend, D. A.; Good, A. C.; Kuntz, I. D. Molecular docking towards drug discovery. *J. Mol. Recognit.* **1996**, *9*, 175–186.
(2) Green, D. V. S. Virtual screening of chemical libraries for drug discovery. *Expert Opin. Drug Discovery* **2008**, *3*, 1011–1026.
(3) Guido, R. V. C.; Oliva, G.; Andricopulo, A. D. Virtual screening and its integration with modern drug design technologies. *Curr. Med. Chem.* **2008**, *15*, 37–46.
(4) Vaque, M.; Ardrevol, A.; Blade, C.; Salvado, M. J.; Blay, M.; Fernandez-Larrea, J.; Arola, L.; Pujadas, G. Protein-ligand docking: A review of recent advances and future perspectives. *Curr. Pharm. Anal.* **2008**, *4*, 1–19.
(5) Vieth, M.; Hirst, J. D.; Dominy, B. N.; Daigler, H.; Brooks, C. L. Assessing search strategies for flexible docking. *J. Comput. Chem.* **1998**, *19*, 1623–1631.
(6) Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L. Assessing energy functions for flexible docking. *J. Comput. Chem.* **1998**, *19*, 1612–1622.
(7) Dias, R.; de Azevedo, W. F. Molecular Docking Algorithms. *Curr. Drug Targets* **2008**, *9*, 1040–1047.
(8) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151–166.
(9) Jain, A. N. Scoring functions for protein-ligand docking. *Curr. Protein Pept. Sci.* **2006**, *7*, 407–420.
(10) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.
(11) Raha, K.; Merz, K. M. A quantum mechanics-based scoring function: Study of zinc ion-mediated ligand binding. *J. Am. Chem. Soc.* **2004**, *126*, 1020–1021.
(12) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of accurate charges in molecular docking: quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.* **2005**, *26*, 915–931.
(13) Cho, A. E. Effect of Quantum Mechanical Charges in Binding Sites of Metalloproteins. *BioChip J.* **2007**, *1*, 70–75.
(14) Cho, A. E. Quantum Mechanical Calculations for Binding Sites of Metalloproteins. *BioChip J.* **2008**, *2*, 148–153.
(15) Cho, A. E.; Rinaldo, D. Extension of QM/MM Docking and its Applications to Metalloproteins. *J. Comput. Chem.* DOI: 10.1002/jcc.21270.
(16) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47*, 2977–2980.
(17) Wang, R.; Fang, X.; Lu, Y.; Yang, C.-Y.; Wang, S. The PDBbind Database: Methodologies and updates. *J. Med. Chem.* **2005**, *48*, 4111–4119.
(18) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a software program for pK (a) prediction and protonation state generation for drug-like molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681–691.
(19) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model* **2009**, *49*, 377–389.
(20) 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.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(21) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions 0.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
(22) *QM-Polarized Ligand Docking*; Schrödinger, LLC: Portland, OR, 2006.
(23) *QSite, version 3.5*; Schrodinger, LLC: Portland, OR, 2004.
(24) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 235–249.
(25) Weisel, M.; Proschak, E.; Kriegl, J. M.; Schneider, G. Form follows function: Shape analysis of protein cavities for receptor-based drug design. *Proteomics* **2009**, *9*, 451–459.
(26) Sakharkar, M. K.; Li, P.; Zhong, Z. W.; Sakharkar, K. R. Quantitative analysis on the characteristics of targets with FDA approved drugs. *Int. J. Biol. Sci.* **2008**, *4*, 15–22.
(27) Bartoli, S.; Fincham, C. I.; Fattori, D. Fragment-based drug design: Combining philosophy with technology. *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 422–429.
(28) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci.* **2007**, *104*, 808–813.

CI900231P