Genomics & Informatics Vol. 5(1) 24-29, March 2007

Evaluation of Advanced Structure-Based Virtual Screening Methods for Computer-Aided Drug Discovery

Hui Sun Lee[1], Jiwon Choi[1] and Sukjoon Yoon[1]*

[1]Department of Biological Sciences, Research Center for Women's Diseases (RCWD), Sookmyung Women's University, Seoul 140-742, Korea

Abstract

Computational virtual screening has become an essential platform of drug discovery for the efficient identification of active candidates. Moleculardocking, a key technology of receptor-centric virtual screening, is commonly used to predict the binding affinities of chemical compounds on target receptors. Despite the advancement and extensive application of these methods, substantial improvement is still required to increase their accuracy and time-efficiency. Here, we evaluate several advanced structure-based virtual screening approaches for elucidating the rank-order activity of chemical libraries, and the quantitative structure-activity relationship (QSAR). Our results show that the ensemble-average free energy estimation, including implicit solvation energy terms, significantly improves the hit enrichment of the virtual screening. We also demonstrate that the assignment of quantum mechanical-polarized (QM-polarized) partial charges to docked ligands contributes to the reproduction of the crystal pose of ligands in the docking and scoring procedure.

Keywords: virtual screening, docking and scoring, QSAR, drug discovery

Introduction

One of the major challenges in drug discovery is to identify novel compounds with biological activity. Computer-aided drug discovery technology has become an essential and powerful platform for the discovery of new lead compounds, as an alternative from, and complement to experimental approaches. As the number of high resolution structures of potential therapeutic targets and small molecules has grown, the significance of *in silico* experimental approaches has become increasingly important as demonstrated in recent studies by making use of public data (Cherkasov *et al.*, 2006; Cleves and Jain, 2006; Yoon *et al.*, 2005a; b).

Virtual high-throughput screening (Klebe, 2006; Oprea and Matter, 2004), which is a method to rapidly identify biologically active compounds *in silico*, can be roughly divided into two categories; ligand-centric and receptor-centric. Ligand-centric methods essentially focus on the comparative analysis of the structural shapes and chemical complementarities between compounds and known ligands. A knowledge of the experimentally selected active compounds is a prerequisite when using this approach (Stahura and Bajorath, 2004). Receptor-centric methods predict the interaction of given compounds with a target receptor, and hence they do not require experimental data about the structure of the ligand. Molecular docking is one of the key methodologies for receptor-centric virtual screening. It is a technique for predicting the best binding mode for a given compound that fits into a target receptor, and evaluating its binding affinity. The docking approach has become a primary technique used in many drug discovery programs (Kitchen *et al.*, 2004; Sousa *et al.*, 2006).

The docking process involves a conformational search for a compound which complements a target binding site, with the aim of identifying the best-matching binding pose. A common computational strategy is to use a suitable scoring function to theoretically evaluate the binding affinities of thousands of molecules in a compound library for a target protein. An accurate rank-ordered prediction of the compound binding affinities using the scoring function is an invaluable step. Most of scoring functions used in docking programs are designed to predict binding affinity by evaluating the interaction between a compound and a receptor. However, it should be noted that ligand-receptor recognition process is determined not only by enthalpic effects but also by entropic effects. Moreover, the scoring functions have a simplified form for the energy function to facilitate high-throughput evaluation of a large number of compounds in a single docking run. These functions may be problematic when used with contemporary docking programs, and can result in a decrease of virtual screening accuracy. To overcome this problem, more precise but time-consuming computational methodologies are necessary.

There have been a number of reports evaluating the efficiency of various virtual screening approaches, including the evaluation of docking programs (Warren *et al.*, 2006), machine-learning methods for ligand-based descriptors

*Corresponding author: E-mail yoonsj@sookmyung.ac.kr
Tel +82-2-710-9415, Fax +82-2-2077-7322
Accepted 17 Jan 2007

(Chen *et al.*, 2007) and comparison of shape-matching with docking (Hawkins *et al.*, 2007). Here, we describe and evaluate several receptor-centric computational methodologies which are applicable for use in drug discovery applications. We focus on accurate docking and rank-ordering for the improvement of the predictability of biologically active compounds.

Overview of the Methodologies Tested in This Study

We applied several computational solutions from the Schrödinger software package (Schrödinger, LLC: Portland, OR). A brief overview of these methodologies is presented.

Glide Docking

We used the Glide program (Friesner *et al.*, 2004) as our docking engine. The Glide docking algorithm performs a series of hierarchical searches for locations of possible ligand affinity within the binding site of a receptor. A rough positioning and scoring algorithm is applied during the initial search step, followed by torsional energy optimization on an OPLA-AA non-bonded potential energy grid for enduring candidate poses. The pose conformations of the very best candidates are further refined by using Monte Carlo sampling. Selection of the final docked pose is accomplished using a Glide score, which is a model energy function that combines empirical and force-field-based terms. The Glide score is a modified and extended version of the ChemScore function (Eldridge *et al.*, 1997).

Multi-Ligand Bimolecular Association with Energetics (eMBrAcE)

The eMBrAcE (*MacroModel v9.1*) program calculates binding energies between ligands and receptors using molecular mechanics energy minimization for docked conformations. eMBrAcE applies multiple minimizations, during which each of the specified pre-positioned ligands is minimized with the receptor. For the energy-minimized structures, the calculation is performed first on the receptor ($E_{protein}$), then on the ligand (E_{ligand}), and finally on the complex ($E_{complex}$). The energy difference is then calculated as:

$$\Delta E = E_{complex} - E_{ligand} - E_{protein}$$

Prime MM-GBSA

This application is used to predict the free binding energy between a receptor and a ligand. MM-GBSA is a method that combines OPLS molecular mechanics energies (E_{MM}), surface generalized Born solvation model for polar solvation (G_{SGB}), and a nonpolar solvation term (G_{NP}). The G_{NP} term comprises the nonpolar solvent accessible surface area and van der Waals interactions. The total free energy of binding is calculated as:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$
$$G = E_{MM} + G_{SGB} + G_{NP}$$

Liaison

The Liaison (*Liaison v4.0*) program is an application for estimating the binding affinities between ligands and receptors, using a linear interaction approximation (LIA) model. The LIA model is an empirical method fitted to a set of known binding free energies. Liaison runs molecular mechanics (MM) simulations for the ligand-receptor complex, and for the free ligand and free receptor using the surface generalized Born (SGB) continuum solvation model. The simulation data and empirical binding affinities are analyzed to generate the Liaison parameters, which are subsequently used to predict binding energies for other ligands with the same receptor. The empirical function used by Liaison for the prediction of binding affinities is as follows.

$$\Delta G = \alpha(\langle U^b_{vdw}\rangle - \langle U^f_{vdw}\rangle) + \beta(\langle U^b_{elec}\rangle - \langle U^f_{elec}\rangle) + \gamma(\langle U^b_{cav}\rangle - \langle U^f_{cav}\rangle)$$

In this equation, $\langle\ \rangle$, *b* and *f* represent the ensemble average, the bound form, and the free form of the ligand, respectively. Parameters α, β and γ are the coefficients. U_{vdw}, U_{elec} and U_{cav} are the van der Waals, electrostatic and cavity energy terms in the SGB model, respectively.

QM-Polarized Ligand Docking (QPLD)

The QM-Polarized Ligand Docking (QPLD) protocol is an improved docking method, which incorporates quantum mechanical and molecular mechanical (QM/MM) calculations (Cho *et al.*, 2005). This method applies the Glide algorithm to generate the best candidate poses for ligand docking. The partial charges on the atoms of the ligand are then replaced with charges derived from QM calculations on the ligand in the field of the receptor for each ligand-receptor complex. The charges are calculated from the electrostatic potential energy surface of the ligand, which is generated from a single-point calculation using the BLYP density function for the QM region. Glide then re-docks each of the ligands with updated atom charges,

and returns the most energetically favorable pose.

Materials and Methods

Preparation of the Receptor

Two receptor co-crystal structures, estrogen receptor α (ERα, PDB entry: 3ERD) and peroxisome proliferator-activated receptor γ (PPARγ, PDB entry: 1KNU), both of which belong to the nuclear receptor superfamily, were used in this study. The coordinates for these proteins were obtained from the RCSB Protein Data Bank (http://www.rcsb.org/pdb).

Preparation of the Ligands

We obtained the SMILES representation of 232 test compounds, for which the experimental binding affinities were taken from published data (Blair *et al.*, 2000; Hong *et al.*, 2002). The experimental binding affinities were represented as log(RBA) (Yoon and Welsh, 2004), where RBA refers to the relative binding affinity and log(RBA) is defined as the logarithm of the percent ratio of the IC_{50} between 17β-estradiol and a test compound. Thus, the RBA of 17β-estradiol is 100, and log(RBA) of 17β-estradiol is 2. We then generated 3D energy-minimized conformations of these compounds using the LigPrep program (*LigPrep v2.0*, Schrödinger, LLC: Portland, OR.). Compounds with atom types that were not recognized by LigPrep were eliminated from the test set. This resulted in the retention of 173 of the original 232 compounds.

The Glide Docking Protocol

For Glide docking, PDB co-crystal structures of ERα and PPARγ were prepared using the Maestro interface of the Schrödinger software package. All water molecules were removed, and multimeric complexes were simplified from the PDB structures. Prior to molecular docking, receptor structures were preprocessed using protein preparation and refinement components in the Glide docking package. Hydrogen atoms were added by applying an all-atom force field. Side chains that were not close to the ligand binding site and did not participate in salt bridges were neutralized. A restrained minimization using the OPLS-AA force field was performed to refine the complex structure. This procedure reorients side-chain hydroxyl groups, and alleviates potential steric clashes. The minimization process was continued until the average RMS deviation of the non-hydrogen atoms reached the specified limit of 0.3 Å. Once receptor grid files had been generated, all compounds were docked to the receptor structures using the standard mode of Glide docking (Glide SP 4.0).

Ligand & Structure-Based Descriptors (LSBD) Protocol

The eMBrAcE, Prime MM-GBSA and Liaison calculations were performed using the Ligand & Structure-Based Descriptors (LSBD) application of the Schrödinger software package. These calculations were applied the ligand-receptor complex structures obtained from Glide docking.

The QM-Polarized Ligand Docking Protocol

The grid for QPLD was set up as a grid file in PPARγ using the grid generation data from the previous Glide standard docking operation. The ligand to be QPLD-docked was prepared from the cognate ligand bound in PPARγ. An energy-minimized conformation of the extracted cognate ligand was generated using LigPrep. The level of quantum-mechanical treatment was set up as Fast mode.

Results and Discussion

One of the key challenges in computer-aided drug discovery is to maximize the capabilities of the method in use for predicting and rank-ordering the binding affinities of compounds for a given target protein. The efficiency of a prediction method is predominantly determined by these capabilities. Various descriptors extracted from the structural information on ligand-receptor complex may provide an advantageous solution to creating a reliable binding-affinity-prediction model. Here, we combined the results obtained from a standard docking protocol with data from three different structure-based descriptors, and then investigated the utility of these descriptors on the virtual screening efficiency for ERα ligands (Fig. 1). The virtual screening efficiency was compared using an analysis of receiver operating characteristic (ROC) curves (Hand *et al.*, 2001). A ROC curve describes the tradeoff between sensitivity and specificity, where the sensitivity is defined as the ability of the model to avoid false negatives, and the specificity relates to its ability to avoid false positives. The area under the ROC curve (AUC) is a measure of the test accuracy. For example, an AUC value of 0.5 represents a random prediction, whereas 1.0 represents a perfect prediction.

For a total of 94 true positives with log(RBA) > -4.0 (corresponding to > 200 μM activity), AUC values ranged narrowly from 0.71 to 0.75, depending on the scoring methods tested. The standard docking scoring (Glide score in the figure) method having 0.75 AUC could slightly enrich the virtual screening, and was better than the other

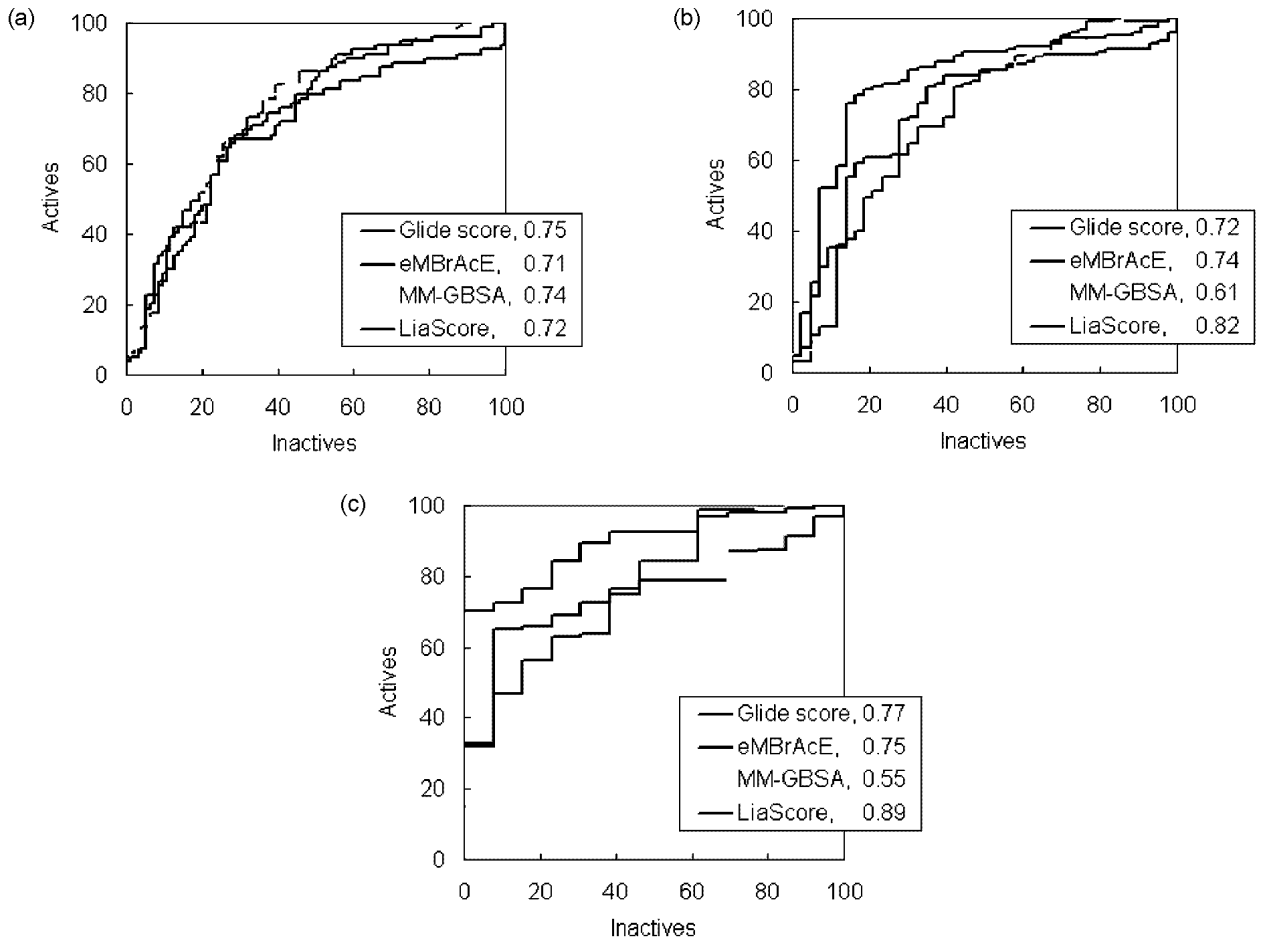

Fig. 1. Virtual screening efficiency of four different scoring methods on ERα ligands; standard Glide docking (Glide score), Multi-Ligand Bimolecular Association with Energetics (eMBrAcE), Prime MM-GBSA calculation (MM-GBSA) and Liaison calculation (LiaScore). (a) 94 true positives with log(RBS) > −4.0, (b) 43 true positives with log(RBS) > −2.0, (c) 13 true positives with log(RBS) > 0.0 out of a total of 173 compounds. The calculated AUC values are included in the insets.

descriptor-combined scoring methods (Fig. 1a). In Figure 1b we applied a more stringent definition of "active" versus "inactive" compounds. Since ~10 μM activity is generally the minimum required to identify initial lead compounds in drug discovery programs, we set log(RBA) > −2 as the cutoff, which corresponds to a 10–100 μM activity for ERα. In this case the Liaison scoring method (LiaScore in the figure) significantly improved the efficiency of virtual screening. On the other hand the Prime MM-GBSA method (MM-GBSA in Figure) showed the lowest level of enrichment of the process. An even more stringent definition of log(RBA) > 0, corresponding to the binding affinity in a nanomolar activity range, can indicate a lead compound as being 'promising' in general drug discovery terms. At this cutoff value only 13 compounds were classified as active. The overall trends in the graphs were similar to the graphs of log(RBA) > −2 cutoff. The virtual screening efficiency was generally improved, except in the case of the MM-GBSA scoring method (Fig. 1c). These results indicate that methodologies with a better prediction precision in binding affinities, though more time-consuming, can provide a significant advantage in prioritizing candidate compounds with high biological activity (low micromolar or nanomolar activity). Among the scoring methods tested, the Prime MM-GBSA method showed a relatively poor prediction capability when screening the compounds with high binding affinities.

Obtaining accurate structural information on the binding pose of a ligand at a binding site is essential to the design of optimized lead compounds in computer-aided drug discovery. An accurate calculation of atomic partial charges of a ligand in the field of the receptor would result in improved docking results. We tested whether charges obtained from the QM/MM calculation for ligand/PPARγ

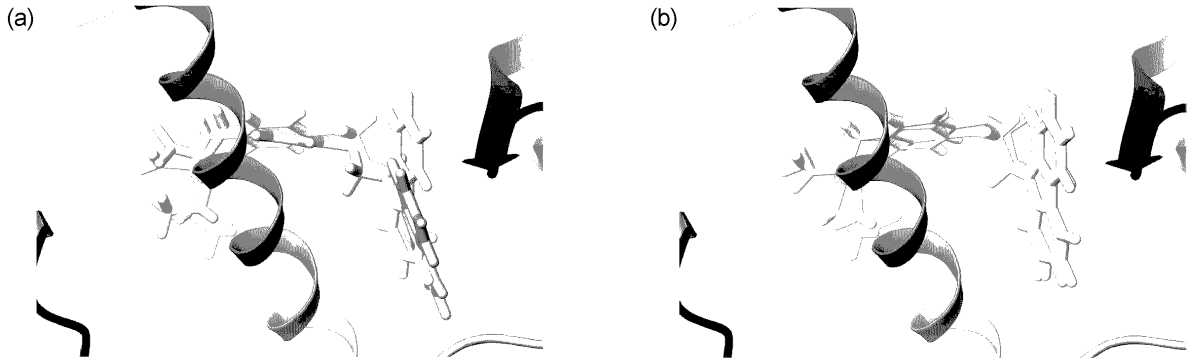

Fig. 2. Overlapped docking poses of a cognate ligand at the binding site of PPARγ obtained from (a) Glide docking (RMSD = 2.17 Å) and (b) QPLD (RMSD = 0.86 Å). In this figure the conformations of the ligand in the co-crystal structure (purple), from Glide docking (brown) and from QPLD (yellow), are represented.

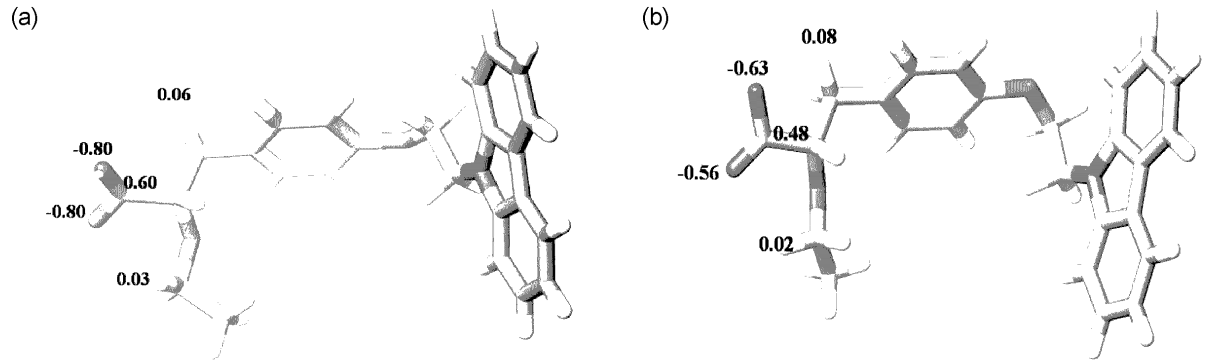

Fig. 3. Comparison between (a) force-field charges and (b) quantum mechanical charges for the PPARγ cognate ligand from standard Glide docking and QPLD docking.

structure would provide a more precise binding pose compared to the standard docking method, which relies on the default force-field charges. The results of a standard docking run and a QPLD run for a cognate ligand of PPARγ are shown in Fig. 2. The RMSD (Root Mean Square Deviation) value between crystal and docked poses of the cognate ligand was 2.17 Å from the standard docking run, while the QPLD method returned a significantly improved RMSD value of 0.86 Å. Atomic charge values in parts of the ligand structure were revealed (Fig. 3). We confirmed the existence of significant changes in partial charges by using the QM/MM calculation of the QPLD run. These results indicate that polarization effects induced by the field of the receptor can significantly affect the final conformation of a ligand bound to PPARγ. We have therefore demonstrated that an additional process of calculating subtle changes in charges, by incorporating environmental polarization effects, considerably improves the accuracy of docking predictions.

In this article we have introduced several advanced computer-aided drug discovery methodologies for receptor-entric virtual screening. We have evaluated their reliability using a set of test ligands and two receptor structures belonging to the nuclear receptor superfamily. Our data suggest that some of these methodologies significantly improve virtual screening efficiency by better prioritizing active compounds, and by more precisely reproducing the crystal pose of cognate ligands. Although the current study does not involve a large number of receptors and test sets of compounds, our evaluation data should add valuable information that may enhance the practice of computer-ased drug discovery.

Acknowledgements

The authors gratefully acknowledge Dr. Weida Tong (National Center for Toxicological Research, U.S. Food & Drug Administration, Jefferson, AR) for providing the experimental data set and 2D structures of Estrogen Receptor compounds which were used in this study. This research was supported by Sookmyung Women's University Research Grants 1-0603-0020.

References

Blair, R. M., Fang, H., Branham, W. S., Hass, B. S., Dial, S. L., Moland, C. L., Tong, W., Shi, L., Perkins, R., and Sheehan, D. M. (2000). The estrogen receptor relative binding affinities of 188 natural and xenochemicals: structural diversity of ligands. *Toxicol. Sci.* 54, 138-153.

Chen, B., Harrison, R. F., Papadatos, G., Willett, P., Wood, D. J., Lewell, X. Q., Greenidge, P., and Stiefl, N. (2007). Evaluation of machine-learning methods for ligand-based virtual screening. *Journal of computer-aided molecular design* 21, 53-62.

Cherkasov, A., Ban, F., Li, Y., Fallahi, M., and Hammond, G. L. (2006). Progressive docking: a hybrid QSAR/docking approach for accelerating in silico high throughput screening. *Journal of medicinal chemistry* 49, 7466-7478.

Cho, A. E., Guallar, V., Berne, B. J., and Friesner, R. (2005). Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.* 26, 915-931.

Cleves, A. E. and Jain, A. N. (2006). Robust ligand-based modeling of the biological targets of known drugs. *Journal of medicinal chemistry* 49, 2921-2938.

Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., and Mee, R. P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of computer-aided molecular design* 11, 425-445.

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., and Shenkin, P. S. (2004). Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* 47, 1739-1749.

Hand, D., Mannila, H., and Smyth, P. (2001). *Principles of Data Mining*. (Cambridge, Massachsetts: The MIT Press).

Hawkins, P. C., Skillman, A. G., and Nicholls, A. (2007). Comparison of shape-matching and docking as virtual screening tools. *Journal of medicinal chemistry* 50, 74-82.

Hong, H., Tong, W., Fang, H., Shi, L., Xie, Q., Wu, J., Perkins, R., Walker, J. D., Branham, W., and Sheehan, D. M. (2002). Prediction of estrogen receptor binding for 58,000 chemicals using an integrated system of a tree-based model with structural alerts. *Environmental health perspectives* 110, 29-36.

Kitchen, D. B., Decornez, H., Furr, J. R., and Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* 3, 935-949.

Klebe, G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today* 11, 580-594.

Oprea, T. I. and Matter, H. (2004). Integrating virtual screening in lead discovery. *Curr. Opin. Chem. Biol.* 8, 349-358.

Sousa, S. F., Fernandes, P. A., and Ramos, M. J. (2006). Protein-ligand docking: current status and future challenges. *Proteins* 65, 15-26.

Stahura, F. L. and Bajorath, J. (2004). Virtual screening methods that complement HTS. *Comb. Chem. High Throughput Screen.* 7, 259-269.

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., Tedesco, G., Wall, I. D., Woolven, J. M., Peishoff, C. E., and Head, M. S. (2006). A critical assessment of docking programs and scoring functions. *Journal of medicinal chemistry* 49, 5912-5931.

Yoon, S., Smellie, A., Hartsough, D., and Filikov, A. (2005a). Computational identification of proteins for selectivity assays. *Proteins* 59, 434-443.

Yoon, S., Smellie, A., Hartsough, D., and Filikov, A. (2005b). Surrogate docking: structure-based virtual screening at high throughput speed. *Journal of computer-aided molecular design* 19, 483-497.

Yoon, S. and Welsh, W. J. (2004). Identification of a minimal subset of receptor conformations for improved multiple conformation docking and two-step scoring. *J. Chem. Inf. Comput. Sci.* 44, 88-96.