

RESEARCH

Open Access

iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis

Kai-Cheng Hsu^{1†}, Yen-Fu Chen^{1†}, Shen-Rong Lin¹, Jinn-Moon Yang^{1,2,3*}

From The Ninth Asia Pacific Bioinformatics Conference (APBC 2011)
Inchon, Korea. 11-14 January 2011

Abstract

Background: Pharmacological interactions are useful for understanding ligand binding mechanisms of a therapeutic target. These interactions are often inferred from a set of active compounds that were acquired experimentally. Moreover, most docking programs loosely coupled the stages (binding-site and ligand preparations, virtual screening, and post-screening analysis) of structure-based virtual screening (VS). An integrated VS environment, which provides the friendly interface to seamlessly combine these VS stages and to identify the pharmacological interactions directly from screening compounds, is valuable for drug discovery.

Results: We developed an easy-to-use graphic environment, *iGEMDOCK*, integrating VS stages (from preparations to post-screening analysis). For post-screening analysis, *iGEMDOCK* provides biological insights by deriving the pharmacological interactions from screening compounds without relying on the experimental data of active compounds. The pharmacological interactions represent conserved interacting residues, which often form binding pockets with specific physico-chemical properties, to play the essential functions of a target protein. Our experimental results show that the pharmacological interactions derived by *iGEMDOCK* are often hot spots involving in the biological functions. In addition, *iGEMDOCK* provides the visualizations of the protein-compound interaction profiles and the hierarchical clustering dendrogram of the compounds for post-screening analysis.

Conclusions: We have developed *iGEMDOCK* to facilitate steps from preparations of target proteins and ligand libraries toward post-screening analysis. *iGEMDOCK* is especially useful for post-screening analysis and inferring pharmacological interactions from screening compounds. We believe that *iGEMDOCK* is useful for understanding the ligand binding mechanisms and discovering lead compounds. *iGEMDOCK* is available at <http://gemdock.life.nctu.edu.tw/dock/igemdock.php>.

Background

Structure-based drug design is widely used to identify lead compounds with the growing availability of protein structures [1-3]. Many tools (*e.g.*, GEMDOCK [4], DOCK [5], AutoDock [6], and GOLD [7]) have been developed for virtual screening (VS) and successfully identified lead compounds for some target proteins.

However, the accuracy of these docking tools remained intensive because of the incomplete understandings of ligand binding mechanisms [1-3]. In addition, most of scoring functions are lack of pharmacological interactions that are essential for ligand binding or biological functions [8]. Recently, some approaches have been proposed to derive pharmacological interactions from known compounds [8-10]. These approaches apparently increase hit rates for identifying the active compounds which are often similar to the known compounds. In addition, these approaches are often unable to be

* Correspondence: moon@faculty.nctu.edu.tw

† Contributed equally

¹Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu, 30050, Taiwan

Full list of author information is available at the end of the article

applied for new targets, which have no known active compounds.

Generally, a VS procedure consists of four main steps: preparations of the target protein and the compound library, docking and post-screening analysis (e.g., clustering compounds and pharmacological interactions). Most docking programs (e.g. DOCK [5] and AutoDock [6]) only provide docked poses or loosely coupled these steps. They often provided limit ability for post-screening analysis. Therefore, a VS framework, providing an easy-to-use graphic and integrated environment, is an emergent task for drug discovery.

To address these issues, we have developed a structure-based VS framework, named *iGEMDOCK*, from preparations through to post-screening analysis. *iGEMDOCK* is an integrated environment, which integrates the heavily modified and enhanced in-house tool GEMDOCK, protein-ligand profiles, pharmacological interactions, and compound clusters. GEMDOCK was comparative to several docking tools (e.g. DOCK [5] and GOLD [7]) and has been successfully applied to identify new inhibitors and new binding sites for some targets [4,8,11-14]. Notably, *iGEMDOCK* derives the pharmacological interactions from screening compounds without using a set of known active compounds. The pharmacological interactions, which often form binding pockets with specific physico-chemical properties of the target protein, can represent conserved interactions between the interacting residues and the screening compounds. We initially validated the pharmacological interactions on three therapeutic protein targets, including estrogen receptor α for antagonists and agonists and thymidine kinase. Our experimental results show that the derived pharmacological interactions are often essential for the ligand binding or maintaining biological functions for these targets. In addition, *iGEMDOCK* provided a post-screening analysis module, which is convenient for clustering compounds and visualizing the pharmacological interactions by interaction profiles. We believe that *iGEMDOCK* is useful for drug discovery and identifying essential residues and interactions for understanding the binding mechanisms.

Methods

Preparations of proteins and compound sets

To initially validate the pharmacological interactions, we selected three therapeutic protein targets, including estrogen receptor α for agonists (ERA, PDB code 1gwr [15]), estrogen receptor α for antagonists (ER, PDB code 3ert [16]), and thymidine kinase (TK, PDB code 1kim [17]) because these proteins were well studied. The catalytic mechanisms, biological functions, key functional residues, and active compounds of the three targets were available in the literatures. Estrogen receptor is an

important therapeutic target for osteoporosis and breast cancer [18], and TK is a drug target for the therapy of herpes simplex virus type-1 [19]. Moreover, we also evaluate the docking and screening accuracy of *iGEMDOCK*. For docking, a highly diverse dataset comprising 305 protein-compound complexes (i.e., CCDC/Astex set [20]) was selected; for screening, we prepared 10 known active compounds and 990 compounds were randomly selected from Available Chemical Directory (ACD) proposed by Bissantz *et al.* [21] for each therapeutic protein target.

Main procedure

iGEMDOCK is an integrated VS environment from preparations through post-screening analysis with pharmacological interactions (Fig. 1A). First, *iGEMDOCK* provides interactive interfaces to prepare both the binding site of the target protein and the screening compound library (Figs. 1B and 1C). Each compound in the library is then docked into the binding site by using the in-house docking tool GEMDOCK [4]. Subsequently, *iGEMDOCK* generates protein-compound interaction profiles of electrostatic (E), hydrogen-bonding (H), and van der Waals (V) interactions. Based on these profiles and compound structures, *iGEMDOCK* infers the pharmacological interactions and clusters the screening compounds for the post-screening analysis (Figs. 1D and 1E). Finally, *iGEMDOCK* ranks and visualizes the screening compounds by combining the pharmacological interactions and energy-based scoring function of GEMDOCK.

Mining pharmacological interactions

iGEMDOCK mines the pharmacological interactions based on protein-compound interaction profiles (Fig. 2). The size of each profile is $N \times 2K$ where N and K are the numbers of screening compounds and interacting residues of the target protein, respectively. Here, an interacting residue is divided into two interacting groups: main and side chains. A profile $P(I)$ with type I (E, H, or V) is given as (Fig. 2A):

$$P(I) = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,2K} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,2K} \end{bmatrix}$$

where $p_{i,j}$ is a binary value (0 or 1) for the compound i interacting to the residue group j . In the E and H profiles, the $p_{i,j}$ is set to 1 (green) if hydrogen-bonding or electrostatic interactions are yielded between the compound i and the residue j (energy ≤ -2.5 kcal/mol); otherwise, $p_{i,j}=0$ (black). For the V profile, $p_{i,j} = 1$ if the interacting energy is less than -4 kcal/mol (Fig. 2A).

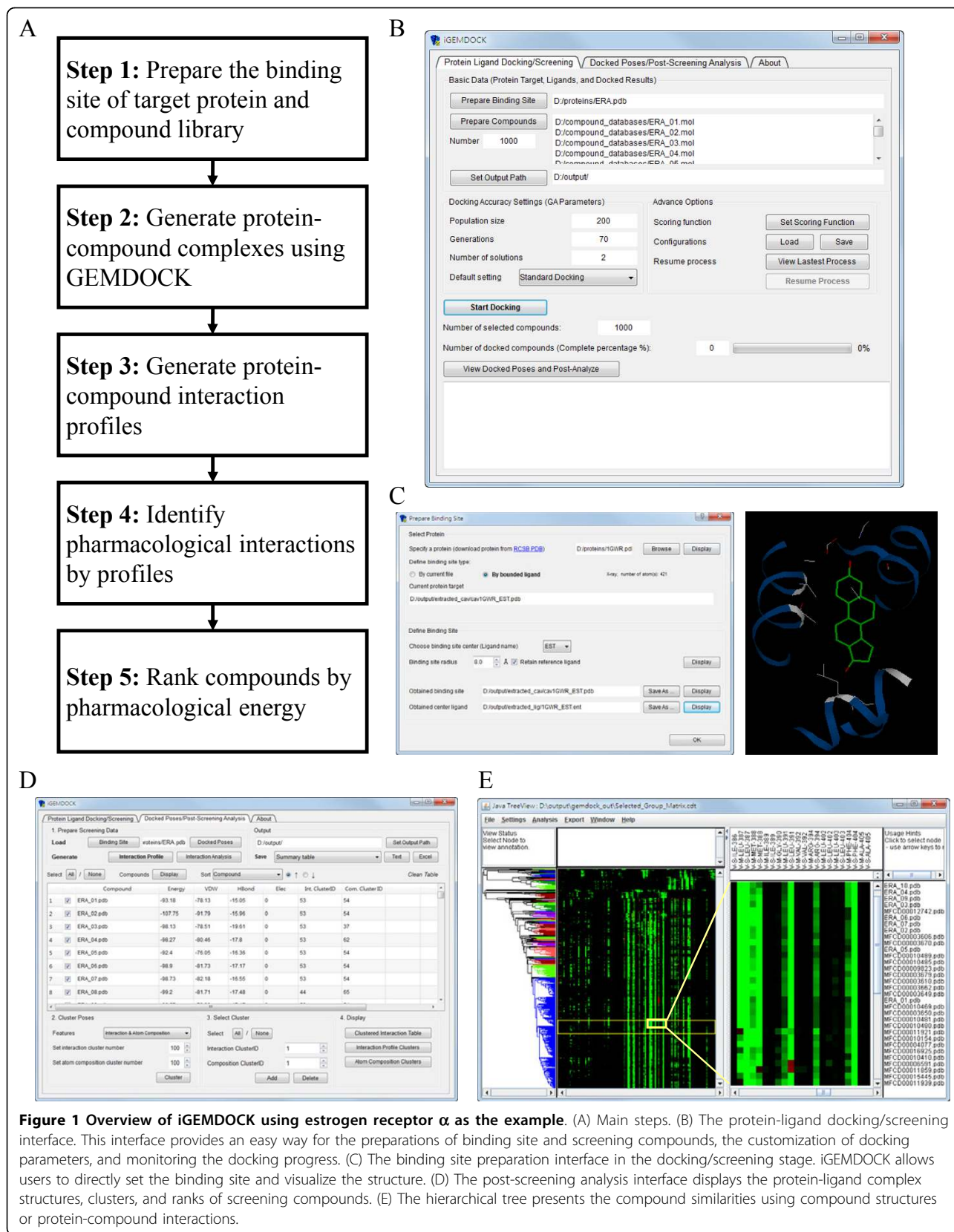
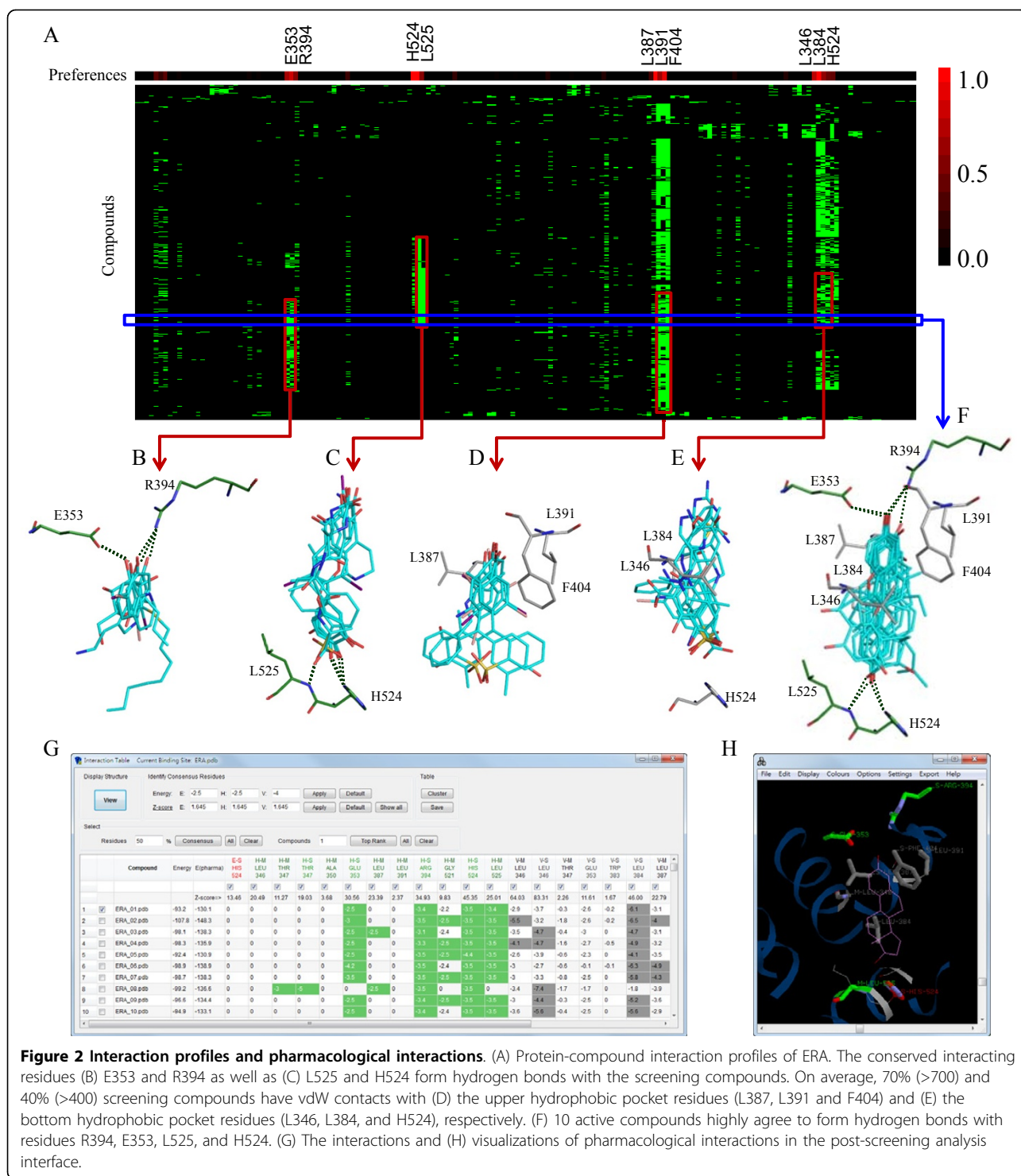


Figure 1 Overview of iGEMDOCK using estrogen receptor α as the example. (A) Main steps. (B) The protein-ligand docking/screening interface. This interface provides an easy way for the preparations of binding site and screening compounds, the customization of docking parameters, and monitoring the docking progress. (C) The binding site preparation interface in the docking/screening stage. iGEMDOCK allows users to directly set the binding site and visualize the structure. (D) The post-screening analysis interface displays the protein-ligand complex structures, clusters, and ranks of screening compounds. (E) The hierarchical tree presents the compound similarities using compound structures or protein-compound interactions.



After the generations of the profiles, we identified the pharmacological interactions. For each interacting residue group, the z-score value is used to measure the interaction conservation between the interacting groups and the screening compounds. For computing the z-scores of interacting groups in a profile, 1000 randomly

shuffled profiles are utilized to obtain the standard deviation (σ) and mean (μ). The z-score of the interacting residue group j is defined as $Z_j = \frac{f_j - \mu}{\sigma}$, where f_j is given as $f_j = \sum_{i=1}^N \frac{p_{ij}}{N}$, where N is the number of

screening compounds. Finally, we normalize the z -score value as follows:

$$W_j = \begin{cases} z_j/z_{max} & \text{if } z_j > 0 \\ 0 & \text{if } z_j \leq 0 \end{cases} \quad (1)$$

where W_j is the interaction conservation of the residue group j related to the largest z -score (z_{max}) among all of interacting groups in the binding site. Here, an interaction conservation is viewed as a pharmacological preference and an interaction is considered as the pharmacological interaction if $W_j \geq 0.4$. For example, for the hydrogen profile of the target ERA, the pharmacological preferences of E353 and R394 are 0.64 and 0.80, respectively; for the V profile, the preferences of L387, L391, and F404 are 1.00, 0.61, and 0.90, respectively (Fig. 2B). In this case, over 300 (>30%) screening compounds form hydrogen bonds with the residues E353 or R394 by polar moieties (e.g., hydroxyl group (27%), carboxyl group (20%), sulfuric acid monoester (9%), ketone (8%), and phosphoric acid monoester (6%)). Moreover, the aromatic rings of the screening compounds are often sandwiched by vdW interacting residues L387, L391, and F404 (Fig. 2D).

Based on the pharmacological interactions, we developed a pharmacological scoring function for identifying the active compounds from thousands of screening compounds. The pharmacological scoring function is given as

$$E_{pharma} = E_{GEMDOCK} + E(E)_{pharma} + 2E(H)_{pharma} + 0.5E(V)_{pharma} \quad (2)$$

where $E_{GEMDOCK}$ is the docked energy of GEMDOCK and $E(E)_{pharma}$, $E(H)_{pharma}$, and $E(V)_{pharma}$ are the pharmacological scores of electrostatics, hydrogen-bonding, and vdW interactions, respectively. The $E(I)_{pharma}$ with interaction type I (i.e., E, H, or V) is defined as

$$E(I)_{pharma} = \sum_{j=1}^{2K} W_j e_j$$

where e_j is the energy obtained by the GEMDOCK scoring function for the residue group j . Finally, *iGEMDOCK* provides the ranks of energy-based and pharmacological scoring functions for all screening compounds.

Implementation of *iGEMDOCK*

iGEMDOCK is an easy-to-use VS environment and includes three main modules (Fig. 1): docking and virtual screening tool (GEMDOCK); post-screening analysis methods; and visualization tools (RasMol [22] and Java Treeview [23]). We employed ERA as an example to present these modules, procedures and features of *iGEMDOCK*.

For protein-ligand docking/screening module, *iGEMDOCK* provides an interactive interface for the preparations of the binding site and compound library; setting docking parameters; and monitoring progress status (Fig. 1B). For most docking tools, users usually need to prepare the binding site structure and compound library through complicated steps (e.g., add hydrogen atoms and generate the grids of the protein). Here, *iGEMDOCK* provides a straightforward method to derive the binding site from the bounded ligand. For example, the binding site of ERA (PDB code 1gwr) was obtained from the estradiol (Fig. 1C). *iGEMDOCK* is able to automatically consider the effects of hydrogen atoms when preparing the binding site and the compound library. In addition, *iGEMDOCK* allows users to visualize and refine the binding site of the target protein. Additionally, *iGEMDOCK* offers the similar way to prepare screening compounds and docking parameters (e.g., the population size and the number of generations).

After the screening process, *iGEMDOCK* utilizes the post-screening analysis module to infer pharmacological interactions and cluster screening compounds based on protein-ligand complexes and compound structures (Fig. 1D). First, *iGEMDOCK* generates interaction profiles and calculates the pharmacological preference (W_j) of each interacting group for deriving the pharmacological interactions (Fig. 2). These pharmacological preferences and interactions are shown in an interactive window (Fig. 2G); furthermore, RasMol displays the pharmacological interactions with conserved interacting residues and functional groups of compounds (Fig. 2H). Additionally, *iGEMDOCK* supports a hierarchical clustering method to cluster screening compounds according to interaction profiles and the atomic composition (Fig. 1E). The atomic composition, which is similar to the amino acid composition of a protein sequence, is useful for measuring compound similarity. *iGEMDOCK* provides an interactive interface for visualizing compound similarity with a hierarchical tree by Java Treeview. Finally, *iGEMDOCK* ranks and visualizes the screening compounds by combining the pharmacological interactions and the energy-based scoring function.

Results and discussion

Pharmacological interactions

The pharmacological interactions derived by *iGEMDOCK* are often involved in biological reactions or essential for ligand binding. We examined the pharmacological interactions on three selected target proteins (ERA, ER, and TK). First, we compared the pharmacological interactions, derived from 1000 screening compounds, to the consensus interactions, derived from 10 active compounds (Table 1 and Fig. 3). Here, the residue

Table 1 Pharmacological interactions and consensus interaction ratio on estrogen receptor α and thymidine kinase

Protein	Predicted pharmacological interactions	Consensus interaction ratio ^a	Related works			
ERA	R394-H ^b (0.80 ^c)	1.0	Form hydrogen bonding networks for ligand binding [26,27]			
	E353-H (0.64)	0.8				
	H524-H (1.00)	1.0				
		L387-V (1.00)	0.9	Form non-polar contacts with A-ring of sterols scaffolds [28,29].		
		L387-H (0.52)	0.2			
		F404-V (0.90)	1.0			
		V346-V (0.98)	0.5			
		L391-V (0.61)	0.9			
		L384-V (0.57)	0.9			
L525-H (0.53)		0.9				
TK		R222-H (1.00)	0.8		Transfer phosphate in the substrate phosphorylation [30,31,36]	
	R222-E (1.00)	0.0				
	R222-V (0.62)	0.4				
	R163-H (0.99)	0.6				
	R163-E (0.40)	0.0				
	R163-V (0.56)	0.8				
	E83-V (0.54)	0.2				
		Y101-H (0.40)	0.2	Form hydrogen bonds with thymidine; activity was decreased over 90% if Q125 mutated [32]		
		Y101-V (0.45)	0.0			
		Q125-H (0.40)	1.0			
		Y172-V (1.00)	1.0		Sandwich the thymine moiety of substrates [33]	
		M128-V (0.58)	1.0			
			W88-V (0.87)		0.9	Constitute a pocket for ligand binding [33]
			H58-V (0.68)		0.9	

^aThe consensus interaction ratio of the residue i is defined as A_i/A , where A_i is the number of active compounds interacting to the residue i and A is total number of active compounds.

^bH, E and V are the interaction types.

^cThe pharmacological preferences (i.e. W_i defined in Equation (1)).

i is considered as “hot spot” if the consensus interaction ratio ≥ 0.5 [9,10,24,25]. Among 10 predicted pharmacological interactions (residues) for ERA, 9 pharmacological interactions (9 of 9 residues) agree with hot spots except the L387 with the hydrogen-bonding interaction. For TK, 8 of 14 pharmacological interactions (7 of 9 residues) are the hot spots. These results indicate the pharmacological interactions (residues) from screening compounds are often essential for the ligand binding. For example, 10 active compounds of TK form stacking interactions with the residue Y172 (vdW preference is 1.0 defined in Equation (1)) that stabilizes the binding of thymine or purine moieties.

We also examined the pharmacological interactions by their biological functions or binding mechanisms. For estrogen receptor α , H524 (hydrogen-bonding preferences are 1.0 and 0.42 for ERA and ER, respectively) is involved in a hydrogen-bonding network [26]; similarly, E353 and R394 (hydrogen-bonding preferences ≥ 0.5 for both ERA and ER) interact the structural water to form the hydrogen bonding network (Table 1 and Fig. 3) [27].

These two hydrogen bonding networks are essential for estrogen receptor modulators to trigger the responses of estrogen receptor α [26,27]. For ER and ERA, hydrophobic interacting residues, L346, L387, F404, and L525 with high vdW interaction preferences, contact with the sterols or flavones scaffolds of the active compounds. These residues contribute the major vdW interactions for the ligand binding of estrogen receptor α [28,29].

For TK, R222 and R163 play major roles for inhibitor and substrate binding [30,31], and their hydrogen-bonding preferences are 1.0 and 0.99, respectively (Table 1). Our method identified the electrostatic interactions of R222 and R163 (preferences are 1.0 and 0.4, respectively), which help to transfer phosphate in the substrate phosphorylation [30]. However, these two electrostatic interactions are not observed from 10 active compounds (Fig. 3). For the residue Q125 (H preference 0.40), the TK activity was decreased over 90% if Q125 mutated into Asp, Glu, or Asn [32]. The residues M128, Y172, H58, R163, and Y88 constitute a pocket to fix the substrate, and their vdW preferences are 0.58, 1.00, 0.68,

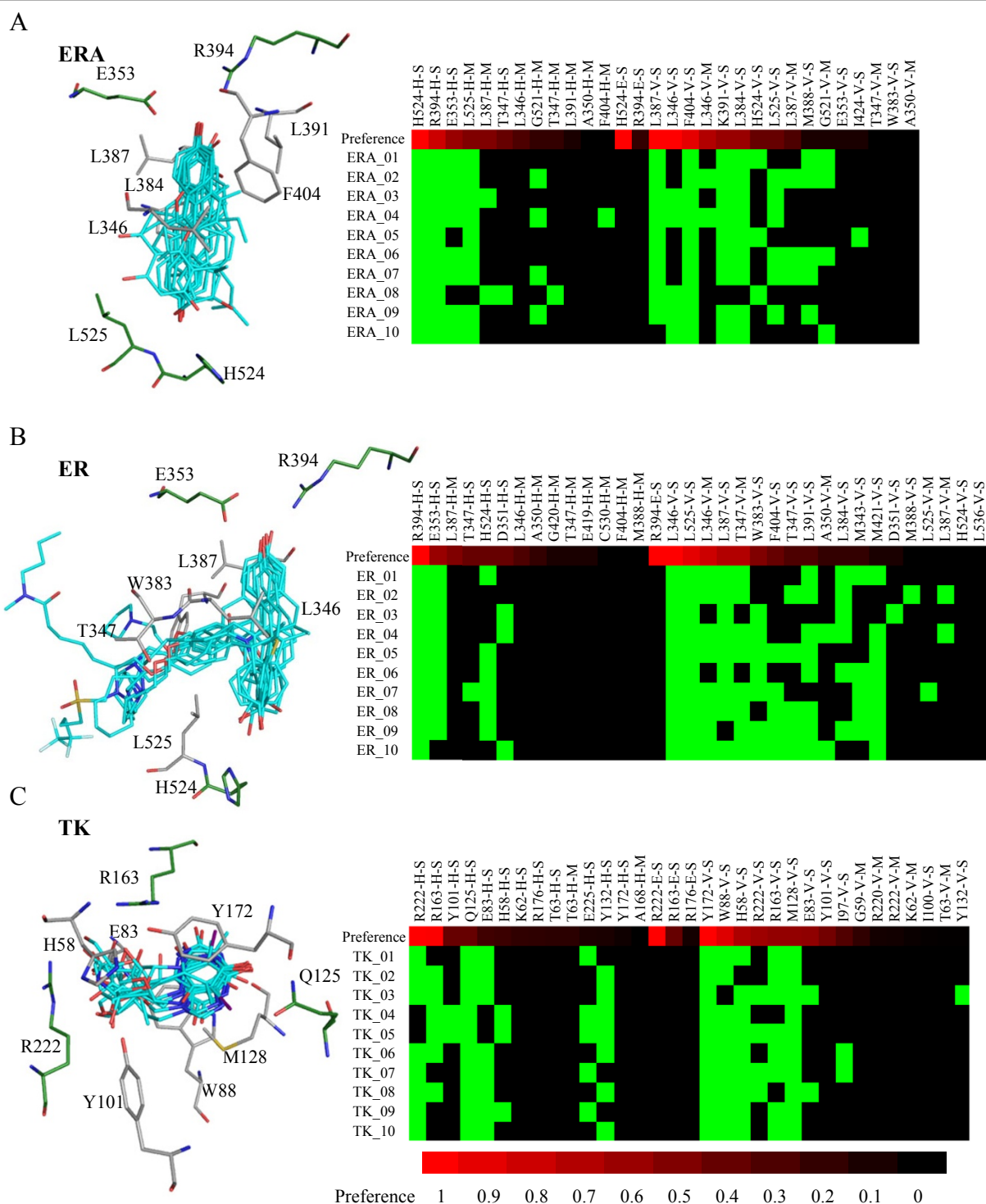


Figure 3 Relationship between the pharmacological interactions and the active compounds of (A) ERA, (B) ER, and (C) TK. The residue with a pharmacological preference ≥ 0.4 is colored by the interaction types [H: green (E353 and R394 in ERA); E: yellow; and V: gray (L391 and F404 in ERA)]. In the profile, the first row presents the pharmacological preferences of the interacting residue groups using the color-coding bar, with red-through-black indicating high-through-low. The following rows show the interactions between the active compounds and the interacting residue groups. H, E, and V indicate the interaction types; M and S indicate the main chain and the side chain of the interacting residue, respectively. The hydrogen-bonding or electrostatic interactions are colored in green if the energy ≤ -2.5 . The vdW interactions are colored in green when the energy is less than -4 .

0.56, and 0.87, respectively (Table 1). For the substrate binding, M128 and Y172 sandwich the thymine moiety and W88 is a part of the quasi-helical motif [33,34]. These results demonstrated that the pharmacological interactions derived by *iGEMDOCK* are often involved in the biological functions and the ligand binding.

Molecular docking and virtual screening

To initially evaluate the utility of *iGEMDOCK* for docking and virtual screening, we selected a highly diverse dataset with 305 protein-ligand complexes (*i.e.*, CCDC/Astex set [20]) and ERA, ER, and TK with 1000 compounds as test sets. Please note that the docking and screening tool of *iGEMDOCK* is GEMDOCK which was well-studied for VS and some applications [4,8,11-14]. In order to compare with previous works, we followed the docking procedures and performance indices proposed by Nissink, *et al.* A docked result was considered as a success solution if the root-mean-square derivation (RMSD) ≤ 2.0 Å between the docked solutions and X-ray crystal structures. For these 305 complexes, the success rates of *iGEMDOCK* and GOLD are 78% and 68%, respectively (Table S1 in additional file 1).

The pharmacological scoring function was then applied to identify the active compounds from the 1000 compounds of ERA, ER, and TK. Furthermore, we compared the screening results with those of using the energy-based scoring function of GEMDOCK. These two approaches were tested on the same datasets. The true hits of the three testing sets were used to access the screening accuracy of the two approaches (Fig. S1 in additional file 1). The screening accuracy was generally improved when *iGEMDOCK* considered the pharmacological interactions.

We compared *iGEMDOCK* (pharmacological scoring function) with three screening methods (DOCK, GOLD, and FlexX) on the ER and TK sets (Table S2 in additional file 1). To compare with previous works, we followed the experiment design and performance indices used by Bissantz *et al.* When true-positive rate is 80%, the false positive rates were 2.3% (*iGEMDOCK*), 13.3% (DOCK), 57.8% (FlexX), and 5.3% (GOLD), for ER. The false positive rates were 7.8% (*iGEMDOCK*), 23.4% (DOCK), 8.8% (FlexX), and 8.3% (GOLD) for TK.

Post-screening analysis

To identify leads from vast amount of docked poses generated during the virtual screening procedure is the key step for the drug discovery. In addition, the top-ranked compounds based on the scoring functions are not advisable since these compounds may be similar in structures or physico-chemical properties. For these two issues, *iGEMDOCK* provides a post-screening analysis module to cluster compounds based on the interactions

profiles and the atomic compositions. Selecting representative compounds from each cluster is able to maintain compound diversity and then reduces the false positives. Further, when active compounds are available, users can choose the similar compounds in the same cluster of the actives based on hierarchical trees (Fig. 1E). Notably, *iGEMDOCK* visualizes the interaction profiles of the compounds, and thereby the top-ranked compounds with pharmacological interactions can be selected according to the interaction table (Fig. 2G).

The post-screening analysis module of *iGEMDOCK* is useful for clustering and selecting compounds based on interaction profiles. We selected a set of compounds, including 10 ERA active compounds and top-ranked 100 compounds based on both the pharmacological and energy-based scoring functions. The hierarchical clustering dendrogram and the profile revealed that the protein-ligand interactions derived from the pharmacological scoring function are significantly different from those derived from the energy-based scoring function (Figs. 4A and B). The compounds with the high pharmacological scores and the active compounds consistently keep the pharmacological interactions; conversely, the compounds derived from the energy-based scoring function are often lack of the pharmacological interactions (Fig. 4B). This result indicates the pharmacological interactions are useful for identifying active compounds. For example, the two active compounds, ERA_03 and ERA_06, were ranked as 187 and 160 using the energy-based scoring function, respectively. When the pharmacological interactions were considered, the ranks of ERA_03 and ERA_06 were 91 and 87, respectively (Fig. 4C).

Some compounds having high pharmacological scores are structurally and chemically similar to the active compounds (Fig. 4C). For example, MFCD00003670 (Tetrahydrocortisol) and MFCD00012742 (Tetrahydropapaveroline) were analogues of the ERA active compounds, and both of them were clustered into the same cluster. In addition, the pharmacological scoring function can reduce the ill-effect of most energy-based scoring functions which often favor high molecular weight or highly-polar compounds [8,35]. For instance, the ranks of two high molecular weight and polar compounds, MFCD00009772 and MFCD00016940, are 267 and 274, respectively (Fig. 4C). To further examine the pharmacological scoring function, we analyzed the relationship between the molecular weights of the active compounds and the rank improvement using the pharmacological scoring function (Fig. 5). The pharmacological scoring function generally improves the screening accuracy when the molecular weights of the active compounds are less than 400.

In summary, *iGEMDOCK* can mine the pharmacological interactions from the screening compounds

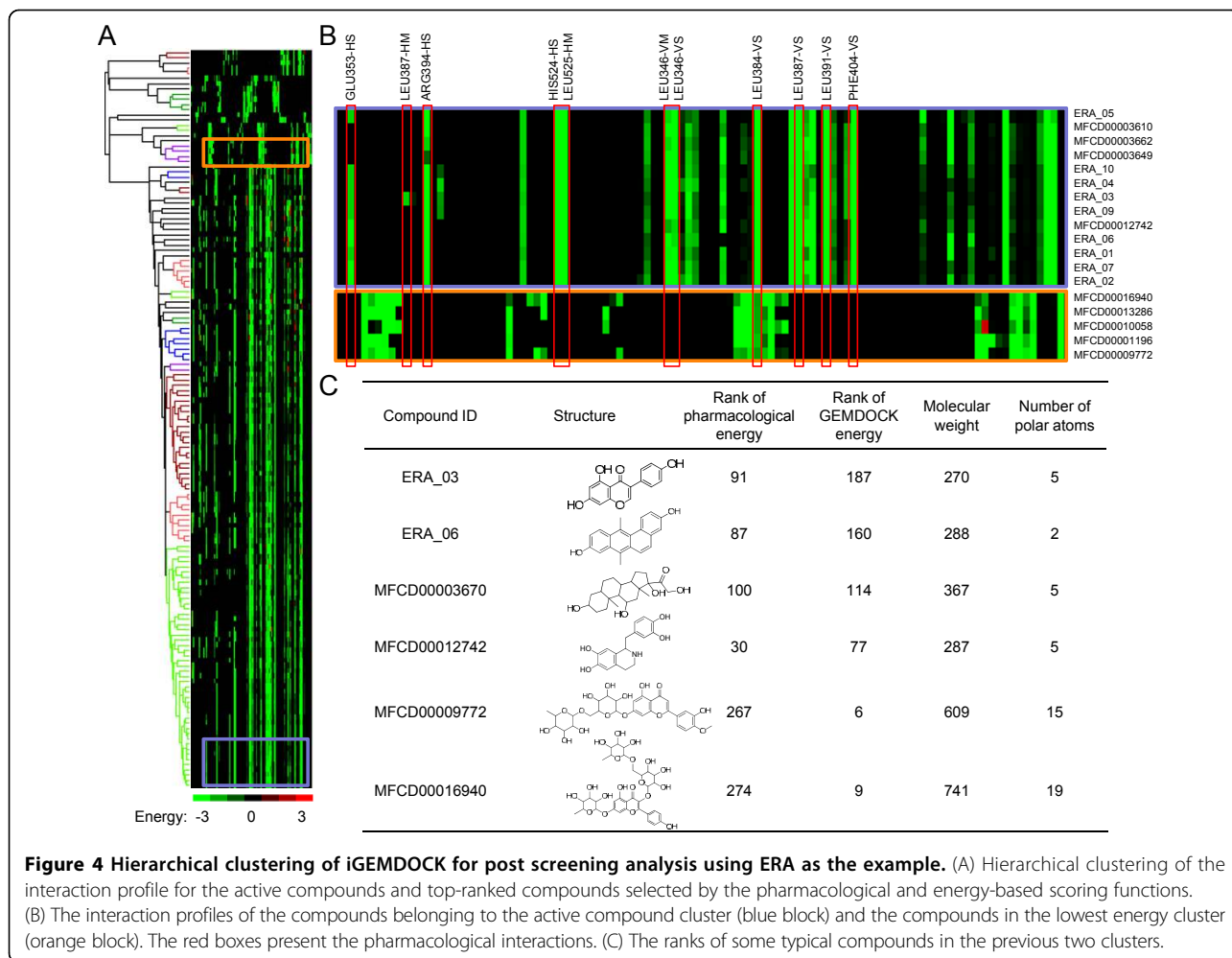


Figure 4 Hierarchical clustering of iGEMDOCK for post screening analysis using ERA as the example. (A) Hierarchical clustering of the interaction profile for the active compounds and top-ranked compounds selected by the pharmacological and energy-based scoring functions. (B) The interaction profiles of the compounds belonging to the active compound cluster (blue block) and the compounds in the lowest energy cluster (orange block). The red boxes present the pharmacological interactions. (C) The ranks of some typical compounds in the previous two clusters.

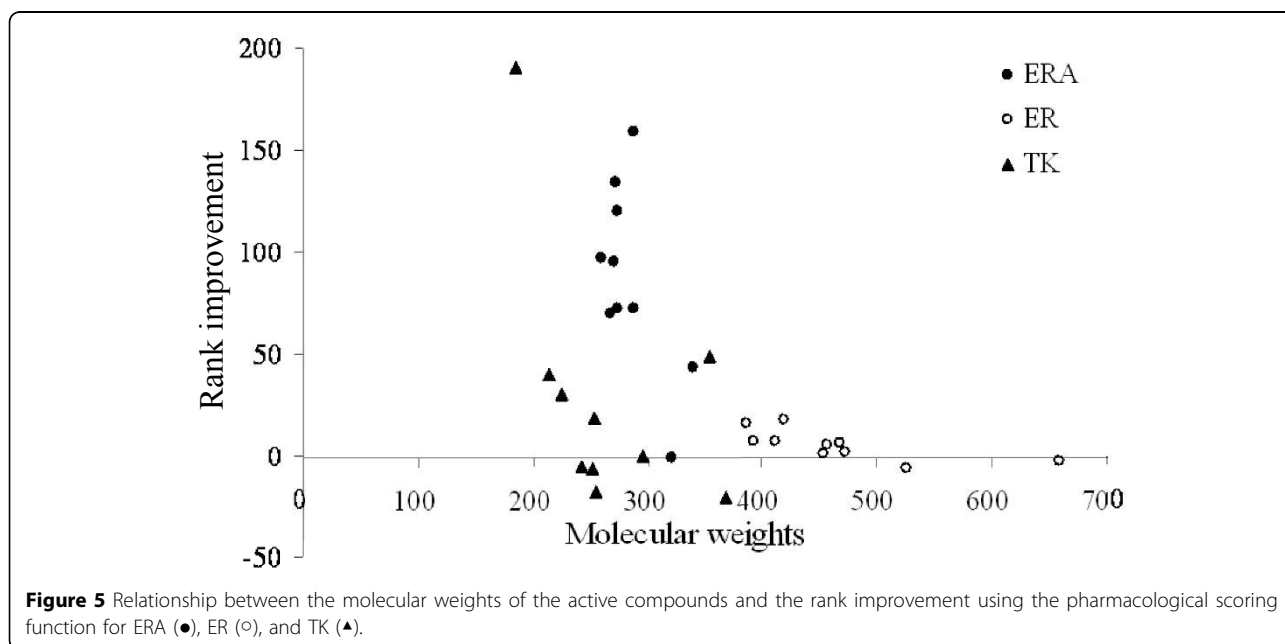


Figure 5 Relationship between the molecular weights of the active compounds and the rank improvement using the pharmacological scoring function for ERA (●), ER (○), and TK (▲).

without known active compounds. Therefore, *iGEMDOCK* can provide a good starting point for deriving pharmacological interactions (residues) and identifying new potential active compounds for a new protein structure. In addition, *iGEMDOCK* offers the visualization of the interaction profiles, pharmacological interactions, and the hierarchical clustering dendrogram. Users are able to easily observe and select compounds for post-screening analysis to enrich accuracies.

Conclusions

We have developed a structure-based VS framework *iGEMDOCK* from the preparations through to the post-screening analysis. *iGEMDOCK* is an integrated and easy-to-use environment which is especially useful for post-screening analysis and inferring pharmacological interactions from screening compounds. The friendly user interface is helpful to biologically oriented nonexperts. The experimental results show that the pharmacological interactions are often essential for the binding of the active compounds and involved in biological mechanisms. The pharmacological interactions can reduce the ill effects of energy-based scoring functions to enrich the hit rates in VS. We believe *iGEMDOCK* is useful for drug discovery and understanding protein-ligand mechanisms.

Additional material

Additional file 1:

Acknowledgements

J-M. Yang was supported by National Science Council and partial support of the ATU plan by MOE. Authors are grateful to both the hardware and the software supports of the Structural Bioinformatics Core Facility at National Chiao Tung University.

This article has been published as part of *BMC Bioinformatics* Volume 12 Supplement 1, 2011: Selected articles from the Ninth Asia Pacific Bioinformatics Conference (APBC 2011). The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2105/12?issue=S1>.

Author details

¹Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu, 30050, Taiwan. ²Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, 30050, Taiwan. ³Core Facility for Structural Bioinformatics, National Chiao Tung University, Hsinchu, 30050, Taiwan.

Authors' contributions

KCH, YFC, and JMY conceived and designed the experiments. YFC, SRL, and JMY implemented the program. KCH, YFC, and JMY performed the experiments and analyzed the data. KCH, YFC, and JMY wrote the paper.

Competing interests

The authors declare that they have no competing interests.

Published: 15 February 2011

References

1. Lyne PD: Structure-based virtual screening: an overview. *Drug Discovery Today* 2002, **7**:1047-1055.
2. Tanrikulu Y, Schneider G: Pseudoreceptor models in drug design: bridging ligand- and receptor-based virtual screening. *Nature Reviews Drug Discovery* 2008, **7**:667-677.
3. Kitchen DB, Decornez H, Furr JR, Bajorath J: Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery* 2004, **3**:935-949.
4. Yang J-M, Chen C-C: GEMDOCK: a generic evolutionary method for molecular docking. *Proteins* 2004, **55**:288-304.
5. Kramer B, Rarey M, Lengauer T: Evaluation of the flexX incremental construction algorithm for protein-ligand docking. *Proteins* 1999, **37**:228-241.
6. Morris GM, Goodsell DS, Huey R, Olson AJ: Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design* 1996, **10**:293-304.
7. Jones G, Willett P, Glen RC, Leach AR, Taylor R: Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* 1997, **267**:727-748.
8. Yang JM, Shen TW: A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. *Proteins* 2005, **59**(2):205-220.
9. Tafi A, Bernardini C, Botta M, Corelli F, Andreini M, Martinelli A, Ortore G, Baraldi PG, Fruttarolo F, Borea PA, et al: Pharmacophore based receptor modeling: the case of adenosine A3 receptor antagonists. An approach to the optimization of protein models. *Journal of Medicinal Chemistry* 2006, **49**:4085-4097.
10. Wolber G, Seidel T, Bendix F, Langer T: Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discovery Today* 2008, **13**:23-29.
11. Yang JM, Chen YF, Tu YY, Yen KR, Yang YL: Combinatorial computational approaches to identify tetracycline derivatives as flavivirus inhibitors. *PLoS One* 2007, **2**:e428.
12. Chin KH, Lee YC, Tu ZL, Chen CH, Tseng YH, Yang JM, Ryan RP, McCarthy Y, Dow JM, Wang AH, et al: The cAMP receptor-like protein CLP is a novel c-di-GMP receptor linking cell-cell signaling to virulence gene expression in *Xanthomonas campestris*. *Journal of Molecular Biology* 2010, **396**:646-662.
13. Hung HC, Tseng CP, Yang JM, Ju YW, Tseng SN, Chen YF, Chao YS, Hsieh HP, Shih SR, Hsu JT: Aurintricarboxylic acid inhibits influenza virus neuraminidase. *Antiviral Research* 2009, **81**:123-131.
14. Yang M-C, Guan H-H, Yang J-M, Ko C-N, Liu M-Y, Lin Y-H, Chen C-J, Mao SJT: Rational design for crystallization of beta-lactoglobulin and vitamin D-3 complex: revealing a secondary binding site. *Crystal Growth & Design* 2008, **8**:4268-4276.
15. Warnmark A, Treuter E, Gustafsson JA, Hubbard RE, Brzozowski AM, Pike AC: Interaction of transcriptional intermediary factor 2 nuclear receptor box peptides with the coactivator binding site of estrogen receptor alpha. *Journal of Biological Chemistry* 2002, **277**:21862-21868.
16. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL: The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 1998, **95**:927-937.
17. Champness JN, Bennett MS, Wien F, Visse R, Summers WC, Herdewijn P, de Clerq E, Ostrowski T, Jarvest RL, Sanderson MR: Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins* 1998, **32**:350-361.
18. Zhou HB, Sheng S, Compton DR, Kim Y, Joachimiak A, Sharma S, Carlson KE, Katzenellenbogen BS, Nettles KW, Greene GL, et al: Structure-guided optimization of estrogen receptor binding affinity and antagonist potency of pyrazolopyrimidines with basic side chains. *Journal of Medicinal Chemistry* 2007, **50**:399-403.
19. Manikowski A, Verri A, Lossani A, Gebhardt BM, Gambino J, Foher F, Spadari S, Wright GE: Inhibition of herpes simplex virus thymidine kinases by 2-phenylamino-6-oxapurines and related compounds: structure-activity relationships and antiherpetic activity in vivo. *Journal of Medicinal Chemistry* 2005, **48**:3919-3929.
20. Nissink JW, Murray C, Hartshorn M, Verdonk ML, Cole JC, Taylor R: A new test set for validating predictions of protein-ligand interaction. *Proteins* 2002, **49**:457-471.

21. Bissantz C, Folkers G, Rognan D: **Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations.** *Journal of Medicinal Chemistry* 2000, **43**:4759-4767.
22. Sayle RA, Milner-White EJ: **RASMOL: biomolecular graphics for all.** *Trends in Biochemical Sciences* 1995, **20**:374.
23. Saldanha AJ: **Java Treeview—extensible visualization of microarray data.** *Bioinformatics* 2004, **20**:3246-3248.
24. Yang J-M, Shen T-W: **A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators.** *Proteins* 2005, **59**:205-220.
25. Fradera X, Knegtel RM, Mestres J: **Similarity-driven flexible ligand docking.** *Proteins* 2000, **40**:623-636.
26. Qin Z, Kastrati I, Chandrasena RE, Liu H, Yao P, Petukhov PA, Bolton JL, Thatcher GR: **Benzothiophene selective estrogen receptor modulators with modulated oxidative activity and receptor affinity.** *Journal of Medicinal Chemistry* 2007, **50**:2682-2692.
27. Manas ES, Xu ZB, Unwalla RJ, Somers WS: **Understanding the selectivity of genistein for human estrogen receptor-beta using X-ray crystallography and computational methods.** *Structure* 2004, **12**:2197-2207.
28. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA, Carlquist M: **Molecular basis of agonism and antagonism in the oestrogen receptor.** *Nature* 1997, **389**:753-758.
29. Shadnia H, Wright JS, Anderson JM: **Interaction force diagrams: new insight into ligand-receptor binding.** *Journal of Computer-aided Molecular Design* 2009, **23**:185-194.
30. Wild K, Bohner T, Folkers G, Schulz GE: **The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue.** *Protein Sci* 1997, **6**:2097-2106.
31. Kussmann-Gerber S, Kuonen O, Folkers G, Pilger BD, Scapozza L: **Drug resistance of herpes simplex virus type 1—structural considerations at the molecular level of the thymidine kinase.** *European Journal of Biochemistry / FEBS* 1998, **255**:472-481.
32. Hinds TA, Compadre C, Hurlburt BK, Drake RR: **Conservative mutations of glutamine-125 in herpes simplex virus type 1 thymidine kinase result in a ganciclovir kinase with minimal deoxypyrimidine kinase activities.** *Biochemistry* 2000, **39**:4105-4111.
33. Pilger BD, Perozzo R, Alber F, Wurth C, Folkers G, Scapozza L: **Substrate diversity of herpes simplex virus thymidine kinase. Impact Of the kinematics of the enzyme.** *The Journal of Biological Chemistry* 1999, **274**:31967-31973.
34. Evans JS, Lock KP, Levine BA, Champness JN, Sanderson MR, Summers WC, McLeish PJ, Buchan A: **Herpesviral thymidine kinases: laxity and resistance by design.** *The Journal of General Virology* 1998, **79**:2083-2092.
35. Pan Y, Huang N, Cho S, MacKerell AD Jr.: **Consideration of molecular weight during compound selection in virtual target-based database screening.** *J Chem Inf Comput Sci* 2003, **43**(1):267-272.
36. Sulpizi M, Schelling P, Folkers G, Carloni P, Scapozza L: **The rational of catalytic activity of herpes simplex virus thymidine kinase. a combined biochemical and quantum chemical study.** *The Journal of Biological Chemistry* 2001, **276**:21692-21697.

doi:10.1186/1471-2105-12-S1-S33

Cite this article as: Hsu et al.: iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. *BMC Bioinformatics* 2011 **12**(Suppl 1):S33.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

