Chapter 3

Validation and Application of GEMDOCK

To evaluate the strengths and limitations of GEMDOCK on molecular docking, we tested the program on a highly diverse data set of 305 protein-ligand complexes proposed by Nissink et al[17]. The validation analysis of GEMDOCK includes accuracy, precision, stability and affected factors. The accuracy of docking processes is assessed by RMSD values which are widely used in estimating the performance of docking tools. The precision and stability of this tool are assessed by the successful rates in different criteria of runs and affected factors of docking performance are considered in various aspects of constitution components of GEMDOCK, such as scoring function and algorithm. The validation results indicate GEMDOCK is a molecular docking tool of distinction. And we also evaluate the ability of GEMDOCK on application of virtual screening for a pharmaceutical interest target, human α-thrombin. Human α-thrombin plays a critical role in regulating blood coagulation and this protein belongs to serine protease which is widely distributed in nature and performs a variety of different function in various organisms. Viral genes code for proteases that cleave the precursor molecules of their coat proteins, bacteria produce many different extracellular proteases to degrade proteins in their surroundings, and higher organisms use proteases for such different functions as food digestion, cleavage of signal peptides, and control of blood pressure and clotting. With this successful rational drug design target, we explore and estimate the performance of GEMDOCK on virtual screening via enrichment factors and false positive rates when true positive rates are from 50% to 100%. These validation processes and results are presented in detail as following.

3.1 Molecular Recognition

As mention above, molecular docking is a valuable technique in structure based molecular design. In the development progress of docking tools, each approach without exception tries to capture the key components of the real protein-ligand binding state in different levels, such as physical-based, empirical-based and knowledge-based scoring functions and various heuristic

algorithms for flexible ligand/protein docking solutions searching. In order to effectively and impartially estimating performances of these docking programs, the benchmark test on a diverse and large test set is needed. In this thesis we evaluated our docking program, GEMDCOK, on such a largely diverse data set namely CCDC/Astex test set. This validation is an useful benchmark test for testing and comparison with other tools, such as GOLD and we use the set to explore and discuss the performance of GEMDCOK

3.1.1 Test Data Set

In our previous works[16,22,23], we have developed an evolutionary approach for flexible ligand docking and our approach, GEMDOCK[16], has been tested and validated on a diverse data set of 100 protein-ligand complexes. In this study, we further evaluated and explored the strengths and limitations of GEMDOCK on a newly large diversity data set of 305 protein-ligand complexes, CCDC/Astex test set (Table 3) which was proposed by Nissink et al[17]. Our evaluation followed our previous works[16] and Jones et al.[9] which analyzed the docking performance by root-mean-square deviation (RMSD) of heavy atom positions between the docked conformation and the crystal structure. Practice has shown that good docking solutions usually have RMSD within 2 ? and we use this index to access the accuracy of docking prediction.

The CCDC/Astex test set is designed for the purpose of validating algorithms that rely on the prediction of protein-ligand interactions[17]. This test set consists of 305 protein-ligand complexes from protein data bank (PDB) and the constituents of these entries are from original GOLD set (134 entries), chemscore set (68 entries, but overlapping with GOLD set in 18 entries) and 123 new entries. Nissink et al[17] have carried out particularly checking to identify improper entries in this set. These checks are described as follow:

1. Assessing the involvement of crystallographically related protein units in ligand binding.
2. Identification of bad clashes between protein side chains and ligand.
3. Assessment of structural errors and/or inconsistency of ligand placement with crystal structure electron density.

And the test set has been modified into several subsets for reasons of diversity and demands of

different structural resolution ranges. Table 3 integrates the whole information of this test set. All entries in this test set consist of 305 complexes. After carrying out structural and crystallographical checks and diversity identification, 81 entries (60 entries with structural deficits and 21 entries with diversity concern) were removed. The rest of 224 entries consist of the clean list. The clean list is further pruned to two subsets with criteria of resolution value 2.5 ? (180 entries) and 2.0 ? (92 entries). This test set is freely available to the academic purpose and the download website is BioXGEM (http://gemdock.life.nctu.edu.tw/bioxgem/).

3.1.2 The Protein/Ligand Preparation and GEMDCOK Parameters

GEMDOCK could accept three kinds of file formats as ligand inputs, such as PDB, MDL MOL and SYBYL MOL2. In our validation process, we transformed all ligands from PDB format into MDL molecule file format with preserving their coordinates. Crystal coordinates of the ligand and protein atoms were obtained from PDB and then separated as files of ligand and receptor. Our program would assign the atom formal charge and atom types (i.e., donor, acceptor, both, or non-polar) for each atom of both the ligand and protein. The bond type (sp^3 and sp^3-sp^3, sp^2-sp^2, or others) of a rotatable bond inside a ligand was also assigned. These variables were used in Equation 3 to calculate the scoring value of a docking solution.

When the coordinates of target protein were prepared, the size and the location of the ligand bind site were determined by extracting atoms of the protein which were located < 8 ? from each ligand atom. Atoms of metals and cofactors in the active site were also included. All structure water molecules were retained in their files, but our program would ignore them in docking procedure. Furthermore, we would evaluate our method in the presence of structure water molecules and compared the results. GEMDOCK then automatically decided the search cube of a binding site based on the maximum and minimum values of coordinates among these selected protein atoms. Table 4 indicates the setting of GEMDOCK parameters in validation procedure, such as initial step sizes, family competition length (L = 2), population size (N = 300), and recombination probability (p_c = 0.3).

3.1.3 Overall Accuracy on 305 Complexes

Table 5 collates validation results for the full validation set of 305 complexes (all entries), and for the clean list and subsets of different resolution thresholds. The results generated by GEMDCOK are compared with GOLD[9] in Table 6. All results are derived from 10 independent docking runs, and the docked lowest energy structure was considered for each test case. On average, GEMDOCK took 177s for a docking run on AMD 1.8GHz personal computer with single processor. The shortest time was 21 seconds for 1ldm and the maximum time was 876 seconds for 2er7.

When the solution at first rank was considered, GEMDCOK achieved 78% success in identifying the experimental binding model (Table 5). The RMSD values of 159 complexes were < 1.0 ?. When structure water molecules in the binding site were retained, the success rates improved to 83%. Figure 3 shows four typical acceptable solutions (the RMSD value <2.0 ?) in which GEMDOCK predicted correct positions for most of the ligand groups and RMSD values of these acceptable conformations were below 1.0 ?. Figure 3A shows glucose docked into glycogen phosphoryase b (2gpb)[24] and in Figure 3A, another form of glucose is docked soybean ß-amylase (1byb)[25]. Figure 3C is that hexadecanesulfonic acid is docked into adipocyte lipid-binding protein (1lic)[26] and Figure 3D shows progesterone is docked into progesterone receptor (1a28)[27]. Figure 3 shows four kinds of ligand, such as small and polar, large and polar, highly flexible, rigid and hydrophobic and GEMDOCK could indeed find out good solutions for these distinct cases.

We examined whether GEMDOCK could yield the correct solution in less than 10 runs for the 243 correct docking conformations of 305 test complexes. Figure 4 shows that GEMDOCK obtained the correct solutions for 43 complexes after one run (i.e., the percentage of success is 100%), while a total of 137 and 205 complexes were correctly predicted after 2 (i.e., = 50%) and 5 (i.e., =20 %) runs. When the structure water molecules were retained, GEMDOCK yielded the correct conformation for 64 complexes in a single run and for 176 and 236 complexes after 2 and 5 runs.

Table 6 is the comparison between GEMDOCK and GOLD in different lists. GEMDOCK had

78% of successful rate on the whole test set and GOLD was only 68%. When removed improper entries from test set, GEMDOCK still obtained 10% more than GOLD in this clean set. We made the comparison with GEMDOCK and GOLD in different resolution thresholds (2.5 ? and 2.0 ?). The average successful rates on these two subsets are 83% and 75.5% for GEMDCOK and GOLD respectively. GEMDOCK maintained successful rates of 83% on resolution 2.5 ? and 2.0 ? subsets and GOLD improved the rates form 73% to 78% when resolution criteria promoted. This analysis indicated that GEMDOCK is less sensitive than GOLD in crystal structure quality and GEMDOCK is superior to GOLD in this benchmark testing. An analysis of success rates in different rotational bond levels is carried out in Figure 5. We compiled statistics of success rate among different rotational bond levels and GEMDOCK reached an average success rate of 77% from 0 to 49 rotational bonds. The performance of GOLD decreased with the number of rotational bonds increasing. Comparing with GOLD, GEMDOCK could maintain good performances (over 65%) with the number of rotational bonds changing and these analyses suggested GEMDOCK is robust and useful in molecular docking.

3.1.4 Characteristics Analysis

This characteristics analysis discussed objective and subjective factors in this validation. Objective factors consider influence elements existing the test set, such as resolutions, structural defects, and structural water molecules. Subjective factors mainly come from the components in our program, such as scoring function and search algorithm. We analyzed and discussed these factors through exploring program performances in different testing conditions and accessing these performances via some estimating indexes, such as RMSD value and success rates. Except serious structural defects, critical structure water molecules losing and force field defects, most of unacceptable docking solutions caused by these factors could get improvement by adjusting parameters of GEMDOCK. The analysis could prove us important information about our program and could be the basis of program improvement.

As shown in Figure 4, GEMDOCK generates 67 unacceptable solutions shown in Table 5 (i.e., RMSD value > 2.0 ?) and these solutions can be divided into five categories. First category contains solutions when critical structure water molecules were removed from binding site

(Figure 6A, 3cla[28]). In second category, the ligand structures are large and highly flexible (Figure 6B, 1rne[29]). Third, the crystal structure complexes have significant clashes between protein and ligand atoms or incorrect ligand representation (Figure 6C, 1ake[30]). The fourth factor is that some specific interactions between protein and ligand have not been token into consideration in our energy model, such as the interaction of I..O in 1eta [31] shown in Figure 6D. The final factor is our scoring function could not distinguish between native and non-native conformations. According to these incorrectly docking solutions and crystal structures, GEMDOCK often inferred more hydrogen bonds than are in the native states to minimize the docking energy based on our energy function (Equation 3). In Table 6, we observed changes of successful rates different in different subsets. Comparing success rates between all entries and clean list, the rate promoted mainly because of removing entries with structural problems, such as clashes between protein and ligand atoms, incorrect ligand representations or dubious ligand geometries. These structural problems would lead to wrong predictions of docking program. As the comparing in Table 6, both GEMDOCK and GOLD were affected by this factor. Different resolution thresholds distinguished the clean list into three different test sets. When we analyzed the variation of success rates in these sets, we found that GEMDOCK was less sensitive on crystal structure qualities because of less changes in success rates. This character of GEMDOCK is an advantage for application on widely and different kinds of structural data sources, such as homology models or imprecisely crystal structures. Structure water molecules are another hard problem for molecular docking. Because considering water effect would substantially increase the computational complexity, GEMDOCK considered water effect depending on water molecules in crystal structures. We removed all water molecules in crystal structures in validation test but some critical water molecules played important roles in practice. The example in Figure 6A is a classical case. Structural water molecules mediated critical binding interactions in the reaction of acyltransferase. In this specific situation, retaining water molecules would correctly simulate the real binding condition and lead to correctly docking predictions. Such as shown in Figure 6A, when taking water molecules into account the RMSD value of the best predicted solution would reach 1.0 ? .

The genetic algorithm and scoring function are the cores of GEMDOCK. In general, a good scoring function often plays the significant role in molecular docking and most analyses focus on this part. Somehow, the search algorithm is often influenced by environmental interferences which involve the protein and flexible ligand. Figure 7 shows the relationships among

performances and these possible factors. In Figure 7A and 7B, the increases of sizes in receptors and ligands followed the increases of RMSD values. From analyzing relationships of RMSD values and sizes, we could found that sizes of cavities and ligands would bring heavy burden to the searching ability of genetic algorithm. Figure 7C indicates that the number of hydrogen bonds affects to the performance of GEMDCOK and this is due to our scoring function was designed for searching more number of hydrogen bond to optimize the fitness of docking solution. Figure 7D shows the flexibility of ligand also played a role in interfering the performance. We discovered the performance was affected by the number of hydrogen bonds and ligand flexibility from Figure 7C and 7D. In order to determinate major factors, we adopted ligand sizes to normalize numbers of hydrogen bonds and single bonds and results which excluded the interference of ligand sizes show on Figure 7E and 7F. From trends in Figure 7B, 7E and 7F, we could found that number of single bonds has more effect than number of hydrogen bonds in docking performance. With these evidences, our analysis of relationships among ligand size, number of hydrogen bond and number of rotational bond showed that the ligand flexibility played the major role in interference of the prediction of GEMDOCK. The trend in Figure 7G proved the conclusion in Table 6. Resolution of crystal structures made rare interference in the prediction of GEMDOCK.

3.2 Virtual Screening of Human a-thrombin

Virtual screening includes many computational techniques in structure-based drug design and the objective of virtual screening is to reduce the size of compound library to manageable quantity against specific protein target. We chose human a-thrombin as our virtual screening target because of its pharmacological importance and biological universality. We evaluated the virtual screening capability of GEMDOCK on this rational drug design target. In following evaluation, we access the performance of virtual screening through hit rate, enrichment factor and false positive rate. These indexes are widely used in estimating the effectiveness and reliability of virtual screening tools. The evaluation procedure is described in detail as following.

3.2.1 Introduction to Virtual Screening

Computer-Aided Molecular Design is becoming an essential component of drug discovery in every pharmaceutical industry. Structure-based virtual screening methods accelerate the process by speeding up the discovery of new chemical entities that may become new drugs. Structure-based virtual screening of compound databases has emerged as one of the most powerful and inexpensive approach to discover novel rational lead compounds for drug development[18,32]. With thrusts of explosions of high-resolution crystal protein structures, advent of the structural proteomics technologies, enriching the hit rate of high-throughput screening (HTS)[32,33] and reducing cost of drug discover, virtual screening is highly increasing used. Virtual screening covers four phases, including target protein modeling, compound database preparation, molecular docking and post-docking analysis.

The computational method of structural-based virtual screening involves two basic critical elements: high-effect molecular docking and a reliable scoring method. A molecular docking method for virtual screening has to be fast enough to screen millions of potential compounds with reasonable accuracy. Nowadays many molecular docking approaches have been developed and these approaches can be roughly categorized as rigid docking[34], flexible ligand docking[9,15], and protein flexible docking. Most current screening methods adopt flexible docking tools, such as incremental and fragment-based approaches (DOCK[35] and FlexX[15]), Monte Carlo simulation approaches (QXP[36]), evolutionary algorithms (GOLD[9], AutoDock[7], and GEMDOCK[16]).

Scoring methods for virtual screening should effectively discriminate between correct binding state and non-native docked conformation during the molecular docking phase and distinguish a small number of active compounds from hundreds of thousands of non-active compounds during the post-docking analysis. The scoring functions that calculate the binding free energy mainly include knowledge-based [2], physics-based [37], and empirical-based [4] scoring functions. The performance of these scoring functions is often inconsistent across different systems from a database search [38,39]. It has been proposed that combining multiple scoring functions (consensus scoring) improves the enrichment of true positives [38,39].

3.2.2 Introduction to Human a-thrombin

Human a-thrombin is involved in many diverse processes, such as cell signaling, memory, and blood coagulation[40]. In these functions of thrombin, the regulation of blood coagulation is discussed most. Thrombin is central to the clotting arm of haemostasis and also plays a role in shutting down the coagulation cascade. Thrombin is the product of an enzymatic amplification network in which inactive zymogen forms of many proteases and cofactors are activated by proteolytic cleavage. Functions of thrombin are mainly based on its protease characteristic. Thrombin cleaves fibrinogen into fibrin to create a fibrous plug and it also amplifies its own production through the activation of factor XI and cofactors V and VIII. Furthermore, thrombin activates the transglutaminase factor XIII that cross-links the polymerized fibrin plug stabilizing the clot. Thrombin also plays an important part in the activation of platelets through the cleavage of the protease-activated receptors (PARs) on the platelet surface[41]. The clinical importance of regulating thrombin activity is emphasized by the purpose of anticoagulant drugs to treat and prevent thrombosis. Anticoagulants include direct and indirect inhibitors of enzymes which involve in the coagulation pathways, mainly thrombin and factor Xa. Direct inhibitors interact with the procoagulant enzyme's active site or an exosite blocking its protease activity. Indirect inhibitors promote the protease inhibitory activity of natural anticoagulants[42]. Although there are many anticoagulants for antithrombic therapy, current therapy still suffers from the risk toward serious bleeding[43,44]. A direct correlation exists between the intensity of anticoagulation and severity of bleeding. The ideal anti-thrombin drug would prevent thrombosis without causing excess bleeding and should be orally available, inexpensive, reversible, and not require monitoring. Rational drug design has already resulted in synthetic compounds that increasingly satisfy the profile of the ideal agent. With this medically important serine protease, thrombin, we intended to estimate the screening ability of our program and mining useful information toward novel drug discovery of thrombin.

3.2.3 Parameters of GEMDOCK

The setting of GEMDOCK parameters for virtual screening is shown in Table VI and the

maximum generation is set as 60 for speeding up screening procedure. The GEMDOCK optimization stops when either the convergence is below certain threshold value or the iterations exceed a maximal preset value which is set as 60. Hence, GEMDOCK generated 1200 solutions in one generation and terminated after exhausting 720000 solutions in the worse case. These parameters were decided after molecular docking experiments guided and these values varied in different cases. GEMDOCK averagely took 242 seconds for a docking run on AMD 1.8 GHz personal computer with single processor.

3.3.4 Target and Database Preparations

For the thrombin screening experiment, the crystal structure of NAPAP in thrombin[45] (PDB code 1dwd) was used as the receptor structure[46]. The binding site of 1dwd was obtained by selecting atoms within 8 ? from the crystal ligand position and the known major binding pockets in the active site were included in this cavity, such as S1-S4 sub-sites and catalytic residues. All structural water molecules in the active site were removed and then GEMDOCK assigned a formal charge and atom type for each protein atom.

For building the test set of thrombin, we tested GEMDCOK on docking 10 active ligands (Figure 8) of human a-thrombin back the complexes with respective crystal structures from Protein Data Bank (PDB) and the binding site of 1dwd. The ligand was systematically named with four letters following three letters. Taking 7kme_BNN as example, the first four characters are PDB code and later three characters denote the ligand name in PDB. We selected and built this compound set according to complexes which conformed with following criteria. First, we found 86 entries from PDB and these entries are crystal structures of human α-thrombin. We removed 28 entries with characteristics below:

1. Protein inhibitors
2. Large peptide inhibitors (over 6 residues)
3. Inhibitors with covalent bonding inhibition to thrombin
4. Non-peptide inhibitors
5. Inhibitors binding on exosites

After removing improper complexes, we further selected 10 inhibitors from 58 complexes as our active ligands in screening test. These active ligands were chosen because of their diverse structures and non-covalent inhibitions to thrombin. As Figure 8 showed, these thrombin ligands are synthesized by peptide and modified amino acid and they all conserved the specific interaction which is recognized by S1 pocket of thrombin. In additional 1000 randomly chosen non-active compounds from the MDL Drug Date Report (MDDR) database and the 10 known active ligands were combined as the virtual screening test set for human a-thrombin. The drug database of MDDR included 132,726 compounds until May, 2004. In the first step of preparing 1000 randomly chosen non-active compounds, we filtered the MDDR with molecular weights between 200 and 650 (113,221 compounds) and removed analogues of thrombin inhibitors (removing 15,039 compounds). After removing small fragments from multi-component records, we randomly selected 1000 compounds from the remainder

Our program automatically decided the formal charge and atom type of each ligand atom in docking procedure. The ligand characteristics (i.e., the numbers of electrostatic atoms, hydrogen donor, and hydrogen acceptor) and the bond types of single bonds inside a ligand were also calculated. In Equation 9, these variables were utilized to calculate the scoring value of a docking solution. GEMDOCK re-ranked and sorted all docked ligand solutions for the post-analysis after completing all docking tasks.

GEMDOCK could utilize the knowledge of binding-site pharmacophores and ligand preference when set of known ligands presented. Figure 9 shows the binding-site pharmacological preferences and interactions by superimposing ten crystal structures of thrombin ligands. Six pharmacological preferences (hot spot atoms) are labeled in Figure 9 corresponding to their positions on active site of 1dwd. According to these observations, we added following pharmacological weights on active site of 1dwd in virtual screening: H57-NE2 and S214-O are hydrogen bonds with weighted value 2.0; D189-OD1, G216-O and G219-O are also hydrogen bonds with weighted value 3.0; D189-OD2 does not restrict interaction types with weighted value 4.0. These weights were used in Equation 1 for calculating the value E_{pharma}. Except mining binding-site pharmacological consensuses from active site of protein, we add the knowledge of known ligands from mining their structural and physical chemistry properties. These ligand preferences improve the screening accuracy by reducing the deleterious effects of ligand molecular weights and ligand structures that are rich in charged or polar atoms. Table 8 shows

values of ligand preferences in detail. The electrostatic parameter values (see Equation 9) for thrombin inhibitors included the maximum number of charged atoms (q_{elec}=4.00), standard derivation of the charged atoms (s_{elec}=0.97) and upper bound number of charged atoms (UB_{elec}=4.97). For the hydrophilic preferences (see Equation 10), the maximum ratio (q_{hb}) was 0.32, the standard derivation (s_{hb}) of the ratios was 0.04, and the upper bound ratio (Ur_{hb}) of polar atoms was 0.36. For molecular weight (see Equation 11), the mean of heavy atoms (m_{mw}) calculated as 33.2 and linear normalization parameter K was 0.17. These ligand preferences would be contributed to E_{ligpre} (Equation 1).

3.3.5 Screening Accuracy Analysis

Some common factors were used to evaluate the screening quality, including coverage (the percentage of active ligands retrieved from the database), yield (the percentage of active ligands in the hit list), false positive (FP) rate, enrichment, and goodness-of-hit (GH). The coverage (true positive rate) is defined as A_h/A (%), A_h/T_h (%) is the yield (hit rate), and the FP rate is defined as $(T_h\text{-}A_h)/(T\text{-}A)$ (%). The enrichment is defined as $(A_h/T_h)/(A/T)$. A_h is the number of active ligands in the hit list, A is the total number of active ligands in the database, T_h is the total number of compounds in the hit list, and T is the total number of compounds in the database. The GH score is defined as [47]

$$GH = \left(\frac{A_h(3A + T_h)}{4T_h A} \right) \left(1 - \frac{T_h - A_h}{T - A} \right) \tag{12}$$

The GH score contains a coefficient to penalize excessive hit list size and, when evaluating hit lists, is calibrated by weighting the score with respect to the yield and coverage. The GH score ranges from 0.0 to 1.0, where 1.0 represents a perfect hit list (i.e., containing all of, and only, the active ligands). In the data sets for screening the ER agonists or ER antagonists, A and T are 10 and 1000, respectively. Here, we also took the averages of hit rates, enrichments, GH scores, and FP rates. For example, the averages of the hit rates and enrichments are defined as $(\sum_{i=1}^{A} i/T_h^i)/A$ and $\left\{\sum_{i=1}^{A}(i/T_h^i)/(A/T)\right\}/A$, respectively, where T_h^i is the number of compounds in a hit list containing i active compounds.

Table 9 summarizes the overall predicted accuracies of GEMDOCK and GOLD when they docked 10 known active ligands into the reference protein. Both GOLD and GEMDCOK used default setting for screening, respectively and the structure with the lowest energy was used to calculated RMSD values for ligand heavy atoms between the docked conformation and the crystal structure. These ligands were docked into reference protein (1dwd). All the docked conformations predicted by GEMDOCK had an RMSD of less than 2.0 Å, but GOLD could only have 3 acceptable conformations at default library screening setting of GOLD. GEMDOCK yielded similar results regardless of whether the pharmacological preferences were considered.

Table 10 compares GEMDOCK with GOLD on screening 1010 compounds with false positive rates. GEMDOCK used the scoring function without pharmacological interactions and ligand preferences (None in Table 10) and with both of E_{pharm} and E_{ligpre} (Both in Table 10). The result showed that GEMDOCK generated lower false positive rates with using pharmacological preferences (3.40%) than without using them (10.70%). GOLD with combination of external *vdw* and external *H-bond* of GoldScore (New in Table 10) performed better than with GoldScore (Ori in Table 10) in false positive rates due to GoldScore giving too large penalty in internal *vdw* for docking solutions. The summary of screening accuracies was shown in Table 11. The average of these measure factors are defined as forward and the accuracy of GEMDOCK with pharmalogical preferences are accessed through its hit rate, enrichment and GH score. The average GH score of GEMDOCK without pharmacological preferences was 0.30, GEMDOCK with pharmacological preferences improved to 0.39 and GOLD with GoldScore and recombinant GoldScore were 0.07 and 0.10, respectively. The average false positive rate showed the pharmacological preferences and ligand preferences indeed improved the screening quality of GEMDOCK and this characteristic is clearly revealed when comparing Table 11 with Table 12. The whole detail of screening ranks was shown in Table 12. In Table 12, GEMDOCK was superior to other approach (GOLD) for screening the thrombin inhibitors, especially with pharmacological preferences. When the true positive rates range from 80% to 100% (Table 10), GEMDOCK with pharmacological preferences had the lowest false positive rate (3.40%) among different methods (GEMDOCK without pharmacological preferences: 10.70%, GOLD with GoldScore: 99.30%, and GOLD with recombination GoldScore: 99.80%). Lower false positive rate is important when screening a large number of compounds against a specific target. In such condition, the low false positive rate would improve the efficiency of hit rate in follow-up experiments. Figure 10 shows

more detail results of GEMDOCK for true hits, GH scores, enrichment factors and false positive rates in different levels of true positive rates. As shown in Figure 10B, GEMDOCK generally recovered all known ligands more efficient than GOLD and GEMDCOK with pharmacological preferences maintained the GH score over 0.3 when true positive rates ranged form 50% to 100%. In Figure 10C, the enrichment factors of using pharmacological preferences (enrichment 22.95 at true positive rate 100%) are higher than that using only original empirical score in GEMDCOK (enrichment 8.63 at true positive rate 100%). The false positive rates for different true positive rates are shown in Figure 10D, and the false positive rates of GEMDOCK (average 4.18% without pharmacological preferences and 1.38% with them) are small than GOLD (average 53.66% with GoldScore and 33.67% with recombinant GoldScore). Figure 10 revealed that GEMDOCK yielded good performance when it used pharmacological preferences in original empirical scoring function. When comparing with GOLD, GEMDOCK generally improves the screening quality by considering both ligand preferences and binding-site pharmacological interactions although we did not attempt to refine any parameters of these combinations. Therefore, the performance of GEMDOCK on this screening test indicates that GEMDOCK has the ability for virtual chemical database screening and the scoring function of GEMDOCK is suitable for molecular docking and it could also prioritize the hits in virtual screening.