Journal of Computer-Aided Molecular Design, **15:** 429–446, 2001.
KLUWER/ESCOM

Protein ligand docking based on empirical method for binding affinity estimation

Peng Tao & Luhua Lai*
Institute of Physical Chemistry, Peking University, Beijing 100871, P.R.China

Received 14 June 2000; accepted 12 January 2001

Key words: binding affinity, docking, molecular recognition, protein-ligand interactions, score

Summary

An empirical protein-ligand binding affinity estimation method, SCORE, was incorporated into a popular docking program, DOCK4. The combined program, ScoreDock, was used to reconstruct the 200 protein-ligand complex structures and found to give good results for the complexes with high binding affinities. A quality assessment method for docking results from ScoreDock was developed based on the whole test set and tested by additionally selected complexes. The method significantly improves the docking accuracy and was shown to be reliable in docking quality assessment. As a docking tool in structural based drug design, ScoreDock can screen out final hits directly based on the predicted negative logarithms of dissociation equilibrium constants of protein-ligand complexes, and can explicitly deal with structure water molecules, as well as metal atoms.

Introduction

Structural Based Drug Design (SBDD), which is based on the three dimension (3-D) protein structures from X-ray crystallographic or nuclear magnetic resonance (NMR) methods, is one of the most successful strategies in computer-aided drug design (CADD) [1–3]. As one advantage for this approach, new specific binding ligands can be rationally produced based on the information of the protein-ligand binding mode. And the native 3-D complex structure is the very start point of the whole process. Although the number of possible ligands for each protein can be large, due to the restriction of X-ray crystallographic and NMR methods, the protein-ligand complexes with known 3-D structure are not many. At present, two prevalent strategies in SBDD are database screening [4–6] and de novo design [7–10]. However, how to find out the most possible binding modes between ligands and proteins, and to rank them based on the binding affinity remain challenging.

*To whom correspondence should be addressed: E-mail: lai@ipc.pku.edu.cn

Kuntz, I.D. and coworkers first introduced and extensively developed a so-called docking method to predict 3-D structure of ligand-receptor complexes [5, 6, 11]. In general, docking process can be divided into two phases. One is the searching algorithm, which finds possible binding geometries of the protein and its ligand. The other is the scoring function, which ranks the searching results and selects out the best binding geometry based on the energies of the complexes or, more theoretical value, ΔG_{bind}, the binding free energy difference between the bound and unbound states of the ligand and protein. The program suite, DOCK, which was distributed by Kuntz, I.D., is one of the most popular docking tools [6]. The basic idea of DOCK searching algorithm is to generate a set of spheres to fill the whole pocket as the 'negative image' of the protein binding site. Then DOCK program matches the ligand to this set of spheres and generates many complementary binding geometries of the ligand into the protein. Besides DOCK, many dock programs based on other searching algorithms, e.g. surface complementary matching [12–14], fragments growing [15–17], random sampling using stochastic processes (Monte Carlo, simulate annealing) [18–22], and genetic algorithm [23–25] have been developed.

Except for receptor flexibility, scoring methodology is more attractive and challenging than other problems in docking studies. Ajay and Murcko have provided a comprehensive review of this area [26]. Similar to the requirement of searching algorithm, accuracy and efficiency have almost equal importance in scoring algorithm application. Although elaborate methods, such as free energy perturbation (FEP) and thermodynamic integration (TI), sound more theoretically convincing, the demand of computational power prevent them from being applied widely in docking and other molecule modeling processes. However they remain the basic methods presently, due to their explicit treatment of calculation ensemble averages and solvent molecules [27–29]. There are still attempts to introduce these methods into real computer-aided drug design process [30].

Most of the docking tools use empirical potential energy functions to calculate the binding energies of protein-ligand complexes. These functions sum up several basic energy terms, such as Van der Waals interaction, electrostatic interaction and hydrogen bonds. The set of parameters and the functional forms used in energy functions is a so-called force field. Nowadays, many force fields [31–37], which differ in the values of corresponding parameters and in the functional forms to deal with particular chemical or biological system, have been widely used in molecular modeling.

Meanwhile, another calculation technique – so-called grid-based energy evaluation has also been introduced into scoring to facilitate the calculation of binding energy [38, 39]. Applying this technique, an evenly divided cubic grid is generated in the box that efficiently encloses the space that ligands are likely to occupy. Then the receptor-dependent terms in the potential function at points of grid are precalculated and stored. During the docking process, the energy of each binding mode is calculated based on the cartesian coordinates of ligand atoms and the precalculated terms, avoiding to calculate through each protein atom repeatedly. To balance between the accuracy and the efficiency, the selection of cubic grid resolution is a dilemma of this method.

Recently, empirical schemes have met with significant interest, with the attempt to circumvent the huge computational load and cover the physical effects as many as possible. Generally, these empirical methods try to partition the binding affinity into several additive terms, which will cover both of enthalpic and entropic effects in binding affinity. The parameters used by this method are often determined by the analysis of binding affinity data set of complexes. And this kind of method is also referred as 'Master Equation' [26]. In early stage of studies, only a series of ligands of one particular protein have been investigated, and the valid scope of the equation is restricted to the complexes of the target protein either [40]. Richard D. Head et al. [41] analyzed a diverse training set of 51 crystal complexes by partial least squares (PLS) statistics and neural network analysis, and obtained models for the general prediction of the binding affinity of complexes with known 3-D structure. Subsequently, several empirical scoring models have been advanced [42–47], in which parameters cover parts of the following chemical and physical effects, e.g. hydrophobic, electrostatic, the number and geometry of hydrogen bonds, ionic interactions between the protein and ligand, the size of the lipophilic contact surface, the flexibility of the ligand, the electrostatic potential in the binding site, structure water, etc. [42]

Recently, another method denoted knowledge-based potential of mean forces (PMFs) has attracted a great deal of interest in ligand-protein complex binding affinity estimation [48–50]. Holger Gohlke et al. [50] developed a new knowledge-based scoring function named DrugScore to predict protein-ligand interactions, and tested their models by re-ranking the docked results of two docking programs. Different from the previous empirical methods, DrugScore focused on the statistic thermodynamic properties of complex crystal structures. Using the Boltzmann format function, DrugScore scored complexes by summing up the distance dependent pair-preferences and solvent-accessible surface (SAS) dependent singlet-preferences of protein and ligand atoms, based on the assumption that only those binding modes are favorable that fit to the maxima of distributions of occurrence frequencies among interatomic contacts between particular atom pairs in experimentally determined structures.

SCORE is an empirical method for estimating the binding affinity for protein-ligand complexes developed in the author's group [51]. Based on the analysis of a diverse training set of 170 protein-ligand complexes, one robust scoring function has been obtained: the standard multivariate regression yielded a squared correlation coefficient (r^2) of 0.777, and a standard deviation of 6.6 kJ/mol in binding free energy. Although the final model has been used to predict the binding affinity of additional 11 protein-ligand complexes to test its predictive ability, its power to discriminate be-

tween the native or near-native binding geometries and those largely deviating from the native structure generated by docking process has not yet been tested. In this paper we applied this scoring method to a popular docking tool, DOCK4, as the ranking function in the second stage of docking process. A test set which contains 200 complexes has been used, and the docking results generated by the combination program of SCORE and DOCK are compared to those generated by original DOCK.

Methods

SCORE function

In SCORE method, the protein-ligand binding affinity is partitioned into several terms, e.g. Van der Waals interaction between the protein and its ligand, metal-ligand bonding, hydrogen bonding, desolvation effect, and deformation of ligand. Each term is described briefly in the following text. For detailed information about this method, please refer to Ref. 51.

(1) Van der Waals (VDW) interaction

The term for VDW interaction is simply a pairwise counting of VDW bumps between the protein and the ligand.

$$K_{vdw} = \sum_i \sum_j VB(d_{ij}) \tag{1}$$

where d_{ij} is the distance between atom i and j, VB is a step function of d_{ij}.

(2) Metal-ligand bonding

The term for metal-ligand bonding in SCORE method is the sum over all metal- O/N bonds,

$$K_{\text{metal}} = \sum_i \sum_j MB(d_{ij}) \tag{2}$$

where d_{ij} refers to the distance between ligand atom i and metal j, and MB is a step function of d_{ij}.

(3) Hydrogen bonding

In this model hydrogen bonds can be seen as two types: one (HB) is completely formed by the atoms from protein and ligand, the other (WH) is involved by water molecule. Based on a step function of the distance of the atoms involved in hydrogen bond, these two kinds hydrogen bonds are classified as strong, moderate, and weak (water-involved) hydrogen bonds with SHB (SWH), MHB (MWH), and WHB (WWH) as the indicator respectively.

Without taking the angle dependence of hydrogen bonding strength, the contribution of hydrogen bonding between the protein and its ligand is calculated as

$$\begin{aligned} K_{hbond} &= K_{SHB}+K_{MHB}+K_{WHB}+K_{SWH} \\ &\quad +K_{MWH}+K_{WWH} \\ &= \sum_i\sum_j SHB(d_{ij}) + \sum_i\sum_j MHB(d_{ij}) + \\ &\quad \sum_i\sum_j WHB(d_{ij}) \\ &\quad + \sum_i\sum_j SWH(d_{ij}) + \sum_i\sum_j MWH(d_{ij}) + \\ &\quad \sum_i\sum_j WWH(d_{ij}) \end{aligned} \tag{3}$$

(4) Desolvation and deformation effect

Since both the protein and its ligand are solvated before complexation, the desolvation effect accompanies the whole binding process. The contribution to the binding affinity of desolvation effect is calculated using the following equation,

$$K_{HM} = \sum_i F_i \times HM_i \tag{4}$$

where HM is an indicator of hydrophobic matching between the ligand atom and its environment. It is set to 1 if ligand atom i is hydrophobic and placed in a hydrophobic environment; otherwise it is set to 0. F_i is the atomic hydrophobic scale of ligand atom i. Here the atom hydrophobic scale (both of protein and ligand) is referred to the results of a previous work [52].

In calculation of deformation effect, only the number of ligand rotatable bonds, i.e. rotors, is used in the term, because the protein is considered as rigid body in this method.

$$K_{RT} = \sum_i 0.5 \times RT_i \tag{5}$$

Here RT_i is the number of rotors in which ligand atom i is involved. The factor 0.5 means that each rotor is split into halves and assigned onto the two atoms involved.

The dissociation constant of a protein-ligand complex can be calculated by summing up all of the terms,

$$\begin{aligned} pK_d = {} & K_0 + c_1 \times K_{VB} + c_2 \times K_{MB} \\ & + c_3 \times K_{SHB} + c_4 \times K_{MHB} \\ & + c_5 \times K_{WHB} + c_6 \times K_{SWH} \\ & + c_7 \times K_{MWH} + c_8 \times K_{WWH} \\ & + c_9 \times K_{HM} + c_{10} \times K_{RT} \end{aligned} \tag{6}$$

K_0 is the regression constant, which may contain the translational and rotational entropy loss upon the binding process, and c_i is the regression parameter for each term.

DOCK process

We used the 4.0 version of DOCK program [6]. The core process of DOCK4 program suite is the program dock, which is written in C and Fortran languages. DOCK4 requires a set of overlapped spheres, which describes geometry of the receptor active pocket, while the ligand structure is input as mol2 file format in preference [53]. Then the ligand atoms are matched to the spheres set. Each matching is one of the possible orientations for ligand in the binding pocket. If ligand flexibility is considered in docking process, the ligand molecule will be taken apart into several rigid fragments, and one appropriate fragment is selected as anchor. Then the selected anchor is matched into the binding site initially. After that, the program searches the position of the remaining fragments and links them to the anchor subsequently. Meanwhile, each fragment is scored and the score is used to direct the further conformation searching. After the whole searching process, the searching results, i.e., ligand conformation and orientation is scored and sorted altogether, and the best scored solutions are output as final results of docking. Figure 1 shows a brief flow chart of the dock process.

The application of SCORE in the scoring stage of DOCK4

The main purpose to develop scoring methods is to predict the binding affinity of a given protein-ligand complex structure. A practical usage of scoring function is to distinguish between native or near native binding modes and those largely deviate from the native structure. Therefore applying it directly in docking process as a criterion to rank the matching results will be a better way to test SCORE method than just re-ranking the output results of docking process, in consideration of some good poses can be discarded by the scoring function used by dock program. In the whole process of dock program, there are two choices for SCORE function to apply: first, SCORE function can be applied as searching driver during ligand conformation searching phase; second, it can act as final scoring function after searching process.

Although the whole equation of SCORE can be expressed in atom addition form [51], all of the parameters are generated by multiple linear regression analysis based on the full structure of each complex in the training set. Therefore these parameters implicitly require the completeness of ligand structure. For instance, the regression constant K_0 in equation reflects the translational and rotational entropy change of the whole binding process. On the other hand, in each calculation process of SCORE, all of protein atoms will be examined one by one, including structure water molecules. This calculation process is quite time-consuming comparing to the grid-based energy calculation method used in DOCK4 program. Based on these analysis, we chose the second stage of docking process to apply the SCORE function, i.e., using SCORE method to rank the ligand poses generated after searching stage. Finally, the top ranking solutions are written to the output file.

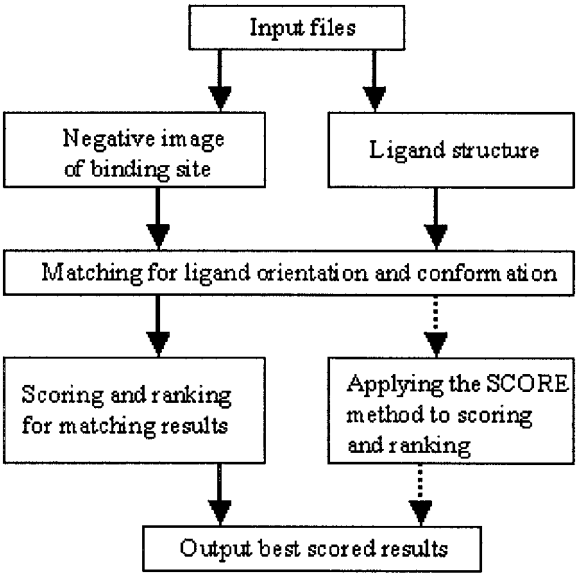

Figure 1. The flowchart of dock in DOCK4 (dashed lines indicate the application of SCORE in DOCK4).

During the modification of DOCK program, the searching algorithm of original program remains intact. The SCORE function is added to the final scoring part as optional for user. If users select SCORE as final scoring method, the whole docking process is the combination of searching engine of dock and the SCORE method. Therefore the qualities of final output results are mainly influenced by the power of searching engine and the discrimination ability of SCORE method. In the following text, we will refer the combination of the two programs as ScoreDock for convenience. The flowchart of ScoreDock is also shown in Figure 1, by

flowing in and out the scoring box through the dashed lines.

Results

Test set and data preparation

To test ScoreDock, we selected 200 complexes[1] crystal structures from Protein Data Bank [54] as the test set. There are 85 complexes taken from the database of human disease-related protein structures [55], while the other 115 complexes in the test set were randomly selected from the ReLibase [56]. All of the 200 crystal structures have resolutions better than 2 Å. The number of non-hydrogen atoms of these 200 ligands range from 4 to 110 with an average value of 28.8. The number of rotatable bonds in these ligands range from 0 to 49 with an average value of 7.1. For all the docking schemes (rigid and flexible), the following boolean parameters, such as 'write_orientations', 'rank_orientations', 'intramolecular_score', 'intermolecular_score', 'bump_filter', 'energy_minimize' are set as yes. The parameter 'maximum_iterations', 'rank_orientation_total', 'bump_maximum' are set to 500, 10 and 3, respectively. For flexible docking scheme, the boolean parameters, 'flexible_ligand', 'anchor_search', 'minimize_anchor', 'peripheral_search', 'torsion_drive', 'torsion_minimize', 'reminimize_ligand' are set as yes. The parameter 'clash_overlap' is set to 0.3. The chemical score and energy score can be used as scoring function by set 'chemical_score' and 'energy_score' as 'yes', respectively [57]. All of other parameters not mentioned here are set as default.

To prepare the data set for Dock process, we extracted the ligand molecules from the complexes, delete all of the water molecules and heterogenous atoms including metal atoms for the protein surface calculation. For the SCORE calculation, we kept all of the atoms except the ligand atoms. All of the structural water molecules can be kept for SCORE function without causing VDW bumps, because the program can automatically select appropriate water molecules for hydrogen bonding calculation and simply disregard others.

Here we choose a generally accepted value to measure the quality of docking results: the root-mean-square deviation (RMSD) of the docking position with respect to the original pose of the ligand in the crystal structure. Considering the restriction of crystal structure imposed by experimental method, such as resolution, the temperature factor etc., a ligand pose with RMSD less than 2.0 Å can be considered as a 'well-docked' solution [50]. To evaluate the power of scoring function to distinguish between the 'well-docked' solutions and those poses largely deviating from the crystal structure, we select the criterion that the scoring function can be regarded as good only if one well-docked pose is scored best among all of the computer generated solutions.

Calculated results of the full test set

For the 200 complexes, ChemicalDock (DOCK using the 'chemical score') ranks a well-docked solution first in 27 cases, while EnergyDock (DOCK program using energy score) and ScoreDock perform well in 105 and 108 cases, respectively. Here we sorted the first ranked RMSD values of the whole test set ascendingly, and plotted the rank number versus the RMSD values (Figure 2). In addition, for all the complexes, the poses with the smallest RMSD value generated by ScoreDock are also selected and plotted in Figure 2, without considering the ranks of these poses. ScoreDock and EnergyDock act at nearly the same good level at low RMSD value (0–2.0 Å), while ScoreDock shows slight improvement in the region of 2.0–4.0 Å. Both ScoreDock and EnergyDock generated much better results than ChemicalDock. In Table 1, a detailed comparison of these methods is presented. As Table 1 shows, the searching method acts rather satisfactorily here – for almost 91% of all the complexes, at least one well-docked solution has been found. So the main factor that limits the selective ability of scoring functions is their own deficiency. In the consideration of the variable size of ligands, the uniform cutoff value (2.0 Å) for 'well-docked' cases may not be suitable. We have plotted the smallest RMSD values among all the gen-

[1] 16pk 181l 1a09 1a27 1a4q 1a8b 1abe 1abf 1afl 1aim 1ake 1amk 1aoe 1aqv 1ax0 1azm 1ba8 1bcd 1bck 1ben 1bhp 1bht 1bji 1bjv 1bk0 1bmb 1bul 1bv3 1bv7 1bz0 1bzm 1bzy 1can 1cbm 1cbs 1cbx 1cg8 1cil 1ckb 1clu 1coy 1cwl 1cyn 1daa 1dae 1dan 1dhj 1dim 1dor 1drf 1dyj 1ecf 1ela 1eus 1fkb 1fkf 1fkg 1frp 1gai 1get 1ghb 1grb 1gua 1gup 1gux 1hcb 1hck 1hew 1hfc 1hgx 1hne 1hrn 1hsb 1hsg 1hsl 1hvr 1hyt 1iak 1icn 1ida 1ir3 1ivd 1jah 1jap 1kbc 1klt 1kvr 1lcp 1ldg 1lhc 1lic 1lkk 1llo 1lna 1lst 1meu 1mh1 1mik 1mld 1mrg 1mrj 1mrk 1mtv 1mwe 1nes 1nsc 1nue 1phf 1php 1poc 1pph 1ppl 1pso 1qa7 1qqp 1rbp 1rdn 1rds 1rnn 1rnt 1rob 1rpj 1ruv 1rxg 1scn 1sdk 1sep 1sha 1shd 1snc 1sre 1srj 1sth 1syb 1tag 1tem 1tgj 1tng 1tnh 1tni 1tnk 1tpb 1tpp 1tyr 1udh 1ukz 1v39 1vbs 1vp3 1vpe 1vps 1vpt 1vwo 1wap 1wyk 1xan 1xid 1xie 1xif 1xym 1xzm 1zfp 21gs 2ak3 2bpv 2clr 2cmd 2cpp 2fal 2fmb 2gbp 2hbe 2ki5 2mth 2qwc 2rth 2shp 2sim 2sns 2tmn 2xis 3bto 3ca2 3chb 3cla 3cpa 3er5 3ptb 3sli 4csc 4est 4fiv 4gal 5abp 5tmn 6abp 6rnt 6rsa 6tmn 7tim

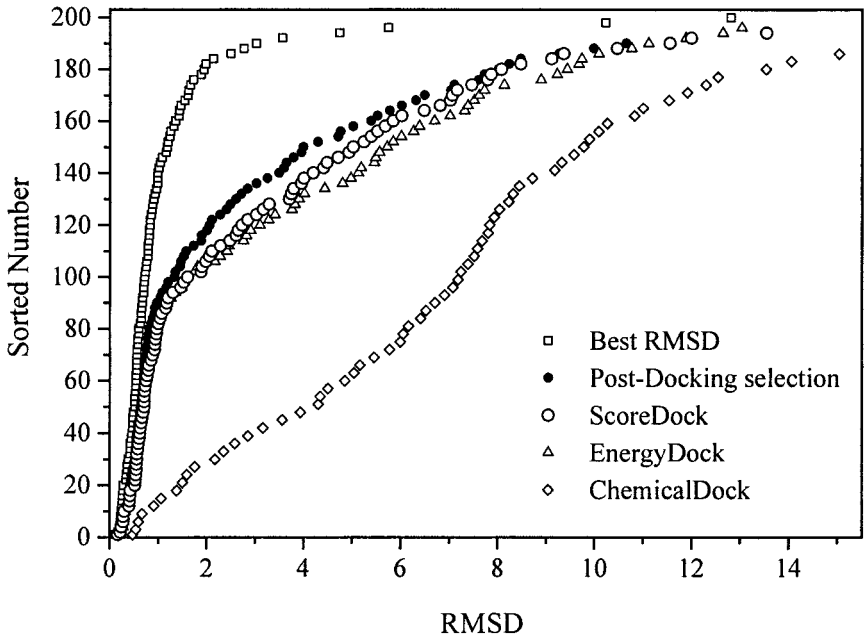

Figure 2. Sorted number of the first ranked solution according to the RMSD value for the three scoring functions. The results after the post-docking selection based on the docking quality estimation are also plotted. 'Best RMSD' indicates the smallest RMSD generated by ScoreDock disregarding the rank number.

erated solutions versus non-hydrogen atom numbers (Figure 3). For most cases in the test set, the ligands are large enough comparing to the best RMSD. Therefore the 2.0 Å cutoff value is reasonable for this test set.

The differences between the RMSD values of first ranked solutions by EnergyDock and ScoreDock, which reflect the fine distinction of these two methods, are shown in Figure 4. For 117 cases, the RMSD difference between these two methods is lower than 1.0 Å. And in 42 cases, the poses selected by ScoreDock are more than 1.0 Å closer to the crystal structure. For the remaining 41 cases, the poses selected by EnergyDock are closer to the crystal structure. Although EnergyDock and ScoreDock generate similar statistical results (see Table 1), the cases, for which these two functions could rank a well-docked solution best, as expected, are not completely identical. There are 82 complexes altogether that both EnergyDock and ScoreDock correctly rank one well-docked solution at first. However each of them can only tell slightly different partial truth of whole story – for 26 cases, ScoreDock could rank a well-docked solution first, while EnergyDock failed, and for the other 23 cases, *vice versa*.

Assessment of ScoreDock running quality

According to the statistic results for the full test set, the probability that ScoreDock can rank a well-docked solution first is 54%. Among the ScoreDock running results, there are 92 cases that ScoreDock failed to rank a well-docked solution first among the final results. In those 92 complexes, the RMSD value of the first ranked solution ranges from 2 Å to 25 Å. We assumed that higher RMSD values indicate lower reliability of ScoreDock. To inspect the relationship between the ligand property and the reliability of ScoreDock, we averaged the number of non-hydrogen atoms, the number of rotors and the RMSD of the first ranked solution. Next, the lowest RMSD case is deleted from the data set and the remaining cases are averaged. The calculation is repeated until the last case with the largest RMSD value remains. The averaged rotor number and non-hydrogen atom number are plotted versus the corresponding average value of RMSD in Figure 5. In general, the larger RMSD, the more rotors and non-hydrogen atoms will the ligands have.

However, to determine the quality of docking results, we need more information in addition to the size of the ligands. To meet this requirement, an assessment method to predict the quality of ScoreDock running results is necessary, especially for cases where

Table 1. Statistics of docking results by applying three scoring functions and post-docking selection on the whole test set

		The number of complexes with RMSD reference to crystal structure			
		<1.0 Å	<1.5 Å	< 2.0 Å	≥ 2.0 Å
All ranks of ScoreDock[a]		140	167	182	18
First rank[b]	Chemical dock	13	20	27	173
	Energy dock	86	99	105	95
	Score dock	82	97	108	92
Post-Docking selection[c]		91	106	117	83

[a]The smallest RMSD generated by ScoreDock is selected for all of 200 complexes.
[b]The first ranked solutions of each complex according to three methods are considered
respectively.
[c]The results after performaning the post-docking selecting on ScoreDock results using
docking quality estimation method.

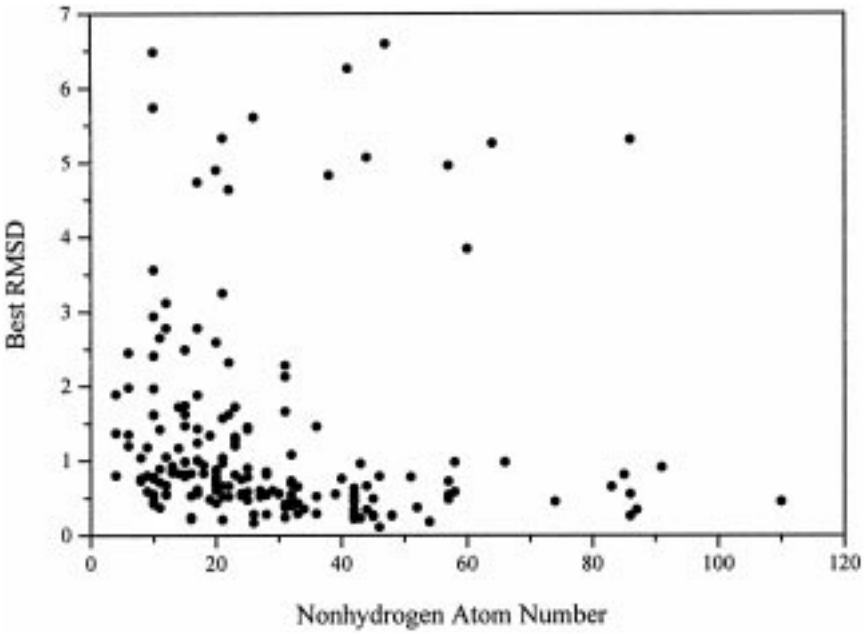

Figure 3. The smallest RMSD value among all the generated solutions for each case versus non-hydrogen atom numbers.

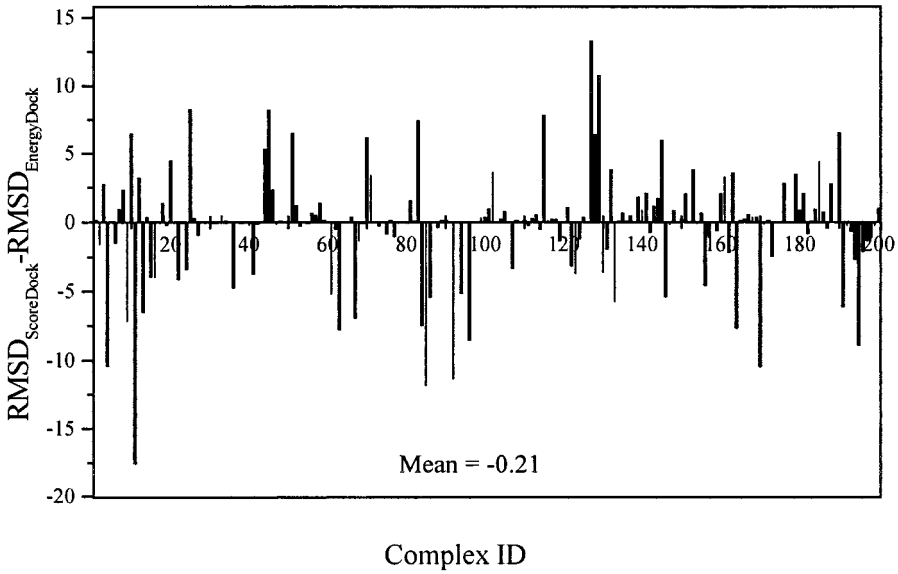

Figure 4. The differences between RMSD value of the first ranked solution by ScoreDock and EnergyDock. The mean value of all differences is −0.21.

the complex crystal structures are unavailable. After analyzing the full test set, we advanced an assessment method to meet this need.

Based on ScoreDock running results, we divided the full test set into two groups: one group contains the well-docked cases, while the remaining cases in the other one. When analyzing the full data set, we found that the first ranked scores in these two groups have distinct distributions (Figure 6). Obviously, the pK_d scores of those cases, for which the ScoreDock fail to rank a well-docked solution first, are generally lower than the well docked cases. The overlaped pK_d range of two groups is approximate from 3 to 8. We averaged the pK_d scores of these two groups respectively: for the well-docked group, the averaged pK_d score is 7.23; for the other group, the averaged value is 4.39. We use these two values as cutoffs to determine the quality of docking results, i.e. if the first ranked pK_d score is lower than 4.39, it is very probable that ScoreDock fail to rank a well-docked solution first, while for the value higher than 7.23, it is likely that we get a well-docked pose ranked first.

In the range from 4.39 to 7.23, the pK_d distributions of the two groups overlap substantially, and more information are needed to distinguish these cases. During the calculation of SCORE, each kind of interaction, e.g., Van der Waals, hydrogen bond, desolvation and deformation effect are counted and summed through all of the ligand atoms [51]. These count numbers contains detail information about the binding mode. The first ranked solutions of those cases with the pK_d score between 4.39 and 7.23 are extracted. The count numbers of those interaction defined in SCORE method are calculated. For the selected cases, we define 2 as docking quality of well-docked cases, while 1 for the remaining cases. Then the multiple linear regression analysis was performed on these data, which yields r^2 of 0.745, F value of 4.90, and N of 73. And the empirical prediction equation has the following additive form,

$$Q = \sum P_i C_i \tag{7}$$

where Q is the calculated quality of docking result, P_i is the regression parameter of each effect term defined in SCORE [51], and C_i is the count number of each effect calculated by program SCORE. The P_i values are listed in Table 2. If the calculated quality Q is larger than 1.5, it is very possible that the docking result is a well-docked solution, i.e. with RMSD lower than 2 Å; otherwise, it is more probable not a well-docked solution.

Using this estimation function, we reselected the ranked solution generated by ScoreDock based on the reliability of each binding solution. To perform this kind of post-docking selection, we calculated the reliabilities of the top 5 ranked solutions, and selected out the most reliable (with the largest Q value) solution as the final result without regarding the rank number. During this process, the rank number is disregarded.

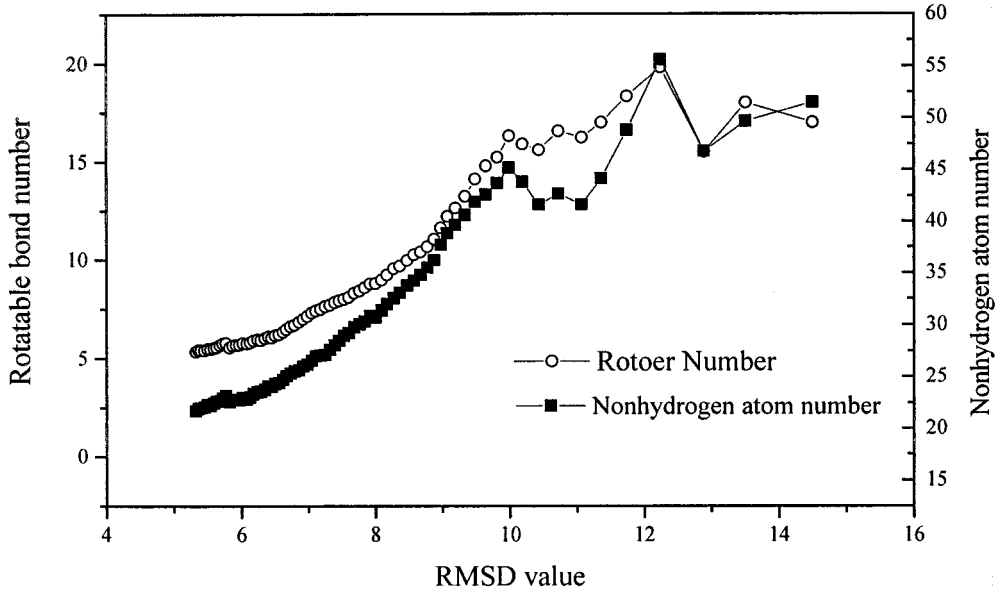

Figure 5. The stepwise averaged numbers of rotor and nonhydrogen atom of ligand plotted versus the corresponding RMSD values exhibited by the first ranked solutions for the cases the ScoreDock fail to rank a well-docked solution first.

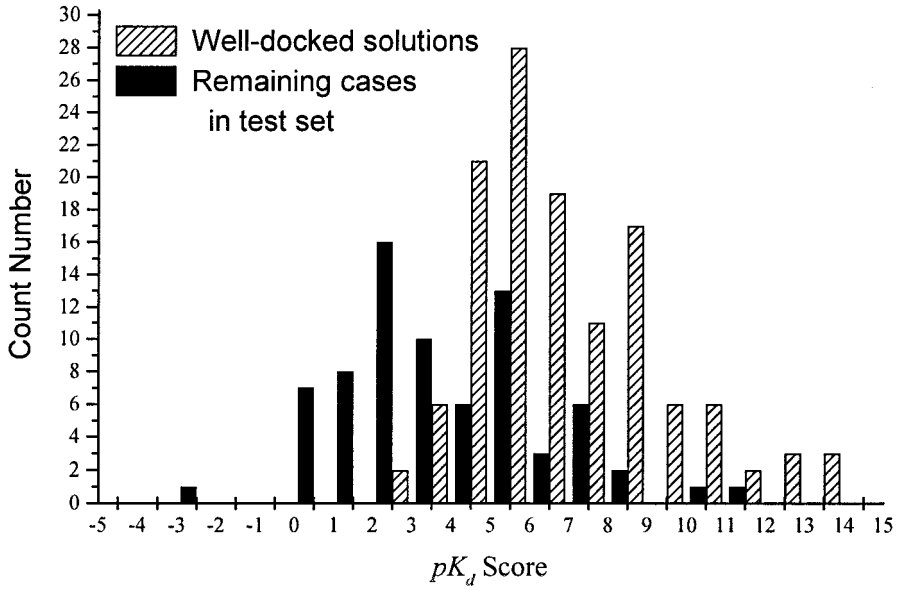

Figure 6. The pK_d distribution of the well-docked cases and the remaining cases in the whole test set (200 cases).

Table 2. The multiple linear regression parameters for ScoreDock running quality estimation[a]

ScoreDock interaction	VB	WHB	MHB	SHB	WWH	MWH	SWH	MB	HM	RT
Regression parameter(P_i)	−0.0860	0.1084	0.1466	0.2253	0.0983	−0.1042	0.1544	0.3349	0.3846	−0.0429

[a]The regression results: $R^2 = 0.745$, $F = 4.90$, $N = 73$.

As the reselected results, there are 117 cases with the RMSD of final solution less than 2 Å. The reselected results are also plotted in Figure 2. Comparing to the EnergyDock and ScoreDock results (Table 1), this estimation method can obviously improve the docking quality. In addition, whether or not use the cutoff values 4.39 and 7.23, i.e., perform the Equation 7 on all the cases or only the cases which fall into the range of 4.39–7.23, the post-docking selecting results are almost the same. This shows the robustness of this method. For all of the 200 cases, there are 154 cases for which this method could predict the quality of docking results correctly. Furthermore, we randomly selected 20 complexes with resolution better than 2 Å from ReLibase [54] to test the reliability of this method (see Table 3). The rotor numbers of the ligands in these complexes range from 0 to 42 with an average value of 6.6, while the number of non-hydrogen atoms range from 9 to 94 with an average value of 24.9. After the reselecting on these complexes, there is one more well-docked case in final results, and the assessment method correctly predicted the quality of the docking results for 19 cases.

By far, one assumption of the above analysis remains untested: does the complex, that ScoreDock fail to rank a well-docked solution first, really have low binding affinity, i.e., be less attractive in structure based ligand design? To answer this question, the docking on complexes with known experimentally determined binding affinity is necessary. In the original test set (200 complexes), there are 30 complexes[2] that the observed binding affinities are available [51]. In addition, we selected out 20 complexes[3] from the SCORE training set, in which, ten cases have observed pK_d lower than 4.39, while the other ten cases have observed pK_d higher than 7.23. After the performance of ScoreDock on the additionally selected 20 cases, we analyzed the distribution of observed pK_d values for these 50 complexes (Figure 7). The distributions of observed pK_d still have the trend we observed in the original test set, i.e., for the complexes with high binding affinity, the ScoreDock shows high reliability on ranking a well-docked solution first. One possible reasonable explanation for this phenomenon is that those ligands with high binding affinities to their receptors have particular binding modes corresponding to the observed binding affinities. For these cases, any different binding modes of ligands, including different conformation and binding sites, will have much lower binding affinities. This is the basic assumption for molecular recognition theory. On the other hand, those ligands in complexes with low binding affinity are not selectively bound to the proteins. Some of these complexes may take form during the crystallization process. Therefore, it is reasonable that ScoreDock can find another binding site for the ligand with near or higher binding affinity, while both binding modes, calculated by ScoreDock and taken by crystal, have low binding affinities, and have no specificity either. For these 50 cases with known pK_d values, the first ranked pK_d scores are plotted versus the experimental values of corresponding crystal complexes in Figure 8. The correlation between the observed and predicted values of the 50 cases yielded correlation coefficient of (r) 0.64, standard deviation (s) of 2.01. For most cases, the binding affinities blindly predicted by ScoreDock, i.e., from the separated structures of ligand and protein, do show the correlation to the observed value of the corresponding crystal structure, although the correlation is not very obvious. After discarding the cases with first ranked pK_d scores lower than 4.39, we replotted the sorted numbers of cases versus RMSD values in Figure 9. Among the 118 cases with best RMSD less than 2.0 Å, EnergyDock succeeds in 89 cases, and ScoreDock detects 97 cases. After the post-docking selection, there are 103 well-docked cases altogether (Table 4).

Furthermore, we performed the post-docking selection on those newly selected 20 cases with known binding affinities. As additional proof, there are 16 cases that the predicted qualities of the running results agree well with the RMSD of the first ranked solution, and there is one more well-docked case in reselecting results. (Data not shown.)

The recognition of crystal structures

For a scoring function with good discriminatory power, the well-docked solutions should obtain better scores than those poses largely deviating from the crystal structure. As a more stringent criterion, the crystal structure (RMSD = 0) should get the best score, i.e., the score for the crystal structure should rank first among all of the poses generated by docking program, considering that the probability for docking search tools to generate exact the crystal structure is extremely low. We scored all of the 200 crystal

[2] 1abe 1abf 1cbx 1drf 1fkb 1fkf 1hsl 1pph 1ppl 1rbp 1rnt 1snc 1tng 1tnh 1tni 1tnk 2ak3 2gbp 2sns 2tmn 2xis 3cla 3cpa 3ptb 5abp 5tmn 6abp 6rnt 6tmn 7tim

[3] 1apt 1apw 1csc 1hpv 1hvi 1hvl 1l83 2ctc 2rnt 3fx2 4dfr 4sga 4tmn 5can 5sga 5tim 6cpa 7can 7est 8xia

Table 3. The post-docking selection results on randomly selected test set

PDB entry[a]	pK_d score	Calculated quality[c]	Predicted quality[b]	Real quality	RMSD (Å)
1aht	5.18	2.27	1	1	1.56
1art	3.92	0.90	0	0	6.15
1ax1	3.23	0.67	0	0	5.55
1bir	5.25	1.65	1	1	0.58
1bu4	7.39	2.10	1	0	4.85
1csi	4.36	1.14	0	0	4.51
1etb	5.27	1.05	0	0	10.19
1hnl	3.73	0.74	0	0	2.50
1mfa	4.27	1.70	1	1	0.92
1nsd	4.13	1.54	1	1	0.84
1rx9	8.34	2.76	1	1	1.09
1xig	3.29	0.52	0	0	2.41
1xzl	3.44	0.46	0	0	3.37
2izk	3.39	0.84	0	0	3.56
2qwf	4.78	1.50	1	1	0.47
2tps	6.65	2.02	1	1	0.45
3vgc	5.77	1.31	0	0	6.27
4sli	3.02	1.02	0	0	5.09
6cpp	10.62	3.05	1	1	0.61
8rsa	4.31	1.15	0	0	2.98

[a]The data of the first ranked solution in each case are listed in the table.
[b]Number 1 represents a well-docked solution, while 0 represents a pose with RMSD larger than 2 Å.
[c]The cutoff values (4.39 and 7.23) are not used for reselecting, i.e., reselection are performed on all the cases, without considering the predicted pK_d values.

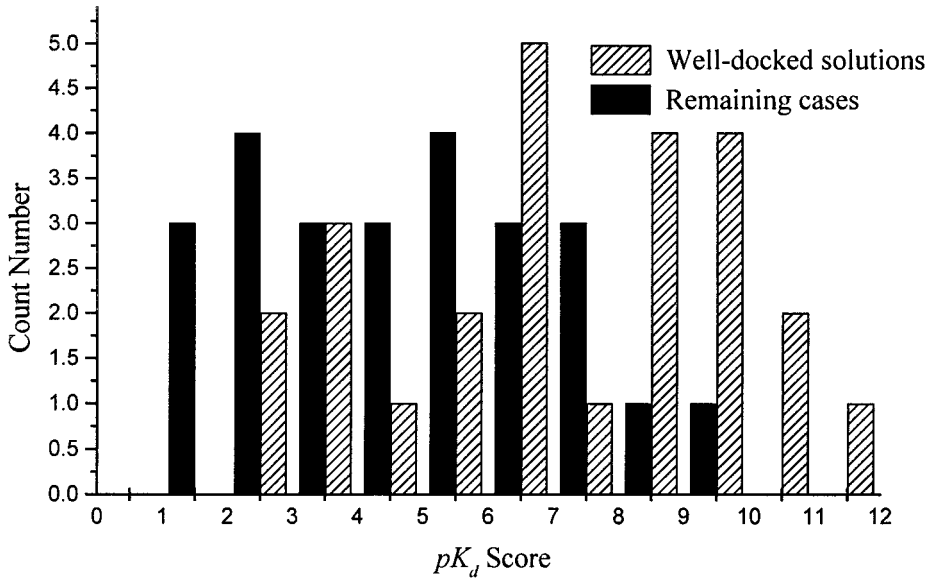

Figure 7. The distributions of experimental pK_d scores for well-docked cases and the remaining cases in 50 complexes based on the running results of ScoreDock.

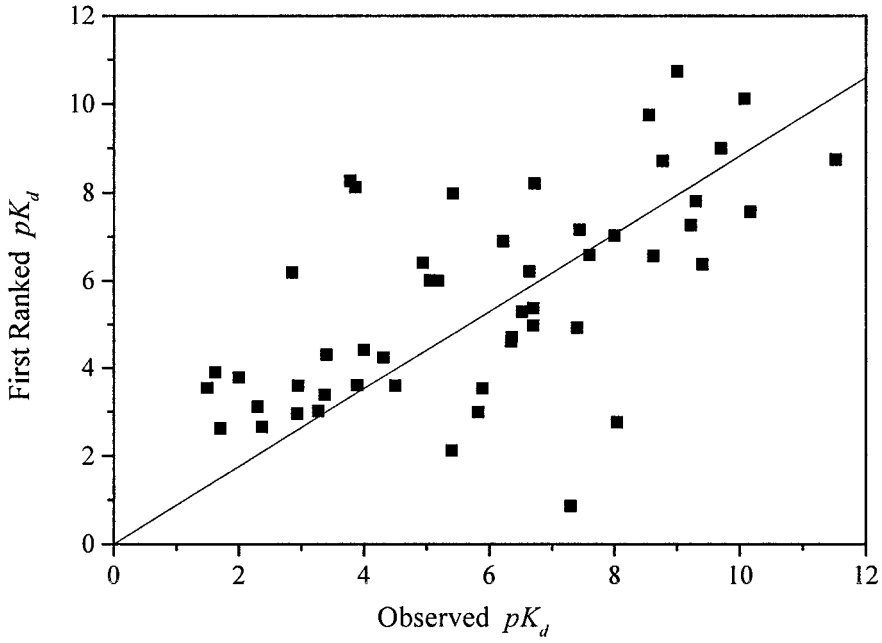

Figure 8. The first ranked pK_d scores versus the experimental pK_d values of 50 complexes $r = 0.64$, $s = 2.01$.

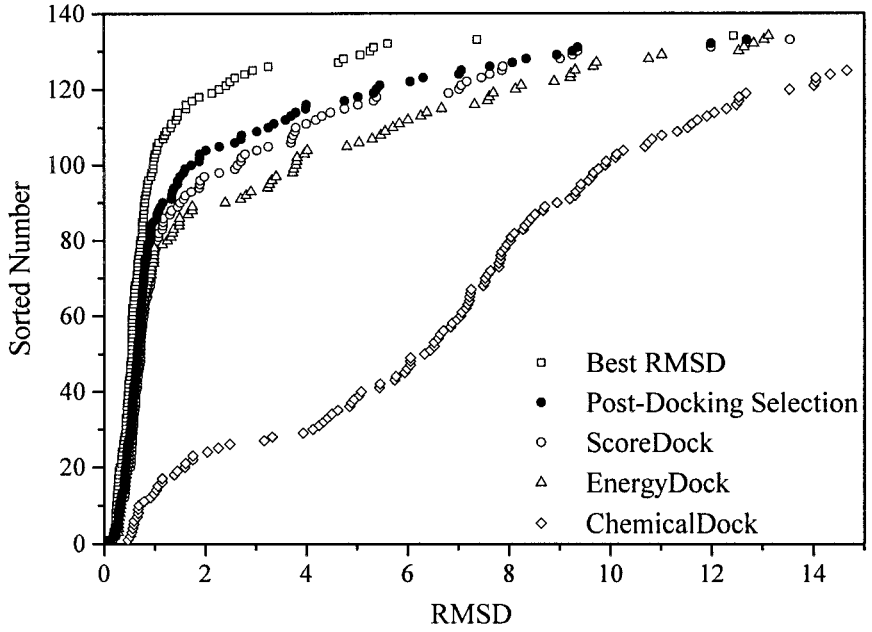

Figure 9. Sorted number versus RMSD value for the first ranked solutions and post-docking selecting results of 137 cases from the origin test set with the first ranked pK_d score higher than 4.39. 'Best RMSD' indicates the smallest RMSD generated by ScoreDock disregarding the rank number.

Table 4. Statistics of docking results by applying the types of scoring functions and post-docking selection on the 137 cases with predicted pK_d higher than 4.39 from the original test set

		The number of complexes with RMSD reference to crystal structure			
		<1.0 Å	<1.5 Å	<2.0 Å	≥2.0 Å
All ranks of ScoreDock[a]		101	114	118	19
First rank[b]	Chemical dock	13	19	23	114
	Energy dock	78	86	89	48
	Score dock	76	89	97	40
Post-docking selection[c]		85	96	103	34

[a]The smallest RMSD generated by ScoreDock is selected for all of 137 complexes.
[b]The first ranked solutions of each complex according to three methods are considered respectively.
[c]The results after performing the post-docking selecting on ScoreDock results using docking quality estimation method.

complexes and ranked them among the ScoreDock generated results. Figure 10 shows the rank number of each complex in the full test set. Indeed, crystal structure score does rank first in 86 complexes. Using a less strict criterion, i.e., those cases, for which the RMSD values of those poses ranked better than crystal structure are all less than 2.0 Å, will still be considered as good discriminatory results, ScoreDock meets this criterion in 125 cases altogether.

Flexible docking test

Although the above test results demonstrate the reliability of ScoreDock for rigid docking, this program still needs to be tested in consideration of the ligand flexibility. After discarding the rigid ligands, for which the rotor number is zero, we randomly selected out 100 complexes[4] from the original test set to perform flexible docking. For the flexible docking performance of ScoreDock, EnergyDock and ChemicalDock, all the boolean parameters in 'Ligand Flexibility' section are set as 'yes', except that 'multiple_anchors' and 'reminimize_anchor' are set as 'no'. All the real and integer parameters use default values, except that 'clash_overlap' is set to 0.3 [57]. In Figure 11, the sorted numbers are plotted versus the corresponding RMSD values. Meanwhile, the solutions with smallest RMSD values according to ScoreDock running results are plotted, too.

After the performance of post-docking selection based on docking quality estimation, the number of well-docked cases increases from 15 to 20 (Figure 11). Since our approach is generated from the rigid docking results, we found this result encouraging. Furthermore we compared the flexible docking results to the application of DrugScore on DOCK [50] in Table 5. The percentage is the test cases found on first rank or post-docking selecting results with RMSD values of < 1.0 Å, < 1.5 Å, < 2.0 Å and ≥ 2.0 Å compared to the best approximating geometry found on any rank. The last column percentages, except the first two lines, are compared to the number of cases with the best RMSD less than 2.0 Å according to the test sets for DrugScore and ScoreDock respectively. With respect to the recognition of well-docked solutions on the first rank, the EnergyDock succeeds in 35%, whereas ScoreDock detects 48%. After the post-dock selecting on the top 5 solutions, the successful percentage is 65%.

[4] 1afl 1aoe 1aqv 1ax0 1azm 1ba8 1bji 1bjv 1bk0 1bul 1bzm 1bzy 1cbx 1cg8 1cil 1ckb 1cwl 1daa 1dae 1dan 1dhj 1dim 1drf 1ela 1fkf 1gai 1get 1ghb 1gup 1hew 1hne 1hsl 1hyt 1icn 1ida 1ir3 1jap 1kvr 1lcp 1lic 1llo 1lna 1meu 1mld 1mrk 1mtv 1nes 1nsc 1phf 1php 1poc 1pph 1ppl 1pso 1rbp 1rdn 1rds 1rnn 1rob 1rpj 1ruv 1scn 1sep 1sha 1snc 1sth 1syb 1tag 1tng 1tnh 1tni 1tnk 1tpb 1tpp 1tyr 1vbs 1vpe 1vwo 1wap 1xid 1xie 1xym 2cmd 2fal 2gbp 2hbe 2mth 2sim 2sns 2tmn 2xis 3bto 3cla 3cpa 4csc 4est 5tmn 6rnt 6rsa 6tmn

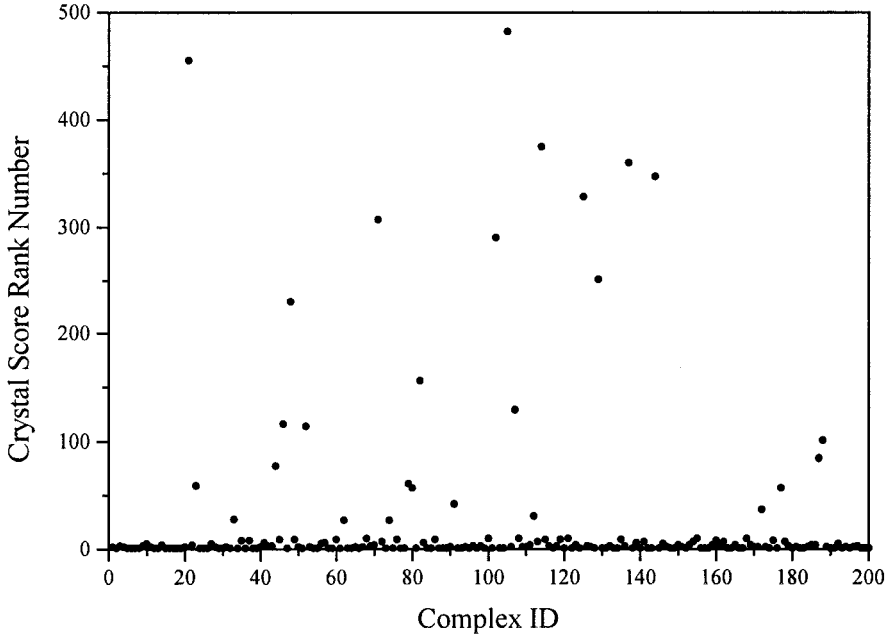

Figure 10. The crystal structure rank number among the ScoreDock generated solutions for whole test set.

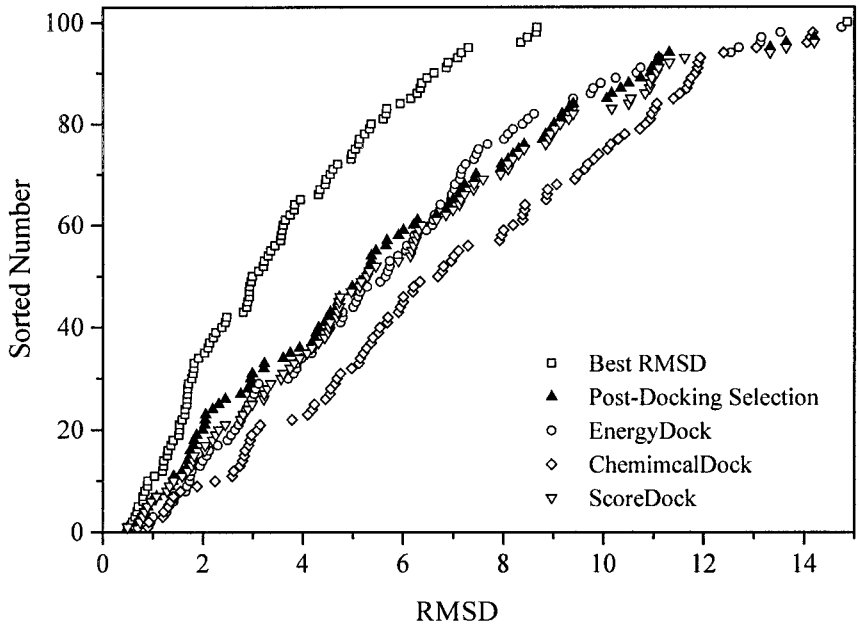

Figure 11. Sorted number versus RMSD value based on the first ranked solutions and post-docking selecting results for the selected 100 cases in consideration of the ligand flexibility. 'Best RMSD' indicates the smallest RMSD generated by ScoreDock disregarding the rank number.

Table 5. Statistics of docking results by applying the types of scoring functions and post-docking selection on the selected 100 complexes in consideration of the ligand flexibility, and comparison with DrugScore[50].

		[a]Percentage of cases number with the reference to the best RMSD cases			
		<1.0 Å	<1.5 Å	<2.0 Å	≥2.0 Å
All ranks of ScoreDock[b]	DrugScore[50]	17	31	43	57
	ScoreDock	10	17	31	[d]69
First rank[c]	DOCK (using chemical score)[50]	18	33	46	54
	ChemicalDock	20	40	29	71
	EnergyDock	20	35	35	65
	DrugScore[50]	41	48	70	30
	ScoreDock	60	59	48	52
Post-Docking Selection		60	65	65	35

[a]The test set used by DrugScore is not same to ours.
[b]The smallest RMSD generated by ScoreDock is selected for all of 100 complexes.
[c]The first ranked solutions of each complex according to the methods are considered respectively.
[d]The percentage in this column except the first two lines are compared to the number of cases with the best RMSD less than 2.0 Å for DrugScore and ScoreDock test sets respectively

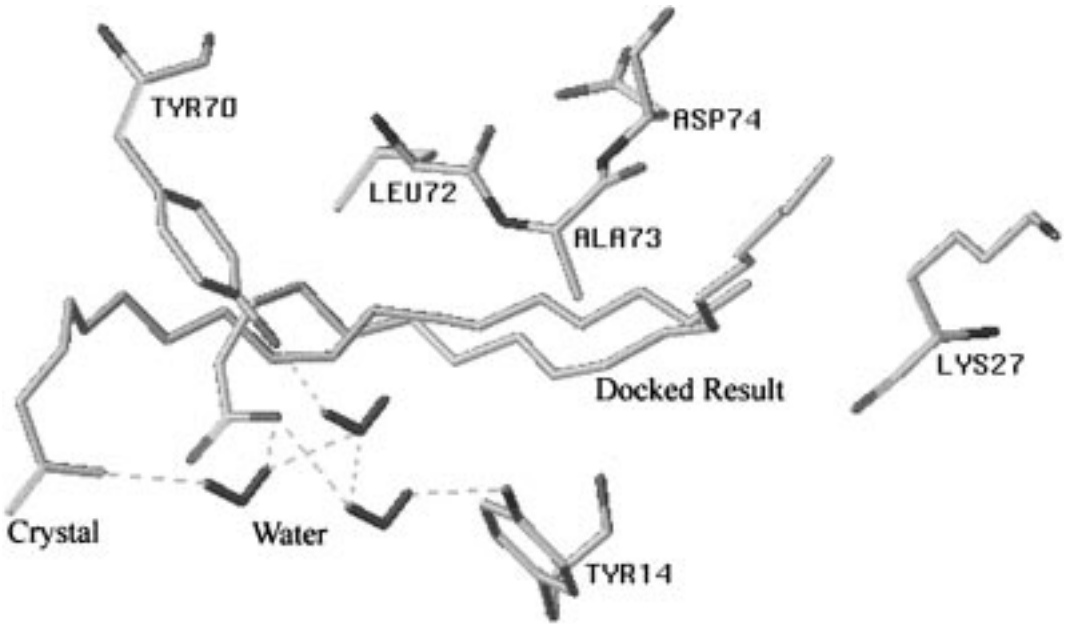

Figure 12. The binding modes of ligand in the crystal structure and the third ranked solution (RMSD 6.30 Å) of 1ICN are displayed together with structure water and related residues. Possible hydrogen bonds among ligands, residues and structure water molecules are depicted as dotted lines.

Discussion

ScoreDock is the combination of DOCK4 searching algorithm and an empirical binding affinities prediction method, SCORE [51]. Recalling that the number of possible orientations generated during the conformation searching stage is very large, including those well-docked solutions and those poses largely deviating from the crystal structures, application in docking process is a real challenge for SCORE method. The final results of rigid docking for 200 complexes and flexible docking for 100 complexes show obvious improvement comparing to the standalone program DOCK4, thus demonstrate the discrimination ability of SCORE method.

The development of SCORE method is based on the analysis of protein ligand complexes crystal structures and their negative logarithms of dissociation equilibrium constants, i.e., pK_d values. The latter information is more attractive in drug design process than the binding energy calculated by force field method, and is more suitable for screening out the final hits of docking process, due to its chemical meaning.

According to the test results, the successful rate of ScoreDock is limited, and even lower when considering the flexibility of ligands. In real docking work, i.e., for those complexes the crystal structures are unknown, to determine the reliability of docking results is rather difficult. Here, we developed a new assessment method to help users to judge the reliability of docking results generated by ScoreDock. The test results of randomly selected complexes demonstrate the robustness of this method. Using the computer-aided prediction method, the estimation of docked solution will be more constant than the direct observation by eye. Tested by the original 200 complexes, randomly selected 20 complexes, additional 20 complexes with known pK_d values, and 100 flexible docking cases, this method shows its reliability on docking quality estimation and robustness to improve the docking results.

Based on the analysis of test results, drug designer will find ScoreDock attractive due to the following features: (i) ScoreDock is more reliable in cases where complex has high pK_d score. (ii) An assessment method is developed to predict the quality of docking results and improve the docking accuracy. (iii) In addition to binding conformation of ligands, ScoreDock can provide detail information about the interaction mode between receptor and ligand [51]. Based on these features, ScoreDock can provide a convenient and reliable start point for drug design circle.

Furthermore ScoreDock can deal with the structure water molecules, i.e., consider the hydrogen bonds contribution involved by water into binding affinity [51]. According to the calculated results of SCORE method, these structure waters can play an important role in binding process. For instance, in the third ranked solution (RMSD 6.30 Å) of 1icn (resolution 1.74 Å), the ligand oleate retreats from the pocket about 5 Å. In this new position, the carboxyl group of oleate can form several hydrogen bonds to three structure water molecules, while the alkyl chain remains the same orientation to the crystal structure (Figure 12). Interestingly, the carboxyl group of oleate is highly disordered in crystal structure. Therefore browsing the less ranked solutions generated by ScoreDock can provide valuable information about binding pocket.

In some docking cases, the crystal structure of unbound receptor is used, in stead of complex structure, which is unknown by far. If using ScoreDock, user can keep all of the structural water in receptor structure for SCORE. On the other hand, because we do not keep any water molecule for the surface calculation, which is necessary in the data preparation, the searching engine can generate all the possible conformation and orientation for the ligand. In SCORE scheme, the appropriate structural water can be considered for hydrogen bonding [51], which may play important role in complex. Furthermore, if no crystal structure is known, the structure generated by other methods can be used for docking purpose. In these cases, we believe that randomly put water molecules around and in binding pocket will help the process of ScoreDock, because SCORE method can deal with these water molecules automatically and appropriately. We need to note that in the original DOCK4.0 manual [57], the authors suggest to delete all of the structure water, unless the user exactly knows which water molecule is essential in binding. Even though, there is one problem remains, that is the space occupied by the reserved water molecule will not be taken by ligand, because the protein surface is calculated based on the given structure in consideration of all atoms. In addition, the scoring functions in DOCK4 can not take structural water into account.

Conclusions

In this paper, we incorporated an empirical protein-ligand complex binding affinity estimation method, SCORE, into a popular docking tool, DOCK4, and tested this scoring method for its ability to discriminate between the near native poses of ligand and those largely deviating from the native structure. The combined docking program shows compatible predictive power with DOCK4 for a test set containing 200 complexes. Furthermore, the new program shows improvement comparing to DOCK4, in consideration of the flexibility of ligands. An empirical docking quality assessment method is developed based on the statistical results of test set to assist the validation of ScoreDock running results. This method is also tested on other selected complexes, and shows its reliability in docking quality prediction. After the reselection based on docking results qualities, the final results show obvious improvements. Meanwhile, the SCORE method shows good ability to recognize the crystal structure out of the computer-generated poses for the full test set.

Acknowledgement

This work was supported by the Science and Technology Ministry of China, the National Natural Science Foundation of China and the Committee of Science and Technology of Beijing. We acknowledge the helpful suggestions of Dr. Renxiao Wang.

References

1. Whittle, P.J. and Blundell, T.L., Annu. Rev. Biophys. Biomol. Struct., 23 (1994) 349.
2. Böhm, H.J., Prog. Biophys. Mol. Biol., 66 (1996) 197.
3. Marrone, T.J., Briggs, J.M. and McCammon, J.A., Annu. Rev. Pharmacol. Toxicol., 37 (1997) 71.
4. Gschwend, D.A., Good, A.C. and Kuntz, I.D., J. Mol. Recogn., 9 (1996) 175.
5. Makino, S. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1812.
6. Ewing, T.J.A. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1175.
7. Roe, D.C. and Kuntz, I.D., J. Comput. Aid. Mol. Des., 9 (1995) 269.
8. Caflisch, A. and Karplus, M., Perspect. Drug Discov. Des., 3 (1995) 51.
9. Pearlman, D.A. and Murcko, M.A., J. Med. Chem., 39 (1996) 1651.
10. DeWitte, R.S. and Shakhnovich, E.I., J. Am. Chem. Soc., 118 (1996) 11733.
11. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., J. Mol. Biol., 161 (1982) 269.
12. Bacon, D.J. and Moult, J., J. Mol. Biol., 225 (1992) 849.
13. Fischer, D., Lin, S.L., Wolfson, H.L. and Nussinov, R., J. Mol. Biol., 248 (1995) 459.
14. Sobolev, V., Wade, R.C., Vriend, G. and Edelman, M., Proteins, 25 (1996) 120.
15. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., J. Mol. Biol., 261 (1996) 470.
16. Rarey, M., Wefing, S. and Lengauer, T., J. Comput. Aid. Mol. Des., 10 (1996) 41.
17. Rarey, M., Kramer, B. and Lengauer, T., J. Comput. Aid. Mol. Des., 11 (1997) 369.
18. Yue, S.Y., Protein Eng., 4 (1990) 177.
19. Hart, T.N. and Read, R.J., Proteins, 13 (1992) 206.
20. Knegtel, R.M.A., Boelens, R. and Kaptein, R., Protein Eng., 7 (1994) 761.
21. Morris, G.M., Goodsell, D.S., Huey, R. and Olson, A.J., J. Comput. Aid. Mol. Des., 10 (1996) 293.
22. Liu, M. and Wang, S., J. Comput. Aid. Mol. Des., 13 (1999) 435.
23. Brodmeier, T. and Pretsch, E., J. Comput. Chem., 15 (1994) 588.
24. Oshiro, C.M., Kuntz, I.D. and Dixon, J.S., J. Comput. Aid. Mol. Des., 9 (1995) 113.
25. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.
26. Ajay and Murcko, M.A., J. Med. Chem., 38 (1995) 4953.
27. Beveridg, D.L. and DiCapua, F.M., Annu. Rev. Biophys. Biophys. Chem., 18 (1989) 431.
28. Reddy, M.R., Varney, M.D., Kalish, V., Viswanadhan, V.N. and Appelt, K., J. Med. Chem., 37 (1994) 1145.
29. Helms, V. and Wade, R.C., J. Am. Chem. Soc., 120 (1998) 2710.
30. Åqvist, J., Medina, C. and Samuelsson, J.E., Protein Eng., 7 (1994) 385.
31. Weiner, S.J. and Kollman, P.A., J. Comput. Chem., 7 (1986) 230.
32. Jorgensen, W.L. and Tirado-Rives, J., J. Am. Chem. Soc., 110 (1988) 1657.
33. Halgren, T.A., J. Am. Chem. Soc., 114 (1992) 7827.
34. Maple, J.R., Hwang, M.-J., Stockfisch, T.P., Dinur, U., Waldman, M., Ewig, C.S. and Hagler, A.T., J. Comput. Chem., 15 (1994) 162.
35. MacKerell, A.D., Jr., Wiórkiewicz-Kuczera J. and Karplus, M., J. Am. Chem. Soc., 117 (1995) 11946.
36. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., F.D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A., J. Am. Chem. Soc., 117 (1995) 5179.
37. Cheatham, T.E.I., Cieplak, P. and Kollman, P.A., J. Biomol. Struct. Dyn., 16 (1999) 845.
38. Meng, E.C., Shoichet, B.K. and Kuntz, I.D., J. Comput. Chem., 13 (1992) 505.
39. Luty, B.A., Wasserman, Z.R., Stouten, P.F.W., Hodge, C.N., Zacharias, M. and McCammon, J.A., J. Comput. Chem., 16 (1995) 454.
40. Viswanadhan, V.N., Reddy, M.R., Wlodawer, A., Varney, M.D. and Weinstein, J.N., J. Med. Chem., 39 (1996) 705.
41. Head, R.D., Smythe, M.L., Oprea, T.I., Waller, C.L., Green, S.M. and Marshall, G.R., J. Am. Chem. Soc., 118 (1996) 3959.
42. Williams, D.H., Searle, M.S., Mackay, J.P., Gerhard, U. and Maplestone, R.A., Proc. Natl. Acad. Sci., 90 (1993) 1172.
43. Verkhivker, G., Appelt, K., Freer, S.T. and Villafranca, J.E., Protein Eng., 8 (1995) 677.

44. Weng, Z., Vajda, S. and Delisi, C., Protein Sci., 5 (1996) 614.
45. Böhm, H.J., J. Comput. Aid. Mol. Des., 12 (1998) 309.
46. Eldridge, M.D., Murray, C.W., Auton, T.R., Paolini, G.V. and Mee, R.P., J. Comput. Aid. Mol. Des., 11 (1997) 425.
47. Murray, C.W., Auton, T.R. and Eldridge, M.D., J. Comput. Aid. Mol. Des., 12 (1998) 503.
48. Mitchell, J.B.O., Laskowski, R.A., Alex, A. and Thornton, J.M., J. Comput. Chem., 20 (1999) 1165.
49. Mitchell, J.B.O., Laskowski, R.A., Alex, A., Forster, M.J. and Thornton, J.M., J. Comput. Chem., 20 (1999) 1177.
50. Gohlke, H., Hendlich, M. and Klebe, G., J. Mol. Biol., 295 (2000) 337.
51. Wang, R., Liu, L., Lai, L. and Tang, Y., J. Mol. Model, 4 (1998) 379.
52. Wang, R., Fu, Y. and Lai, L., J. Chem. Inf. Comput. Sci., 37 (1997) 615.
53. SYBYL 6.2, Tripos Assoc. Inc., St. Louis, MO, 1995.
54. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
55. Tao, P. and Lai, L., unpublished results.
56. WWW URL: http://relibase.ebi.ac.uk/
57. T. Ewing, DOCK Version 4.0 Manual Regents of the University of California, San Francisco, CA., 1997.