PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

An iterative knowledge-based scoring function for protein–protein recognition

Sheng-You Huang[1,2,3] and Xiaoqin Zou[1,2,3]*

[1] Department of Physics and Astronomy, University of Missouri, Columbia, Missouri 65211

[2] Department of Biochemistry, University of Missouri, Columbia, Missouri 65211

[3] Dalton Cardiovascular Research Center, University of Missouri, Columbia, Missouri, 65211

ABSTRACT

Using an efficient iterative method, we have developed a distance-dependent knowledge-based scoring function to predict protein–protein interactions. The function, referred to as ITScore-PP, was derived using the crystal structures of a training set of 851 protein–protein dimeric complexes containing true biological interfaces. The key idea of the iterative method for deriving ITScore-PP is to improve the interatomic pair potentials by iteration, until the pair potentials can distinguish true binding modes from decoy modes for the protein–protein complexes in the training set. The iterative method circumvents the challenging reference state problem in deriving knowledge-based potentials. The derived scoring function was used to evaluate the ligand orientations generated by ZDOCK 2.1 and the native ligand structures on a diverse set of 91 protein–protein complexes. For the bound test cases, ITScore-PP yielded a success rate of 98.9% if the top 10 ranked orientations were considered. For the more realistic unbound test cases, the corresponding success rate was 40.7%. Furthermore, for faster orientational sampling purpose, several residue-level knowledge-based scoring functions were also derived following the similar iterative procedure. Among them, the scoring function that uses the side-chain center of mass (SCM) to represent a residue, referred to as ITScore-PP(SCM), showed the best performance and yielded success rates of 71.4% and 30.8% for the bound and unbound cases, respectively, when the top 10 orientations were considered. ITScore-PP was further tested using two other published protein–protein docking decoy sets, the ZDOCK decoy set and the RosettaDock decoy set. In addition to binding mode prediction, the binding scores predicted by ITScore-PP also correlated well with the experimentally determined binding affinities, yielding a correlation coefficient of $R = 0.71$ on a test set of 74 protein–protein complexes with known affinities. ITScore-PP is computationally efficient. The average run time for ITScore-PP was about 0.03 second per orientation (including optimization) on a personal computer with 3.2 GHz Pentium IV CPU and 3.0 GB RAM. The computational speed of ITScore-PP(SCM) is about an order of magnitude faster than that of ITScore-PP. ITScore-PP and/or ITScore-PP(SCM) can be combined with efficient protein docking software to study protein–protein recognition.

Proteins 2008; 72:557–579.

Key words: Protein–protein interaction; scoring function; knowledge-based potentials; protein–protein docking; molecular recognition.

INTRODUCTION

Interactions between proteins are crucial for the existence of life, given the importance of protein–protein recognition in many biological processes such as signal transduction, cell regulation, and other macromolecular assemblies.[1–5] With the rapid development of the proteomics project, a large number of protein structures are being determined by various experimental techniques and deposited in the Protein Data Bank (PDB).[6] Moreover, the vast and yet growing amount of sequences can also be used for *in silico* prediction of many additional protein structures based on known structural data. However, compared with known single protein structures, the number of known protein–protein complex structures is very limited, because structural determination on protein–protein complexes is much more challenging than structural determination on their isolated components. Therefore, given the importance of protein–protein recognition and the small number of determined protein–protein complexes, there is a pressing need to develop computational methods for predicting protein–protein complex structures.

Grant sponsor: NIH; Grant number: DK61529; Grant sponsor: CFF; Grant number: ZOU07I0; Grant sponsor: The University of Missouri Research Board; Grant number: RB-07-32; Grant sponsors: Federal Earmark NASA Funds for Bioinformatics Consortium Equipment, Dell, SGI, Sun Microsystems, TimeLogic, and Intel.
*Correspondence to: X. Zou, Dalton Cardiovascular Research Center, University of Missouri, Columbia, MO 65211. E-mail: zoux@missouri.edu
Received 6 August 2007; Revised 30 November 2007; Accepted 4 December 2007
Published online 4 February 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21949

For years, various protein–protein docking algorithms have been developed to predict the binding modes of protein–protein complexes.[1–5] There are two challenging aspects in protein–protein docking: searching for possible docking configurations and assessing each configuration with a scoring function. According to their search methods, current protein docking algorithms can be classified into three broad categories: global search, local shape matching, and randomized search.[4] The first type of search algorithms perform a global search of all possible binding configurations, either in the real space (e.g., SOFTDOCK and BiGGER[7,8]) or in the transformation space using Fast Fourier Transform (FFT)[9] (e.g., 3D-DOCK, GRAMM, DOT, ZDOCK, and MolFit[10–16]). The second type of search algorithms are based on matching local shape features, such as the distance geometry algorithm,[17–20] and Geometric Hashing algorithms.[21–26] The last type, randomized search, uses genetic algorithm (e.g., GAPDOCK and AutoDock[27,28]) or Monte Carlo methods (e.g., AutoDock, ICM, Rosetta, and ATTRACT[27,29–31]).

The second aspect in protein–protein docking, scoring, is equally challenging. The success of protein–protein docking relies on the development of a rapid and accurate energy scoring function, which is used to evaluate the energies of protein–protein docking poses so as to identify the pose with the lowest energy as the predicted binding mode.[32] The scoring function is required to be precise so as to discriminate the correct binding modes from decoy complex structures. The function is also required to be computationally efficient so as to evaluate the huge number of possible solutions for a protein–protein complex.[3] Despite important advances, this goal has not yet been achieved satisfactorily. Currently, there are several different scoring methods for predicting protein–protein interactions. The first type of scoring methods use physical force field parameters such as those provided in DOCK.[19,33–35] The second type of methods are based on shape complementarity between proteins plus some additional descriptors such as desolvation and electrostatics terms.[25,36,37] Because of is simplicity and fast speed, shape complementarity-based scoring methods are commonly used in the initial filtering stage of many docking programs such as ZDOCK and 3D-DOCK.[10,15] The third type of methods are empirical scoring functions, which are usually a linear combination of energy-related quantities such as conformational entropy, hydrophobic and hydrophilic surface areas, number of hydrogen bonds, etc., and are obtained by fitting the experimental binding data.[38–40]

The fourth type of scoring methods for protein–protein docking are referred to as knowledge-based potentials,[41–49] which are derived from known structures of protein–protein complexes in a database such as PDB, etc.[6,50–52] The interaction potentials are directly converted from the statistical occurrences observed in the complex database by using an inverse Boltzmann relation:

$$u_{ij}(r) = -k_B T \ln[g_{ij}(r)], \quad g_{ij}(r) = \rho_{ij}(r)/\rho^*_{ij}(r) \qquad (1)$$

where k_B denotes the Boltzmann constant, T stands for the system "temperature", i and j stand for the atoms of type i and type j from different partners in a complex, and $g_{ij}(r)$ is the pair distribution function. $\rho_{ij}(r)$ and $\rho^*_{ij}(r)$ are the number densities of the atom ij pairs at distance r in the experimental structures and in the reference state, respectively. Here, the reference state stands for a state where the interatomic interactions are zero. $u_{ij}(r)$ is the true potentials for atom type pair ij extracted from structural information. The resulted knowledge-based energy score is the sum of all interatomic interaction energies:

$$\text{energy score} = \sum_{\text{R}-\text{L atom pair}} u_{ij}(r) \qquad (2)$$

where R and L stand for the receptor protein and ligand protein, respectively.

Despite its simplicity, the knowledge-based scoring function is perceived to be general because the potentials are derived from a large number of diverse known complex structures rather than fitting limited known affinities. The pairwise feature of the potentials also allows the scoring function to be rapidly assessed. Therefore, in addition to being used during the docking process,[53] the knowledge-based potentials are also extensively used in the postdocking stage of docking programs (e.g., 3D-DOCK, BiGGER) for rapid pose refinement.[8,10,54]

Despite significant achievements, one hurdle in deriving knowledge-based scoring functions is the determination of the "reference state" $\rho^*_{ij}(r)$ in Eq. (1). This is because the reference state, in which the interactions between any atoms are zero, is not achievable for complicated systems such as proteins.[55,56] Several intelligent approximations have been introduced to calculate the number of atomic pair density in the reference state $[\rho^*_{ij}(r)]$.[57–59] Yet, the reference state remains a challenge for the accuracy of knowledge-based scoring functions. In our recent work,[60,61] the reference state problem was circumvented by using a novel iterative method[62,63] in the derivation of the knowledge-based potentials for protein–ligand interactions. The idea of the method is to adjust pair potentials by iteration until the scoring function distinguishes native binding modes from decoy ligand poses. Extensive tests demonstrated that our scoring function (ITScore) derived from this iterative method yielded very good performances on predictions of ligand binding modes and affinities and on virtual screening of compound databases.[60,61] To extend the iterative approach from protein–ligand to protein–protein structure prediction, in this work we have derived an efficient

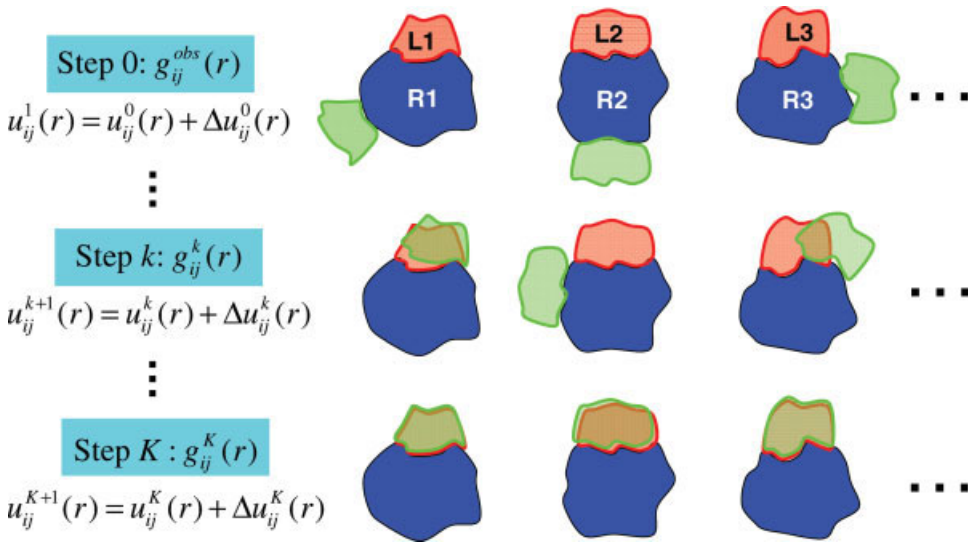

Figure 1

A cartoon illustration of the iterative method for deriving ITScore-PP. Here, only three complexes (R1-L1, R2-L2, and R3-L3) in the training set are shown for illustration purpose. The native receptor and ligand protein structures are colored in blue and red, respectively. The ligand orientation that corresponds to the lowest energy for each receptor (i.e., the predicted binding mode) calculated with the current trial potentials $\{u_{ij}^{k}(r)\}$ is colored in green. The iterative step k = 0, 1, 2, . . ., K. At the beginning (k = 0), the native and predicted structures are far different. At the end of the iteration (k = K), all native structures in the training set are correctly predicted.

energy scoring function for protein–protein interactions using the similar iterative procedure.[60] An advantage of protein–protein systems over protein–ligand systems is that the atom types for proteins are simpler than those for ligands. Therefore, we can derive pair potentials for atom types that are specific for proteins, in order to obtain a more accurate scoring function specific for protein–protein docking studies.

We have tested the new scoring function, referred to as ITScore-PP, on diverse sets of protein–protein complexes. Both bound and unbound docking were performed. Bound docking refers to redocking one of the protein partners into its cocrystalized protein structure. Unbound docking, which is more realistic and yet more challenging, reconstructs the complex structure using individual unbound conformations (i.e., conformations in free solution or extracted from different complexes). We have examined our predictions on both binding modes and affinities in these docking studies.

MATERIALS AND METHODS

The iterative method

As aforementioned, the basic idea of our method to circumvent the unachievable reference state is to iteratively improve the interatomic pair potentials $\{u_{ij}(r)\}$ by comparing the experimentally observed structures and the predicted binding modes, using a training set of known protein–protein complex structures. The idea can be represented by the following iterative equation:

$$u_{ij}^{k+1}(r) = u_{ij}^{k}(r) + \Delta u_{ij}^{k}(r),$$
$$\Delta u_{ij}^{k}(r) = \frac{1}{2} k_B T [g_{ij}^{k}(r) - g_{ij}^{\mathrm{obs}}(r)] \quad (3)$$

Here, k stands for the iterative step. i and j represent the types of a pair of atoms in the receptor protein and ligand protein. $g_{ij}^{\mathrm{obs}}(r)$ is the pair distribution function for atom pair ij calculated from the ensemble of experimentally measured protein–protein complex structures given in the training set. $g_{ij}^{k}(r)$ is the pair distribution function calculated from the ensemble of the binding modes of these protein–protein complexes predicted with the trial potentials $\{u_{ij}^{k}(r)\}$ at the k-th step. $\{u_{ij}^{k+1}(r)\}$ are the improved potentials from $\{u_{ij}^{k}(r)\}$. The formulas for calculating the pair distribution functions $g_{ij}^{\mathrm{obs}}(r)$ and $g_{ij}^{k}(r)$ are described later. Without loss of generality, $k_B T$ was set to unit 1 in the present study.

Therefore, given a set of initial potentials $\{u_{ij}^{0}(r)\}$, the improved potentials $\{u_{ij}^{1}(r)\}$ can be calculated from $\{u_{ij}^{0}(r)\}$ using Eq. (3), and then $\{u_{ij}^{2}(r)\}$ from $\{u_{ij}^{1}(r)\}$, $\{u_{ij}^{3}(r)\}$ from $\{u_{ij}^{2}(r)\}$, ... The iterative process will continue until all the native structures are discriminated from decoys by the current trial potentials $\{u_{ij}^{k}(r)\}$, which are set as the final pair potentials for the knowledge-based scoring function.

An illustration of our iterative method is plotted in Figure 1. The details of the method are described as follows.

- At the beginning of the iteration (i.e., $k = 0$), the experimentally observed pair distribution functions were calculated for the ensemble of the native complex structures in the training set as

$$g_{ij}^{\text{obs}}(r) = \rho_{ij}^{\text{obs}}(r)/\rho_{ij,\text{bulk}}^{\text{obs}} \tag{4}$$

where $\rho_{ij}^{\text{obs}}(r)$ and $\rho_{ij,\text{bulk}}^{\text{obs}}$ are the number densities of atom pair ij occurring in a spherical shell of radius from $r - \Delta r/2$ to $r + \Delta r/2$ and in a reference sphere with a radius of $R_{\max}$. Here, the bin size Δr was set to 0.2 Å, and the radius of the reference sphere, $R_{\max}$, was set to 15 Å. The atomic pair densities $\rho_{ij}^{\text{obs}}(r)$ and $\rho_{ij,\text{bulk}}^{\text{obs}}$ were calculated as

$$\rho_{ij}^{\text{obs}}(r) = \frac{1}{M}\sum_{m}^{M}\frac{n_{ij}^{m}(r)}{4\pi r^2\Delta r} \quad \text{and} \quad \rho_{ij,\text{bulk}}^{\text{obs}} = \frac{1}{M}\sum_{m}^{M}\frac{N_{ij}^{m}}{V(R_{\max})} \tag{5}$$

where M is the number of protein–protein complexes in the training database, $n_{ij}^{m}(r)$ is the number of atom pair ij in the spherical shell for the m-th native complex structure, $V(R_{\max})$ is the volume of the reference sphere and equals $4\pi R_{\max}^3/3$, and N_{ij}^{m} is the total number of atom pair ij in the reference sphere for the m-th native complex structure and equals to $\sum_{r=0}^{r=R_{\max}} n_{ij}^{m}(r)$.
- To start the iteration, initial pair potentials $\{u_{ij}^{0}(r)\}$ were required. Theoretically speaking, $\{u_{ij}^{0}(r)\}$ may be arbitrarily set. However, because of the rugged landscape in the multidimensional parameter space, a good set of initial potentials can help fast convergence of the iterative process.
  Following the strategy used in our previous study on ligand–protein interactions,[60] $u_{ij}^{0}(r)$ was set as the combination of a van der Waals (VDW) potential and a potential of mean force, $w_{ij}(r)$. In the present study, $w_{ij}(r)$ was calculated as

$$w_{ij}(r) \equiv -k_{\text{B}}T\ln[g_{ij}^{\text{obs}}(r)/f_{ij}^{\text{obs}}(r)] \tag{6}$$

where $f_{ij}^{\text{obs}}(r)$ is an excluded volume correction factor similar to that in Refs. 57 and 58 (see Appendix A). The inclusion of the correction factor $f_{ij}^{\text{obs}}(r)$ makes the potential of mean force calculation more reasonable,[57,58] and thus makes the iterative procedure more efficient.
  Then, using the method in our previous work,[60] the initial potential was calculated according to the following combination of the potential of mean force and the VDW potential:

$$u_{ij}^{0}(r) = \begin{cases} w_{ij}(r) & \text{for hydrogen-bond pairs} \\ \frac{v_{ij}(r)e^{-v_{ij}(r)} + w_{ij}(r)e^{-w_{ij}(r)}}{e^{-v_{ij}(r)} + e^{-w_{ij}(r)}} & \text{otherwise} \end{cases} \tag{7}$$

where $v_{ij}(r)$ is the Lennard-Jones 6–12 potentials with the VDW radii taken from the AMBER force field.[33,34] The well depths of $v_{ij}(r)$ were set to six times of the corresponding AMBER values, so that $v_{ij}(r)$ and $w_{ij}(r)$ yielded comparable potential minima for typical hydrophobic interactions.[60]
  Notice that atom pair occurrences in the training set are zero for short r, because of VDW repulsions. To avoid logarithm divergence in the potential of mean force calculations and to avoid waste of computational time on highly repulsive VDW potentials at short distances, the maximum (repulsive) potential for $u_{ij}^{0}(r)$ was set to 100 kcal/mol in the present work unless otherwise specified.
- Using the above initial pair potentials $\{u_{ij}^{0}(r)\}$, the binding mode for each protein–protein complex in the training set was predicted following the procedures described later. First, 1000 putative ligand protein poses were generated around the receptor protein, using the protein–protein docking program ZDOCK 2.1, a program that uses shape complementarity only to evaluate different ligand poses.[14] The default parameters except the number of output predictions (set to 1000, here) in ZDOCK 2.1 were used in the present work. The ligand orientations were further minimized by using the SIMPLEX optimization algorithm to remove atomic clashes. The scoring function for minimization is a pure van der Waals energy term using the AMBER force field parameters.[33,34,60] This process on generating putative ligand poses is a one-time calculation. The optimized 1000 poses plus the native structure for each complex were used for the whole iterative procedure. Next, the energy score for each ligand pose was calculated using the current potentials $\{u_{ij}^{0}(r)\}$:

$$\text{energy score} = \sum_{\text{R-L atom pair}}^{r<r_{\text{cut}}} u_{ij}^{0}(r) \tag{8}$$

where R and L stand for the receptor protein and ligand protein in a putative complex, respectively. Here, the cutoff distance r_{cut} for the potentials was set to 10 Å. Namely, the potentials were set to zero for $r > 10$ Å. The introduction of the cutoff[57,60] was to remove the fluctuations of the initial potentials at large distances, which inherently result from the statistical uncertainties. The ligand pose with the lowest energy score was predicted as the binding mode. Usually, at the beginning, the predicted binding modes were far from the native (i.e., true) binding modes (Fig. 1) because the initial potentials were inaccurate and needed to be improved by iteration.
- In the first iteration, the pair potentials were improved according to Eq. (3). Specifically, to calculate $\Delta u_{ij}^{0}(r)$,

the predicted pair distribution function $g_{ij}^0(r)$ was first calculated using the following formula:

$$g_{ij}^0(r) = \rho_{ij}^0(r)/\rho_{ij,\mathrm{bulk}}^0 \tag{9}$$

where $\rho_{ij}^0(r)$ and $\rho_{ij,\mathrm{bulk}}^0$ are the predicted number densities of atom pair ij in the spherical shell (with radius r and thickness of Δr) and the reference sphere (with radius $R_{\max}$) at the first iterative cycle, respectively. $\rho_{ij}^0(r)$ and $\rho_{ij,\mathrm{bulk}}^0$ were calculated as the Boltzmann-weighted pair frequencies over different decoy structures:

$$\rho_{ij}^0(r) = \frac{1}{M \cdot L}\sum_m^M \sum_l^L \frac{n_{ij}^{ml}(r)e^{-\beta U_{ml}}}{4\pi r^2 \Delta r} \quad \text{and}$$
$$\rho_{ij,\mathrm{bulk}}^0 = \frac{1}{M \cdot L}\sum_m^M \sum_l^L \frac{N_{ij}^{ml}e^{-\beta U_{ml}}}{V(R_{\max})} \tag{10}$$

where $\beta = 1/k_B T$ and was set to 1 as aforementioned. $n_{ij}^{ml}(r)$ and N_{ij}^{ml} are the numbers of atom pair ij in the spherical shell and the reference sphere for the l-th decoy ligand pose of the m-th complex, respectively. U_{ml} is the energy score of this ligand pose calculated from Eq. (8). L is the total number of putative ligand poses for each complex (including the native binding mode), and equals 1001 in the present study.

Then, with the calculated $g_{ij}^{\mathrm{obs}}(r)$ and $g_{ij}^0(r)$, the potential correction $\Delta u_{ij}^0(r)$ was calculated using Eq. (3), and the new potential $u_{ij}^1(r)$ for the next iteration was calculated as

$$u_{ij}^1(r) = u_{ij}^0(r) + \Delta u_{ij}^0(r) \tag{11}$$

- Using similar iterative procedure, improved potentials $\{u_{ij}^2(r)\}$ were derived from the previous $\{u_{ij}^1(r)\}$, and then $\{u_{ij}^3(r)\}$ from $\{u_{ij}^2(r)\}$, ... Step by step, the iterative potentials $\{u_{ij}^k(r)\}$ became more and more accurate and were able to identify more and more native binding modes in the training set, as illustrated in Figure 1. The iterative procedure will continue until the following convergence criterion was satisfied:

$$\eta \equiv \frac{1}{M}\sum_m^{\mathrm{rmsd}_m<2.5} 1 \geq \eta_0 \tag{12}$$

where rmsd_m is the root mean square deviation between the best-scored ligand pose and the native binding mode for the m-th complex. The convergence parameter η represents the success rate of identifying native-like binding modes, defined as rmsd $<$ 2.5 Å. In this study, the convergence threshold η_0 was set to 100%.
- The final pair potentials were treated with the smoothing algorithm proposed in Ref. 64 to account for inaccuracies in experimental atomic coordinates (also see Ref. 60). Specifically, the potential at the distance of the i-th bin was set to the weighted average of 1 : 2 : 4 : 2 : 1 of the potentials from bins $(i - 2)$ to $(i + 2)$.

Our iterative method is efficient, and normally converges within 60 iterative steps.

Preparation of the training database

In the present study, a nonreductant training set of known protein–protein complex structures with true biological interfaces were extracted from the PDB[6] to derive our scoring function. Specifically, the training set contained only crystal dimeric structures at resolutions of 2.5 Å or better. Additional selection criteria included: each chain in the dimeric structure should have at least 10 amino acids. The number of interacting residue pairs should be more than 30, where interacting residues are defined as a pair of residues from different chains that have at least one pair of heavy atoms within 4.5 Å. The protein–protein interface should contain no special interacting residues other than the standard 20 amino acids. To avoid redundancy, if two complexes have >70% sequence identity in the receptor–receptor pair and ligand–ligand pair, only the one with better resolution should be kept. Here, the homology scores were calculated using the FASTA2 program ALIGN by Myers and Miller.[65] Finally, only the dimeric structures with true biological interfaces were kept. To do this, we manually checked the header lines of the remaining pdb files and removed those nonbiological dimeric structures that are artifacts of crystallization. The filtered set contained 851 protein–protein complexes, including 655 homodimers and 196 heterodimers. A complex is regarded as a homodimer if its two chains have >70% sequence similarity; otherwise, the complex is regarded as a heterodimer. The PDB codes of the complex structures in the training set are listed in Appendix B.

Test sets for ITScore-PP

Benchmarks by Weng and coworkers

The first test set for binding mode prediction was a combination of the protein–protein complexes from the three benchmarks (Benchmark 0.0, 1.0, and 2.0) prepared by Weng and coworkers.[66,67] If the same PDB entry appears in two or more benchmarks, only the one in the later benchmark was kept. Thus, a total of 122 protein–protein complexes remained, in which at least one of the two partners in the complex has an unbound structure. These 122 complexes can be grouped into three categories: rigid-body, medium difficulty, and difficult cases.[67] Following Ref. 39, only the rigid and medium difficulty cases were kept, because the difficult cases involve very large conformational changes and are therefore more

suitable for protein flexibility studies than for scoring tests. Lastly, the complexes that contain nonstandard residues at their interface or overlap with the complexes in our training set were removed. This yielded a set of 91 protein–protein complexes, which were used to test ITScore-PP for binding mode prediction (see Table III in the Results section).

It should be noted that this test set shares nine homologous complexes with the training set. The homologous complexes in the test set are 1AVX, 1GRN, 1KXP, 1MEL, 1SPB, 2PTC, 2SNI, 2TEC, 7CEI, which have >60% sequence similarity in both the receptor–receptor pair and ligand–ligand pair with 1AVW, 2NGR, 1MA9, 1RJC, 1SCJ, 1BZX, 1LW6, 1TEC, 1UJZ in the training set, respectively. However, these nine homologous complexes were not excluded from the test set in the present study. The reason is that compared to the large and diverse training set of 851 complexes, the above nine homologous complexes are a very small portion and are thus not expected to significantly change the statistics of the training set data for our potential derivation. Therefore, our test set is expected to be unbiased. The view has been verified by the fact that the success rates on binding mode prediction are very similar for the test set with or without the nine homologous complexes, as shown in Section on Test on Benchmarks from Weng and Coworkers.

ZDOCK decoy set

The second test set for binding mode prediction is the ZDOCK decoy set.[14,15] The decoys were generated from the Benchmark 0.0 of 49 protein–protein complexes by Weng and coworkers, using unbound docking via the ZDOCK software. Each protein–protein complex contains 2000 decoys. In the present study, two ZDOCK decoy sets were used: one set was generated through ZDOCK 2.1 and the other through ZDOCK 2.3. The difference between the two ZDOCK versions is their scoring functions. ZDOCK 2.1 uses a pairwise shape-complementarity scoring function, whereas ZDOCK 2.3 also includes the desolvation and electrostatic energies in addition to the pairwise shape complementarity. The ZDOCK 2.1 decoy set and ZDOCK 2.3 decoy set can be downloaded from the ZDOCK home page (http://zlab.bu.edu/zdock/).

RosettaDock decoy set

The third test set used in the present study for binding mode prediction is the RosettaDock unbound perturbation decoy set.[30] The decoy set was generated through the RosettaDock software suite for 54 protein–protein complexes in the Benchmark 1.0 prepared by Weng and coworkers.[66] Specifically, for each complex, the unbound protein partners were superimposed on the bound complex. Then, 1000 ligand decoys were generated using the RosettaDock protocol for perturbed native structures, followed by simultaneous optimization of backbone displacement and side-chain conformations using Monte Carlo minimization. The scoring function in RosettaDock is a linear combination of 11 weighted energy terms, including an attractive van der Waals score, a repulsive van der Waals score, an implicit solvation score, a surface area-based solvation term, a hydrogen bonding score, a rotamer probability term, a residue–residue pair probability term, etc. The performance of RosettaDock on the decoy set is available from the Internet (http://graylab.jhu.edu/docking/decoys/), from which the success rates of binding mode prediction can be calculated for comparison.

Zhang et al.'s test set

In addition to the test sets for binding mode prediction, a test set was also used for binding affinity prediction of ITScore-PP. The test set was taken from the protein–protein complexes prepared by Zhang *et al.*,[59] which includes 84 protein–protein complexes with known binding affinities and experimentally determined structures. The complexes in the test set that overlap with the complexes in our training set were removed. The remaining 74 protein–protein complexes were used for binding affinity prediction.

RESULTS AND DISCUSSION

Extracted atom-level, distance-dependant pair potentials for ITScore-PP

As aforementioned, the pair potentials were extracted from a training set of 851 nonredundant protein–protein dimeric structures collected from the Protein Data Bank, as described in Materials and Methods. The ligand decoys for each complex in the training set were generated with the protein docking program ZDOCK 2.1 by Chen and Weng.[14] In other words, for every complex listed in Appendix B, a total of 1001 ligand protein poses (1000 decoys generated from ZDOCK plus one native binding mode) were used to derive the pair potentials following the iterative procedure described in Materials and Methods. According to our previous study on protein–ligand interactions, the derived potentials with the iterative strategy are not sensitive to different decoy sets of the training set.[60] Because of the similarity between protein–ligand and protein–protein interactions, the same conclusion is expected to be held for the present study.

The atom types for the pair potentials were defined by classifying the atoms in the standard 20 amino acid residues according to their environment in the proteins. Comparing with the atom types used in our previous work for protein–ligand interactions,[60] the atom types proposed for this work are more specific to proteins. A total of 20 atom types were included in the present scoring function, which are listed in Table I.

Table I
List of the 20 Atom Types that are Used in ITScore-PP for the Heavy Atoms in the 20 Standard Amino Acid Residues

No.	Symbol	Description	Atom name
1	C2+	Carbons bonded to a positively charged nitrogen	ARG_CZ
2	C2−	Carbons bonded to a negatively charged nitrogen	*_C (on terminal residues), ASP_CG, GLU_CD
3	C2M	Carbonyl carbons on the mainchain	*_C (not on terminal residues)
4	C2S	Carbonyl carbons on sidechains	ASN_CG, GLN_CD
5	Car	Aromatic carbons	HIS_CD2, HIS_CE1, HIS_CG, PHE_CD1, PHE_CD2, PHE_CE1, PHE_CE2, PHE_CG, PHE_CZ, TRP_CD1, TRP_CD2, TRP_CE2, TRP_CE3, TRP_CG, TRP_CH2, TRP_CZ2, TRP_CZ3, TYR_CD1, TYR_CD2, TYR_CE1, TYR_CE2, TYR_CG, TYR_CZ
6	C3C	Aliphatic carbons bonded to carbons or hydrogens only	ALA_CB, ARG_CB, ARG_CG, ASN_CB, ASP_CB, GLN_CB, GLN_CG, GLU_CB, GLU_CG, HIS_CB, ILE_CB, ILE_CD1, ILE_CG1, ILE_CG2, LEU_CB, LEU_CD1, LEU_CD2, LEU_CG, LYS_CB, LYS_CD, LYS_CG, MET_CB, PHE_CB, PRO_CB, PRO_CG, SER_CB, THR_CG2, TRP_CB, TYR_CB, VAL_CB, VAL_CG1, VAL_CG2
7	C3A	C_α atoms	*_CA
8	C3X	Aliphatic carbons on sidechains bonded to a polar atom	ARG_CD, CYS_CB, LYS_CE, MET_CE, MET_CG, PRO_CD, THR_CB
9	N2N	Amide nitrogens with one hydrogen	*_N (not on terminal residues)
10	N2X	Amide nitrogens with two hydrogens	*_N (on terminal residues), ASN_ND2, GLN_NE2
11	Nar	Aromatic nitrogens	HIS_ND1, HIS_NE2, TRP_NE1
12	N2+	Guanidine nitrogens with two hydrogens	ARG_NH1, ARG_NH2
13	N21	Guanidine nitrogen with one hydrogen	ARG_NE
14	N3+	Nitrogen with three hydrogens	LYS_NZ
15	O2M	Oxygens in carbonyl groups of the mainchain	*_O (not on terminal residues)
16	O2S	Oxygens in carbonyl groups of sidechains	ASN_OD1, GLN_OE1
17	O3H	Oxygens in hydroxyl groups	SER_OG, THR_OG1, TYR_OH
18	O2−	Oxygens in carboxyl groups	*_O or *_OXT (on terminal residues), ASP_OD1, ASP_OD2, GLU_OE1, GLU_OE2
19	S31	Sulfur with one hydrogen	CYS_SG
20	S30	Sulfur without hydrogens	MET_SD

"*" stands for any residue.

The 20 atom types result in a total of 210 possible atomic pair potentials, because the potential for atom pair *ij* cannot be distinguished from the potential for *ji* for protein–protein binding partners. The huge number of atom pair occurrences for most atom pairs in the large training set warrants sufficient statistics to derive the pair potentials for these atom pairs. For example, the number of occurrences is 6,129,200 for a typical atom pair, C3C-C3C. To ensure good statistics, only the interaction potentials of the atom type pairs with more than 2000 occurrences were retained, yielding 204 pairs of effective interaction potentials.

To show the effectiveness of our iterative method on deriving the pair potentials, Figure 2 plots the convergence parameter η, defined in Eq. (12), as a function of the iterative step. A second function plotted in the figure is the average RMSD between the predicted ligand modes and the native modes. It can be seen that the convergence parameter (or success rate) rapidly approached 100% as the iteration went on, indicating that all the native binding modes of the protein–protein complexes in the training set were found at the end of the iteration. The fast convergence of the iterative procedure is a sign of the efficacy of our method in extracting effective interaction potentials.

Figure 3 shows three representative interaction potentials of ITScore-PP. Several notable properties can be seen from the figure, which are consistent with experi-

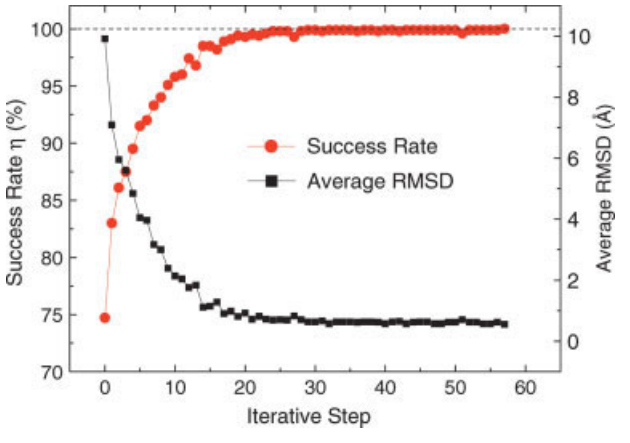

Figure 2
The success rate η (circles) and the average RMSD (squares) of the predicted binding modes as a function of the iterative step. The dashed line stands for the success rate of 100%. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

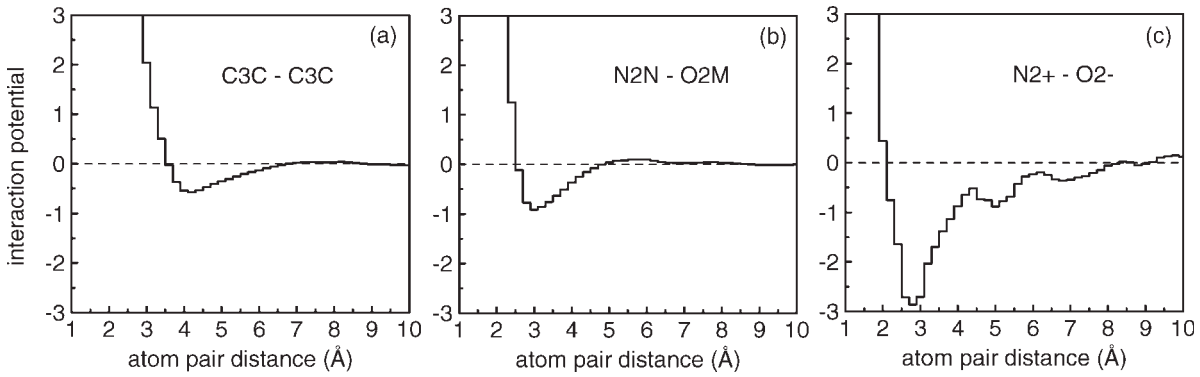

Figure 3
A selected set of pair potentials for ITScore-PP. The dashed line in each panel is y = 0.

mental findings. For the atom pair C3C-C3C, there exists a potential minimum around the distance of 4.2 Å, representing hydrophobic interactions within the atom pair. For the atom pairs N2N—O2M and $N2^+$—$O2^-$, the potential minimums are shifted to the distances between 2.7 Å and 3.0 Å, which are consistent with the hydrogen-bonding interactions between these atom types. It can also be seen that the well of potential curve for the $N2^+$—$O2^-$ pair is deeper than that for the N2N—O2M pair. This can be understood by the fact that in addition to the hydrogen-bonding interactions, the oppositely charged $N2^+$—$O2^-$ pair also experience strong electrostatic interactions.

The extracted pair potentials enable us to calculate the energy score of a given protein–protein complex structure, using Eq. (2). The resulting scoring function, ITScore-PP, was then tested on diverse sets of protein–protein complexes in terms of binding mode and affinity predictions, as shown in the following text.

Binding mode prediction of ITScore-PP

Criteria for prediction quality

In the present work, the prediction quality was judged with the criteria used in the CAPRI (Critical Assessment of PRedicted Interactions) contest,[68,69] which involve three parameters, f_{nat}, L_{rms}, and I_{rms}. The f_{nat} parameter is the number of native residue-residue contacts in the predicted complex divided by the number of residue contacts in the crystal structure. A pair of residues are defined as in contact if they are from different binding partners and are within 5 Å from each other. L_{rms} is the ligand RMSD between the predicted mode and the native structure after the corresponding receptors are optimally superimposed according to their backbone atoms. I_{rms} is the RMSD of the interface region between the predicted and native structures after optimal superimposition of the backbone atoms of the interface residues. The interface is defined as follows: if any atom of a residue in one partner of a complex is within 10 Å from the other partner, this residue is defined to belong to the interface of the complex.

Based on different combinations of these three parameters, the qualities of the predictions can be divided into four categories: high accuracy, medium accuracy, acceptable accuracy, and incorrect, which are listed in Table II. In the present study, we followed the definition that the binding mode of a complex is successfully predicted if the best-scored (top 1) orientation/conformation has an acceptable accuracy. This is the default success criterion unless otherwise specified.

Table II
Quality of Predictions Based on the Combination of Three Parameters f_{nat}, L_{rmsd}, and I_{rmsd}[68,69]

Quality	Combination of three parameters
High accuracy	$f_{nat} \geq 0.5$ and ($L_{rmsd} \leq 1.0$ or $I_{rmsd} \leq 1.0$)
Medium accuracy	($0.3 \leq f_{nat} < 0.5$) and ($L_{rmsd} \leq 5.0$ or $I_{rmsd} \leq 2.0$) or ($f_{nat} \geq 0.5$ and $L_{rmsd} > 1.0$ and $I_{rmsd} > 1.0$)
Acceptable accuracy	($0.1 \leq f_{nat} < 0.3$) and ($L_{rmsd} \leq 10.0$ or $I_{rmsd} \leq 4.0$) or ($f_{nat} \geq 0.3$ and $L_{rmsd} > 5.0$ and $I_{rmsd} > 2.0$)
Incorrect	$f_{nat} < 0.1$ or ($L_{rmsd} > 10.0$ and $I_{rmsd} > 4.0$)

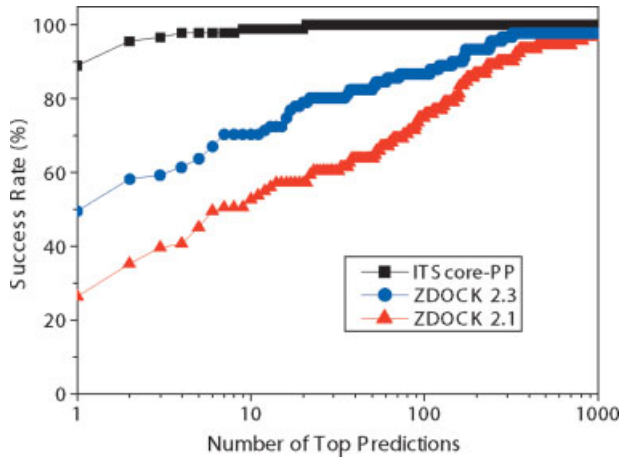

Figure 4

The success rates of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 as a function of the number of top ranked orientations for the bound test cases of 91 protein–protein complexes. The definition of a successful prediction is given in Materials and Methods. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Test on benchmarks from Weng and coworkers

The bound test cases. ITScore-PP was first tested for bound docking, namely, whether the scoring function is able to discriminate the true binding mode from decoy modes, based on the conformations of the protein partners extracted from their cocrystalized dimeric structure. Bound docking serves as a good test for scoring functions, because the complicated protein flexibility problem, which is a separate issue, is avoided. In other words, the induced conformational changes are exhibited in the cocrystalized structures, leaving the accuracy of a scoring function a crucial factor for binding mode prediction in bound docking.

Here, a ligand was defined as the smaller protein in the complex. The other binding partner, referred to as the receptor, was fixed. About 2000 decoy ligand orientations were generated for each protein–protein complex by using ZDOCK 2.1 with default parameters.[14] Together with the native ligand orientation, 2001 ligand poses for each complex were evaluated with ITScore-PP. The SIMPLEX algorithm was used to optimize the ITScore-PP scores for each orientation.[70] For every complex, all the orientations were ranked according to their ITScore-PP scores. The ranked orientations were then clustered. For two orientations with rmsd < 5 Å, only the orientation with lower ITScore-PP score was kept. For consistency, the same clustering strategy was also applied to the ranked orientations when calculating the success rates for ZDOCK 2.1 and ZDOCK 2.3.

Figure 4 shows the success rates of ITScore-PP as a function of the number of the top ranked ligand orientations (referred to as top predictions) for a diverse test set of 91 protein–protein complexes. For comparison, the scoring results of two other well-known models, ZDOCK 2.1[14] and ZDOCk 2.3,[15] are also plotted in the figure. ZDOCK 2.1 evaluates only shape complementarity, and ZDOCK 2.3 also accounts for the effects of desolvation and electrostatics in its scoring function. It can be seen from the figure that ITScore-PP shows a significant improvement in binding mode prediction as compared with the scoring functions in ZDOCK 2.1 and ZDOCK 2.3. If only the top ranked orientation was considered, ITScore-PP yielded a success rate of 89%, compared with 26.4% for ZDOCK 2.1 and 49.5% for ZDOCK 2.3. ITScore-PP identified all the native binding modes and yielded a success rate of 100% when the top 21 predictions were considered. The high success rate suggests the efficiency of ITScore-PP in binding mode prediction.

Figure 4 also shows that ZDOCK 2.3 performs significantly better that ZDOCK 2.1. This is because ZDOCK 2.3 uses a more sophisticated and physical scoring function than ZDOCK 2.1. The trade-off is the computational time. For example, ZDOCK 2.3 takes about ~3 h to dock a complex, compared with ~1 h for ZDOCK 2.1.

To further investigate the accuracy of ITScore-PP, the qualities and rankings of the first successful predictions on individual test complexes using ITScore-PP, ZDOCK 2.1 and ZDOCk 2.3 are compared and listed in Table III. It can be seen from the table that the first binding modes successfully predicted by ITScore-PP all have high-accuracy quality according to the CAPRI categories given in Table II. For ZDOCK 2.1 and ZDOCK 2.3, overall the predicted modes have a lower quality with many medium-accuracy or acceptable predictions.

As mentioned before, the test set of 91 complexes shares nine homologous complexes with our training set of 851 protein–protein complexes. To investigate the effect of these homologous complexes on the results, we calculated the success rates when including or excluding these nine complexes. The results are shown in Table III. It can be seen that both cases gave almost the same success rates for ITScore-PP (<1.2%). These results are consistent with our previous perspective. Namely, a few homologous complexes in the large training set will not bias the derived potentials toward these homologous complexes. For comparison, we also recalculated the success rates of ZDOCK 2.1 and ZDOCK 2.3 for the test set without the nine homologous complexes (see Table III); the changes in the success rates were insignificant. The conclusion on the relative performance of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 remains unchanged.

The unbound test cases. Next, ITScore-PP was tested on the unbound cases of 91 complexes in which at least one partner of the target complex has an unbound struc-

Table III
Comparison of the Docking Performance of ITScore-PP with those of ZDOCK 2.1 and ZDOCK 2.3 on the Bound Test Cases of 91 Protein–Protein Complexes[66,67]

Complex		ZDOCK 2.1		ZDOCK 2.3		ITScore-PP				
PDB code	Category[a]	Quality[b]	Rank[c]	Quality	Rank	Quality	Rank	L_{rmsd}	I_{rmsd}	f_{nat}
1AVX[d]	E	3	2	2	1	1	1	0.64	0.23	0.95
1BRC	E	1	2	1	1	1	1	0.41	0.17	0.98
1BRS	E	1	1	1	1	1	1	0.45	0.23	1.00
1CHO	E	2	5	2	2	1	1	0.93	0.33	0.97
1D6R	E	2	12	2	5	1	1	0.67	0.26	1.00
1DFJ	E	3	5	2	1	1	1	0.68	0.35	0.95
1E6E	E	1	1	1	1	1	1	1.24	0.60	0.98
1EAW	E	1	1	1	1	1	1	0.16	0.07	1.00
1EZU	E	1	1	1	1	1	2	0.37	0.19	1.00
1F34	E	1	1	1	1	1	1	0.39	0.19	0.99
1FSS	E	1	1	1	1	1	1	0.26	0.11	1.00
1HIA	E	1	1	1	1	1	1	0.37	0.16	0.98
1JTG	E	1	1	1	1	1	1	0.34	0.14	1.00
1MAH	E	1	1	1	1	1	1	0.56	0.25	0.97
1STF	E	1	3	1	1	1	1	0.46	0.21	0.96
1TAB	E	1	6	1	2	1	1	0.76	0.29	0.92
1UDI	E	2	1	2	1	1	1	0.66	0.29	1.00
2KAI	E	1	1	1	1	1	1	0.38	0.14	0.97
2MTA	E	1	193	1	5	1	1	0.67	0.29	1.00
2PTC[d]	E	2	1	2	1	1	1	0.52	0.21	0.96
2SIC	E	2	4	2	1	1	1	0.34	0.14	1.00
2SNI[d]	E	1	1	1	1	1	1	0.61	0.22	0.99
2TEC[d]	E	1	2	1	1	1	1	0.34	0.13	0.99
4HTC	E	1	1	1	1	1	1	0.49	0.21	0.99
7CEI[d]	E	3	3	2	1	1	1	0.53	0.23	1.00
1AHW	A	2	6	1	2	1	1	0.41	0.19	0.96
1BGX	A	1	1	1	1	1	1	0.47	0.24	0.97
1BQL	A	1	6	1	4	1	1	0.59	0.28	1.00
1BVK	A	3	63	3	61	1	1	0.40	0.19	0.98
1DQJ	A	1	1	1	1	1	1	0.38	0.19	1.00
1E6J	A	1	22	1	2	1	1	0.53	0.25	1.00
1EO8	A	1	90	2	6	1	1	0.73	0.36	1.00
1FBI	A	1	2	1	1	1	1	0.54	0.29	1.00
1IAI	A	1	2	1	1	1	1	0.67	0.35	0.95
1JHL	A	2	101	1	17	1	1	0.45	0.20	1.00
1JPS	A	1	38	1	7	1	1	0.31	0.15	1.00
1KXT	A	1	14	1	16	1	1	0.49	0.21	0.97
1KXV	A	1	11	1	53	1	1	0.37	0.19	0.98
1MEL[d]	A	2	7	1	2	1	1	0.63	0.32	0.98
1MLC	A	1	265	1	51	1	1	0.39	0.18	1.00
1QFU	A	1	66	1	1	1	1	0.44	0.22	1.00
1VFB	A	1	160	1	173	1	1	0.38	0.16	0.96
1WEJ	A	3	81	3	13	1	1	0.37	0.18	1.00
2VIR	A	3	336	3	241	1	1	0.40	0.18	1.00
2VIS	A	*	—	*	—	1	9	0.50	0.20	1.00
1BJ1	AB	1	1	1	1	1	1	0.40	0.20	1.00
1FSK	AB	1	5	1	1	1	1	0.87	0.32	0.98
1I9R	AB	2	325	2	167	1	2	0.41	0.21	0.98
1IQD	AB	1	10	1	1	1	1	0.31	0.15	1.00
1K4C	AB	1	5	1	1	1	1	0.28	0.12	1.00
1KXQ	AB	1	3	1	1	1	1	0.43	0.22	1.00
1NCA	AB	1	23	1	1	1	1	0.39	0.19	1.00
1NSN	AB	3	89	2	72	1	1	0.79	0.31	0.94
1QFW	AB	1	56	1	20	1	1	0.51	0.19	1.00
2JEL	AB	3	112	2	172	1	1	0.43	0.20	0.96
2QFW	AB	3	53	2	36	1	1	0.65	0.29	1.00
1A0O	O	1	168	2	6	1	1	0.56	0.27	1.00
1A2K	O	2	33	1	2	1	2	0.34	0.17	1.00
1AK4	O	3	160	3	236	1	1	0.48	0.18	0.98
1AKJ	O	2	13	2	2	1	1	0.36	0.17	1.00
1AVZ	O	1	3	1	1	1	1	0.45	0.22	0.97
1B6C	O	2	10	2	1	1	1	1.19	0.48	1.00
1BUH	O	1	76	1	145	1	1	0.69	0.31	0.98

(*Continued*)

Table III
(Continued)

Complex		ZDOCK 2.1		ZDOCK 2.3		ITScore-PP				
PDB code	Category[a]	Quality[b]	Rank[c]	Quality	Rank	Quality	Rank	L_{rmsd}	I_{rmsd}	f_{nat}
1F51	O	3	37	2	7	1	1	0.29	0.16	1.00
1FC2	O	2	93	1	37	1	1	0.63	0.32	1.00
1FQJ	O	3	226	1	6	1	1	0.28	0.13	1.00
1GCQ	O	1	6	1	3	1	1	0.33	0.17	1.00
1GLA	O	2	444	1	278	1	2	0.48	0.23	0.97
1GP2	O	*	—	1	16	1	1	0.56	0.26	1.00
1GRN[d]	O	1	1	1	1	1	1	0.43	0.20	1.00
1HE1	O	1	1	1	1	1	1	0.40	0.17	0.98
1HE8	O	2	354	2	122	1	4	0.54	0.27	1.00
1I2M	O	1	1	1	1	1	1	0.41	0.20	0.99
1I4D	O	1	22	1	4	1	1	0.45	0.23	1.00
1IB1	O	3	148	1	7	1	1	0.41	0.17	0.98
1IGC	O	1	52	2	108	1	1	0.23	0.12	1.00
1IJK	O	3	130	1	1	1	1	0.89	0.40	1.00
1K5D	O	1	1	1	1	1	1	0.30	0.15	1.00
1KAC	O	3	229	1	17	1	1	0.64	0.32	1.00
1KLU	O	*	—	*	—	1	1	0.42	0.18	1.00
1KTZ	O	3	176	2	18	1	2	0.34	0.16	1.00
1KXP[d]	O	1	2	1	1	1	1	0.33	0.16	0.99
1M10	O	3	820	1	2	1	1	0.28	0.14	1.00
1MDA	O	3	157	3	320	1	3	1.47	0.79	0.90
1ML0	O	1	2	1	1	1	1	0.51	0.24	0.99
1N2C	O	1	1	1	1	1	1	0.30	0.14	0.99
1QA9	O	3	702	1	22	1	21	0.63	0.23	0.95
1RLB	O	3	127	1	12	1	2	0.41	0.17	0.98
1SPB[d]	O	1	1	1	1	1	1	0.27	0.09	0.96
1WQ1	O	1	1	1	1	1	1	0.54	0.25	0.98
2BTF	O	1	2	1	1	1	1	0.40	0.19	0.96
Top 1 (Top 10)[e]		26.4% (52.7%)		49.5% (70.3%)		89.0% (98.9%)				
Top 1 (Top 10)[f]		24.4% (47.6%)		45.1% (67.1%)		87.8% (98.8%)				

[a]The complex categories are labelled with E (Enzyme/Inhibitor or Enzyme/Substrate), A (Antibody/Antigen), AB (Antigen/Bound Antibody), and O (Others), respectively.
[b]The quality of the predicted binding mode is represented by 1 (high accuracy), 2 (medium accuracy), 3 (acceptable accuracy), and * (incorrect), respectively.
[c]The rank of the first orientation among the ranked orientations that has at least acceptable accuracy (i.e., a successful prediction). The lower the rank is, the better the prediction is. "—" indicates that no hit was found.
[d]These nine complexes have homologous complexes in the training set.
[e]The success rates for the test set of 91 protein–protein complexes when considering the top 1 (top 10) orientation(s).
[f]The success rates when the nine homologous complexes were excluded from the test set.

ture, and the unbound structure is used for docking. The unbound structure may be in free form or extracted from another complex.[66,67] Unbound docking is normally thought to be more realistic, and protein flexibility is a challenging issue for unbound cases because induced conformational changes are absent in unbound structures. Therefore, unbound docking is also a challenge for scoring functions, especially if protein flexibility is not explicitly considered. In such cases, a scoring function is required to tolerate certain conformational changes in additional to its accuracy to make reasonable binding mode predictions. In the present unbound test study, ITScore-PP was set to a softer scoring function by attenuating the repulsive pair potentials at short distances, allowing for a closer approach between the ligand and the receptor. Namely, the maximum (repulsive) pair potentials were set to 10 kcal/mol for short interatomic distances so that ITScore-PP has some tolerance on severe atomic clashes in unbound cases.

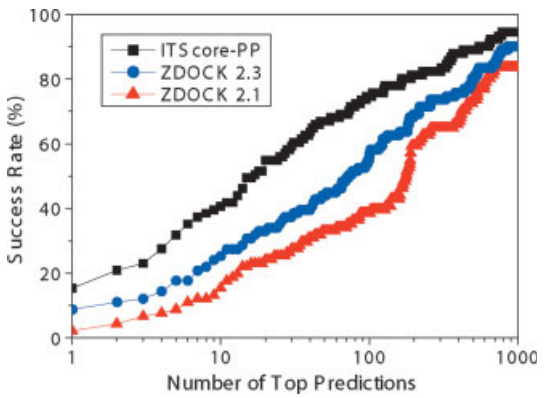

Figure 5
The success rates of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 as a function of the number of top ranked orientations for the unbound test cases of 91 protein–protein complexes. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table IV
Comparison of the Docking Performance of ITScore-PP with those of ZDOCK 2.1 and ZDOCK 2.3 on the unbound Test Cases of the Benchmark of 91 Protein–Protein Complexes (cf., Table III)

Complex		ZDOCK 2.1		ZDOCK 2.3		ITScore-PP				
PDB code	Category[a]	Quality[b]	Rank[c]	Quality	Rank	Quality	Rank	L_{rmsd}	I_{rmsd}	f_{nat}
1AVX[c]	E	3	169	3	7	3	1	8.28	4.11	0.38
1BRC	E	3	11	2	4	2	1	3.72	1.35	0.79
1BRS	E	3	10	2	52	2	18	3.55	1.99	0.51
1CHO	E	2	1	2	1	2	1	3.27	1.32	0.82
1D6R	E	3	39	3	77	3	15	9.39	5.11	0.14
1DFJ	E	2	2	2	1	2	4	3.66	1.89	0.74
1E6E	E	*	—	2	102	2	5	1.68	1.09	0.77
1EAW	E	2	3	2	5	2	2	2.96	1.22	0.69
1EZU	E	3	77	2	467	3	64	7.70	4.71	0.18
1F34	E	2	87	2	42	3	4	6.51	2.72	0.33
1FSS	E	2	253	2	41	1	28	1.30	0.95	0.69
1HIA	E	3	23	3	95	3	6	8.70	3.89	0.21
1JTG	E	1	2	1	1	3	1	5.02	2.14	0.38
1MAH	E	2	35	2	2	1	2	1.49	0.87	0.79
1STF	E	1	9	1	1	1	1	0.95	0.44	0.96
1TAB	E	3	11	3	9	1	1	2.08	0.77	0.85
1UDI	E	3	102	3	18	3	3	6.30	3.26	0.31
2KAI	E	2	140	2	26	3	47	7.90	3.14	0.11
2MTA	E	3	614	3	21	3	2	7.38	3.96	0.26
2PTC[d]	E	3	20	2	15	3	5	8.93	4.10	0.12
2SIC	E	3	14	2	11	2	20	3.06	1.21	0.70
2SNI[d]	E	2	71	2	11	2	2	4.27	1.65	0.62
2TEC[d]	E	2	6	2	1	1	1	1.42	0.54	1.00
4HTC	E	1	6	1	7	2	1	2.21	1.05	0.74
7CEI[d]	E	2	30	2	2	2	6	3.30	1.30	0.71
1AHW	A	3	34	3	9	2	31	1.31	1.03	0.66
1BGX	A	*	—	*	—	*	—	0.00	0.00	0.00
1BQL	A	3	43	1	14	1	15	0.97	0.71	0.86
1BVK	A	3	183	3	100	3	14	6.80	4.00	0.15
1DQJ	A	*	—	*	—	3	11	9.87	6.02	0.20
1E6J	A	3	192	3	70	3	10	8.59	5.08	0.12
1EO8	A	3	445	1	744	1	3	2.02	0.75	0.86
1FBI	A	3	188	3	31	3	164	7.76	4.02	0.20
1IAI	A	3	5	3	3	3	41	8.56	4.11	0.17
1JHL	A	3	13	3	8	1	35	0.93	0.52	0.83
1JPS	A	2	495	2	437	1	43	1.30	0.69	0.73
1KXT	A	2	211	2	66	1	6	0.56	0.47	0.91
1KXV	A	3	162	3	259	3	125	8.05	3.53	0.40
1MEL[d]	A	2	28	1	5	1	40	1.42	0.91	0.90
1MLC	A	3	140	3	7	1	20	1.50	0.96	0.79
1QFU	A	2	192	2	117	1	20	1.49	0.73	0.95
1VFB	A	3	60	3	61	3	9	6.57	3.75	0.19
1WEJ	A	3	221	2	181	1	37	0.95	0.64	0.93
2VIR	A	*	—	1	774	1	55	0.64	0.52	0.86
2VIS	A	3	190	3	380	3	740	11.83	2.60	0.33
1BJ1	AB	2	170	2	66	1	1	1.75	0.86	0.90
1FSK	AB	1	14	1	1	1	1	1.62	0.82	0.91
1I9R	AB	2	701	3	516	2	7	4.29	2.42	0.62
1IQD	AB	3	538	3	10	1	8	1.47	0.85	0.59
1K4C	AB	3	228	3	490	3	105	9.71	5.56	0.11
1KXQ	AB	3	10	3	4	2	1	2.79	1.37	0.77
1NCA	AB	2	176	2	43	1	5	0.73	0.42	0.88
1NSN	AB	*	—	*	—	1	228	0.53	0.45	0.86
1QFW	AB	3	189	3	34	1	96	0.77	0.79	0.85
2JEL	AB	2	179	3	132	3	27	5.21	3.05	0.29
2QFW	AB	3	600	3	487	2	17	1.44	1.35	0.83
1A0O	O	2	3	2	15	2	5	3.09	2.17	0.62
1A2K	O	*	—	2	84	3	75	9.54	4.91	0.36
1AK4	O	3	12	3	48	3	366	9.09	4.12	0.30
1AKJ	O	3	48	3	17	3	431	5.45	2.80	0.35
1AVZ	O	2	162	*	—	*	—	0.00	0.00	0.00

(*Continued*)

Table IV
(Continued)

Complex		ZDOCK 2.1		ZDOCK 2.3		ITScore-PP				
PDB code	Category[a]	Quality[b]	Rank[c]	Quality	Rank	Quality	Rank	L_{rmsd}	I_{rmsd}	f_{nat}
1B6C	O	2	190	2	27	2	4	2.91	2.05	0.82
1BUH	O	3	415	3	545	2	4	3.55	1.69	0.55
1F51	O	*	—	2	534	3	73	9.85	5.20	0.15
1FC2	O	*	—	*	—	2	127	3.17	2.35	0.51
1FQJ	O	3	259	3	101	*	—	0.00	0.00	0.00
1GCQ	O	3	457	*	—	2	15	2.15	1.14	0.76
1GLA	O	3	419	3	160	3	2	13.05	4.89	0.32
1GP2	O	*	—	3	678	3	7	6.66	3.28	0.17
1GRN[d]	O	*	—	3	187	2	340	2.78	1.77	0.42
1HE1	O	3	7	3	5	3	13	9.85	3.95	0.25
1HE8	O	3	595	2	329	3	329	9.96	5.34	0.43
1I2M	O	3	528	3	713	3	81	8.59	4.78	0.11
1I4D	O	3	651	3	858	3	88	9.32	4.11	0.13
1IB1	O	3	396	*	—	3	641	9.28	3.89	0.28
1IGC	O	2	86	3	199	1	30	1.63	0.92	0.96
1IJK	O	*	—	3	218	2	13	2.12	1.13	0.54
1K5D	O	*	—	3	727	3	192	6.52	2.93	0.15
1KAC	O	3	161	3	259	3	546	9.64	4.90	0.28
1KLU	O	*	—	*	—	*	—	0.00	0.00	0.00
1KTZ	O	*	—	2	217	1	166	1.11	0.45	0.80
1KXP[d]	O	3	410	2	25	2	1	2.16	1.53	0.60
1M10	O	3	368	3	182	3	782	9.85	4.37	0.16
1MDA	O	3	685	3	123	3	14	5.06	3.28	0.35
1ML0	O	2	4	2	1	2	1	4.50	2.03	0.60
1N2C	O	*	—	*	—	*	—	0.00	0.00	0.00
1QA9	O	3	16	3	98	3	356	6.87	3.59	0.37
1RLB	O	3	537	3	123	3	305	5.27	2.64	0.41
1SPB[d]	O	1	1	2	1	1	1	0.89	0.45	0.93
1WQ1	O	2	178	2	185	3	651	3.90	1.89	0.29
2BTF	O	3	129	2	73	3	25	9.86	4.56	0.38
Top 1 (Top 10)[e]		2.2% (15.4%)		8.8% (25.3%)		15.4% (40.7%)				
Top 1 (Top 10)[f]		1.2% (14.6%)		7.3% (22.0%)		12.2% (36.6%)				

[a]The complex categories are labelled with E (Enzyme/Inhibitor or Enzyme/Substrate), A (Antibody/Antigen), AB (Antigen/Bound Antibody), and O (Others), respectively.
[b]The quality of the predicted mode is represented by 1 (high accuracy), 2 (medium accuracy), 3 (acceptable accuracy), and * (incorrect), respectively.
[c]The rank of the first orientation among the ranked orientations that has at least acceptable accuracy. The lower the rank is, the better the prediction is. "—" indicates that no hit was found.
[d]These nine complexes have homologous complexes in the training set.
[e]The success rates for the test set of 91 protein–protein complexes when considering the top 1 (top 10) orientation(s).
[f]The success rates when the nine homologous complexes were excluded from the test set.

Similar to the bound cases, ZDOCK 2.1 was used to generate 2000 decoy ligand orientations for each protein–protein complex, based on the unbound structures taken from the benchmarks prepared by Weng and coworkers.[66,67] The decoy orientations and the original unbound orientation were scored by ITScore-PP and clustered following the procedures described in last subsection. Figure 5 shows a comparison on the success rates of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 for the unbound cases of the aforementioned 91 protein–protein complexes. Again, the figure demonstrates a significant improvement for ITScore-PP over ZDOCK 2.1 and ZDOCK 2.3 on binding mode predictions. If the top ranked orientation was taken as the predicted binding mode, ITScore-PP yielded a success rate of 15.4%, compared to 2.2% for ZDOCK 2.1 and 8.8% for ZDOCK 2.3. The success rate of ITScore-PP increased to 21% if the top two predictions were considered. Figure 5 also shows that ZDOCK 2.3 performs better than ZDOCK 2.1, as expected.

Comparing Figures 4 and 5, it can be seen that although the relative performances of the three models (ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3) remain the same, the success rates are much higher for the *bound* cases than for the *unbound* cases, indicating the impact of protein flexibility. The qualities and rankings of the first successfully predicted orientations are also much better for the bound cases than for the unbound cases, as seen from Tables III and IV.

In common practices of protein–protein docking, the top 10 predictions are often used for binding mode anal-

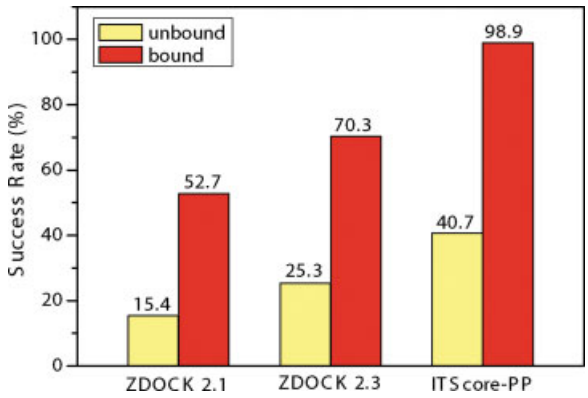

Figure 6

The success rates of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 for both the bound and unbound test cases of 91 protein–protein complexes when the top 10 ranked orientations were considered. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

ysis.[68] Therefore, the success rates of ITScore-PP at the top 10 predictions for the bound and unbound cases are plotted in Figure 6. The corresponding success rates of ZDOCK 2.1 and ZDOCk 2.3 are also plotted in the figure. As seen from the figure, ITScore-PP showed an encouraging performance, having 40.7% and 98.9% success rates for the unbound and bound cases on the test set, higher than those of ZDOCK 2.1 (15.4% and 52.7%) and ZDOCK 2.3 (25.3% and 70.3%).

Similar to the bound cases, exclusion of the nine homologous complexes from the test set had no significant effect on the success rates of ITScore-PP, ZDOCK 2.1, and ZDOCK 2.3 in the unbound cases (see Table IV).

ITScore-PP with different protein representations. In addition to the atom-level potentials, the reduced, residue-level potentials are often used in protein structure prediction to increase the computational efficiency of conformational space sampling. Therefore, it is valuable to consider the residue-level representation of proteins for an energy scoring function.[48] Following Ref. 48 by Zhou and coworkers, three types of united centers were used in this study to represent a residue of proteins, C_α, C_β, and side-chain center of mass (SCM). Accordingly, three types of united-residue potentials, ITScore-PP(C_α), ITScore-PP(C_β), and ITScore-PP(SCM), were derived by using the similar iterative procedure described in the Materials and Methods section. If no residue pair was observed at a short distance, the corresponding pair potential at the distance was set to 100/10 = 10 kcal/mol, considering the fact that the average number of heavy atoms in a residue is about 10. The derived residue-level scoring functions were tested on both the bound and unbound cases.

Figure 7 plots the success rates of ITScore-PP in atom- and residue-level representations as a function of the number of the top ranked modes for the bound test cases. It can be seen from the figure that the order for the performances is ITScore-PP > ITScore-PP(SCM) > ITScore-PP(C_β) > ITScore-PP(C_α). If the top 10 ranked modes were considered, the success rates were 98.9%, 71.4%, 45.1%, and 26.4% for ITScore-PP, ITScore-PP(SCM), ITScore-PP(C_β), and ITScore-PP(C_α), respectively (see Fig. 9).

Similar trends were also observed for the success rates in the unbound test cases, except that the performances of ITScore-PP(SCM) and ITScore-PP(C_β) were close, as shown in Figure 8. The success rates for the top 10 ranked modes were 40.7%, 30.8%, 30.8%, and 22.0% for ITScore-PP, ITScore-PP(SCM), ITScore-PP(C_β), and ITScore-PP(C_α), respectively (see Fig. 9). It is noticeable that the differences in performance for the atom-level scoring function and residue-level scoring functions were less dramatic for unbound cases than for bound cases (Fig. 7 cf. Fig. 8). One possibility is that the impact of protein flexibility reduces the performance of a scoring function in unbound cases. Another possibility is that the united center feature of the residue-level potentials, at the sacrifice of accuracy, allows more tolerance on severe atomic clashes that are common in unbound cases without explicit considerations of protein flexibility.

Furthermore, Figures 7 and 8 show that overall the residue-level potentials of ITScore-PP yielded comparable

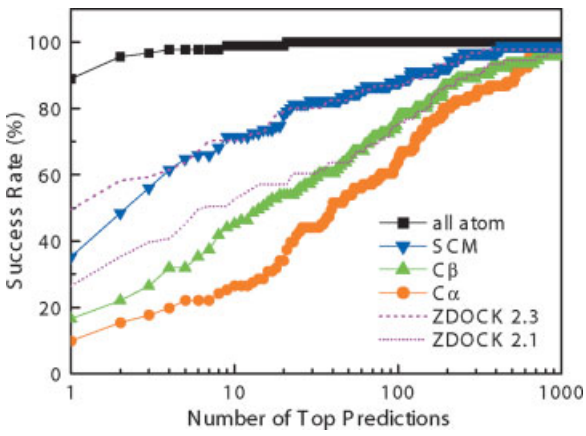

Figure 7

Comparison of the success rates of ITScore-PP with the atom-level potentials and three types of residue-level potentials (C_α, C_β, and SCM) on the bound test cases of 91 protein–protein complexes. For comparison, the results for ZDOCK 2.1 and ZDOCK 2.3 are also displayed. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

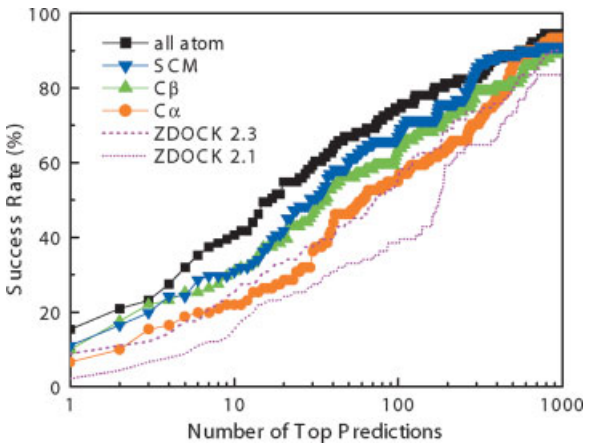

Figure 8

Comparison of the success rates of ITScore-PP with the atom-level potentials and three types of residue-level potentials (C_α, C_β, and SCM) on the unbound test cases of 91 protein–protein complexes. For comparison, the results for ZDOCK 2.1 and ZDOCK 2.3 are also displayed. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

or higher success rates than ZDOCK 2.1. The conclusion also held for comparisons with ZDOCK 2.3 if five or more top predictions were kept for each complex. For example, if the top 10 predictions were considered, ITScore-PP(SCM) yielded 71.4% and 30.8% success rates in the bound and unbound cases, compared with 52.7% and 15.4% for ZDOCK 2.1 and 70.3% and 25.3% for ZDOCk 2.3 (Figs. 6 cf. 9). This finding is encouraging because a residue-level representation of proteins involves much fewer atoms in calculations than an atom-level representation. Therefore, using the residue-level ITScore-PP such as ITScore-PP(SCM) can dramatically increase the computational efficiency of protein–protein docking (see later).

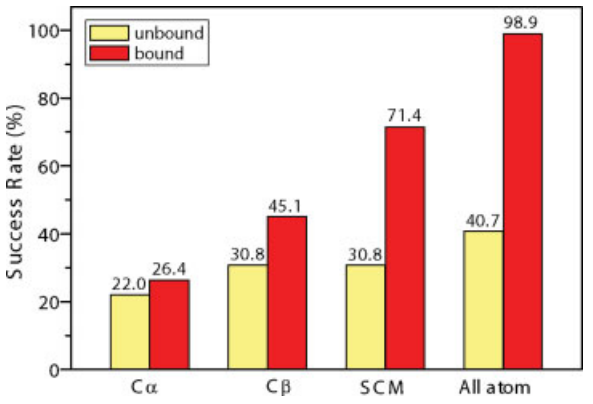

Figure 9

Comparison of the success rates of ITScore-PP with the atom-level potentials and three types of residue-level potentials (C_α, C_β, and SCM) on the both bound and unbound test cases of 91 protein–protein complexes when the top 10 ranked orientations were considered. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table V

Comparison of the Docking Performance of ITScore-PP and RDOCK on the ZDOCK2.1 Decoy Set

	ZDOCK2.1		RDOCK		ITScore-PP	
Complex	Rank[a]	I_{rmsd}	Rank[a]	I_{rmsd}	Rank	I_{rmsd}
1CGI	4	2.41	8	2.24	56	2.17
1CHO	1	1.67	1	1.28	1	0.98
2PTC	1655	2.11	2	1.12	416	1.31
1TGS	3	1.83	8	2.03	1	0.64
2SNI	—	—	—	—	—	—
2SIC	241	1.68	1	1.17	62	1.11
1CSE	1537	1.17	1	1.17	43	0.57
2KAI	1399	2.47	141	2.36	—	—
1BRC	173	1.58	3	2.41	1	2.10
1ACB	25	1.33	1	1.86	1	0.76
1BRS	61	2.23	13	1.23	117	1.60
1JTG	3	1.52	13	1.54	1	2.04
1MAH	849	1.49	1	0.91	1	0.77
1UGH	305	2.37	1	2.08	587	2.42
1DFJ	37	2.48	1	2.48	31	2.49
1FSS	731	1.52	42	1.52	174	0.72
1AVW	45	2.07	2	2.00	1	1.73
1PPE	1	0.59	1	0.69	1	0.12
1TAB	65	1.21	10	0.76	1	0.73
1UDI	31	1.19	3	1.06	1	0.81
1STF	1	0.88	1	1.04	1	0.39
2TEC	1	0.39	1	0.83	1	0.35
4HTC	1	1.79	1	1.46	1	0.92
1MLC	1106	2.10	2	1.65	432	2.33
1WEJ	1396	1.07	4	0.91	560	0.61
1AHW	26	1.57	1	1.61	16	1.08
1DQJ	1341	2.45	952	2.45	—	—
1BVK	974	1.89	1314	1.64	1213	1.44
1FBI	1786	2.15	53	2.15	1851	1.91
2JEL	112	1.82	301	1.70	521	2.50
1BQL	172	1.18	1	1.18	134	0.34
1JHL	404	1.46	41	0.88	193	1.98
1NCA	2	0.85	8	0.83	16	0.25
1NMB	693	1.10	1	1.11	196	0.35
1MEL	12	1.19	1	1.37	140	0.26
2VIR	476	1.03	80	1.19	162	0.74
1EO8	—	—	—	—	—	—
1QFU	407	1.17	29	0.95	190	0.77
1IAI	—	—	—	—	1076	2.39
2PCC	—	—	—	—	—	—
1WQ1	5	1.37	16	1.91	390	2.45
1AVZ	—	—	—	—	—	—
1MDA	—	—	—	—	—	—
1IGC	22	1.20	21	1.18	27	0.62
1ATN	360	0.80	1	0.80	1	0.62
1GLA	—	—	—	—	—	—
1SPB	1	0.61	1	0.70	1	0.54
2BTF	32	0.68	1	0.95	1	0.30
1A0O	833	2.46	11	2.46	7	2.00
Top 1[b]	12.2%		36.7%		32.7%	

[a]The rank of the first successful prediction that has an interface rmsd $\leq$ 2.5 Å.
[b]The success rate when the top predicted mode was considered.

Table VI
Comparison of the Docking Performance of ITScore-PP, RDOCK, and EMPIRE on the ZDOCK2.3 Decoy Set

			EMPIRE[a]							
	ZDOCK2.3		Original		Sidechain		Minimization		ITScore-PP	
Complex	Rank[b]	I_{rmsd}	Rank[b]	I_{rmsd}	Rank[b]	I_{rmsd}	Rank[b]	I_{rmsd}	Rank	I_{rmsd}
1CGI	4	2.41	107	1.54	48	2.02	1	2.18	50	2.23
1CHO	3	1.57	1	1.26	1	1.01	1	1.57	1	0.97
2PTC	193	1.83	8	1.03	1	0.44	1	0.44	309	1.42
1TGS	3	2.22	10	2.46	4	1.55	3	1.85	1	0.63
2SNI	1262	2.22	425	2.22	6.17	2.22	92	2.22	740	2.26
2SIC	11	2.37	2	2.06	3	2.06	3	1.04	68	1.00
1CSE	198	2.20	1	0.50	5	1.10	4	1.24	32	0.51
2KAI	388	1.61	151	2.30	3	1.69	28	1.69	555	1.60
1BRC	24	2.32	21	1.21	1	1.73	1	2.30	1	2.01
1ACB	18	1.33	2	1.44	14	1.44	4	0.93	1	0.80
1BRS	65	2.13	20	1.30	26	1.97	15	2.29	97	1.65
1JTG[c]	1	1.52	n/a	n/a	n/a	n/a	n/a	n/a	1	2.02
1MAH	24	1.29	238	1.78	104	0.84	1	0.89	1	0.72
1UGH	8	2.25	1069	1.60	66	1.13	1	1.60	555	2.38
1DFJ	1	2.48	517	2.38	1	1.70	1	1.70	23	2.49
1FSS	50	1.52	54	1.04	1	1.07	2	1.05	106	0.67
1AVW	3	2.07	1	1.89	12	1.48	1	1.53	1	1.63
1PPE	1	0.90	1	0.52	1	1.46	1	0.87	1	0.12
1TAB	79	1.21	1	0.51	1	1.56	1	1.56	1	0.71
1UDI	5	1.19	12	1.06	1	0.94	1	0.79	1	0.78
1STF	1	0.88	1	0.80	1	1.42	1	1.01	1	0.34
2TEC	1	0.76	1	0.68	1	1.25	1	0.92	1	0.36
4HTC	3	2.46	45	1.40	1	0.69	1	0.69	1	0.99
1MLC	128	1.65	46	2.46	395	2.46	338	2.46	286	2.44
1WEJ	183	1.04	5	0.91	12	0.57	62	0.57	170	0.68
1AHW	7	1.82	25	1.41	7	1.75	4	1.23	20	0.80
1DQJ	—	0.00	—	—	—	—	—	—	—	0.00
1BVK	821	2.34	672	2.34	450	2.34	419	2.34	—	0.00
1FBI	642	2.03	1593	2.18	534	2.18	447	2.18	1744	2.48
2JEL	233	1.46	598	1.90	20	1.16	1	1.09	1218	1.97
1BQL	13	1.07	14	0.68	11	0.84	9	0.84	181	0.32
1JHL	333	1.37	121	1.16	9	1.16	50	1.85	263	1.19
1NCA	1	1.06	8	1.51	56	0.83	2	1.93	29	0.28
1NMB	135	0.98	1	0.99	427	0.99	337	1.13	255	0.35
1MEL	3	1.19	2	1.36	3	1.01	1	1.07	120	0.50
2VIR	1101	1.03	79	1.03	527	1.03	521	1.19	245	0.40
1EO8	1497	0.96	55	0.94	607	0.94	72	0.94	12	0.91
1QFU	388	1.14	21	0.75	92	0.78	1	0.78	205	0.77
1IAI	997	1.70	52	1.47	106	1.47	429	1.70	180	1.60
2PCC	—	0.00	—	—	—	—	—	—	—	0.00
1WQ1	15	1.31	121	2.23	10	1.88	9	1.20	451	2.32
1AVZ	—	0.00	—	—	—	—	—	—	—	0.00
1MDA	—	0.00	—	—	—	—	—	—	—	0.00
1IGC	153	1.20	141	1.18	785	1.20	227	1.18	44	0.68
1ATN	7	0.80	1	0.56	1	0.52	1	0.80	1	0.29
1GLA	—	0.00	—	—	—	—	—	—	—	0.00
1SPB	1	0.61	2	0.61	1	0.61	1	0.95	1	0.50
2BTF	2	0.95	1	0.65	1	1.02	1	0.83	1	0.23
1A0O	284	2.45	21	2.45	13	2.25	427	2.45	11	1.83
Top 1[d]	14.3%		20.8%		29.2%		41.7%		32.7%	

[a]Three types of tests for EMPIRE on the ZDOCK 2.3 decoy set: direct application (original), after sidechain optimization (sidechain), and after further CHARMM energy minimization (minimization).
[b]The rank of the first successful prediction that has an interface rmsd $\leq$ 2.5 Å.
[c]No data available for EMPIRE on this complex.
[d]The success rates when the top predicted mode was considered.

Test on the ZDOCK decoy sets

Next, the ITScore-PP scoring function was used to evaluate 2000 orientations for each complex in the ZDOCK 2.1 decoy set and ZDOCK 2.3 decoy set, respectively.[14,15] For the evaluation process, a rigid-body minimization was performed for each orientation via the SIMPLEX optimization algorithm to remove the atomic clashes. The orientations were then ranked from low to high according to their ITScore-PP scores. To keep the consistency with the literature,[71,72] the success criterion for binding mode prediction used in the ZDOCK studies was adopted for our ZDOCK decoy tests.[14,15] Namely, a binding mode was defined as a successful prediction if the interface rmsd from the native structure (I_{rmsd}) is less than or equal to 2.5 Å. Also for consistency, no clustering was performed for the ranked orientations. The ranks and rmsd of the first successful predictions from ITScore-PP were listed in Tables V and VI for ZDOCK 2.1 decoy set and ZDOCK 2.3 decoy set, respectively. The success rates for the top orientation were shown in Figure 10.

For comparison, the results for RDOCK and EMPIRE extracted from their original articles[71,72] were also shown in the above figure and tables. Here, RDOCK is a refinement algorithm for the orientations generated by ZDOCK.[72] The refinement algorithm first applies 50 steps of minimization for VDW interactions to remove atomic clashes, followed by 60 steps of optimization for polar interactions and then 20 steps of optimization for charge interactions. During the first 50 steps of minimization of RDOCK, the proteins are treated as flexible bodies and all atom are free to move. EMPIRE is a recently developed empirical scoring function for protein–protein interactions, which introduces a reference energy to remove the unrealistic dependence of binding affinity of docking decoys on the buried solvent accessible surface area of interface.[71] The EMPIRE scoring function can be used to evaluate binding modes at three levels, direct application, sidechain optimization, and further CHARMM energy minimization.

It can be seen from Figure 10 and Tables V and VI that ITScore-PP achieved significant improvement on mode prediction over the original ZDOCK scoring functions. The success rate of ITScore-PP was 32.7% for both the ZDOCK 2.1 decoy set and the ZDOCK 2.3 decoy set when the top orientation was considered, compared to 12.2% for the ZDOCK 2.1 scoring function and 14.3% for ZDOCK 2.3. These results are consistent with our previous tests on the benchmarks by Weng and coworkers. Comparing ITScore-PP with RDOCK, ITScore-PP (32.7%) yielded a slightly lower success rate than RDOCK (36.7%) on the ZDOCK 2.1 decoy set; however, ITScore-PP (32.7%) achieved significantly higher success rate than RDOCK (20.4%) on the ZDOCK 2.3 decoy set. The difference between the success rates of RDOCK indicates that in this case the performance of RDOCK was sensitive to different decoys for the same 49 protein–protein complexes. The same performance of ITScore-PP on both the ZDOCK 2.1 decoy set and the ZDOCK 2.3 decoy set suggests the robustness of the ITScore-PP scoring function. Compared to EMPIRE, for the ZDOCK 2.3 decoy set, ITScore-PP (32.7%) had higher success rate than the direct application (20.8%) of EMPIRE and after sidechain optimization (29.2%), and lower success rate than the further CHARMM energy minimization (41.7%) of EMPIRE on the ZDOCK 2.3 decoy set. No direct comparison was done for the the ZDOCK 2.3 decoy set because the corresponding data for EMPIRE were not published. Considering that ITScore-PP was only applied to the decoy sets by a rigid body optimization, it is encouraging that ITScore-PP reached a reasonable success rate for unbound docking. The success of a combination of EMPIRE, sidechain optimization and CHARMM energy minimization suggests that sidechain optimization and accurate energy minimization may be coupled with ITScore-PP in our future development. However, this would be done at the cost of substantial increase of computational time, for example, a few weeks of CPU run time for each complex.[71]

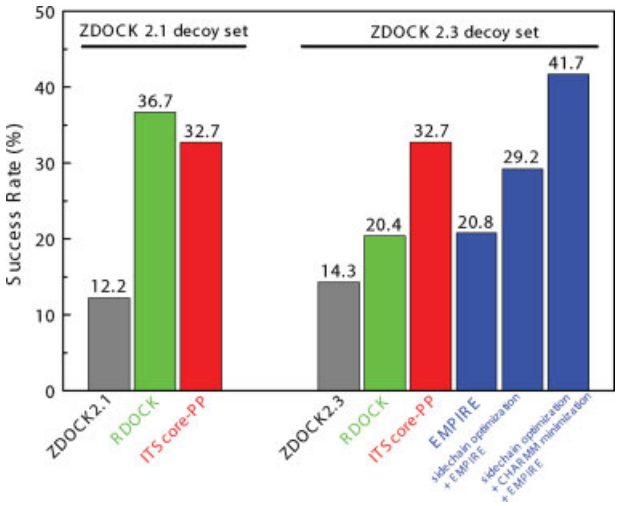

Figure 10

Comparison of the success rates of ITScore-PP, ZDOCK, RDOCK, and EMPIRE on the ZDOCK 2.1 decoy set and the ZDOCK 2.3 decoy set when the top ranked orientation was considered. No data available for EMPIRE on the ZDOCK 2.1 decoy set. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Test on RosettaDock decoy set

The ITScore-PP scoring function was tested with the RosettaDock unbound perturbation decoy set for binding

Table VII
The Docking Performance of ITScore-PP in the RosettaDock Unbound Decoy Set

	RosettaDock					ITScore-PP				
Complex	Quality[a]	Rank[b]	L_{rmsd}	I_{rmsd}	f_{nat}	Quality	Rank	L_{rmsd}	I_{rmsd}	f_{nat}
1ACB	3	3	9.08	3.41	0.13	3	4	9.04	3.39	0.11
1AOO	2	1	10.29	4.18	0.57	2	1	10.29	4.18	0.57
1AHW	2	1	6.41	2.37	0.51	3	4	6.46	3.00	0.49
1ATN	3	1	5.83	2.79	0.49	3	1	9.49	4.70	0.38
1AVW	2	1	6.02	2.09	0.67	2	1	5.07	1.81	0.71
1AVZ	3	37	8.08	4.05	0.29	3	22	11.47	5.55	0.35
1BQL	1	1	1.57	0.86	0.64	1	5	1.81	0.72	0.65
1BRC	2	4	3.77	1.21	0.75	2	1	3.77	1.21	0.75
1BRS	2	1	4.78	1.73	0.64	3	1	8.68	4.46	0.33
1BTH	3	4	18.21	5.54	0.30	3	2	5.60	2.35	0.42
1BVK	3	1	7.91	3.93	0.20	3	1	7.18	3.54	0.20
1CGI	2	2	3.79	1.86	0.50	3	8	6.01	2.37	0.42
1CHO	3	1	6.31	2.19	0.46	3	1	10.32	3.76	0.18
1CSE	2	6	10.10	3.12	0.56	2	1	8.81	2.66	0.71
1DFJ	2	1	5.69	2.66	0.59	2	1	5.69	2.66	0.59
1DQJ	3	1	6.71	3.35	0.31	3	5	5.00	2.12	0.34
1EFU	3	44	5.98	3.78	0.16	3	26	7.83	4.21	0.10
1EO8	3	1	10.73	5.54	0.31	3	31	6.12	3.36	0.15
1FBI	2	1	2.79	1.37	0.54	2	1	3.64	1.86	0.51
1FIN	3	364	9.88	5.38	0.12	3	109	8.26	4.19	0.12
1FQ1	3	1	11.37	5.43	0.31	3	1	9.33	5.09	0.41
1FSS	1	1	3.07	0.97	0.74	2	1	4.46	1.42	0.46
1GLA	2	4	5.96	1.85	0.65	3	7	11.96	3.83	0.35
1GOT	3	12	7.86	3.95	0.19	3	4	9.71	4.26	0.19
1IAI	3	8	6.60	3.42	0.22	2	1	4.18	1.61	0.62
1IGC	1	1	2.15	0.63	0.85	1	1	2.15	0.63	0.85
1JHL	3	3	8.56	4.51	0.26	2	2	6.09	2.66	0.56
1MAH	2	1	3.72	1.19	0.70	3	1	8.66	2.77	0.26
1MDA	3	1	8.77	3.59	0.19	3	1	9.84	3.92	0.26
1MEL	2	1	8.47	2.62	0.50	2	4	8.47	2.62	0.50
1MLC	2	7	4.91	1.37	0.52	3	1	15.36	3.45	0.20
1NCA	2	1	3.06	1.53	0.61	1	1	1.24	0.64	0.75
1NMB	1	1	0.90	0.44	0.80	1	6	2.66	0.76	0.85
1PPE	1	1	1.38	0.52	0.73	3	1	7.42	2.39	0.28
1QFU	2	1	3.02	1.10	0.64	1	4	2.93	1.00	0.69
1SPB	1	1	1.06	0.62	0.68	1	1	1.47	0.70	0.69
1STF	1	1	1.91	0.68	0.89	1	2	1.57	0.54	0.91
1TAB	2	1	4.16	1.30	0.74	2	1	4.39	1.37	0.76
1TGS	2	1	2.48	1.44	0.59	2	1	2.26	1.38	0.64
1UDI	2	1	3.35	1.43	0.63	2	1	2.08	1.01	0.74
1UGH	1	1	1.78	0.86	0.67	2	1	4.64	1.90	0.46
1WEJ	3	15	9.28	2.92	0.44	2	2	6.90	2.55	0.62
1WQ1	3	1	5.53	2.46	0.34	2	1	3.38	1.92	0.39
2BTF	3	1	10.03	3.23	0.22	1	1	1.52	0.60	0.75
2JEL	2	1	4.82	2.22	0.52	2	1	6.55	1.98	0.65
2KAI	1	2	2.02	0.97	0.67	2	2	2.48	1.05	0.70
2PCC	3	2	9.19	3.39	0.24	3	1	9.38	3.84	0.29
2PTC	1	4	0.82	0.44	0.80	2	4	5.86	1.82	0.70
2SIC	2	1	5.18	1.53	0.85	2	1	5.18	1.53	0.85
2SNI	2	1	6.70	2.01	0.60	1	1	2.88	0.99	0.73
2TEC	1	1	2.46	0.81	0.74	1	1	2.59	0.85	0.78
2VIR	3	1	7.53	4.19	0.26	2	2	5.74	2.17	0.70
3HHR	3	50	9.84	3.95	0.26	3	3	8.17	4.03	0.30
4HTC	2	1	3.81	1.54	0.61	3	1	5.95	2.25	0.42
Top 1 (Top 10)[c]			66.7% (88.9%)					59.3% (92.6%)		

[a]The quality of the predicted binding mode is represented by 1 (high accuracy), 2 (medium accuracy), 3 (acceptable accuracy) defined in Table II.
[b]The rank of the first conformation with at least acceptable accuracy.
[c]The success rates when the top 1 (top 10) conformation(s) was considered.

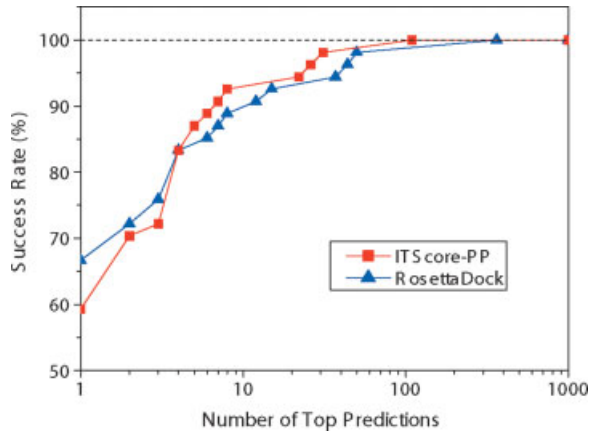

Figure 11

The success rates of ITScore-PP and RosettaDock as a function of the number of top ranked orientations on the RosettaDock decoy set. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

mode predictions. The decoy set consists of 54 protein–protein complexes and each complex has 1000 decoys.[30] No optimization was performed for ITScore-PP on this test set because the RosettaDock decoys have already been optimized and contain no severe atom clashes.

To evaluate the efficiency of the scoring function on the RosettaDock decoy set, we followed the same success criterion for mode prediction that was used for the previous benchmarks of 91 complexes. Namely, a prediction was defined as a success if the predicted mode has an acceptable quality according the categories in Table II. It should be noted that we did not use the funnel criterion proposed in the RosettaDock study because the aim of this criterion is to test the efficiency of the combination of the search algorithm and the scoring function.[30] As mentioned before, the RosettaDock decoy set was constructed on-the-fly through the Monte Carlo search by using the empirical scoring function in the RosettaDock program.[30] Due to incomplete sampling of the ligand conformations (here, 1000 decoys per complex), the constructed (locally optimized) decoy set would bias toward the scoring function used in RosettaDock. Therefore, in the absence of more sufficient sampling, because searching algorithm development is beyond the scope of this work, we focused on evaluating the scoring function alone by using the aforementioned CAPRI success rate criterion.

Table VII lists the results of the first successful prediction for ITscore-PP. As a reference, the results of RosettaDock extracted from Ref. 30 were also listed in the table. The success rates of ITscore-PP and RosettaDock as functions of the number of top predictions were plotted in Figure 11. It can be seen from the table and the figure that although ITScore-PP (59.3%) yielded a lower success rate than RosettaDock (66.7%) for top 1 prediction, ITScore-PP (92.6%) achieved a higher success rate than RosettaDock (88.9%) for top 10 predictions. Overall, the performance of ITScore-PP is close to the performance of RosettaDock for this test set. These results are encouraging because ITScore-PP is computationally efficient and because ITScore-PP can be improved in the future by including intramolecular energies for side chains. It is noted that a novel empirical scoring function, EMPIRE, was recently published and tested on the RosettaDock decoy set by Liang *et al.*[71] Their study showed that the introduction of a reference energy term combined with sidechain modeling and CHARMM energy minimization can substantially improve the performance in terms of the funnel criterion. In this work, ITScore-PP was not compared with EMPIRE/refinement because the rank and rmsd data are not provided in Ref. 71 and the success rates therefore cannot be extracted. It is emphasized that the RosettaDock decoy set was used only to validate ITScore-PP on binding mode predictions.

Binding affinity prediction of ITScore-PP

In addition to binding mode prediction, ITScore-PP was also tested for affinity prediction on the test set of 74 protein–protein complexes prepared by Zhang *et al.*[59] Figure 12 shows the binding energy scores predicted by ITScore-PP as a function of the experimental binding data. The calculated correlation coefficient (R) is 0.71, suggesting the efficiency of ITScore-PP on binding affinity prediction.

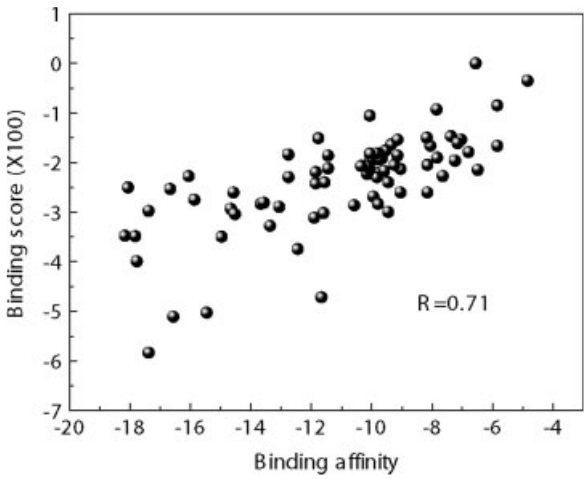

Figure 12

Correlation between the experimentally determined binding affinities and the calculated binding scores with ITScore-PP on a test set of 74 protein–protein complexes given in Table VIII.

Table VIII
Test Set of 74 Protein–Protein Complexes for Binding Affinity Predictions[59]

1A0O	1ABI	1AHW	1AK4	1ATN	1AVZ	1B05	1B32	1B3F	1B3G
1B3L	1B40	1B46	1B4Z	1B51	1B52	1B58	1B5I	1B5J	1B9J
1BQL	1BRS	1BTH	1CHO	1DFJ	1DKG	1DKZ	1EBP	1EFN	1FDL
1FSS	1GLA	1GUA	1HBS	1HWG	1IGC	1JET	1JEU	1JEV	1JHL
1LCJ	1LCK	1MDA	1MEL	1MLC	1NMB	1NSN	1OLA	1QKA	1QKB
1SPS	1STF	1TBQ	1TCE	1TPA	1VFB	1WEJ	2ER6	2JEL	2KAI
2OLB	2PCC	2PLD	2PTC	2SNI	2TPI	3HFL	3HFM	3HHR	3SGB
3SSI	3TPI	4HTC	4TPI						

Computational efficiency

The pairwise feature of the ITScore-PP potentials allows for rapid calculations of the energy scores. The use of the grid-based algorithm further accelerates the score calculations by precalculating the receptor contributions on a grid.[35,73,74]

On average, evaluating (including optimization) 2001 ligand orientations for a target using the atom-level scoring function, ITScore-PP, took about 1 min on a personal computer with 3.2 GHz Pentium IV CPU and 3.0 GB RAM. The same calculations took only several seconds if using the residue-level scoring function, for example, ITScore-PP(SCM). The efficiency of ITScore-PP and ITScore-PP(SCM) makes it possible to use them to evaluate large numbers of ligand orientations during the protein–protein docking process or in the postdocking stage for rescoring purpose.

CONCLUSION

We have derived an efficient knowledge-based energy scoring function for protein–protein interactions (ITScore-PP) by using an iterative method to extract pairwise interaction potentials from 851 protein–protein complex structures in the Protein Data Bank. The iterative method circumvents the challenging reference state problem in deriving knowledge-based scoring functions, and allows for extraction of realistic interaction potentials.

The atomic-level, distant-dependent ITScore-PP was tested for binding mode prediction on a diverse set of 91 protein–protein complexes, which do not overlap with the complexes in the previous training set. If the top 10 ranked orientations were considered,[68] ITScore-PP yielded success rates of 98.9% and 40.7% for the *bound* and *unbound* cases, respectively, compared with 70.3% and 25.3% for ZDOCK 2.3, and 52.7% and 15.4% for ZDOCK 2.1. The performance of ITScore-PP on binding mode prediction was also validated on two other test sets, the ZDOCK decoy set and the RosettaDock decoy set. In addition, ITScore-PP was tested for binding affinity prediction on a set of 74 protein–protein complexes with known affinities, yielding a good correlation coefficient of $R = 0.71$. The pairwise nature of ITScore-PP makes it fast to calculate, about 1 min (including optimization) for 2001 ligand orientations on a personal computer with 3.2 GHz Pentium IV CPU and 3.0 GB RAM. That is equivalently to roughly 0.03 s per orientation.

Several residue-level knowledge-based scoring functions were also derived using the same iterative method. It was found that when each residue was represented by the side-chain center of mass (SCM), the resulted scoring function yielded comparable or better performance than the scoring functions of ZDOCK 2.1 and ZDOCK 2.3. The residue-level ITScore-PP(SCM) is about one order of magnitude faster than the atom-level ITScore-PP, and may be applied to time-consuming processes such as protein conformational sampling.

Both ITScore-PP and ITScore-PP(SCM) can be combined with efficient protein docking software such as ZDOCK for the study of protein–protein interactions.

It is emphasized that all the tests performed in the present study are used for validation of our scoring functions. The results do not serve for competing interests.

ACKNOWLEDGMENTS

We thank Dr. Zhiping Weng, Dr. Julie C. Mitchell, and Mr. Brian Pierce for helpful discussions.

REFERENCES

1. Wodak SJ, Janin J. Computer analysis of protein–protein interaction. J Mol Biol 1978;124:323–342.
2. Smith GR, Sternberg MJ. Prediction of protein–protein interactions by docking methods. Curr Opin Struct Biol 2002;12:28–35.
3. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins 2002;47:409–443.
4. Schneidman-Duhovny D, Nussinov R, Wolfson HJ. Predicting molecular interactions in silico. II. Protein–protein and protein-drug docking. Curr Med Chem 2004;11:91–107.
5. Gray JJ. High-resolution protein–protein docking. Curr Opin Struc Biol 2006;16:183–193.

6. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res 2000;28:235–242.
7. Jiang F, Kim SH. Soft docking: matching of molecular surface cubes. J Mol Biol 1991;219:79–102.
8. Palma PN, Krippahl L, Wampler JE, Moura JJ. BiGGER: a new (soft) docking algorithm for predicting protein interactions. Proteins 2000;39:372–384.
9. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 1992;89:2195–2199.
10. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 1997;272:106–120.
11. Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. Protein docking using continuum electrostatics and geometric fit. Protein Eng 2001;14:105–113.
12. Vakser IA. Evaluation of GRAMM low-resolution docking methodology on the hemagglutininantibody complex. Proteins 1997;Suppl 1:226–230.
13. Chen R, Weng ZP. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47:281–294.
14. Chen R, Weng ZP. A novel shape complementarity scoring function for protein–protein docking. Proteins 2003;51:397–408.
15. Chen R, Li L, Weng ZP. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87.
16. Heifetz A, Katchalski-Katzir E, Eisenstein M. Electrostatics in protein–protein docking. Protein Sci 2002;11:571–587.
17. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromoleculeligand interactions. J Mol Biol 1982; 161:269–288.
18. Shoichet BK, Kuntz ID. Protein docking and complementarity. J Mol Biol 1991;221:327–346.
19. Lorber DM, Udo MK, Shoichet BK. Protein–protein docking with multiple residue conformations and residue substitutions. Protein Sci 2002;11:1393–1408.
20. Shoichet BK, Kuntz ID. Predicting the structure of protein complexes: a step in the right direction. Chem Biol 1996;3:151–156.
21. Wolfson HJ, Lamdan Y. Geometric hashing: a general and efficient model-based recognition scheme. In the Proceedings of the IEEE International Conference on Computer Vision, Tampa, FL; 1988. pp 238–249.
22. Norel R, Fischer D, Wolfson H, Nussinov R. Molecular surface recognition by a computer vision based technique. Protein Eng 1994;7:39–46.
23. Fischer D, Lin SL, Wolfson HL, Nussinov R. A geometry based suite of molecular docking processes. J Mol Biol 1995;248:459–477.
24. Norel R, Petrey D, Wolfson H, Nussinov R. Examination of shape complementarity in docking of unbound proteins. Proteins 1999;35:403–419.
25. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nuclei Acid Res 2005;33:W363–W367.
26. Bordner AJ, Gorin AA. Protein docking using surface matching and supervised machine learning. Proteins 2007;68:488–502.
27. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and and empirical binding free energy function. J Comput Chem 1998;19:1639–1662.
28. Gardiner EJ, Willett P, Artymiuk PJ. Protein docking using a genetic algorithm. Proteins 2001;44:44–56.
29. Abagyan R, Totrov M, Kuznetsov D. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 1994;15:488–506.
30. Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D. Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. J Mol Biol 2003;331:281–299.
31. Zacharias M. Protein–protein docking with a reduced protein model accounting for side-chain flexibility. Protein Sci 2003;12: 1271–1282.
32. Tobi D, Bahar I. Optimal design of protein docking potentials: efficiency and limitations. Proteins 2006;62:970–981.
33. Weiner SJ, Kollman PA, Case DA. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc 1984;106:765–784.
34. Weiner SJ, Kollman PA, Nguyen DT, Case DA. An all atom force field for simulations of proteins and nucleic acids. J Comput Chem 1986;7:230–252.
35. Meng EC, Shoichet BK, Kuntz ID. Automated docking with grid-based energy approach to macromolecule-ligand interactions. J Comput Chem 1992;13:505–524.
36. Cheng TM, Blundell TL, Fernandez-Recio J. pyDock: electrostatics and desolvation for effective scoring of rigid-body protein–protein docking. Proteins 2007;68:503–515.
37. Bertonati C, Honig B, Alexov E. Poisson-Boltzmann calculations of nonspecific salt effects on protein–protein binding free energies. Biophys J 2007;92:1891–1899.
38. Vajda S, Sippl M, Novotny J. Empirical potentials and functions for protein folding and binding. Curr Opin Struct Biol 1997;7:222–238.
39. Pierce B, Weng Z. ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 2007;67:1078–1086.
40. Andrusier N, Nussinov R, Wolfson HJ. FireDock: fast interaction refinement in molecular docking. Proteins 2007;69:139–159.
41. Tanaka S, Scheraga HA. Medium- and long-range interaction parameters between amino acids for predicting three-dimensional structures of proteins. Macromolecules 1976;9:945–950.
42. Miyazawa S, Jernigan RL. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. Macromolecules 1985;18:534–552.
43. Sippl MJ. Calculation of conformational ensembles from potentials of mean force. J Mol Biol 1990;213:859–883.
44. Moont G, Gabb HA, Sternberg MJE. Use of pair potentials across protein interfaces in screening predicted docked complexes. Proteins 1999;35:364–373.
45. Glaser F, Steinberg D, Vakser I, Ben-Tal N. Residue frequencies and pairing preferences at protein–protein interfaces. Proteins 2001; 43:89–102.
46. Jiang L, Gao Y, Mao FL, Liu ZJ, Lai LH. 2002. Potential of mean force for protein–protein interaction studies. Proteins 2002;46:190–196.
47. Lu H, Lu L, Skolnick J. Development of unified statistical potentials describing protein–protein interactions. Biophys J 2003;84:1895–1901.
48. Zhang C, Liu S, Zhou H, Zhou Y. An accurate residue-level pair potential of mean force for folding and binding based on the distance-scaled ideal-gas reference state. Protein Sci 2004;13:400–411.
49. Li X, Liang J. Knowledge-based energy functions for computational studies of proteins. In: Xu Y, Xu D, Liang J, editors. Computational methods for protein structure prediction and modeling, vol 1. New York: Springer; 2006. pp 71–124.
50. Keskin O, Tsai C-J, Wolfson H, Nussinov R. A new, structurally non-redundant, diverse dataset of protein–protein interfaces and its implications. Protein Sci 2004;13:1043–1055.
51. Douguet D, Chen H-C, Tovchigrechko A, Vakser IA. Dockground resource for studying protein–protein interfaces. Bioinformatics 2006;22:2612–2618.

52. Kundrotas PJ, Alexov E. PROTCOM: searchable database of protein complexes enhanced with domain-domain structures. Nucleic Acid Res 2007;35:D575–D579.
53. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 2006; 65:392–406.
54. Zhang C, Liu S, Zhou Y. Docking prediction using biological information, ZDOCK sampling technique and clustering guided by the DFIRE statistical energy function. Proteins 2005;60:314–318.
55. Koppensteiner WA, Sippl MJ. Knowledge-based potentials—back to the roots. Biochemistry (Moscow) 1998;63:247–252.
56. Thomas PD, Dill KA. Statistical potentials extracted from protein structures: how accurate are they? J Mol Biol 1996;257:457–469.
57. Muegge I, Martin YC. A general and fast scoring function for protein–ligand interactions: a simplified potential approach. J Med Chem 1999;42:791–804.
58. Muegge I. A knowledge-based scoring function for protein–ligand interactions: probing the reference state. Perspect Drug Discov Des 2000;20:99–114.
59. Zhang C, Liu S, Zhu Q, Zhou Y. A knowledge-based energy function for protein–ligand, protein–protein and protein-DNA complexes. J Med Chem 2005;48:2325–2335.
60. Huang S-Y, Zou X. An iterative knowledge-based scoring function to predict protein–ligand interactions. I. Derivation of interaction potentials. J Comput Chem 2006;27:1865–1875.
61. Huang S-Y, Zou X. An iterative knowledge-based scoring function to predict protein–ligand interactions. II. Validation of the scoring function. J Comput Chem 2006;27:1876–1882.
62. Thomas PD, Dill KA. An iterative method for extracting energy-like quantities from protein structures. Proc Natl Acad Sci USA 1996;93:11628–11633.
63. Almarza NG, Lomba E. Determination of the interaction potential from the pair distribution function: an inverse Monte Carlo technique. Phys Rev E 2003;68:011202(1–6).
64. Mitchell JBO, Laskowski RA, Alex A, Thornton JM. BLEEP—potential of mean force describing protein–ligand interactions. I. Generating potential. J Comput Chem 1999;20:1165–1176.
65. Myers EW, Miller W. A software tool for finding locally optimal alignments in protein and nucleic acid sequences. Comput Appl Biosci 1989;4:11–17. Available at: ftp://ftp.virginia.edu/pub/fasta/win32_fasta/.
66. Chen R, Mintseris J, Janin J, Weng Z. A protein–protein docking benchmark. Proteins 2003;52:88–91.
67. Mintseris J, Wiehe K, Pierce B, Anderson R, Chen R, Janin J, Weng Z. Protein–protein docking benchmark 2.0: an update. Proteins 2005;60:214–216.
68. Janin J, Henrick K, Moult J, Ten Eyck L, Sternberg MJE, Vajda S, Vasker I, Wodak SJ. CAPRI: a critical assessment of predicted interactions. Proteins 2003;52:2–9.
69. Mendez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. Proteins 2005;60:150–169.
70. Nelder JA, Mead R. A simplex method for function minimization. Computer J 1965;7:308–313.
71. Liang S, Liu S, Zhang C, Zhou YQ. A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. Proteins 2007;69:244–253.
72. Li L, Chen R, Weng Z. RDOCK: refinement of rigid-body protein docking predictions. Proteins 2003;53:693–707.
73. Huang S-Y, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins 2007;66:399–421.
74. Huang S-Y, Zou X. Efficient molecular docking of NMR structures: application to HIV-1 protease. Protein Sci 2007;16:43–51.

APPENDIX A

Calculation of the Factor for excluded volume correction, $f_{ij}^{obs}(r)$

For the system of protein–protein complexes, the excluded volume issue affects the calculation of the potential of mean force (PMF). For example, for an atom of type i in the ligand protein, any atom in the receptor protein cannot penetrate the volume occupied by the ligand (i.e., the excluded volume). In other words, the effective volume occupied by the observed atom type j of the receptor in a spherical shell (radius r, thickness Δr, centered at the atom i of the ligand) is less than $4\pi r^2\Delta r$. Ignorance of the excluded volume effect leads to the under-estimation of the number density of ij pairs and thus the corresponding effective pair distribution function too. Because PMF is proportional to the logarithm of the effective pair distribution function, ignoring excluded volume results in errors in PMF. Therefore, a volume correction term should be introduced to obtain a reasonable PMF, according to the thorough studies by Muegge and Martin.[57,58]

In the present work, the volume correction factor $f_{ij}^{obs}(r)$ was used in the calculation of the PMF [see Eq. (6)]. $f_{ij}^{obs}(r)$ was estimated using the following approximations:

For an atom of type i in the ligand, the ligand volume correction factor was approximated as

$$g_i^{obs}(r) = \sum_j \rho_{ij}^{obs}(r) \Big/ \sum_j \rho_{ij,bulk}^{obs} \qquad (13)$$

Similarly, for an atom of type j in the receptor, the receptor volume correction factor was approximated as

$$g_j^{obs}(r) = \sum_i \rho_{ij}^{obs}(r) \Big/ \sum_i \rho_{ij,bulk}^{obs} \qquad (14)$$

To account for the excluded volume effect for both the ligand and the receptor, the average volume correction factor was calculated as follows:

$$f_{ij}^{obs}(r) = \max\left\{[g_i^{obs}(r) + g_j^{obs}(r)]/2, 0.1\right\} \qquad (15)$$

where the arbitrary parameter of 0.1 was introduced to make $f_{ij}^{obs}(r)$ nonzero, because $f_{ij}^{obs}(r)$ is a denominator in Eq. (6).

APPENDIX B

The training set of 851 protein–protein complexes used for deriving ITScore-PP. The set includes 655 homodimers and 196 heterodimers, respectively.

Homodimers (655 complexes):

11AS	137L	1A0G	1A4I	1A4X	1A64	1A78	1AA7	1AC1	1AC6	1AD1	1ADE	1AE1	1AFS	1AG1
1AH8	1ALV	1AMH	1AN9	1ANJ	1AOJ	1AQ6	1AQY	1AUO	1AWP	1AZ3	1B14	1B43	1B78	1B7G
1B8A	1B8Z	1B9N	1BBH	1BDM	1BDY	1BHT	1BIS	1BIW	1BJF	1BJM	1BKP	1BKZ	1BMZ	1BO4
1BRW	1BSL	1BWW	1BXG	1BXK	1BYK	1C02	1C1B	1C8U	1C94	1C9U	1CD0	1CDO	1CDT	1CKI
1CM7	1CM9	1CMV	1COZ	1CPJ	1CQ3	1CQK	1CYY	1CZ3	1D0Q	1D6J	1D9C	1DC3	1DDZ	1DEL
1DFP	1DJ0	1DLE	1DOR	1DOS	1DP4	1DPG	1DQP	1DQZ	1DR0	1DV1	1E0B	1E19	1E2H	1E5L
1E6R	1E7N	1E8A	1E8V	1EAJ	1EBF	1EBH	1EBL	1EEJ	1EEQ	1EHI	1EKF	1EN7	1EO6	1EOG
1EPA	1EQT	1ERN	1ETK	1EV7	1EVK	1EVX	1EYM	1EYV	1EZ2	1EZG	1F05	1F14	1F46	1F4N
1F4Q	1F5M	1F6B	1F6Y	1F75	1F89	1FA8	1FBT	1FBY	1FIC	1FJH	1FL1	1FN9	1FON	1FP3
1FQW	1FR8	1FUX	1FX5	1FZV	1G0H	1G0S	1G0Z	1G2Q	1G58	1G8E	1G8Q	1G99	1GAR	1GDH
1GFL	1GK6	1GKL	1GNX	1GQP	1GQY	1GT4	1GU2	1GV3	1GWN	1GWW	1GXJ	1GXM	1GXR	1GYJ
1H03	1H1B	1H1Y	1H3F	1H3I	1H3T	1H6P	1H7E	1H7S	1H8X	1HA4	1HFY	1HGX	1HP0	1HRK
1HRN	1HSI	1HSJ	1HT9	1HUL	1HXQ	1I0H	1I0R	1I2N	1I36	1I3M	1I45	1I4J	1I4S	1I5C
1I8T	1ICW	1ID1	1IGU	1IH8	1IHO	1IHR	1IJY	1IME	1IPE	1IPI	1IQ6	1IRQ	1ISA	1ISH
1ISY	1ITQ	1ITV	1IU8	1IUJ	1IVY	1J1C	1J3N	1J3P	1J7L	1JAX	1JB2	1JCZ	1JE5	1JEZ
1JFL	1JHE	1JK6	1JL0	1JLW	1JLY	1JM6	1JMJ	1JMK	1JP3	1JR8	1JU9	1JUO	1JXH	1JYS
1K0Z	1K20	1K2W	1K3O	1K66	1K6R	1K6Z	1KCF	1KEO	1KET	1KEU	1KGZ	1KJN	1KOB	1KPA
1KPT	1KSO	1KTJ	1KTN	1KWS	1KZB	1KZQ	1L2U	1L4I	1L6R	1L7J	1LBQ	1LBV	1LGQ	1LHP
1LJ9	1LM5	1LQ9	1M0S	1M0W	1M1F	1M2B	1M38	1M4D	1M4R	1M6J	1M98	1MBY	1MD0	1MJF
1MJH	1MJV	1MK4	1MKB	1MKF	1MMI	1MN6	1MO0	1MP9	1MR8	1MU4	1MVP	1MY5	1MY6	1MYK
1MZG	1MZH	1MZJ	1N19	1N1E	1N2O	1N3O	1N4O	1N6M	1N9W	1NA6	1NA8	1NCF	1NCH	1NFS
1NMQ	1NNW	1NOZ	1NP6	1NQD	1NVT	1NW1	1NWC	1NXM	1NXZ	1NZI	1O0W	1O0Y	1O4T	1O4U
1O5K	1O6J	1O80	1O9N	1OAB	1OAS	1OBB	1OFZ	1OGX	1OHQ	1OI3	1OI4	1OIA	1OIV	1OKI
1OKR	1OMO	1ON2	1OOE	1OR4	1OTV	1OV9	1OVN	1OX8	1P0C	1P4K	1P5H	1P60	1P6Z	1P9E
1PA0	1PCZ	1PE0	1PL5	1PM2	1PM7	1PN9	1PP2	1PRG	1PRX	1PSR	1PSU	1PVM	1PVY	1PZM
1Q0Q	1Q18	1Q2W	1Q5T	1Q98	1QB2	1QD1	1QF8	1QFE	1QFH	1QG7	1QKR	1QL0	1QLV	1QMH
1QMJ	1QO2	1QO7	1QOR	1QPX	1QQG	1QSD	1QVW	1QX4	1QXH	1QYA	1QYC	1R0S	1R28	1R2F
1R59	1R5P	1R8J	1RB7	1RCM	1REG	1RK4	1RQ2	1RQI	1RQL	1RTR	1RW0	1RYA	1RZN	1S2P
1S3O	1S66	1S9Q	1SES	1SFL	1SFN	1SGM	1SH8	1SHN	1SJI	1SMT	1SMX	1SOX	1SQE	1SQI
1SQU	1SR7	1SW0	1SYI	1SZB	1T06	1T11	1T33	1T3A	1T47	1T4B	1T6F	1T6N	1T6X	1T8P
1T8T	1TC1	1TD2	1TE5	1TE6	1TKL	1TKS	1TLU	1TMK	1TO6	1TO9	1TU1	1TVD	1TVN	1TW7
1TWN	1TXG	1TYO	1U0K	1U0M	1U0V	1U1X	1U20	1U5K	1U6E	1U6R	1U8S	1UC8	1UCR	1UDV
1UI6	1UII	1UIM	1UJN	1ULH	1ULK	1UPG	1USC	1UT4	1UTY	1UZY	1V1Q	1V2I	1V3U	1V3Z
1V4V	1V5V	1V5X	1V6Z	1V7B	1V8G	1V8H	1V96	1V9A	1V9C	1VB5	1VCV	1VDE	1VE2	1VGM
1VGT	1VGY	1VH4	1VH5	1VHI	1VHM	1VIO	1VJ2	1VJL	1VJQ	1VKH	1VL4	1VLJ	1VM7	1VQU
1VSC	1W5F	1W5R	1W6O	1W9A	1W9C	1WGT	1WIW	1WKV	1WLT	1WN1	1WNF	1WNG	1WNL	1WOV
1WPM	1WPN	1WR8	1WSR	1WU9	1WUE	1WX1	1WXD	1WY2	1WY5	1WZ3	1X13	1X2I	1X6I	1X7I
1X94	1XAO	1XEC	1XFC	1XFF	1XG7	1XGE	1XHE	1XHK	1XI3	1XI8	1XLY	1XML	1XRE	1XSE
1XSO	1XSS	1XUQ	1XZN	1XZW	1Y0H	1Y0U	1Y0Z	1Y3T	1Y71	1Y7I	1Y7Y	1Y9B	1Y9W	1YBX
1YCO	1YDV	1YDY	1YEM	1YHA	1YKD	1YKJ	1YKW	1YLA	1YLM	1YLQ	1YOC	1YOZ	1YSJ	1YTT
1YUX	1YVV	1YXK	1YYQ	1Z4E	1Z72	1Z7U	1Z85	1Z9U	1ZC6	1ZEE	1ZH1	1ZIK	1ZK8	1ZKD
1ZKI	1ZO2	1ZOP	1ZTD	1ZV1	1ZWJ	1ZX2	1ZY4	2A2J	2A35	2A6A	2A72	2AA4	2AAX	2AB5
2ANS	2ARC	2AVD	2AWP	2AXP	2BM4	2BMI	2BNX	2BQP	2BSH	2BTM	2NAC	2SCP	2SPC	2UTG
3DAP	3LYN	3RP2	3SDP	4AKE	4CHA	4MDH	4TS1	5HPG	5RUB					

Heterodimers (196 complexes):

1A2X	1A7N	1ACB	1AKS	1AN1	1AQK	1AUI	1AVW	1AXT	1AY7	1B2F	1B34	1BI2	1BLX	1BPL
1BVN	1BXI	1BZ7	1BZX	1C1Y	1C40	1C5M	1C5W	1CDM	1CGI	1CLV	1CR9	1CSE	1CSO	1CXZ
1D4T	1DKF	1DP5	1DQM	1DS6	1DSF	1E96	1EE5	1EGP	1EJ4	1EMU	1ETR	1EUV	1F2S	1F2T
1F3V	1F4W	1F5R	1F60	1FGV	1FIW	1FLE	1FM0	1FS0	1FT1	1FXW	1G9I	1GL4	1H2V	1H9H
1HDM	1HX1	1IAR	1IQ5	1ITB	1IWQ	1J19	1J2J	1J7D	1JAT	1JD5	1JDH	1JHF	1JKG	1JQL
1JSM	1JTD	1JV5	1JWI	1K3A	1K9O	1KA9	1KFU	1KGC	1KLJ	1KSG	1KY7	1L4D	1L6X	1LEW
1LW6	1M1E	1M45	1M46	1MA9	1MCV	1MG9	1MIE	1MQK	1MTP	1MZW	1N1J	1N98	1NGQ	1NGY
1NLB	1NLN	1NMD	1NQ7	1NTV	1NW9	1O5E	1OC0	1OE9	1OL5	1OP9	1OPG	1OPH	1ORY	1OZ7
1P4U	1P5V	1PDK	1PJM	1PPE	1PPF	1PQ1	1Q62	1QAV	1QGE	1QGK	1QGR	1QTX	1QZ7	1R0R
1R8O	1RJC	1RKE	1RKG	1SB2	1SBN	1SBW	1SCJ	1SDZ	1SFI	1SKO	1SMF	1SPP	1SQ2	1SQK
1SVX	1T01	1T0H	1T3L	1T63	1T7M	1TEC	1TFO	1TGS	1TMQ	1U0S	1U58	1U6H	1UGH	1UGQ
1UJZ	1US7	1USU	1UTI	1UVT	1WKW	1WMH	1WQJ	1X8S	1XEW	1XG2	1XKA	1XT9	1Y2A	1YCS
1YDI	1Z0J	1ZBX	1ZOT	2A5F	2BKH	2BNU	2DLF	2F5A	2KIN	2NGR	2STA	3CAA	4ER4	4SGB
7FAB														