PROTEINS: Structure, Function, and Bioinformatics 67:1078–1086 (2007)

ZRANK: Reranking Protein Docking Predictions With an Optimized Energy Function

Brian Pierce and Zhiping Weng*
Bioinformatics Program, Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

ABSTRACT **Protein–protein docking requires fast and effective methods to quickly discriminate correct from incorrect predictions generated by initial-stage docking. We have developed and tested a scoring function that utilizes detailed electrostatics, van der Waals, and desolvation to rescore initial-stage docking predictions. Weights for the scoring terms were optimized for a set of test cases, and this optimized function was then tested on an independent set of nonredundant cases. This program, named ZRANK, is shown to significantly improve the success rate over the initial ZDOCK rankings across a large benchmark. The amount of test cases with No. 1 ranked hits increased from 2 to 11 and from 6 to 12 when predictions from two ZDOCK versions were considered. ZRANK can be applied either as a refinement protocol in itself or as a preprocessing stage to enrich the well-ranked hits prior to further refinement. Proteins 2007;67: 1078–1086.**

Key words: protein docking; rigid body docking; reranking; refinement

INTRODUCTION

Protein–protein interactions are an essential part of many biological processes. Understanding the mode of interaction between two proteins is important for identifying drug targets and optimizing or eliminating protein–protein interactions via site-directed mutagenesis. While there are many crystallized structures of protein complexes available in the Protein Data Bank,[1] it has become increasingly useful to complement these known structures with *in silico* predictions of protein interactions using protein–protein docking.

Several reviews have been published on protein–protein docking, including those by Smith and Sternberg,[2] and Halperin et al.[3] An overview of protein–protein interactions can be found in Jones and Thornton.[4] There have been many exciting developments in protein–protein docking algorithms, but it is still generally performed in two sequential stages, because of the complexity of the problem. The initial stage, which treats proteins as rigid bodies and generates many predictions (10,000 or more), is followed by the refinement stage, which performs any combination of detailed scoring, energy minimization, side chain searches and clustering on these predictions.

A variety of approaches have been used in initial-stage docking, with respect to both searching and scoring. The programs FTDOCK,[5] GRAMM,[6] and ZDOCK[7] all use grid-based spatial searches that are sped up with a Fast Fourier Transform (FFT), a method first applied by Katchalski-Katzir et al.[8] The program HEX[9] also utilizes the FFT, but in this case it is to speed up a rotational search, using spherical harmonics. Other approaches for initial-stage docking searches include Monte Carlo based searching[10,11] and geometric hashing.[12] The scoring functions for initial-stage docking all employ some measure of shape complementarity, and they generally include electrostatics and desolvation as well; more about initial-stage docking scoring can be found in the reviews mentioned earlier.

Docking refinement has also made progress with the recent development of several algorithms. Examples include the web server ClusPro,[13] that performs clustering of rigid body docking predictions from DOT,[6,14] GRAMM,[6] and ZDOCK.[7] The program RosettaDock[11] performs docking refinement on its predictions, using a Monte Carlo based approach, optimizing side chain positions, and rigid body position. It employs an energy function that includes terms for van der Waals, hydrogen bonding, electrostatics, pair potential, and desolvation. The refinement program MultiDock uses side chain rotamers and rigid body minimization to relax the interfaces of rigid body docking predictions.[15]

The initial-stage docking program ZDOCK has been proven effective both against a docking benchmark[7] and in several rounds of the CAPRI experiment.[16,17] ZDOCK is a grid-based docking algorithm that performs a systematic search in 6D and typically can output 3600 or 54,000 predictions, depending on the sampling density in the rotational space (15° or 6° sampling, respectively). It has been shown that 6° sampling yields more near-native predictions in the top 2000, and the refinement of these predictions can improve their ranking.[7]

The refinement program RDOCK was shown to be successful in refining ZDOCK predictions.[18] It uses the

Grant sponsor: NSF; Grant numbers: DBI-0078194, DBI-0133834, and DBI-0116574.

*Correspondence to: Zhiping Weng, Bioinformatics Program, Department of Biomedical Engineering, Boston University, Boston, MA 02215. E-mail: zhiping@bu.edu

Received 10 June 2006; Revised 28 October 2006; Accepted 13 December 2006

Published online 20 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21373

CHARMM program[19] to perform energy minimization on the top ZDOCK predictions (the top 2000 is recommended), and reranks these predictions using desolvation and electrostatics. While it is successful, there are some limitations to RDOCK. The energy minimization step takes roughly 1 min per test case and therefore RDOCK is only feasible for a limited subset of ZDOCK predictions. Also, following the recommended usage, RDOCK success is limited by the number of near-native structures produced by ZDOCK in the top 2000.

To accommodate this, we have developed the program ZRANK (Zlab Rerank) that quickly and accurately reranks the rigid body docking results from ZDOCK. It uses a more detailed potential than ZDOCK but is fast enough to quickly process and rerank the 54,000 predictions that are produced by the ZDOCK 6° sampling search. It significantly improves the success rate of ZDOCK when tested against the newly released protein–protein docking Benchmark 2.0.[20] Thus, it can be used to rerank predictions on its own, or else can be used as a preprocessing step for refinement programs such as RDOCK.

MATERIALS AND METHODS

Scoring Function

ZRANK utilizes a scoring function that can be quickly computed and effectively employed to discriminate hits from non-hits. The scoring function is a linear weighted sum of van der Waals attractive and repulsive energies, electrostatics short and long range attractive and repulsive energies, and desolvation:

$$\begin{aligned} Score = {} & w_{\mathrm{vdW\_a}}E_{\mathrm{vdW\_a}} + w_{\mathrm{vdW\_r}}E_{\mathrm{vdW\_r}} + w_{\mathrm{elec\_sra}}E_{\mathrm{elec\_sra}} \\ & + w_{\mathrm{elec\_srr}}E_{\mathrm{elec\_srr}} + w_{\mathrm{elec\_lra}}E_{\mathrm{elec\_lra}} \\ & + w_{\mathrm{elec\_lrr}}E_{\mathrm{elec\_lrr}} + w_{\mathrm{ds}}E_{\mathrm{ds}} \end{aligned}$$

Here is a more detailed explanation of the energy components:

van der Waals

This is calculated using the Lennard–Jones 6–12 potential, which is calculated for all atoms i and j for interatom distance $r_{ij} < 8.0$ Å:

$$E_{\mathrm{vdW}}(i,j) = \varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$$

The coefficients for the well depth ϵ_{ij} and width σ_{ij} are from the CHARMM 19 polar hydrogen potential.[21] For distances $r_{ij} < 0.6\ \sigma_{ij}$, a repulsive linearization is used, as in Gray et al.[11]

Electrostatics

For electrostatics interactions, the Coulomb equation is used, with a $1/r$ distance dependent dielectric:

$$E_{\mathrm{elec}}(i,j) = 332\frac{q_i q_j}{r_{ij}^2}$$

For short-range electrostatics ($r_{ij} < 5.0$ Å), partial charges from the CHARMM 19 polar hydrogen potential[21] are used. The minimum distance used in the calculation is the width of the van der Waals energy well between the atom pair, ϵ_{ij}, to avoid extraordinarily large values due to clash.

For electrostatics interaction at distances greater than 5.0 Å, only fully charged side chain atoms are used, with charges assigned as in Gray et al.[11] The long-range electrostatics (>5.0 Å) is represented by a separate term, so that short-range interactions (namely hydrogen bonding) can be weighted separately.

Desolvation

Pairwise Atomic Contact Energy[22] (ACE) is used to calculate the desolvation energy:

$$E_{\mathrm{ds}}(i,j) = a_{ij}$$

The term a_{ij} is the ACE term for the atom pair of types i an j; this is zero if the atoms are greater than 6.0 Å apart and for hydrogens.

There are several major differences that distinguish this scoring function from that of RosettaDock. One is that the short-range electrostatics is determined using polar hydrogen partial charges. This allows for the hydrogen bonding and polar forces to be calculated within the electrostatics. In RosettaDock, the hydrogen bonding potential is in a separate term (only full charges are in the electrostatics terms). Another notable difference is the use of pairwise ACE in the desolvation calculation. Also our simplified scoring function has many fewer terms and thus can be rapidly evaluated for a large number of structures.

This scoring function also differs from that of RDOCK. While they both use CHARMM 19 and pairwise ACE, ZRANK breaks down the electrostatics terms into components and applies weights to them separately. Additionally, the van der Waals energy is included in the scoring function rather than just being a filter. Finally, RDOCK computes its final scores only after several rounds of CHARMM minimization, whereas for ZRANK only the input structure is scored, allowing for faster evaluation.

ZDOCK Predictions

To generate initial rigid-body docking predictions, ZDOCK was used with 6° rotational sampling and randomized start positions. At this sampling density, ZDOCK generates 54,000 predictions by keeping the best translational solution for each rotational angle set. For the Antibody/Antigen test cases, the search was restricted to the CDR portions of the antibodies.

For the test set used in this study, two different versions of ZDOCK were used: ZDOCK 2.1 and ZDOCK 2.3. ZDOCK 2.1 uses Pairwise Shape Complementarity[23] (PSC) to score predictions. ZDOCK 2.3[7] uses a combination of PSC, electrostatics, and desolvation in its scoring function; this has been shown to improve performance against a docking benchmark.

Evaluation of Docking Predictions

The predictions from ZDOCK were evaluated by calculating the Root Mean Square Deviation (RMSD) between the bound and unbound interface alpha carbon (Cα) atoms, as previously described.[7] Hits are predictions with interface Cα RMSD less than 2.5 Å.

Addition of Hydrogen Atoms

The Benchmark 2.0 PDB files do not have hydrogens, so it was necessary to add polar hydrogens to the structures prior to rescoring. To accomplish this, the Rosetta program[11] was used to add hydrogens to the receptor and ligand PDBs separately. These structures were then superposed onto the ZDOCK predictions for these cases. We used CHARMM to add the hydrogens to two cases that were too large to be imported into Rosetta.

Training Weights Using Benchmark 1.0 Test Cases

An important aspect of this scoring function is a set of weights that highlight the relative values of the terms of the scoring function. This was particularly crucial because of the fact that the predictions being reranked were from rigid-body docking for which a certain degree of softness in the scoring function was essential.

Optimal weights for the terms were obtained by training the function with a subset of Benchmark 1.0 cases.[24] Benchmark 2.0 cases[20] were used for testing. To ensure independent training and testing sets, all Benchmark 1.0 cases that were homologous with Benchmark 2.0 cases, through SCOP superfamily (non-antibody antigen cases) and manual inspection (antibody antigen cases), were filtered out. After this filtering, 15 cases remained for training; all of these were then docked using ZDOCK 2.3, generating 54,000 predictions per test case. All predictions for these cases were then classified as hits or non-hits, and scored using the seven terms described earlier.

After obtaining and scoring the ZDOCK predictions, a downhill simplex minimization algorithm was used to determine the optimal set of weights. The downhill simplex algorithm allows for quick minimization in a multidimensional space. Briefly described, for N-dimensional minimization it selects N points around a starting point (by taking a given offset from the point in each dimension) and successively moves in the direction of the lowest valued point. In addition to its speed and ease of use, the simplex was an attractive option due to its lack of need for taking derivatives, and in the case of our training criterion (rank of first hit) the function is not necessarily smooth.

The downhill simplex performed minimization in 7D (corresponding the each of the seven energy terms), and simultaneously optimized the ranks of the first hits of the 14 cases (1 of the 15 nonredundant cases did not have any hits from ZDOCK). The target function was the sum of the top ranking hit for each case, and all ranks over 50 were set to 51, to avoid overoptimizing poor rankings (changing this cutoff did not significantly affect the resultant weights). To avoid local minima, five random restarts were used when the minimum was found by the simplex (as outlined in Press[25]), by reinitializing the simplex at each minimum (taking N points around the minimum, as described earlier). Finally, the simplex was run using 12 separate times, using different starting positions. The weights corresponding to the top set of rankings from all of the simplex runs were kept. Performing the minimization using different random seeds led to little difference in the resultant weights.

The weights that were obtained and then applied to reranking the Benchmark 2.0 complexes are as follows:

- van der Waals attractive: 1.0
- van der Waals repulsive: 0.009
- Electrostatics short range attractive: 0.31
- Electrostatics short range repulsive: 0.34
- Electrostatics long range attractive: 0.44
- Electrostatics long range repulsive: 0.50
- Desolvation: 1.02

Using Weighted Function to Rescore Predictions

The scoring function with weights optimized using ZDOCK predictions for Benchmark 1.0 was applied for reranking of ZDOCK predictions for Benchmark 2.0. For the test set, all cases that had ZDOCK hits for Benchmark 2.0 were used, amounting to 62 cases (out of 76 rigid and medium difficulty cases). The scoring function was applied to all 54,000 predictions obtained from the 6° sampling for each case, and the reranked predictions were then analyzed for the ranks of the hits.

RESULTS

Reranking Benchmark 2.0 Predictions

The performance results of ZRANK on Benchmark 2.0 test cases are presented in Table I. There is a clear improvement in the number of highly ranked cases for both ZDOCK versions. Although the weights were trained on ZDOCK 2.3 predictions (which uses a different scoring function than ZDOCK 2.1), it is encouraging to see improvements in performance for predictions from the latter as well.

The count of No. 1 hits for ZDOCK 2.3 reranking increases from 6 to 12 test cases, and for ZDOCK 2.1 it increases from 2 to 11. In addition for ZDOCK 2.3 reranking there are 22 cases that have hits in the top 20 (in bold in Table I).

ZRANK made notable improvements for many cases. The test case 1RLB (Transthyretin/Retinol binding protein) had a significant improvement in both the number of hits and the rank of the first hit over ZDOCK 2.3 (ZDOCK 2.1 had no hits for this test case). Interestingly, for the medium difficulty test case 1WQ1 (Ras GTPase/Ras GAP), ZRANK improved the ranking of the ZDOCK 2.1 predictions (with a reranked hit in the top 10 predictions), however such improvement was not seen for ZDOCK 2.3. This is most likely due to the conformational change between the

TABLE I. Results From ZDOCK 2.1 and ZDOCK 2.3 Before and After Reranking

	ZDOCK 2.1		ZDOCK 2.1 ZR		ZDOCK 2.3		ZDOCK 2.3 ZR		
Test case[a]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	RMSD[d]
Rigid body									
Enzyme–Inhibitor									
1AVX	**0**	**2863**	**44**	**7**	**7**	**449**	**40**	**9**	**2.17**
1AY7	0	5584	24	111	0	11,358	12	134	2.20
1BVN	**13**	**502**	**33**	**14**	**52**	**23**	**48**	**14**	**1.79**
1CGI	9	145	11	23	0	2423	3	34	2.33
1D6R	0	2951	2	984	0	3538	2	1312	2.30
1DFJ	**40**	**9**	**16**	**1**	**73**	**1**	**18**	**1**	**1.89**
1E6E	**0**	**22,643**	**6**	**3**	**18**	**103**	**67**	**4**	**1.51**
1EAW	62	3	37	62	108	13	22	77	1.65
1EWY	2	259	34	65	11	113	28	129	2.38
1EZU	3	1100	0	4323	0	5267	0	20,347	1.44
1F34	13	5	8	69	16	26	4	160	1.42
1MAH	**9**	**92**	**73**	**1**	**71**	**1**	**85**	**1**	**0.89**
1PPE	**218**	**1**	**217**	**1**	**324**	**1**	**193**	**1**	**1.13**
1TMQ	11	314	3	342	21	126	5	389	1.90
1UDI	4	258	2	673	7	13	0	4259	2.19
2MTA	0	—	0	—	0	35,227	1	1722	2.37
2PCC	0	—	0	—	0	12,916	1	1037	1.45
2SIC	**24**	**173**	**69**	**1**	**45**	**39**	**55**	**1**	**1.15**
2SNI	0	17,906	0	2868	0	5079	6	300	1.99
7CEI	**24**	**106**	**115**	**1**	**186**	**1**	**157**	**1**	**1.09**
Other									
1AKJ	0	4872	5	425	28	96	6	40	2.37
1B6C	**2**	**1717**	**12**	**1**	**6**	**168**	**18**	**1**	**2.26**
1BUH	0	14,556	7	475	0	22,962	4	903	1.65
1E96	**0**	**3094**	**7**	**16**	**2**	**399**	**16**	**16**	**2.02**
1F51	**4**	**230**	**10**	**2**	**17**	**11**	**25**	**3**	**1.76**
1FQJ	0	9889	0	5551	0	17,028	0	6057	2.15
1GCQ	0	24,339	1	429	0	23,148	2	762	2.29
1HE1	0	4672	2	349	1	1146	2	258	1.82
1KAC	0	2896	3	72	1	1523	2	94	2.42
1KTZ	0	53,599	0	4251	0	10,395	7	804	1.28
1KXP	**1**	**1734**	**14**	**2**	**33**	**7**	**20**	**1**	**1.99**
1ML0	**21**	**36**	**50**	**1**	**60**	**1**	**56**	**1**	**1.43**
1QA9	0	5672	2	850	0	13,606	1	1502	1.59
1RLB	**0**	**—**	**0**	**—**	**2**	**302**	**38**	**4**	**2.37**
1SBB	0	—	0	—	0	17,342	0	6089	1.32
2BTF	0	—	0	—	0	13,881	7	96	2.17
Antibody-Antigen									
1AHW	**21**	**268**	**24**	**8**	**46**	**56**	**24**	**10**	**0.98**
1BVK	0	3970	0	9314	0	5390	0	10,315	2.41
1DQJ	0	2287	0	7904	0	10,257	0	34,026	2.26
1E6J	**34**	**15**	**121**	**1**	**77**	**19**	**115**	**2**	**1.91**
1JPS	9	171	18	53	19	181	20	71	1.07
1MLC	**12**	**110**	**39**	**5**	**33**	**7**	**38**	**5**	**1.19**
1VFB	0	2734	1	60	0	4969	1	104	1.79
1WEJ	**8**	**465**	**40**	**2**	**39**	**102**	**55**	**1**	**1.01**
2VIS	0	2747	22	37	1	369	17	60	1.83
Antibody-Antigen: UBB									
1BJ1	**49**	**129**	**61**	**2**	**67**	**18**	**52**	**4**	**1.05**
1FSK	**105**	**1**	**89**	**1**	**163**	**1**	**78**	**1**	**1.14**
1I9R	41	50	16	20	37	90	14	24	1.44
1IQD	5	612	29	73	21	99	33	98	1.16
1K4C	0	20,806	3	359	1	1575	34	345	1.47
1KXQ	**13**	**212**	**28**	**1**	**13**	**301**	**27**	**1**	**1.51**
1NCA	47	14	28	5	50	4	17	166	0.86
1NSN	5	185	1	923	1	445	1	1775	1.81
1QFW	7	257	10	94	23	74	11	61	1.20

TABLE I. (Continued)

	ZDOCK 2.1		ZDOCK 2.1 ZR		ZDOCK 2.3		ZDOCK 2.3 ZR		
Test case[a]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	Hits[b]	Rank[c]	RMSD[d]
2JEL	33	45	6	632	23	86	2	975	1.81
2QFW	**3**	**832**	**41**	**1**	**18**	**179**	**63**	**1**	**1.49**
Medium									
Other									
1GRN	2	1704	3	1661	6	807	1	1523	2.26
1HE8	0	—	0	—	0	47,386	0	11,558	2.36
1I2M	0	—	0	—	0	34,162	0	34,492	2.43
1IJK	0	52,731	0	8688	0	6357	9	205	1.52
1K5D	0	—	0	—	0	6012	8	134	2.29
1WQ1	2	1101	10	6	5	27	0	2899	1.52

[a]Only Benchmark 2.0 test cases with ZDOCK hits in the top 54,000 predictions are listed. Test cases in bold had a hit in the top 20 predictions for the reranked ZDOCK 2.3 predictions.
[b]The number of hits in the top 2000 predictions.
[c]Rank of the first hit. "—" notes that no hits were produced.
[d]Interface Cα RMSD of the first hit.

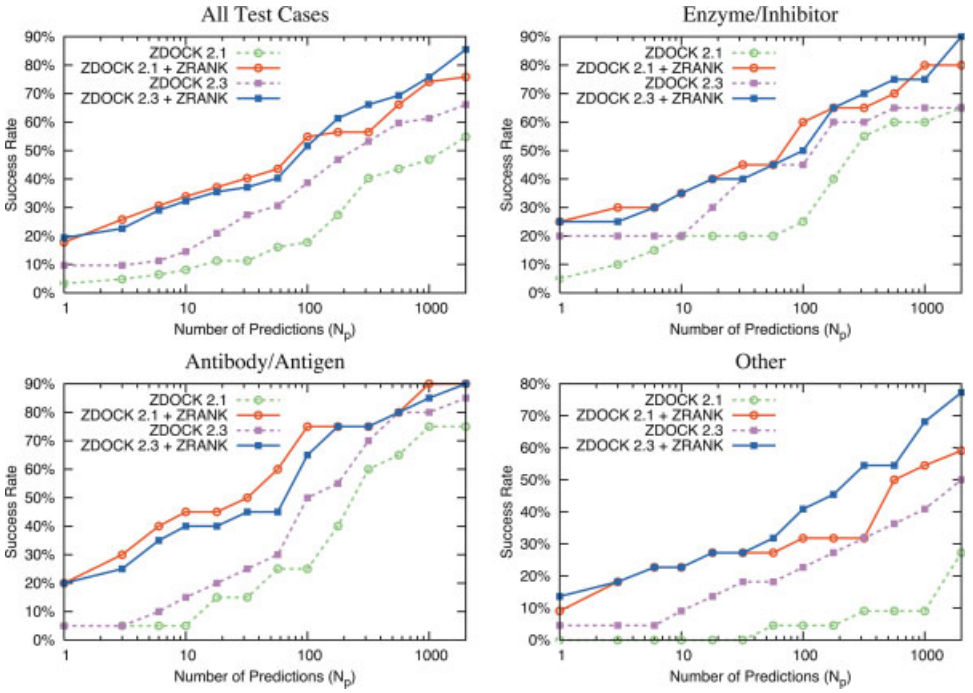

Fig. 1. Success rate versus the number of predictions for reranked predictions and the original predictions from ZDOCK 2.1 and 2.3. Only Benchmark 2.0 test cases with ZDOCK hits in 54,000 predictions are considered in the success rate. All 54,000 predictions were reranked for each test case.

unbound and bound structures leading to unfavorable energies because of the particular arrangements between the two proteins of the hits among the ZDOCK 2.3 predictions.

Success Rate

To provide a clearer overall view of the performance of the original and reranked ZDOCK results, success rate plots are shown in Figure 1. The plot for all test cases reiterates the points mentioned earlier. Both success rates for ZRANK show significant improvements, with the greatest improvement seen for ZDOCK 2.1 reranking, and the greatest overall success rate for the ZDOCK 2.3 reranking.

In addition, the success rates are shown for each individual category of cases. While improvements can be seen for every category, the best improvements can be seen for the *Other* category, which contains the most heterogeneous, the most difficult, and arguably the most biologically interesting test cases of the three categories.

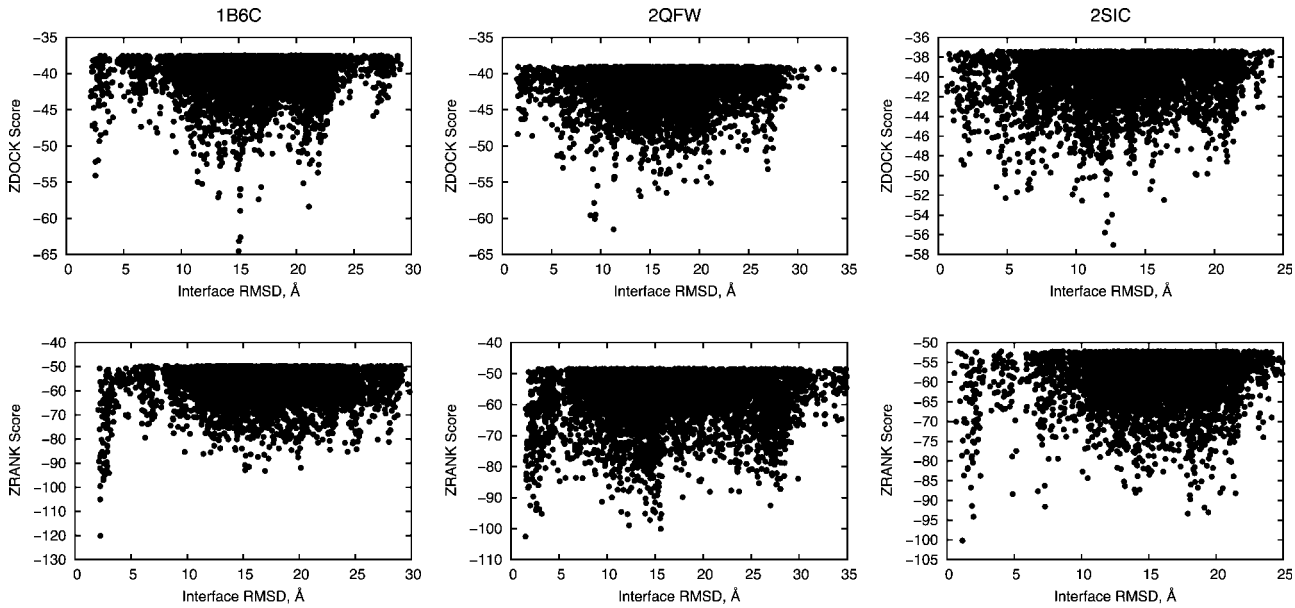

Fig. 2. Score versus interface RMSD plots for the top 2000 ZDOCK 2.3 (top) and reranked (bottom) predictions for several test cases. The ZDOCK score is negated to facilitate comparison between the plots.

Score Versus RMSD Plots

While ranking of the top hit and the number of hits are very important statistics, it is also useful to examine the score versus RMSD plots. If the scoring function is accurate enough, one would expect to see an "energy funnel." The score/RMSD plots for three cases (one *Enzyme Inhibitor*, one *Antibody-Antigen*, and one *Other*) that improved with ZRANK are shown in Figure 2. For all three cases, there is an energy funnel after ZRANK, which did not exist or was not as prominent as for ZDOCK. This supports the improved accuracy of the more detailed scoring function of ZRANK over ZDOCK alone.

Reranking Different Numbers of Predictions

While ZRANK has been discussed in terms of reranking 54,000 predictions, it was also examined to see if the performance changes when reranking the top 5000, 10,000, and 20,000 predictions from ZDOCK 2.3 [Fig. 3(a)]. It can be seen that the success rate improves moderately when reranking more predictions; this is encouraging, as the increased number of false positives subject to reranking does not negatively affect the success. In addition, when $N_P > 100$, the success rate curves diverge and it is clear that the more predictions subject to reranking, the better the success rate becomes.

Examination of the Scoring Potential

As electrostatics is a sensitive term, we tested four formulations of the short-range electrostatics; these are summarized in Table II. All of the scoring functions shown (apart from ZDOCK and PolH vdW) employed a minimum short-range electrostatics distance of 3.0 Å. The ZDOCK Charges function used the same (non-hydrogen) partial charges as in ZDOCK 2.3. The Full Charges function used only fully charged side chain residues, mimicking the treatment of the long-range electrostatics. The PolH 3.0 used the CHARMM 19 polar hydrogen partial charges with a 3.0 distance cutoff. The PolH vdW function is the one selected in this article, and is described in detail in the Methods.

For each formulation, all seven terms was optimized using Benchmark 1.0 predictions as described in the Methods section. The success rate results are shown in Figure 3(b). The Full Charges reranking performed approximately the same as ZDOCK 2.3, while the ZDOCK Charges and PolH 3.0 each performed successively better. In fact the PolH 3.0 function behaved approximately the same in terms of success after the top 30 predictions. As the polar hydrogens used by this scoring function represent an inclusion of more detailed electrostatics (namely hydrogen bonding), this shows that such a term is helpful in discriminating hits. However the PolH 3.0 function did not have as many No. 1 hits as the PolH vdW function given in this article, thus showing that the electrostatics cutoff based on the vdW minima may help to reduce noise that is introduced by having a constant distance for a minimum cutoff.

Success Rate Versus the Quality of Input Predictions

While there is much improvement in the success rate given by the reranking potential, a few cases were reranked less effectively. We hypothesized that the quality of the input docking predictions would affect the outcome of the reranking. To examine this, we binned the test cases according to the lowest RMSD of the ZDOCK predictions in the top 2000. The success rate for each set of cases was

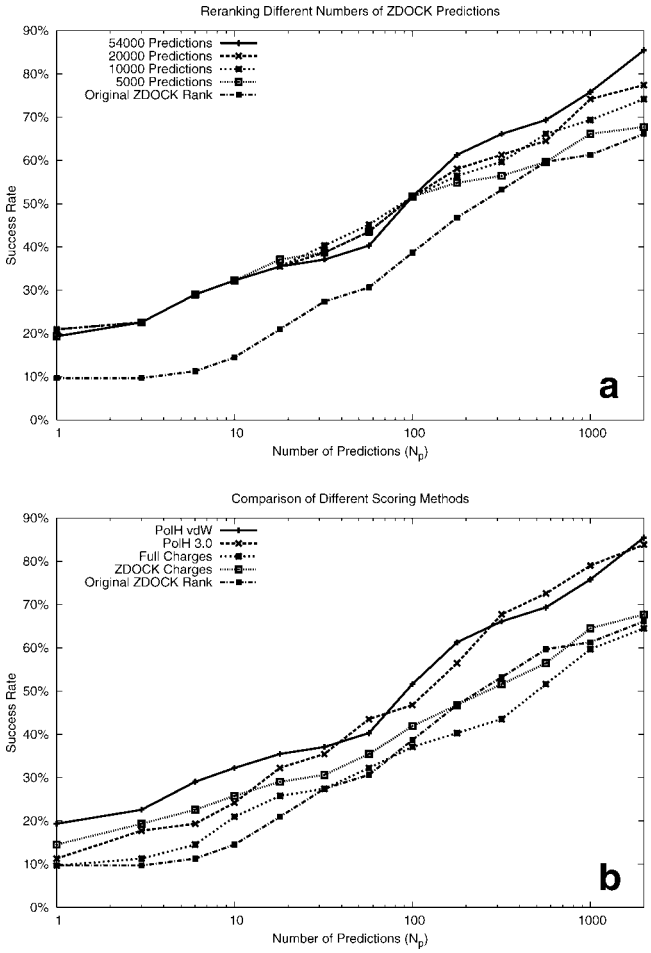

Fig. 3. (**a**) Success rate plots for reranking various numbers of ZDOCK predictions: 5000; 10,000; 20,000; and all 54,000 predictions. (**b**) Success rate comparison for various short-range electrostatics formulations. PolH vdW is ultimately adopted in this article, PolH 3.0 uses a constant 3.0 Å cutoff for minimum electrostatics distance, Full Charges uses this cutoff but with only fully charged side chains and no partial charges, and ZD Charges uses partial charges with no polar hydrogens (with partial charge terms from ZDOCK 2.3).

TABLE II. Charge Terms Used for the Electrostatics Functions

Potential name	Partial charge source	Minimum distance cutoff
PolH vdW	CHARMM 19	vdW minimum
PolH 3.0	CHARMM 19	3.0 Å
Full charges	RosettaDock electrostatics	3.0 Å
ZDOCK charges	ZDOCK 2.3 (no hydrogens)	3.0 Å

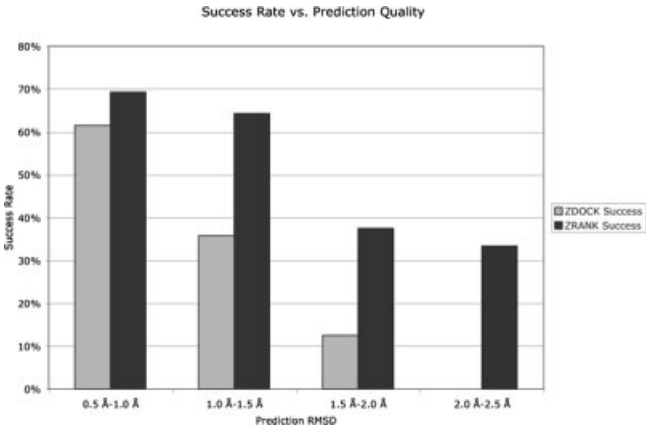

Fig. 4. Success rate as a function of the quality of the predictions. Success is defined as a hit in the top 20 for that test case. Test cases are binned according to the lowest interface RMSD for of the predictions in the first 2000 ZDOCK predictions. No test cases had a hit below 0.5 Å in the top 2000 predictions.

then evaluated for ZDOCK 2.3 and the reranking, as shown in Figure 4. Success rate is computed as the percentage of test cases in the given bin with a hit in the top 20 predictions. As expected, the success rate declines as the quality of the predictions decreases, for both the reranking and ZDOCK. The reranking has significantly greater success than ZDOCK in all the bins, and the improvement is even greater for lower-quality predictions.

Computational Time

One important aspect of ZRANK is that it is possible to rerank all 54,000 predictions from ZDOCK, given its relatively low computational cost. It performs efficiently because the hydrogens only need to be generated once for the initial structure and simple energy calculations are performed on each rigid-body prediction. Reranking all 54,000 predictions takes an average of 180 min on an Intel Pentium III 2.0 GHz machine, or about five predictions per second. This process can be easily parallelized using the Message Passing Interface.

DISCUSSION

We have implemented and tested an algorithm, ZRANK, for quickly and effectively reranking rigid-body docking predictions. It has been shown to increase the success rate across a benchmark set of cases from the initial ZDOCK rankings, and dramatic improvement in the success rate and hit count was seen for several cases. We have also explored variants of this scoring scheme and amount of docking predictions to be reranked [Fig. 3(a,b)], to determine how the ranking success on a benchmark set of cases changes according to different protocols.

One notable feature of ZRANK is that the success rates after reranking ZDOCK 2.3 and ZDOCK 2.1 predictions are quite similar (Fig. 1) when considering all test cases, although ZDOCK 2.3 alone significantly outperforms ZDOCK 2.1. This demonstrates that the improvement for reranking ZDOCK 2.1 is much greater. Nonetheless, there is a slight edge for the reranked ZDOCK 2.3 predictions, most likely due to the higher initial hit counts from ZDOCK 2.3. This is in contrast to the performance from

RDOCK, where the overall success was higher for reranking the ZDOCK 2.1 predictions.[18]

Furthermore, the *Other* category of cases seems to have improved most using ZRANK from the initial ZDOCK ranking. This may be because the PSC function (used in ZDOCK 2.1 and 2.3), which is effective in discriminating the large pockets of binding sites of enzyme/inhibitor complexes,[23] has slightly less success in scoring the non enzyme/inhibitor interfaces. Thus the van der Waals based shape complementarity of the reranking function could provide an advantage in evaluating these cases, which represent a larger percentage of the test cases in Benchmark 2.0.

While the basic energy function used by ZRANK is simple, the weights allow for the high degree of discrimination between hits and non-hits within a set of rigid-body docking predictions. The van der Waals repulsive term is significantly smaller than the van der Waals attractive term; this allows for the softness required when evaluating the rigid body predictions (in addition to the short range linearization used for the repulsive term). Thus, hits with some degree of clash (but compensated by favorable energies in the other terms) can be ranked well.

Another notable part of the weights is that the repulsive electrostatics terms are greater than the attractive terms. This may help to filter out predictions with unfavorable (repulsive) electrostatics; electrostatic repulsion could preclude the formation of the encounter complex between the proteins in that configuration. In addition, a recent study has noted that there is an "asymmetric screening" of charged spheres in the presence of a dielectric,[26] wherein the repulsive force is seen to be stronger than the attractive force. This supports the greater weight of the repulsive electrostatics terms obtained for ZRANK, and the effectiveness of the weighted scoring function when reranking. The long-range electrostatic terms are larger than their short-range counterparts, this may be due to the necessity of long-range electrostatics to form an encounter complex, and it may also be due to more noise in the short-range terms (e.g., due to side chain movements upon binding).

The ZRANK algorithm was developed to explore the limits of scoring rigid-body predictions, to see if the energy terms from initial-stage docking predictions can allow hits to be discriminated and well-ranked. Recent studies have indicated that certain portions of interfaces in protein–protein interactions do not change significantly between bound and unbound conformations. Camacho et al. have found that there are "anchor residues" and corresponding recognition motifs on the binding partner that are very important in the binding process.[27] Another study performed molecular dynamics simulations on a set of protein–protein interfaces, and found that many interfaces have a core region that is less mobile than the periphery of the interface.[28] It is therefore possible that through careful examination of the rigid-body docking predictions these important characteristics can be discriminated, without using side chain and/or backbone movement of the predictions.

As with any scoring program, the success of ZRANK is limited by the quality of the input. This is illustrated to some extent in Figure 4, with regard to the RMSDs of the predictions from ZDOCK. In addition, conformational change upon binding (which is essentially the lower bound of the rigid-body prediction RMSD) of the docking cases limits the success of the initial-stage searching; only about half of the medium difficulty cases in Benchmark 2.0 had any hits from ZDOCK. However, it can be seen both in the benchmark and in the CAPRI docking experiment that many proteins exhibit small conformational change upon binding and are thus within the reach of rigid-body docking. As for the remaining cases, work is being performed to modify ZDOCK to allow for better predictions of cases with conformational change. This includes addition of a statistical potential recently developed using a nonredundant set of transient protein interfaces.[29]

While the scoring scheme presented in this article has been developed and tested for reranking protein docking predictions, it is possible that other uses can be employed. For instance, the scores of a set of predictions can be analyzed in terms of whether there is any near-native prediction at all. Preliminary results indicate that this would require additional information, such as binding energy of the proteins, and a classification tool (e.g., a Support Vector Machine), but such a method, if effective, would certainly be useful for analyzing protein docking output. In addition to this, the ZRANK scoring can also be combined with clustering to help to eliminate redundancy and false positive predictions. One group has recently used ZDOCK scores to help determine the $\Delta\Delta G$'s from mutations at a protein–DNA interface[30]; it is possible that ZRANK scores can be used in this context as well.

The performance of ZRANK indicates that it is possible to greatly increase the ranking results of a set of rigid-body docking predictions without structural refinement. Thus the number of well-ranked hits can be enriched after initial-stage docking, prior to further refinement and/or analysis of the predictions. This is particularly useful before running detailed refinement programs, which can take hours to refine single predictions. Recent research in protein–protein docking includes explicitly modeling flexible protein backbones[31]; as refinement using these methods is computationally intensive and would be limited in the amount of predictions that can be accommodated, this underscores the need for retaining near-native predictions from initial-stage docking among the top ranks.

Future work includes combining ZRANK with side chain repacking and/or energy minimization of predictions. It can be applied in an iterative manner to rerank the predictions before structural refinement (to increase the number of near-native structures to be refined) and after the structural refinement to perform a final reranking. This will allow for even greater success in producing accurate structural models of protein complexes.

ACKNOWLEDGMENTS

We are grateful to the scientific computing facilities at Boston University and the Advanced Biomedical Computing Center at NCI, NIH for computing support.

REFERENCES

1. 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.
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. Jones S, Thornton JM. Principles of protein–protein interactions. Proc Natl Acad Sci USA 1996;93:13–20.
5. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 1997;272:106–120.
6. Vakser IA. Protein docking for low-resolution structures. Protein Eng 1995;8:371–377.
7. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87.
8. 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.
9. Ritchie DW, Kemp GJ. Protein docking using spherical polar Fourier correlations. Proteins 2000;39:178–194.
10. 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.
11. 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.
12. Fischer D, Lin SL, Wolfson HL, Nussinov R. A geometry-based suite of molecular docking processes. J Mol Biol 1995;248:459–477.
13. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 2004;20:45–50.
14. 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.
15. Jackson RM, Gabb HA, Sternberg MJ. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 1998;276:265–285.
16. Chen R, Tong W, Mintseris J, Li L, Weng Z. ZDOCK predictions for the CAPRI challenge. Proteins 2003;52:68–73.
17. Wiehe K, Pierce B, Mintseris J, Tong WW, Anderson R, Chen R, Weng Z. ZDOCK and RDOCK performance in CAPRI rounds 3, 4, and 5. Proteins 2005;60:207–213.
18. Li L, Chen R, Weng Z. RDOCK: refinement of rigid-body protein docking predictions. Proteins 2003;53:693–707.
19. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983; 4:187–217.
20. 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.
21. Neria E, Fischer S, Karplus M. Simulation of activation free energies in molecular systems. J Chem Phys 1996;105:1902–1921.
22. Zhang C, Vasmatzis G, Cornette JL, DeLisi C. Determination of atomic desolvation energies from the structures of crystallized proteins. J Mol Biol 1997;267:707–726.
23. Chen R, Weng Z. A novel shape complementarity scoring function for protein–protein docking. Proteins 2003;51:397–408.
24. Chen R, Mintseris J, Janin J, Weng Z. A protein–protein docking benchmark. Proteins 2003;52:88–91.
25. Press WH. Numerical recipes in C: the art of scientific computing. New York: Cambridge University Press; 2002.xxvi,994 p.
26. Doerr TP, Yu Y-K. Electrostatics of charged dielectric spheres with application to biological systems. Phys Rev E: Stat Nonlinear Soft Matter Phys 2006;73:061902–061914.
27. Rajamani D, Thiel S, Vajda S, Camacho CJ. Anchor residues in protein–protein interactions. Proc Natl Acad Sci USA 2004; 101:11287–11292.
28. Smith GR, Sternberg MJ, Bates PA. The relationship between the flexibility of proteins and their conformational states on forming protein–protein complexes with an application to protein–protein docking. J Mol Biol 2005;347:1077–1101.
29. Mintseris J, Weng Z. Optimizing protein representations with information theory. Genome Inform 2004;15:160–169.
30. Fanelli F, Ferrari S. Prediction of MEF2A-DNA interface by rigid body docking: a tool for fast estimation of protein mutational effects on DNA binding. J Struct Biol 2006;153:278–283.
31. Bonvin AM. Flexible protein–protein docking. Curr Opin Struct Biol 2006;16:194–200.