Consensus scoring for enriching near-native structures from protein-protein docking decoys

Shide Liang[a,b,c], **Samy O. Meroueh**[b], **Guangce Wang**[c], **Chao Qiu**[c], and **Yaoqi Zhou**[a,b,*]

a *Indiana University School of Informatics, Indiana University-Purdue University, Indianapolis, Indiana 46202, USA*

b *Department of Biochemistry and Molecular Biology and Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA*

c *Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, P. R. China*

Abstract

The identification of near native protein-protein complexes among a set of decoys remains highly challenging. A strategy for improving the success rate of near native detection is to enrich near native docking decoys in a small number of top ranked decoys. Recently, we found that a combination of three scoring functions (energy, conservation and interface propensity) can predict the location of binding interface regions with reasonable accuracy. Here, these three scoring functions are modified and combined into a consensus scoring function called ENDES for enriching near native docking decoys. We found that all individual scores result in enrichment for the majority of 28 targets in ZDOCK2.3 decoy set and the 22 targets in Benchmark 2.0. Among the three scores, the interface propensity score yields the highest enrichment in both sets of protein complexes. When these scores are combined into the ENDES consensus score, a significant increase in enrichment of near-native structures is found. For example, when 2000 dock decoys are reduced to 200 decoys by ENDES, the fraction of near-native structures in docking decoys increases by a factor of about six in average. ENDES was implemented into a computer program that is available for download at http://sparks.informatics.iupui.edu.

Keywords

Protein Docking; Near Native Decoy Selection; Energy Score; Interface Propensity; Conservation Score

INTRODUCTION

While significant progress has been accomplished in accurately docking small molecules to biological receptors, the accurate prediction of the three-dimensional structure of protein-protein complexes starting from the structures of the binding partners remains highly challenging[1–3]. The docking process consists of sampling a large number of possible conformations of the protein-protein complex (docking decoys), until the structure that most closely resembles that of the native structure (near-native structure) is identified. The

identification of these near-native structures remains the most challenging aspect of protein-protein docking. It is usually accomplished through the use of scores. The efficacy of these scores is measured by their ability to discriminate near-native structures from decoys that are generated during molecular docking.

One obvious approach to improve near-native identification is to produce more near-native structures by sampling around known binding regions[1,4–6]. This would rely on the availability of methods that can accurately identify protein-protein binding sites (interface residues) from known unbound monomer structures[7–10], such that the docking process can be restricted to those sites. This will significantly reduce the number of degrees of freedom and hence improve the accuracy. However, binding-site predictors[11] are not yet sufficiently reliable to be used directly in protein-protein docking.

An alternative approach for improving near-native identification is to use predicted binding sites as a filter(scoring) to screen docking decoys[11–17]. These filters are mostly based on evolution information obtained from multiple sequence alignment or calculated desolvation energy upon binding. Recently, we have developed a method referred to as PINUP[10], which predicts binding residues based on three properties of amino acid residues: residue energy, conservation and interface propensity. PINUP was found to be one of the most accurate methods in a recent study[11]. In this work, we introduce a scoring function called ENDES (Empirical Near-native Docking-decoy Enrichment Score) that uses a combination of the three aforementioned properties for enriching near-native docking decoys.

METHODS

The scoring function

The proposed scoring function for near-native enrichment is a consensus score consisting of a linear combination of residue-energy score E_{residue}, residue-conservation score E_{consv} and residue-interface propensity score $E_{\text{propensity}}$,

$$E = E_{\text{residue}} + w_c E_{\text{consv}} + w_p E_{\text{propensity}}, \tag{1}$$

where w_c and w_p are to-be-determined weight factors.

The conservation score for a given sequence position i is given by

$$E_{\text{consv}}(i) = \begin{cases} M_{ir} - B_{rr}, & M_{ir} - B_{rr} > 0, \\ 0, & M_{ir} - B_{rr} \leq 0, \end{cases} \tag{2}$$

where M_{ir} is the self-substitution score in the position-specific substitution matrix generated from PSIBLAST[18] for the residue type r at sequence position i and B_{rr} is the diagonal element of BLOSUM62[19] for residue type r. This conservation score[10] is averaged for all the interface residues. Interface residues are surface residues of monomers whose solvent accessible surface areas are decreased by $>1\text{Å}^2$ upon complex formation.

The residue energy score is defined as

$$E_{\text{residue}}(i) = \begin{cases} 0, & E_r^{\text{sidechain}}(i) = Min(E_j^{\text{sidechain}}(i)), \\ 1, & E_r^{\text{sidechain}}(i) \neq Min(E_j^{\text{sidechain}}(i)), \end{cases} \tag{3}$$

where $E_j^{\text{sidechain}}(i)$ is the sidechain energy when a residue type j is placed in sequence position i. The exact expression for $E_j^{\text{sidechain}}(i)$ can be found in Eq. (3) in PINUP[10]. It is calculated from the sidechain energies of all possible rotamers for a given residue type at a sequence position while other sequence positions have native residue types and observed atomic coordinates. The residue energy at sequence position i, $E_{\text{residue}}(i)$, is 1 if the native residue type (r) does not have the lowest sidechain energy. The energy scores of all the interface residues are calculated and averaged. Here, we assume that the residues are structurally important if native residue type has lower side-chain energy than other residue types. Otherwise, the residues could be of functional importance and interface residues[20].

We define interface propensity as follows:

$$E_{\text{propensity}} = \frac{\sum_i S_i \ln P_r^{\text{interface}} / P_r^{\text{surface}}}{\sum_i S_i} \tag{4}$$

where S_i is the accessible surface area of interface residue r at the sequence position i and $P_r^{\text{interface}}$ and P_r^{surface} are the contribution of residue type r to the interface area and to the protein surface area, respectively. The above scoring function is summed over all interface residues of both proteins in a complex structure (decoy or native).

In PINUP, the values of $\ln P_r^{\text{interface}} / P_r^{\text{surface}}$ for the 20 amino acid residues are derived from 75 complexes collected previously[21]. These complexes contain antigen-antibody complexes and homologous chains. In this work, we obtain $P_r^{\text{interface}}$ and P_r^{surface} by statistical analysis of 62 complexes of enzyme/inhibitor and other proteins in protein docking benchmark 2.0[22]: 1AVX, 1AY7, 1BVN, 1CGI, 1D6R, 1DFJ, 1E6E, 1EAW, 1EWY, 1EZU, 1F34, 1HIA, 1MAH, 1PPE, 1TMQ, 1UDI, 2MTA, 2PCC, 2SIC, 2SNI, 7CEI, 1A2K, 1AK4, 1AKJ, 1B6C, 1BUH, 1E96, 1F51, 1FC2, 1FQJ, 1GCQ, 1GHQ, 1HE1, 1I4D, 1KAC, 1KLU, 1KTZ, 1KXP, 1ML0, 1QA9, 1RLB, 1SBB, 2BTF, 1ACB, 1KKL, 1GP2, 1GRN, 1HE8, 1I2M, 1IB1, 1IJK, 1K5D, 1M10, 1N2C, 1WQ1, 1ATN, 1DE4, 1EER, 1FAK, 1FQ1, 1H1V, and 1IBR. Results are shown in Table 1. The correlation coefficient between this set and the set from the 75 protein-protein complexes used in PINUP is 0.92. This suggests that excluding antigen-antibody and homologous complexes from the calculation of residue interface propensity only leads to small changes.

Obtaining interface propensity from this small number of 62 protein-protein complexes is a result of the limited number of transient complex structures that are available from the protein databank. Transient complex structures are defined as those complexes whose unbound monomeric structures are also available. We also obtained residue interface propensity with slightly different structural databases to explore the dependence of the ENDES score on interface propensity (See below).

The docking decoy set

The ZDOCK2.3 decoy set of Benchmark1.0, which consists of 48 protein-protein complexes is downloaded from http://zlab.bu.edu/zdock/decoys.shtml. Each protein complex contains 2000 docking decoys[23]. These decoy sets were obtained from (i) docking unbound structures to each other or (ii) docking unbound to bound structures. The translation step is set to 1.2Å and rotation step is set to 6 degrees. Since our protein binding site prediction method relies on conservation information and is not suitable for predicting antigen-antibody interfaces, 16 antigen-antibody targets were excluded in this work. Among the remaining 32 targets, 4 targets with no near-native decoys are also not considered, resulting in a set of 28 protein-protein

complexes for training and cross validation. It is worth mentioning that a near-native structure is a structure with an interface rmsd less than 2.5Å. An interface rmsd between a docking decoy and a native complex structure is based on the C_a atoms of interface residues.

In addition, we downloaded ZDOCK2.3 docking decoys for Benchmark 2.0 (http://zlab.bu.edu/zdock/decoys.shtml) to further validate the ENDES consensus score. Among the 62 complexes in Benchmark 2.0 (excluding antigen-antibody), 42 complexes possess a sequence identity of 50% or less with any of the 28 targets from the aforementioned training and cross-validation decoy set. Among the remaining 42 complexes, only 15 have near native decoys among the top 2000 decoys and hence only these complexes can be used as a test set for ENDES. Thus, to further expand the number of complexes for testing, we also examine the top 20000 decoys, rather than top 2000 decoys, for Benchmark 2.0. This leads to an additional seven proteins with near-native structures.

Assessment and weight optimization

We assess the performance of the ENDES score based on an the enrichment factor $F_{\text{enrichment}}$ that is defined as

$$F_{\text{enrichment}} = \frac{[N_{\text{near-native}}^{\text{final}}/N_{\text{decoy}}^{\text{final}}]}{[N_{\text{near-native}}^{\text{initial}}/N_{\text{decoy}}^{\text{initial}}]}$$

(5)

where $N_{\text{decoy}}^{\text{initial}}$ and $N_{\text{decoy}}^{\text{final}}$ are the number of initially assessed decoys and the number of finally selected decoys using the ENDES score, respectively, and $N_{\text{near-native}}^{\text{initial}}$ and $N_{\text{near-native}}^{\text{final}}$ are the number of near native structures in the initial and final sets of docking decoys, respectively.

The ENDES score [Eq. (1)] is used to rank the docking decoys for each target. Two weight factors (w_c and w_p) in Eq. (1) are optimized to yield the highest enrichment factor with $N_{\text{decoy}}^{\text{initial}}=2000$ and $N_{\text{decoy}}^{\text{final}}=200$. The two weights are obtained by a simple grid search with a step size of 0.2. We did not use a finer step size of 0.1 because the difference is small between the results from two neighboring weight factors.

RESULTS

Training and cross validation

All complexes from the 28 targets are scored using either individual scores, such as energy, conservation and propensity, or consensus scores, which consist of the linear combination of the individual scores. The complexes are then ranked based on each of these scores. In each case, the top 200 scoring complexes are selected and used to determine the enrichment factors of near native decoys as shown in Table 2. The data reveals that among the individual scores, the interface propensity exhibits the highest enrichment factor of 5.9, followed by the conservation score ($F_{\text{enrichment}} = 5.0$). It appears that the energy score leads to significantly less enrichment. This was also observed when the energy score is combined with other scoring functions, as evidenced by the small change in the enrichment factor when the term is dropped from the consensus score. When two weight factors are optimized for all 28 targets, an enrichment factor of 6.9 is achieved. w_c and w_p are found to be 0.6 and 1.8, respectively.

It is of interest to identify how enrichment factors vary within a given number of top decoys. As Table 3 shows, the change of enrichment factors is small for energy scores, while the conservation score yields higher enrichment factors than the propensity score (assuming that the top 10 or 20 decoys are selected). This is mainly due to 2SNI (Table 4). For 2SNI, there is

only 1 near native decoy in the total 2000 decoys. The near native decoy is ranked 34th by propensity score and 3rd by conservation score. The enrichment factor will be either 0 or 200 if top 10 decoys are selected. This will affect average enrichment factor significantly. The three-term combined score, on the other hand, consistently produces the highest enrichment factor at any given number of top decoys.

The performance of the consensus ENDES score and ZDOCK for individual protein-protein complexes is illustrated in Table 4. The largest possible enrichment factor (10), in which all near natives are within top 200 decoys, is achieved for 11 complexes, compared to 0 by ZDOCK 2.3. The enrichment factor by ENDES is larger than 1 for all complexes except two, namely 1IGC and 1TAB. In the case of 1TAB, which corresponds to the crystal structure of trypsin-Bowman-Birk, the lack of enrichment can be attributed to the missing domains within the crystal structure. For ZDOCK 2.3, the enrichment factors of eight complexes are less than 1.

The increase in enrichment was achieved with ENDES through the optimization of two weight factors based on data from 28 complexes. It is unlikely that this training resulted in over-training in light of the small number of free parameters. To illustrate this fact, we performed a jack-knife test. This consisted of using 27 protein complexes to optimize two weight factors and then used the resulting scoring function to calculate the enrichment factor for the remaining protein complex. We find that the average enrichment factor is decreased only slightly from 6.9 to 6.4. In the leave-one-out-experiment, the best weights are 0.6–0.8, and 1.8–2.0 in 25 cases for w_c and w_p, respectively. This is similar to the weight factors of 0.6 and 1.8 that were obtained when all 28 protein complexes in the training step. We therefore conclude that the parameters that were obtained as well as the performance of the scoring function are robust.

A final check is performed to determine the effects of the more than 50% sequence identify between the 28 targets of the training set and some of proteins that were used to derive residue interface propensity. We find that when those monomers with more than 50% sequence identity with the monomers training set are removed during the calculation of the interface propensity, the enrichment factor by interface propensity alone decreases slightly from 5.9 to 5.4. This confirms that, the results are largely unaffected by the homologous nature of the proteins used to derive values for the residue interface propensity.

Testing

Table 5 shows the number of near-native decoys and enrichment factors given by individual and combined ENDES scores in the top 200 decoys for 15 protein complexes in Benchmark 2.0. Results of ZDOCK 2.3 are also shown in the table. ZDOCK 2.3 gives no enrichment for 8 complexes with an average enrichment factor of 1.8. By comparison, the number of protein complexes that did not result in enrichment are 8 by ENDES energy score, 10 by conservation score, 5 by interface propensity, and 3 by the consensus score. Hence, the interface propensity exhibits the highest enrichment capability as before, while the performance of the conservation score is lower than that of the energy score for this set of protein complexes. Nevertheless, the consensus score continues to show superior performance over individual scores for enrichment. In addition, the average enrichment factor of 5.8 is reasonably close to 6.4, which is the average enrichment factor from jack-knife cross validation of the 28 training complexes. If the weights of the three terms were trained with the 15 testing proteins, the enrichment factor was slightly improved from 5.8 to 6.3. Thus the weights trained using Benchmark 1.0 are reasonable even though the performance of single terms was somewhat different for the two sets of targets.

This test, however, is not completely an independent test because some monomers of the 15 complexes are used in generating the interface propensity. However, as we have demonstrated previously for the set of 28 targets, recalculated interface propensity with non-homologous proteins makes a small change to its enrichment capability. Indeed, the enrichment factor

decreases only somewhat from 4.6 to 3.8 for single term of interface propensity and from 5.8 to 5.4 for the combined score after the 15 complexes are removed from the calculation of interface propensity.

To further test the proposed scoring function, we examined the top 20000 decoys, rather than top 2000 decoys, for Benchmark 2.0 so that there are seven additional proteins with near-native structures. The results of the enriching factors from 20000 decoys to 2000 decoys for all 22 (15+7) protein complexes are shown in Table 6. For seven additional complexes, it is a difficult set because the number of near-native structures in 20000 decoys is less than 10 for six of the seven proteins. None of these near-native decoys was selected by the ZDOCK score function in top 2000 decoys (enrichment of zero). ENDES with three combined terms continues to achieve a high average enrichment factor of 4.45 while the enrichment from individual scoring terms ranges from 1.8 for conservation score, 2.7 for propensity and 3.0 for energy score. This difficult set confirms the robustness of the proposed ENDES scoring function for near-native enrichment while the overall enrichment factor is 5.6 for all 22 protein complexes, similar to the enrichment factor from 2000 to 200 decoys.

DISCUSSION

In this work, the scoring functions used in PINUP[10] are modified for enriching near native docking decoys. We showed that all individual scores including energy, interface propensity, and conservation are capable of enriching near native decoys. Among these, interface propensity leads to the highest enrichment. We find that the average enrichment factor is around 6 for the consensus score, either from Jackknife cross validation in 28 targets or from independent sets of 15 and 22 targets.

The average enrichment factor that we obtain is significantly higher than those obtained in a recent study[13]. The authors of that study employed conservation score, shape complementarity, residue pair potential, and desolvation score to filter out from 56% to 86% generated decoys. Forty eight out of the 55 targets with near-native decoys in the original decoy set have an enrichment factor of 1 or above. There are only 4 targets with an enrichment factor 4 or above. The comparison, however, is only approximate because different docking decoys are involved. It is difficult for us to compare with other studies because most do not report detail on the intermediate step of filtering.

We further applied ENDES to the Rosetta docking decoy sets[24]. The average enrichment factor was 1.4 for the 32 targets of enzyme-inhibitor group and other proteins in selecting top 100 decoys from a total of 1000 decoys. For six difficult targets with significant binding-induced conformational changes, the average enrichment factor was 1.9. This lower enrichment factor when compared to that obtained for ZDOCK decoys is in part due to large number of near native decoys in the Rosetta docking decoy sets. Rosetta docking decoys were generated by perturbation around the native conformation. On average, 31% decoys (global rmsd<10Å) for the 32 targets are near native (12% are near native for the 6 difficult targets). By comparison, an average of 3.2% decoys is near native decoys in ZDOCK 2.3 decoys because they were generated by global search (without any knowledge about native binding interfaces).

Less enrichment for higher quality decoy sets is also true for ZDOCK 3.0 decoys[25] (http://zlab.bu.edu/zdock/decoys.shtml). The enrichment factors from 20000 to 2000 decoys are only 1.4, 1.7, and 0.65 for the single term of energy score, conservation score, and interface propensity, respectively. This result is based on 20 targets with near native structures in top 20000 decoys given by ZDOCK 3.0. By comparison, the enrichment factor given by ZDOCK 3.0 is 3.1. Interestingly, interface propensity has a negative enrichment effect (enrichment factor<1) for the 20 target. This is likely due to the fact that ZDOCK3.0 uses a statistical

potential based on observed atomic contacts between two partner proteins and expected atomic contacts at protein surfaces in sampling[25]. That is, the decoys generated by ZDOCK 3.0 have already been filtered with consideration of "interface propensity" and thus, further filtering with our interface propensity leads to negative enrichment. Indeed, the optimized weights for three scores are 1, 1.4, and −1 for energy, conservation, and residue interface propensity scores, respectively. The resulting average enrichment factor is 3.0, comparable to 3.1 given by ZDOCK 3.0. Without interface propensity, the enrichment factor is 2.3 for a combination of energy score and conservation score (weights=1,1.4). Nevertheless, ZDOCK2.3+ENDES finds more near native poses in the enriched 2000 decoys of 12 targets, less of 6 than in top 2000 decoys generated by ZDOCK 3.0. The prediction results are the same for the remaining two targets. On the other hand, for 15 targets in Table 5, ZDOCK2.3+ENDES finds more near native poses in the enriched 200 decoys of 5 targets, less of 8 than in top 200 decoys generated by ZDOCK 3.0. The prediction results are the same for the remaining two targets.

One interesting question is whether the consensus scoring function (ENDES) can be used for selecting near natives directly. We have optimized the relative weights of the three terms to maximize the success rate in ranking one of the near-native decoys at the top. The optimized success rate is only 46% for the 28 targets, compared with 61% of EMPIRE[26]. Thus, ENDES, a coarse-grained scoring function which does not consider the interaction energy between two monomers, is suitable to use for enrichment only. To further illustrate this, we compare the best ranks of near native structures for 21 complexes given by ZRANK[27] and by ENDES in Table 7. Clearly, ZRANK ranks significantly more near native structures as number 1 than ENDES (6 versus 1). On the other hand, ENDES has only 3 targets whose best ranks of near-native structures are greater than 200, compared to 5 by ZRANK.

If ENDES can not be used for ranking near-native structures directly, is its enrichment useful for near native selections? We apply EMPIRE[26] to re-rank 200 top-ranked docking decoys by the ENDES scoring function. Re-ranking is performed following sidechain modeling and energy minimization, as done previously[26]. The success rate of ranking one of the near natives at the lowest binding free energy, increases from 17 out 28 (61%) (without enrichment) to 21 out 28 (75%) (with enrichment).

Another interesting question is the usefulness of the residue energy score. In Benchmark 1.0, the combination of conservation and propensity alone would yield an average enrichment factor of 6.7, essentially the same as 6.9 with the addition of residue energy score. However, the residue energy score plays a more important role in Benchmark 2.0. The enrichment factor from 20000 to 2000 (Table 6) will reduce from 5.6 to 4.9 without the energy score. This fluctuation is likely due to small number of targets in Benchmarks. Nevertheless, its poor performance relative to interface propensity may suggest the limitation of residue-level energy score function.

Acknowledgements

References

1. Mendez R, Laplae R, Lensink MF, Wodak SJ. Assessment of capri predictions in rounds 3–5 shows progress in docking procedures. Proteins 2005;60:150–169. [PubMed: 15981261]

2. Gray JJ. High-resolution protein-protein docking. Curr Opin Struct Biol 2006;16:183–193. [PubMed: 16546374]

3. Lensink MF, Mendez R, Wodak SJ. Docking and scoring protein complexes: CAPRI 3rd edition. Proteins 2007;69:704–718. [PubMed: 17918726]

4. Aloy P, Querol E, Aviles FX, Sternberg MJ. Automated structure-based prediction of functional sites in proteins: applications to assessing the validity of inheriting protein function from homology in genome annotation and to protein docking. J Mol Biol 2001;311:395–408. [PubMed: 11478868]

5. Zhang C, Liu S, Zhou Y. Docking prediction using biological information, ZDOCK sampling technique and clusterization guided by the DFIRE statistical energy function. Proteins 2005;60:314–318. [PubMed: 15981255]

6. deVries SJ, vanDijk AD, Bonvin AM. Whiscy: what information does surface conservation yield? application to data-driven docking. Proteins 2006;63:479–489. [PubMed: 16450362]

7. Jones S, Thornton JM. Prediction of protein-protein interaction sites using patch analysis. J Mol Biol 1997;272:133–143. [PubMed: 9299343]

8. Zhou H-X, Shan Y. Prediction of protein interaction sites from sequence profile and residue neighbor list. Proteins 2001;44:336–343. [PubMed: 11455607]

9. Neuvirth H, Raz R, Schreiber G. ProMate: a structure based prediction program to identify the location of protein-protein binding sites. J Mol Biol 2004;338:181–199. [PubMed: 15050833]

10. Liang S, Zhang C, Liu S, Zhou Y. Protein binding site prediction with an empirical scoring function. Nucleic Acids Res 2006;34:3698–3707. [PubMed: 16893954]

11. Zhou H-X, Qin S. Interaction-site prediction for protein complexes: a critical assessment. Bioinformatics 2007;23:2203–2209. [PubMed: 17586545]

12. Gottschalk KE, Neuvirth H, Schreiber G. A novel method for scoring of docked protein complexes using predicted protein-protein binding sites. Protein Eng Des Sel 2004;17:183–189. [PubMed: 15007163]

13. Duan Y, Reddy BV, Kaznessis YN. Physicochemical and residue conservation calculations to improve the ranking of protein-protein docking solutions. Protein Sci 2005;14:316–328. [PubMed: 15659366]

14. Heusear P, Bau D, Benkert P, Schomburg D. Refinement of unbound protein docking studies using biological knowledge. Proteins 2005;61:1059–1067. [PubMed: 16208723]

15. Tress M, deJuan D, Grana O, Gomez MJ, Gomez-Puertas P, Gonzalez JM, Lopez G, Valencia A. Scoring docking models with evolutionary information. Proteins 2005;60:275–280. [PubMed: 15981273]

16. Chelliah V, Blundell TL, Fernandez-Recio J. Efficient restraints for protein-protein docking by comparison of observed amino acid substitution patterns with those predicted from local environment. J Mol Biol 2006;357:1669–1682. [PubMed: 16488431]

17. 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. [PubMed: 14693807]

18. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Aci Res 1997;25:3389–3402.

19. Henikoff S, Henikoff J. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A 1992;89:10915–10919. [PubMed: 1438297]

20. Liang S, Zhang J, Zhang S, Guo H. Prediction of the interaction site on the surface of an isolated protein structure by analysis of side chain energy scores. Proteins 2004;57:548–557. [PubMed: 15382230]

21. Conte CL, Chothia L, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285:2177–2198. [PubMed: 9925793]

22. 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. [PubMed: 15981264]

23. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87. [PubMed: 12784371]

24. Gray J, Moughan S, Wang C, Schueler-Furman O, Kuhlman B, Rohl C, Baker D. Protein-protein docking with simultaneous optimization of rigid body displacement and side chain conformations. J Mol Biol 2003;331:281–299. [PubMed: 12875852]

25. Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, Weng Z. Integrating statistical pair potentials into protein complex prediction. Proteins 2007;69:511–520. [PubMed: 17623839]

26. Liang S, Liu S, Zhang C, Zhou Y. A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. Proteins 2007;69:244–253. [PubMed: 17623864]

27. Pierce B, Weng Z. ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 2007;67:1078–1086. [PubMed: 17373710]

Table 1

The values of $\ln P_r^{\text{interface}}/P_r^{\text{surface}}$ (lnP) for twenty amino acid residues.

Amino Acid	lnP	Amino Acid	lnP
ALA	−0.31	LYS	−0.49
ARG	0.12	MET	0.78
ASN	−0.1	PHE	0.75
ASP	−0.31	PRO	−0.29
CYS	0.47	TRP	0.72
GLN	−0.17	VAL	0.37
GLU	−0.51	SER	−0.22
GLY	0.06	THR	−0.07
ILE	0.58	TYR	0.87
LEU	0.42	HIS	0.25

Table 2

Comparison of averaged enrichment factor of the near native decoys in top 200 decoys given by individual and combined scores. Weight factors are optimized by all 28 targets.

Weight Factors			
Energy	Conservation	Propensity	Enrichment
1			1.35
	1		5.0
		1	5.9
1	1.8		5.2
1		3.6	5.9
	1.8	3.2	6.7
1	0.6	1.8	6.9

Table 3

The average enrichment factors for a given number of selected top decoys obtained from different scoring functions.

# Top Decoys	Energy	Conservation	Propensity	Combined
10	1.1	13.5	8.3	22.3
20	0.9	10.3	9.0	18.5
50	1.1	8.6	10.3	13.8
100	1.2	6.8	7.9	10.0
200	1.35	5.0	5.9	6.9
500	1.6	3.1	3.3	3.6
1000	1.5	1.8	1.9	1.9

Table 4

The number of near-native structures and the enrichment factor for individual protein complexes given by ZDOCK and the combined ENDES score.

	#Near Natives (Enrichment)[a]		
PDB ID	All[b]	ZDOCK	ENDES
1WQ1	54	19(3.5)	19(3.5)
1IGC	3	1(3.3)	0(0)
1ATN	24	5(2.1)	24(10)
1SPB	112	53(4.7)	25(2.2)
2BTF	35	16(4.6)	35(10)
1A0O	4	0(0)	3(7.5)
1CGI	76	10(1.3)	15(2.0)
1CHO	99	28(2.8)	99(10)
2PTC	48	1(0.21)	40(8.3)
1TGS	109	34(3.1)	51(4.7)
2SNI	1	0(0)	1(10)
2SIC	52	12(2.3)	52(10)
1CSE	29	1(0.34)	29(10)
2KAI	16	0(0)	14(8.8)
1BRC	54	11(2)	54(10)
1ACB	93	15(1.6)	93(10)
1BRS	21	2(0.95)	11(5.2)
1MAH	28	8(2.9)	23(8.2)
1UGH	20	4(2)	20(10)
1DFJ	51	33(6.5)	31(6.1)
1FSS	15	1(0.67)	9(6)
1AVW	52	17(3.3)	52(10)
1PPE	393	153(3.9)	144(3.7)
1TAB	50	3(0.6)	0(0)
1UDI	35	12(3.4)	35(10)
1STF	83	46(5.5)	57(6.9)
2TEC	185	81(4.4)	141(7.6)
4HTC	57	19(3.3)	18(3.2)
	Mean[c]	(2.5)	(6.9)

[a] Number of near native decoys (enrichment factor) in top 200 decoys ranked by ZDOCK and the ENDES combined score.

[b] Number of near native decoys in all 2000 decoys.

[c] Enrichment factors averaged over all complexes.

Table 5

The number of near native decoys and enrichment factors given by individual and combined ENDES scores as well as by ZDOCK in top 200 decoys for 15 protein complexes in Benchmark 2.0.

PDB ID	All[a]	Energy	Conserv.	Propensity	Combined	ZDOCK
				# Near-Natives (Enrichment)[b]		
1F51	17	0(0)	14(8.2)	6(3.5)	13(9.7)	8(4.7)
7CEI	185	0(0)	0(0)	76(4.1)	0(0)	73(3.9)
1B6C	6	0(0)	0(0)	5(8.3)	2(3.3)	3(5)
1BVN	52	19(3.7)	47(9.0)	51(9.8)	52(10)	7(1.3)
1E6E	18	18(10)	0(0)	0(0)	17(9.4)	1(0.56)
1KAC	1	1(10)	0(0)	0(0)	0(0)	0(0)
1EWY	11	7(6.4)	1(0.9)	5(4.5)	7(6.4)	1(0.91)
1TMQ	21	2(0.95)	20(9.5)	21(10)	21(10)	2(0.95)
1F34	16	0(0)	0(0)	5(3.1)	8(5)	8(5)
1AKJ	28	24(8.6)	2(0.7)	0(0)	10(3.6)	3(1.1)
1ML0	59	0(0)	0(0)	30(5.1)	12(2.0)	25(4.2)
1HE1	1	1(10)	1(10)	0(0)	1(10)	0(0)
1E96	2	0(0)	0(0)	2(10)	2(10)	0(0)
1GRN	6	0(0)	1(1.7)	0(0)	0(0)	0(0)
1RLB	2	1(5)	0(0)	2(10)	2(10)	0(0)
Ave.[c]		(3.6)	(2.7)	(4.6)	(5.8)	(1.8)

[a] Number of near native decoys in all 2000 decoys.

[b] Number of near native decoys (enrichment factor) in top 200 decoys ranked by individual and combined scores.

[c] Enrichment factors averaged over all 15 complexes.

Table 6

The number of near native decoys and enrichment factors given by individual and combined scores in top 2000 decoys for 22 protein complexes in Benchmark 2.0.

PDB ID	All[a]	# Near-Natives (Enrichment)[b]			
		Energy	Conserv.	Propensity	Combined
1F51	46	0(0)	41(8.9)	19(4.1)	35(7.6)
7CEI	306	0(0)	0(0)	135(4.4)	1(0.033)
1B6C	24	0(0)	6(2.5)	23(9.6)	16(6.7)
1BVN	132	37(2.8)	122(9.2)	128(9.7)	131(9.9)
1E6E	86	80(9.3)	2(0.23)	26(3)	78(9.1)
1KAC	5	4(8)	0(0)	0(0)	0(0)
1EWY	87	55(6.3)	3(0.34)	37(4.3)	49(5.6)
1TMQ	73	2(0.27)	68(9.3)	73(10)	73(10)
1F34	34	2(0.59)	0(0)	15(4.4)	21(6.2)
1AKJ	69	61(8.8)	10(1.4)	0(0)	28(4.1)
1ML0	116	0(0)	0(0)	79(6.8)	38(3.3)
1HE1	9	8(8.9)	9(10)	1(1.1)	9(10)
1E96	18	1(0.56)	2(1.1)	18(10)	18(10)
1GRN	41	0(0)	34(8.3)	0(0)	6(1.5)
1RLB	45	8(1.8)	0(0)	43(9.6)	39(8.7)
1QA9	3	0(0)	0(0)	0(0)	0(0)
1IJK	16	2(1.2)	0(0)	16(10)	16(10)
1FQJ	1	1(10)	0(0)	0(0)	0(0)
1KSD	8	3(3.8)	2(2.5)	1(1.2)	8(10)
1SBB	1	0(0)	0(0)	0(0)	0(0)
2PCC	2	1(5)	0(0)	1(5)	1(5)
1KTZ	8	1(1.2)	8(10)	2(2.5)	5(6.2)
Ave.	(All)[c]	(3.1)	(2.9)	(4.4)	(5.6)
Ave.	(Seven New)[d]	(3.0)	(1.8)	(2.7)	(4.45)

[a]Number of near native decoys in all 20000 decoys.

[b]Number of near native decoys (enrichment factor) in top 2000 decoys ranked by individual and combined scores.

[c]Enrichment factors averaged over all 22 complexes.

[d]Enrichment factors averaged over seven new complexes.

Table 7

The best ranks of near native structures given by ZRANK[27] and by ENDES for 6 targets in Benchmark 1.0 and 15 targets in Benchmark 2.0.

PDB	ZRANK[a]	ENDES[b]
1WQ1	2899	33
1CGI	34	10
2SNI	300	2
2SIC	1	1
1MAH	1	47
1DFJ	1	3
1F51	3	10
7CEI	1	> 200
1B6C	1	110
1BVN	14	4
1E6E	4	32
1KAC	94	> 200
1EWY	129	54
1TMQ	389	10
1F34	160	54
1AKJ	40	75
1ML0	1	76
1HE1	258	19
1E96	16	32
1GRN	1523	> 200
1RLB	4	12

[a]ZDOCK2.3 + ZRANK[27]. ZRANK was tested on Benchmark 2.0 and top 54000 decoys from ZDOCK 2.3. The first six targets are contained in Benchmark 1.0.

[b]The decoy sets of Benchmark 1.0 (first 6 targets) and Benchmark 2.0 (next 15 targets) with ZDOCK2.3 hits in the top 2000 predictions were enriched by ENDES to a small set of 200 decoys.