

Protein–Protein Docking: Overview and Performance Analysis

Kevin Wiehe, Matthew W. Peterson, Brian Pierce, Julian Mintseris, and Zhiping Weng

Summary

Protein–protein docking is the computational prediction of protein complex structure given the individually solved component protein structures. It is an important means for understanding the physicochemical forces that underlie macromolecular interactions and a valuable tool for modeling protein complex structures. Here, we report an overview of protein–protein docking with specific emphasis on our Fast Fourier Transform-based rigid-body docking program ZDOCK, which is consistently rated as one of the most accurate docking programs in the Critical Assessment of Predicted Interactions (CAPRI), a series of community-wide blind tests. We also investigate ZDOCK's performance on a non-redundant protein complex benchmark. Finally, we perform regression analysis to better understand the strengths and weaknesses of ZDOCK and to suggest areas of future development for protein-docking algorithms in general.

Key Words: Protein–protein docking; ZDOCK; RDOCK; Fast Fourier Transform; benchmark; CAPRI; shape complementarity; electrostatics; desolvation energy; regression analysis.

1. Introduction

Protein–protein interactions play a central role in biochemistry. This can be seen in cell-signaling cascades, enzyme catalysis, the immune response by means of antibody–antigen interactions, and the large-scale motions of organisms. These interactions are also implicated in many diseases.

From: *Methods in Molecular Biology*, Vol. 413: *Protein Structure Prediction*, Second Edition
Edited by: M. Zaki and C. Bystroff © Humana Press Inc., Totowa, NJ

While experimental techniques such as yeast two-hybrid system and mass spectrometry are able to determine the existence of protein–protein interactions, the structure of the macromolecular complex of two interacting proteins can provide additional information about their interaction, such as the specific residues involved in the interaction and the degree of conformational change undergone by the proteins upon binding.

X-ray crystallography and nuclear magnetic resonance have provided us with the structures of many complexes, but numerous structures still remain unsolved because of time and experimental limitations. This leads to a need for computational methods to understand the nature of protein–protein interactions, one of which is protein–protein docking.

This chapter is divided into three sections. The first section provides an overview of protein–protein docking and describes some of the available algorithms for docking. The second describes the ZDOCK suite of programs in detail, and the third describes an analysis of the performance of ZDOCK.

1.1. Protein–Protein Docking: An Overview

Protein–protein docking is defined as the prediction of the structure of two proteins in a complex, given only the structure of the interacting proteins. The “docking problem” can be broken down into two types of docking: bound docking, in which a complex is separated and reassembled, and unbound docking, where the structure of the complex is found from the individually solved structures of the interacting proteins. Obviously, bound docking has little applicable value, but it is often used for testing and verification purposes.

Unbound docking is much more difficult than bound docking because the proteins involved can change conformation upon binding. A study of conformational changes in protein complexes (*1*) showed that while the general model for protein–protein recognition is an induced fit model where the proteins must change conformation in order to bind, the amount of conformational change was small enough such that binding could be modeled as a “lock-and-key” mechanism as a first approximation. This allows for successful docking results even when there are noticeable changes in the conformation of the interacting proteins. This “rigid-body” approximation has been invaluable in the advancement of the protein–protein docking field. However, modeling induced fit by flexible docking remains a central challenge, and a large portion of current docking research is focused in this area.

There are two main challenges in the development of methods for protein–protein docking. The first is the construction of a scoring function that allows for the discrimination between correct or near-correct predictions and incorrect predictions. The second is the development of an algorithm that quickly searches and scores all possible orientations of the proteins to be docked. The most

obvious way to dock two proteins would be to simulate the molecular dynamics, as this would allow the complex to reach its native state with time. Unfortunately, the computational power necessary for such a simulation makes this currently intractable.

Protein–protein docking is often carried out in two stages. The initial stage treats the proteins as rigid bodies, allowing for an efficient search of the six-dimensional (6-D) space (three dimensions of translational freedom and three dimensions of rotational freedom). The 6-D space is searched for regions of high shape and biochemical complementarity, using a “soft” scoring function that allows for some clashes between atoms. A critical component of docking research has been the development of novel techniques for increasing the speed of the search. One of the most popular methods is the Fast Fourier Transform (FFT) (2), used in ZDOCK (3), FTDock (4), and GRAMM (5) to search translational space and in HEX (6) to search angular space. Other search methods that have been used include representing the proteins using grids of bits (7), Monte Carlo sampling (8,9), genetic algorithms (10), and geometric hashing (11).

Many docking algorithms have a refinement and re-ranking stage. This involves making small changes to the highest-scoring predictions from the initial stage using techniques such as 6-D rigid-body movements, molecular dynamics, and the clustering of similar predictions. Often, a more advanced scoring function, designed to increase the rank of near-native structures and decrease the rank of false positives, is introduced. This allows for a more descriptive approximation of biochemical properties such as desolvation free energy, electrostatics, and hydrogen bonding. Table 1 provides a list of current docking methods, along with their methodologies.

1.2. Measuring the Accuracy of Predicted Complexes

Once a prediction has been created, it is useful to evaluate it in a quantitative fashion. This is most often done using root mean square deviation (RMSD) between the atoms (using all atoms, backbone atoms, or C α atoms) of the prediction and the complex. This is done by first aligning the predicted structure with the crystallized complex in a manner that minimizes RMSD. RMSD between the predicted (p) and actual (a) C α atoms is calculated as follows (with n being the total number of atoms):

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n \left\{ [p_x(i) - a_x(i)]^2 + [p_y(i) - a_y(i)]^2 + [p_z(i) - a_z(i)]^2 \right\}} \quad (1)$$

Two of the most often used metrics for measuring the accuracy of a predicted structure are interface RMSD (iRMSD) and ligand RMSD (lRMSD). iRMSD

Table 1
A Summary of Docking Tools

Name	Method	Server/DL ^a	Reference
AutoDOCK	Flexible docking using Monte Carlo search and incremental construction	N/Y	12
BiGGER	Global search using bit mapping, rescored with multiple filters	N/N	7
ClusPro	Docking with DOT/ZDOCK, clustering	Y/N	13
DOCK	Global search with grid-based energy function, flexible docking with random search and incremental construction	N/Y	14
DOT	FFT global search using shape complementarity and electrostatics	Y/N	15
FTDock	FFT rigid-body search	N/Y	4
GRAMM	FFT with clustering and rescoring	Y/Y	5
HADDOCK	Rigid-body minimization, semi-flexible simulated annealing, rescoring using biochemical data	N/Y	16
HEX	Fourier correlation of spherical harmonics, explicit translational search	Y/N	6
ICM	Docking by combining pseudo-Brownian potential and torsional steps with local gradient minimization	N/N ^b	8
PatchDock/SymmDock	Geometric hashing and pose-clustering to score shape complementarity	Y/N	17
RosettaDock	Optimization of side-chain conformation with rigid-body Monte Carlo minimization	N/Y	9
ZDOCK	FFT search using shape complementarity, desolvation, and electrostatics. Refinement and re-ranking with RDOCK	Y/Y	18

FFT, Fast Fourier Transform.

^a Availability of download to academic users.

^b Browser can be downloaded; docking component must be purchased.

is defined as the C α RMSD of those residues having at least one atom within a distance cutoff of the interacting partner; IRMSD is calculated by superposing the receptor of the predicted structure with the known structure, performing the same transformation on the ligand, and calculating the C α RMSD of the ligand. An advantage of using iRMSD is that unlike IRMSD, it is not affected by conformational change in domains that do not include the binding site.

Often, a prediction is classified as a “hit” if the iRMSD and IRMSD are below a threshold. Unfortunately, this hard cutoff does not take into account many nuances. Another method of evaluating the accuracy of docking predictions is the fraction of native and non-native contacts (f_{nat} and $f_{\text{non-nat}}$). Contacts are defined as residue pairs with less than 5Å distance between the receptor and ligand. f_{nat} is a measure of the number of contacts correctly predicted, and $f_{\text{non-nat}}$ measures the number of incorrectly predicted contacts. $f_{\text{non-nat}}$ serves as an indication of atomic clash between the interface residues in the predicted complex and also as a proxy for conformational change, as residues may move into the interface upon binding.

1.3. The Critical Assessment of Predicted Interactions Experiment

The CAPRI (Critical Assessment of Predicted Interactions) experiment was created to compare the performance of docking algorithms of various groups (19). CAPRI was modeled after Critical Assessment of Structural Prediction (CASP), which started in 1994 to compare the performance of protein-folding algorithms (20).

CAPRI is a blind competition, so the participating groups do not receive the complex structure until after all predictions have been made. Each group submits 10 predictions, ranked by confidence. The predictions are then evaluated based on various factors and assigned a score [incorrect, acceptable (one star), medium (two stars), and high (three stars)] based on their accuracy. The CAPRI metrics for these scores are described by the Boolean expressions below:

$$\begin{aligned}
 \textit{High} &= (f_{\text{nat}} \geq 0.5) \cap [(IRMSD \leq 1.0) \cup (iRMSD \leq 1.0)] \\
 \textit{Medium} &= \{(f_{\text{nat}} \geq 0.3) \cap (f_{\text{nat}} < 0.5)\} \cap [(IRMSD \leq 5.0) \cup (iRMSD \leq 2.0)] \cup \\
 &\quad \{(f_{\text{nat}} \geq 0.5) \cap (IRMSD > 1.0) \cap (iRMSD > 1.0)\} \\
 \textit{Acceptable} &= \{(f_{\text{nat}} \geq 0.1) \cap (f_{\text{nat}} < 0.3)\} \cap [(IRMSD \leq 10.0) \cup (iRMSD \leq 4.0)] \cup \\
 &\quad \{(f_{\text{nat}} \geq 0.3) \cap (IRMSD > 5.0) \cap (iRMSD > 2.0)\} \quad (2)
 \end{aligned}$$

We have made predictions for all CAPRI targets, and **Table 2** summarizes our performance. As an example, **Fig. 1** shows the close resemblance between

Table 2
ZDOCK/RDOCK Performance in the CAPRI Experiment, Rounds 1–5

Target	Protein complex	Accuracy ^a	Contact % ^b	Rank ^c	Success rate ^d (%)
1	Hpr kinase-HPr	Incorrect	7.7	9	3.19
2	Rotavirus VP6-Fab	Medium	96.1	3	2.29
3	Hemagglutinin-Fab	Incorrect	59.7	7	2.19
	HC63				
4	Alpha-amylase-camelide	Incorrect	0.0	1	0.0
	AMD10				
5	Alpha-amylase-camelide	Incorrect	6.3	2	0.0
	AMB7				
6	Alpha-amylase-camelide	Incorrect	27.6	5	8.96
	AMD9				
7	T cell Receptor V(BETA)2-exotoxin A1	Medium	83.8	1	16.91
8	Nidogen-laminin	Medium	47	1	7.45
9	LicT homodimer	Incorrect	8	2	0.20
10	TBEV trimer	Incorrect	11	3	0.97
11	Cohesin-dockerin (model)	Acceptable	13	1	9.12
12	Cohesin-dockerin	High	84	1	10.95
13	SAG1-FAB	High	87	1	7.18
14	MYPT-PP1	High	53	8	16.06
18 ^e	GH11	Medium	91	1	2.15
	Xylanase-TAXI				
19	Ovine Prion-FAB	Medium	57	8	4.52

CAPRI, Critical Assessment of Predicted Interactions.

^a Accuracy, as scored by the CAPRI evaluation team based on interface root mean square deviation (RMSD), ligand RMSD, and percentage of correct contacts predicted.

^b Percentage of correct interface residue contact pairs predicted.

^c Rank, as assigned by ZDOCK team, of best prediction out of the 10 submission for that target.

^d A metric used to evaluate the success of predictions across all groups in CAPRI (21).

^e Targets 15–17 were canceled.

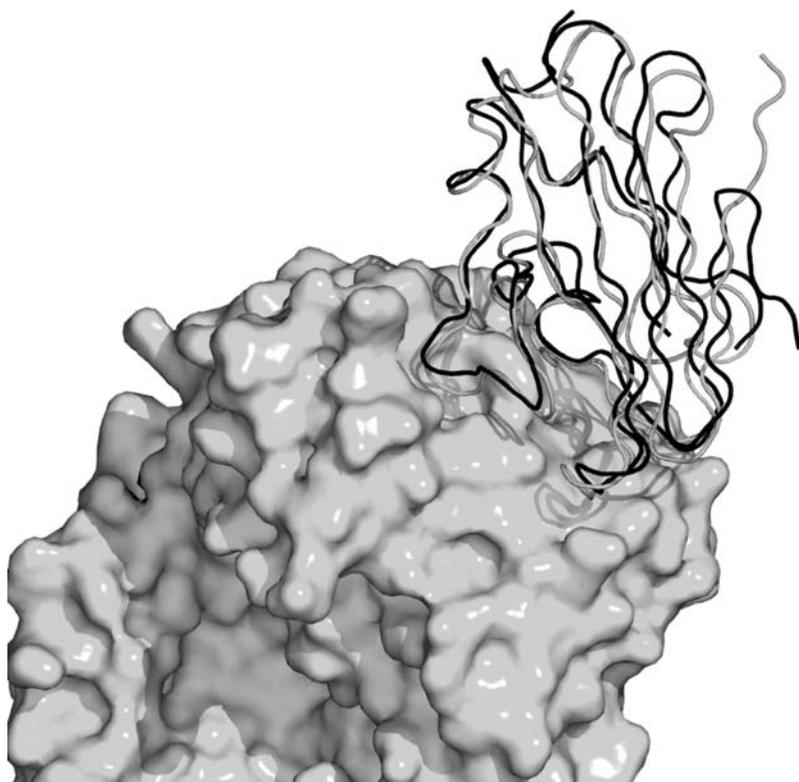


Fig. 1. Prediction of the structure of the SAG1–antibody complex [Critical Assessment of Predicted Interactions (CAPRI) Target 13]. The antibody of the prediction was superposed onto the crystal structure; the predicted SAG1 is in gray loops, whereas the crystal structure SAG1 is shown in black loops (the antibody is shown using surface representation). The non-binding domain of the SAG1 molecule is not shown. Pymol (22) was used to generate this figure.

our predicted structure and the crystal structure for Target 13 (SAG1–antibody complex).

1.4. A Benchmark for Protein–Protein Docking

In order to provide the docking community with a standard set of test cases to test docking algorithms, we developed two protein–protein docking benchmarks. The first benchmark, Benchmark 1.0 (23), contained 59 test cases, consisting of 22 enzyme–inhibitor complexes, 19 antibody–antigen complexes, 11 other complexes, and 7 difficult complexes. Of these complexes, 31 are

unbound–unbound, and 28 are bound–unbound. A number of groups have used this benchmark to test the performance of their docking algorithms (9,24–26).

A newer version of the docking benchmark, Benchmark 2.0 (27), has been created. It includes 84 test cases and was designed to focus on unbound–unbound test cases. Structural classification of proteins (SCOPs) (28) was used to avoid redundancy in the benchmark. This benchmark is classified by docking difficulty, based on the amount of conformational change undergone by the interacting proteins. Complexes classified as rigid and medium fall into the realm of rigid-body docking, whereas complexes classified as difficult would require algorithms that explicitly search backbone conformations.

2. The ZDOCK/RDOCK/M-ZDOCK Approach

2.1. ZDOCK: An FFT-Based Initial Stage Docking Algorithm

ZDOCK is an initial-stage docking algorithm that uses an FFT to find the three-dimensional (3-D) structure of a protein complex. The ZDOCK algorithm optimizes three parameters: shape complementarity, electrostatics, and desolvation free energy.

ZDOCK takes Protein Data Bank (PDB) (29) files as input. The larger of the two interacting proteins is considered the receptor (R), whereas the smaller of the two is considered the ligand (L). These PDB files are first parsed through the supplied program *mark_sur*, which measures the amount of accessible surface area (ASA) of each atom using a water probe of radius 1.4 Å. If an atom has an ASA of more than 1 Å², it is marked as a surface atom. *mark_sur* also marks the atom type for each atom in the structure, based on the 18 atom types based on atomic contact energy (ACE) (30). For any given rotational orientation, the L and R are both discretized onto a 3-D grid of size $N \times N \times N$ with a spacing of 1.2 Å. N must be large enough such that the grid can cover the sum of the maximal spans of R and L, plus 1.2 Å, and it is often set at 128.

2.1.1. The Fast Fourier Transform

As previously mentioned, the FFT is a popular method for quickly searching 3-D translational space. A diagram of the general FFT docking approach is detailed in Fig. 2. The search is performed by randomly perturbing both the receptor and ligand to avoid starting from a near-native state, and then discretizing them into discrete functions [$R(x, y, z)$ and $L(x, y, z)$ for the receptor and ligand, respectively] onto separate 3-D grids. ZDOCK searches rotational space explicitly by rotating the ligand in either 15° or 6° steps, which result in 3600 and 54,000 total angles, respectively. For each angle, only the

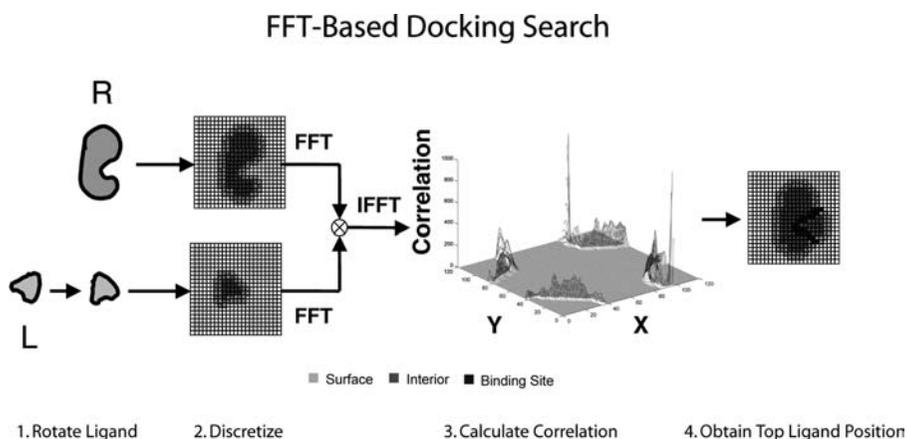


Fig. 2. The steps involved in a Fast Fourier Transform (FFT)-based docking search. For each ligand rotation, it is discretized and this discretization is then correlated with the discretized receptor to obtain the top-scoring ligand position. These steps are repeated to cover all ligand rotations in three dimensions, if necessary. In the case of ZDOCK, this involves 3600 iterations for 15° sampling and 54,000 iterations for 6° sampling.

top-scoring translation is found. To find the highest-scoring translation, we performed a cross-correlation. The correlation for a particular x, y, z translation (i, j, k) is found by taking the complex conjugate of the one of the functions, offsetting the grids, and multiplying the overlapping grid points together, with the sum of these products representing the score for that translation.

$$S(i, j, k) = \sum_{x, y, z} L^*(x + i, y + j, z + k) R(x, y, z) \quad (3)$$

Cross-correlations can be performed globally in a single step by working in the frequency domain. This is done using the Discrete Fourier Transform (DFT) and Inverse Fourier Transform (IFT):

$$S(i, j, k) = \frac{1}{N^3} IFT \{ IFT [L(x, y, z)]^* DFT [R(x, y, z)] \} \quad (4)$$

The FFT is a method for computing the DFT and IFT efficiently. Each FFT is $O(\log_2(N^3))$, whereas the multiplication of the grids is $O(N^3)$. Therefore, using the FFT to perform the translational search reduces the computational complexity of the search from $O(N^6)$ to $O(N^3 \log_2(N^3))$.

ZDOCK uses a combination of three physical and biochemical properties to describe ligand and receptor: shape complementarity, desolvation free energy, and electrostatics.

2.1.2. Shape Complementarity

The physical basis for shape complementarity comes from the van der Waals (vdW) potential. Atoms are subject to an attractive force at long distances, and a repulsive force at short distances, caused by the overlap of electronic orbitals. Most often, this is approximated by the Lennard–Jones 6–12 potential, shown below:

$$V_{L-J} = \frac{A}{r^{12}} - \frac{B}{r^6} \quad (5)$$

The r^6 term represents the attractive energy, whereas the r^{12} term represents the repulsive energy. The minimum of the vdW potential is found at the sum of the vdW radii, which can be thought of as the effective sizes of the interacting atoms.

Early versions of ZDOCK used a shape complementarity function known as grid-based shape complementarity (GSC) (3). Here, two discrete functions, R_{GSC} (GSC function for the receptor) and L_{GSC} (GSC function for the ligand), are used to describe the geometric characteristics of the two proteins as follows:

$$R_{GSC} = \begin{cases} 1 & \text{solvent-accessible surface} \\ 9_i & \text{solvent-excluding surface} \\ 9_i & \text{core} \\ 0 & \text{open space} \end{cases}$$

$$L_{GSC} = \begin{cases} 0 & \text{solvent-accessible surface} \\ 1 & \text{solvent-excluding surface} \\ 9_i & \text{core} \\ 0 & \text{open space} \end{cases} \quad (7)$$

The solvent-excluding surface layer is defined by the grid points marked as surface atoms by *mark_sur*, whereas the core is defined as the atoms not on the surface. The solvent-accessible surface layer is an additional layer of grid points surrounding the surface of the protein.

The current version of ZDOCK uses a complementarity function known as pairwise shape complementarity (PSC) (31). PSC is composed of a favorable term and a penalty term. The favorable term calculates the number of atom pairs between R and L within a distance cutoff D, whereas the penalty component of PSC is proportional to the number of overlapping grid points between

R and L, much like GSC. Whereas the GSC function results in grid spaces with purely real or imaginary values, the PSC function is complex. L_{PSC} and R_{PSC} are shown below.

$$\begin{aligned} \Re [L_{\text{psc}}] &= \begin{cases} 1 & \text{if nearest grid point to ligand atom} \\ 0 & \text{otherwise} \end{cases} \\ \Re [R_{\text{psc}}] &= \begin{cases} \text{Number of receptor atoms within } D = +\text{vdW radius} \\ \text{of nearest atom} & \text{open space} \\ 0 & \text{otherwise} \end{cases} \\ \Im [L_{\text{psc}}] = \Im [R_{\text{psc}}] &= \begin{cases} 3 & \text{solvent-excluding surface} \\ 9 & \text{core} \\ 0 & \text{open space} \end{cases} \end{aligned} \quad (8)$$

The use of PSC rather than GSC for scoring shape complementarity was shown to greatly increase the number of near-native predictions for Benchmark 1.0 during initial stage docking (31).

2.1.3. Desolvation Free Energy and Electrostatics

ACE (30) is used by ZDOCK to estimate desolvation free energy. ACE is defined as the change in free energy resulting from the breaking of two atom–water contacts and the formation of an atom–atom contact and a water–water contact. This is also referred to as the hydrophobic effect, which is known to play a critical role in protein–protein binding. ZDOCK introduces two discrete functions, L_{DE} and R_{DE} , to describe the desolvation energy of the ligand and receptor:

$$\begin{aligned} \Re [L_{\text{DE}}] = \Re [R_{\text{DE}}] &= \begin{cases} \text{PSC + ACE scores of all nearby atoms} & \text{open space} \\ 0 & \text{otherwise} \end{cases} \\ \Im [L_{\text{DE}}] = \Im [R_{\text{DE}}] &= \begin{cases} 1 & \text{if nearest grid point to atom} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (9)$$

The electrostatics energy term for ZDOCK can be expressed as a correlation between the electric potential generated by the receptor with the charges of the ligand atoms. ZDOCK adopts the Coulombic formula used by Gabb et al. (4) but incorporates partial charges using the CHARMM19 parameters from the CHARMM molecular mechanics program (32).

2.1.4. ZDOCK Scoring Function

There are two ZDOCK versions that use PSC to describe shape complementarity: ZDOCK 2.1 simply uses PSC as the scoring function, whereas ZDOCK

2.3 uses a linear combination of the shape complementarity-electrostatics score and the desolvation score. ZDOCK 2.3 incorporates PSC and electrostatics into single complex functions ($R_{\text{PSC+ELEC}}$ and $L_{\text{PSC+ELEC}}$) to improve computation time. These functions are described below:

$$\begin{aligned} \mathfrak{N}[L_{\text{PSC+ELEC}}] &= \mathfrak{N}[R_{\text{PSC+ELEC}}] = \begin{cases} 3.5 & \text{solvent-excluding surface} \\ 3.5^2 & \text{core} \\ 0 & \text{open space} \end{cases} \\ \mathfrak{S}[R_{\text{PSC+ELEC}}] &= \begin{cases} \beta^* \text{ electric potential of all R atoms} & \text{open space} \\ 0 & \text{otherwise} \end{cases} \\ \mathfrak{S}[L_{\text{PSC+ELEC}}] &= \begin{cases} -1^* \text{ atom charge} & \text{grid point closes to ligand atom} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (10)$$

2.2. RDOCK: Refining ZDOCK Predictions

The refinement stage of protein docking with ZDOCK is carried out using an algorithm known as RDOCK (33). Because of the soft scoring function in ZDOCK, many of the top-scoring predictions are false positives (not near-native). RDOCK refines these output structures through energy minimization. This is carried out in three steps, using CHARMM (32).

1. Removal of clashes by minimization of vdW and internal energies.
2. Minimization of total (Coulombic electrostatics, vdW, internal) energy, constraining non-hydrogen atoms, and keeping ionic side chains in their neutral states.
3. Minimization of total energy with no restrictions.

Once energy minimization has been performed, the minimized structures are re-ranked. Any complexes that still exhibit clashes (those that have vdW energy of 10 kcal/mol or greater) after minimization are discarded. Electrostatics and desolvation energy for the complexes are calculated using CHARMM and ACE, respectively. The RDOCK scoring function, $\Delta G_{\text{binding}}$, is a linear combination of desolvation score (ΔG_{ACE}) and electrostatic energy (ΔE_{elec}).

$$\Delta G_{\text{binding}} = \Delta G_{\text{ACE}} + 0.9^* \Delta E_{\text{ELEC}} \quad (11)$$

2.3. M-ZDOCK: Symmetric Multimer Docking with ZDOCK

The ZDOCK algorithm has been modified to predict the structure of C_n multimer complexes, in which two or more identical proteins interact, resulting in a ring-shaped complex. M-ZDOCK (34) reconstructs the multimer based on the optimal position of two adjacent monomers in a single plane. This leads to a reduction in computational time due to the reduced search space, as well as an increase in performance when compared with docking C_n multimers with ZDOCK.

2.4. ZDOCK Performance on Benchmark 2.0

ZDOCK was tested against version 2.0 of the docking benchmark using ZDOCK 2.3 and ZDOCK 2.1, with 6° and 15° angular sampling.

2.4.1. Prediction Evaluation

To evaluate the structure predictions produced by ZDOCK, we used the RMSD of the interface C α atoms. Interface C α atoms were identified by selecting residues that had any atom within 10 Å of the other molecule in the bound complex. A hit was defined as a prediction with an iRMSD ≤ 2.5 Å.

Two measures are defined to evaluate the average performance of a docking algorithm over the entire benchmark. *Success rate* is defined as the percentage of test cases that have a hit in the top N predictions. *Average hit count* is the number of hits for all test cases in the top N predictions, divided by the number of test cases.

2.4.2. Running ZDOCK

Several considerations were taken before and while running ZDOCK. To remove bias from the starting positions (the Benchmark 2.0 unbound test cases are by default aligned to the bound proteins, to facilitate the evaluation of predicted structures), we used a different random seed to rotate the ligand for each case. In addition, the antibodies (apart from the camelid 1KXQ) had most of their non-complementarity-determining region (non-CDR) loops blocked to avoid false-positive predictions. The CDRs of the antibodies were identified using their sequences (loops L1, L2, L3, H1, H2, and H3) and by examination of the structures (loops L4 and H4 and the N-termini).

2.4.3. Success Rate and Hit Count

Figure 3 shows the success rate for ZDOCK when run against all rigid-body cases from Benchmark 2.0. It can be seen that ZDOCK 2.3 performs better overall than ZDOCK 2.1 in terms of success rate. This is because the scoring function used in ZDOCK 2.3 is better at discriminating hits against incorrect predictions across the benchmark. Also, for both ZDOCK 2.1 and ZDOCK 2.3, the 15° sampling has a higher success rate than the 6° sampling. This indicates that for more predictions (i.e., finer sampling), there are more false positives introduced that reduce the rank of the first hit in some of the test cases. However, the 6° sampling is superior with regard to the number of hits, indicated by the hit count plot.

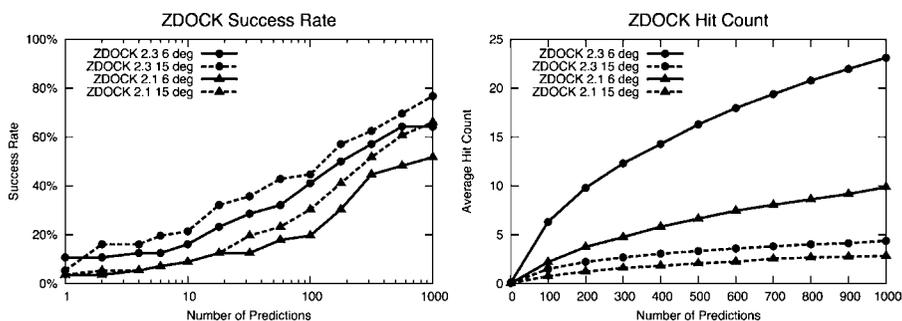


Fig. 3. ZDOCK success rate (left) and hit count (right) for the rigid-body test cases of Benchmark 2.0, for $N = 1$ through 1000 predictions. The success rate is the percentage of test cases with a hit in the top N , whereas the hit count is the number of hits for all test cases within the top N divided by the number of cases. Hits are defined as predictions that have an interface root mean square deviation (RMSD) $\leq 2.5 \text{ \AA}$, as described in the text.

Also in **Fig. 3** is the average hit count for the four ZDOCK modes tested. In this plot, it is clear that ZDOCK 2.3 with 6° sampling is the best, followed by ZDOCK 2.1 with 6° sampling. The greater number of hits produced by ZDOCK 2.3 with 6° sampling make this the best option for following up with re-ranking and refinement of the top predictions (e.g., $N = 2000$), as suggested by us earlier (33).

2.4.4. ZDOCK Performance by Test Case Category

Figure 4 gives the success rate curves for ZDOCK across the three types of test cases in Benchmark 2.0: Enzyme/Inhibitor, Antibody/Antigen, and Others (the latter is defined as those cases that fall into neither of the first two categories).

ZDOCK has the best success rate for the Antibody/Antigen test cases, with the success rate at 1000 predictions with 95% success for ZDOCK 2.3 15° sampling. This may be partly because approximately half of the Antibody/Antigen cases use bound forms of the antibody; thus the interface conformational change of these cases is on average smaller. In addition, the search space is reduced by blocking the non-CDR portions of the antibody.

The Enzyme/Inhibitor cases did not match the Antibody/Antigen cases in terms of success rate at 1000 predictions, although for the top predictions (i.e., small N), ZDOCK performed better for this category. Most notable is

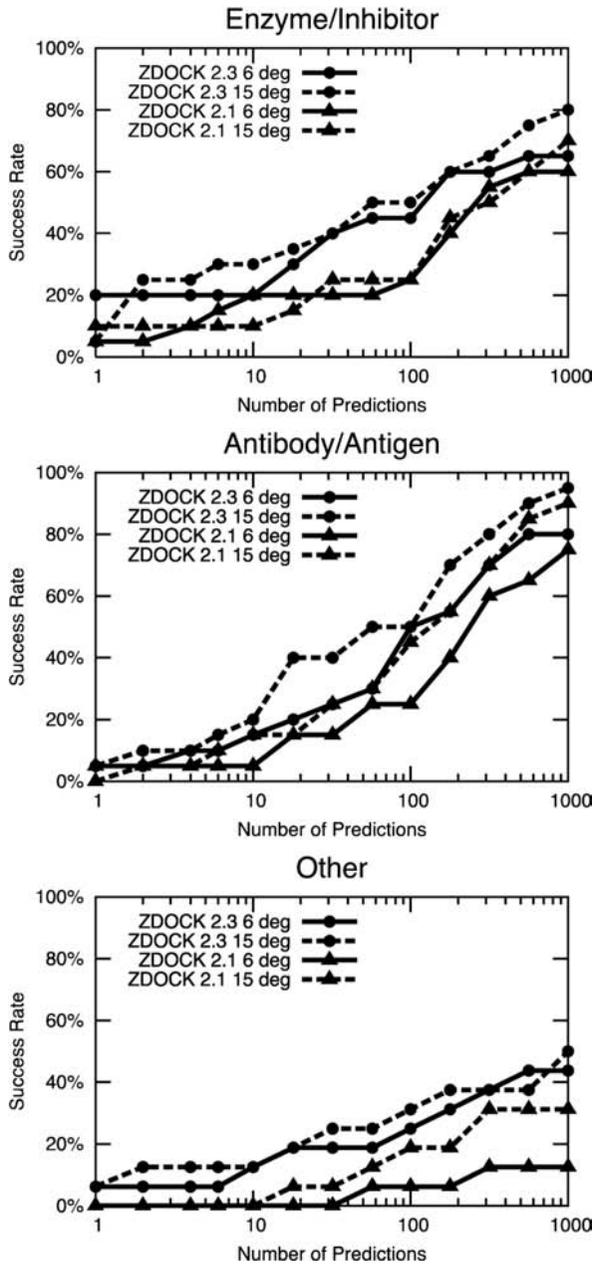


Fig. 4. ZDOCK success rate for Benchmark 2.0 cases, broken down by category: Enzyme/Inhibitor, Antibody/Antigen, and Other.

the 20% success for ZDOCK 2.3 15° sampling for the top prediction (four of 20 cases). This may be due to the PSC scoring function, which when combined with desolvation and electrostatics (as in ZDOCK 2.3) is well suited to identify the pocket-shaped binding sites on the enzymes.

ZDOCK did not perform quite as well on the Others test cases; this was also seen when ZDOCK was run against Benchmark 1.0 Others test cases. Of the four ZDOCK options tested, both sampling levels of ZDOCK 2.3 performed better than ZDOCK 2.1. In fact, at $N = 1000$, ZDOCK 2.3 still performed better than ZDOCK 2.1, whereas for Enzyme/Inhibitor and Antibody/Antigen cases, the ZDOCK 2.1 15° sampling performed better than ZDOCK 2.3 6° sampling. This trend may indicate that shape complementarity (which is the only scoring metric used for ZDOCK 2.1) is less important (versus electrostatics and desolvation) for the Others cases than for the Enzyme/Inhibitor and Antibody/Antigen cases.

2.5. Docking Overview: Summary

Protein–protein docking has evolved to the point where it is possible to predict the structures of many protein complexes based on their unbound proteins. This is demonstrated above using a protein-docking benchmark and the rigid-body-docking algorithm ZDOCK. However, based on the success rate plots of [Fig. 3](#), it is evident that not all cases are successfully predicted within the top few thousand docking predictions, and for a few cases, no hits are found. What leads to this variation in docking success across a set of cases? The final section of this chapter takes an in-depth look at how various properties of proteins impact the ability of docking to successfully predict the complex structure.

3. The Relationships Between ZDOCK Performance and Protein Complex Characteristics

The performance of ZDOCK is dependent on both the accuracy of the energy function and the comprehensiveness of the search algorithm. Both of these are in turn dependent on the many physicochemical characteristics of the protein–protein complex that ZDOCK is attempting to predict. For example, in any particular complex, the exact shape of the protein–protein interface will undoubtedly have an effect on how high shape complementarity is scored in the energy function. Protein–protein complexes with planar interfaces may prove to be the most challenging for ZDOCK. Thus, it is important to examine how ZDOCK performs with respect to differing interface shapes in order to gauge the effectiveness of the shape complementarity term. Knowing how

ZDOCK performs with respect to a vast array of different protein–protein complex characteristics provides an understanding of what types of complexes ZDOCK can be expected to excel in predicting. It also can help lead to more focused improvements in the development of protein docking by identifying the strengths and weaknesses of the algorithm. In addition, it may be possible to extend the conclusions drawn from such an examination to other FFT-based docking algorithms.

3.1. Near-Native Prediction Definitions

In order to objectively and systematically evaluate the performance of the ZDOCK algorithm, it is necessary to compare the near-native docking orientations produced by ZDOCK to the space of orientations available given a particular complex within the rigid-body FFT framework. While the fields of protein structure prediction and docking commonly make use of “decoys” to evaluate algorithm performance, here we adopt an alternative approach. We estimate the space of potential near-native conformations using a newly designed program called HitFinder. This space is reasonably limited under the assumption of rigid-body docking, and therefore focus was placed on the 64 rigid-body cases from the protein-docking benchmark (27).

Using the core framework of the ZDOCK algorithm, HitFinder maps the complex components onto a 1.2 Å grid and uses a 6° Euler angle set (18) to perform FFT search for orientations that would represent near-native hits. HitFinder iterates over the same set of angles and translations as ZDOCK but uses a simple RMSD filter instead of a docking scoring function. For every potential ligand-docking orientation where the ligand overlaps with the native ligand orientation, the docking orientation is retained for further processing if the ligand C α RMSD is less than or equal to 10 Å. Following this initial search, potential docking hits are further defined using a more nuanced protocol based on the CAPRI prediction accuracy criteria. As in CAPRI, these hit definitions rely on the combination of RMSD and native contact fraction criteria. Here two kinds of hits are classified: high quality and medium quality. They are defined by the following Boolean relationships:

$$\begin{aligned} \text{High-quality hits} = & \left[iRMSD \leq (iRMSD_{\text{superposed unbound complex}} + 1 \text{ \AA}) \right] \\ & \cap (f_{\text{nat}} > 0.5) \cap (f_{\text{non-nat}} < 0.5) \end{aligned} \quad (12)$$

$$\begin{aligned} \text{Medium-quality hits} = & \left[iRMSD \leq (iRMSD_{\text{superposed unbound complex}} + 1 \text{ \AA}) \right] \\ & \cap (f_{\text{nat}} > 0.3) \cap (f_{\text{non-nat}} < 0.7) \end{aligned} \quad (13)$$

Because HitFinder does not include the shape complementarity functions that are normally a part of the ZDOCK algorithm, there is no control over potential ligand/receptor clash for those orientations where they come too close to each other. Therefore, this study uses the more strictly defined space of high-quality hits (or three-star hits; **Eq. 2**) as a guide and eliminates all hits with clash significantly greater than the average three-star hit. Clashes are defined as the number of interface contacts within 3 Å. All docking orientations with a clash total greater than the mean number of clashes for the three-star hits plus 2 standard deviations are eliminated. Finally, if an orientation meets all the required hit criteria, it is labeled a “potential hit” and all such structures are recorded for a complex.

3.2. Measuring ZDOCK Success

To examine the success of ZDOCK, a metric for protein–protein docking accuracy is needed. Measuring ZDOCK accuracy per complex could be accomplished by merely counting the number of medium- or high-quality hits the algorithm achieves out of a certain number of predictions. However, because the number of potential hits is inherent to each particular protein–protein complex (see **Fig. 5A and B**), this measure would not reflect precisely how well ZDOCK performs. As an example, if complex A has 100 potential hits and complex B has 1000 potential hits, ZDOCK’s accuracy is not equivalent if it finds one hit for both complexes. Complex B is easier to predict because it possesses some characteristics that allow for a greater number of hits possible. Further discussion as to what characteristics these may be will follow in Section 3. It may make sense to simply take the percentage of hits predicted out of the number of potential hits as a metric for docking accuracy. In this metric, ZDOCK makes successful predictions at a 1% rate for complex A and only a 0.1% rate for complex B and thus clearly performs better on complex A. Yet there is a flaw to this measure as well. As explained previously, the ZDOCK algorithm only keeps the highest-scoring translation for every rotation angle searched. This means that if multiple hits exist in the same rotational angle, ZDOCK will at best only select one of them. In the example of complex A versus complex B, complex A has 100 potential hits, but hypothetically could have 99 in one rotational angle. In that case, the highest number of hits an optimal ZDOCK search could find would only be 2. Thus, accuracy as defined as a percentage of potential hits would reach the upper limit at 2%. It is necessary then to introduce another definition, that of the “hit angle.” A hit angle is defined as any rotational angle in a ZDOCK search that has at least one translation that results in a potential hit (see **Fig. 5C and D**). Using this

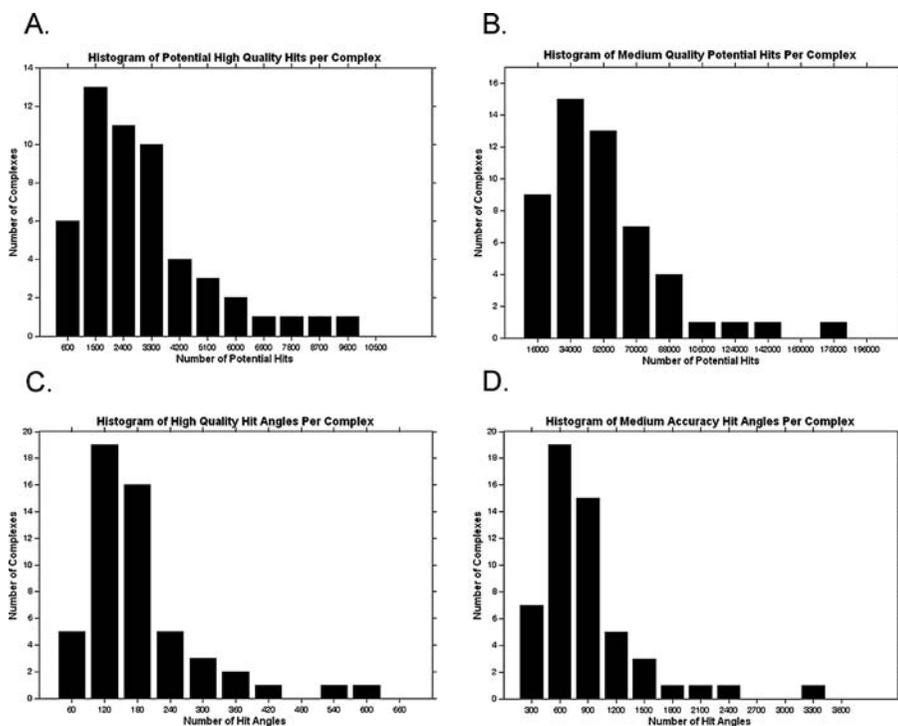


Fig. 5. Medium- and high-quality potential hits and hit angles.

definition, we found complex A has two hit angles because it has 100 potential hits but only two rotational angles in which those potential hits can be found. Finally, the accuracy of ZDOCK performance can robustly be measured as the percentage of hits predicted out of the possible number of hit angles. In the above example, if ZDOCK finds one hit for complex A and there are only two hit angles possible, the accuracy rate on that particular complex is 50%. Thus, for complex A, ZDOCK is operating at half of its maximum performance level. The distribution of accuracy rates for medium-quality and high-quality hits are shown in [Fig. 6](#).

If the accuracy rate of ZDOCK on complex B is much lower than the 50% of complex A, it leads to questions of why ZDOCK performs better on complex A. What are the characteristics of complex A that make it more suitable for creating good ZDOCK predictions? What are the characteristics of complex B that are associated with ZDOCK missing many good predictions? In the next section, these questions are examined for many complexes with myriad

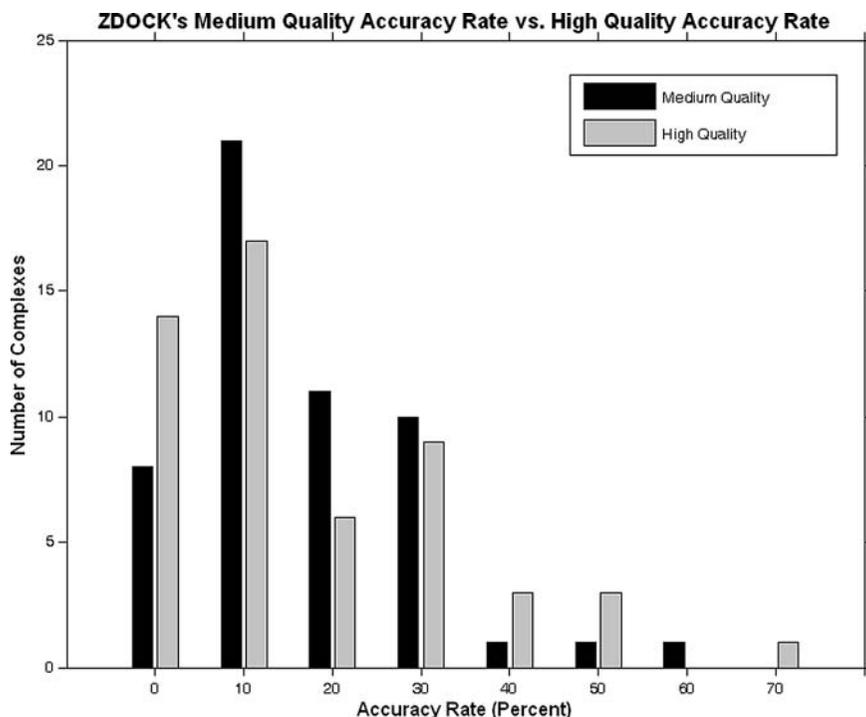


Fig. 6. ZDOCK Performance as measured by accuracy rate of medium- and high-quality predictions. Accuracy rate varies substantially across the 53 complexes of the regression data set.

physicochemical characteristics by employing regression analysis. The goal of the analysis is to get a better understanding about what types of attributes lead to successes and failures in protein–protein docking predictions with ZDOCK.

3.3. Regression Analysis of ZDOCK Performance

To begin to examine by regression analysis which types of attributes of protein–protein complexes are important to the success of ZDOCK, it is necessary to have a large data set of complexes. Fortunately, the protein–protein benchmark represents the largest data set of protein–protein complexes in the docking field and includes 84 such transient complexes. Some paring down of that original data set is required in order for the study to control for factors already known to affect docking results. Previously, 20 of the 84 complexes

have been characterized as undergoing large conformational change upon binding. These complexes were removed from this study in order to focus on docking performance in rigid-body cases. In addition, 11 benchmark complexes are antibody–antigen complexes in which the antibody structure is only solved in the complex. These “unbound–bound” complexes were removed in order to best represent true docking performance. The remaining 53 complexes comprise the data set used in the regression analysis (*see* [Table 3](#)).

A comprehensive list of protein–protein complex attributes is needed to establish what characteristics influence the performance of ZDOCK. One hundred twelve such attributes were used in the regression analysis, some of which are closely related. For brevity, only the general attributes are included in [Table 4](#).

3.4. Simple Linear Regression Approach

Simple linear regression was first employed to determine which single attributes could be associated with ZDOCK performance. Accuracy rate is used as the response variable in the regression. Accuracy rate can be broken down according to the two categories of hits: high and medium quality. Also, accuracy

Table 3
Protein Data Bank Codes in the Regression Analysis
Data Set

1A2K	1E6E	1HE8	1RLB
<i>1AHW</i>	<i>1E6J</i>	1HIA	1SBB
1AK4	1E96	1I4D	1TMQ
1AKJ	1EAW	<i>1JPS</i>	1UDI
1AVX	1EWY	1KAC	<i>1VFB</i>
1AY7	1EZU	1KLU	<i>1WEJ</i>
1B6C	1F34	1KTZ	2BTF
1BUH	1F51	1KXP	2MTA
<i>1BVK</i>	1FC2	1MAH	2PCC
1BVN	1FQJ	1ML0	2SIC
1CGI	1GCQ	<i>1MLC</i>	2SNI
1D6R	1GHQ	1PPE	2VIS
1DFJ	1HE1	1QA9	7CEI
<i>1DQJ</i>			

Antibody–antigen complexes are in italics.

Table 4
General Attributes of Protein–Protein Complexes

Attributes

Surface

Shape

Volume

Weight

Curvature

Interface size

Side-chain conformational change

Backbone conformational change

Number of interface hydrogen bonds

Number of native complex clashes

Hydrophobic character of interface

Polar character of interface

Charged character of interface

Of the 112 attributes used in the regression analysis, only the general attributes are listed here. Most attributes are expanded to include separately their values for complex, receptor, and ligand as well as the unbound and bound states of each.

rate is dependent on the number of predictions made, and this analysis uses all hits from the top 54,000 predictions, corresponding to one prediction per rotational angle at 6° sampling density.

As expected for an intricate system such as protein–protein docking, most of the simple linear regression models in this analysis fail to establish good relationships between single independent protein complex attributes and the outcomes investigated. Only simple linear regression on the accuracy rate for medium-quality predictions resulted in predictors with highly significant correlations ($p < 0.001$). Curvature of the interface has the strongest correlation ($R^2 = 0.36$) with medium-quality accuracy rate (see Fig. 7). It is a positive linear relationship, and thus ZDOCK performance tends to increase as curvature of the interface also increases. Interface curvature is calculated by first fitting a plane to the atoms of the interface. The RMSD from this plane is the curvature score (35). The ZDOCK scoring function relies heavily on shape complementarity for computing the energy of predicted complexes, and thus the importance of that energy term to the performance of ZDOCK is apparent from this correlation.

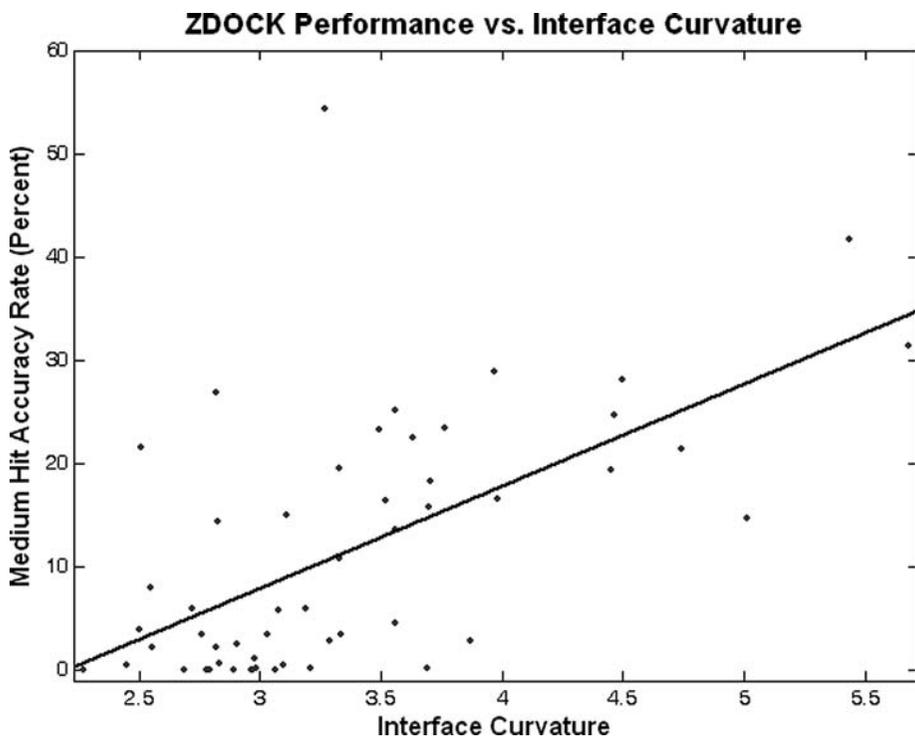


Fig. 7. ZDOCK performance versus interface curvature of the test case. Interface curvature is the strongest correlated predictor in simple linear regression for medium-quality accuracy rate ($R^2 = 0.36$).

The four remaining predictors that showed statistically meaningful correlations all were related to the size of the interface. The strongest correlation among these ($R^2 = 0.30$) was the difference in accessible solvent area between the complex and its constituents, referred to as dASA. The linear relationship between interface size and ZDOCK performance is positive, meaning ZDOCK performs well on protein complexes with larger interfaces. Scores representing larger interfaces are more statistically significant than scores representing smaller interfaces, which lead to better discriminatory power of the algorithm and hence better docking performance. In addition, it suggests that the sensitivity to a few bad contacts is lowered in larger interfaces because they make up a smaller percentage of the overall interface. By contrast, in smaller interfaces, one or two mispositioned side chains could proportionately contribute enough

high energy to the overall docking score to sufficiently lower the rank of a near-native structure such that it is not included in the final prediction set.

3.5. Multiple Linear Regression Approach

Whereas simple linear regression is an important first look at which single characteristics of protein complexes are relevant to the performance of ZDOCK, a more comprehensive approach should involve the employment of multiple linear regression analysis. Finding the relationship of combined attributes to ZDOCK sampling and accuracy gives a better indication of what to expect in terms of successes and failures depending on the type of complex involved in the prediction. In multiple linear regression, it is important to avoid overfitting the data caused by using a small ratio of outcome variables to predictor variables. Therefore in this study, only sets of four attributes were considered for the regression with 53 complexes. It was computationally tractable to do the regression on all permutations of four attributes and thus avoid the pitfalls associated with a stepwise regression approach.

3.5.1. Medium Quality Predictions

Multiple linear regression analysis was computed for the response variables of accuracy rate for medium- and high-quality hits with four predictors. For medium-quality accuracy rate, the four attributes with the highest correlation (R^2 adjusted = 0.53) were: curvature of the interface, size of the ligand interface relative to the size of the ligand, ligand side-chain conformational change, and the hydrophobicity of atoms that are completely buried upon binding (see [Table 5](#)).

The inclusion of curvature of interface in the top correlated set of attributes suggests the importance of shape complementarity just as it did in the simple linear regression for medium-quality accuracy rate. It is possible to exactly determine how important interface curvature or any other predictor is to the overall correlation by looking at the coefficients of partial determination for the regression model. A coefficient of partial determination in this analysis measures the proportionate reduction in variation in ZDOCK performance when a particular predictor is included in the regression model. With the above four attributes, the coefficient of partial determination for the inclusion of interface curvature in the regression model is 0.41. This explains quantitatively that interface curvature accounts for a 41% reduction in the regression error when it is added to the three-attribute model of interface hydrophobicity, ligand side-chain conformational change, and ligand interface size relative to the size of the ligand. Thus, interface curvature is highly important to the multiple linear relationship between these four predictors and ZDOCK performance.

Table 5
The Highest Correlated Regression Models

Attribute	Coefficient	Partial Coefficient of Determination
Medium-quality accuracy rate (R^2 -adjusted: 0.53)		
Interface Curvature	0.58	0.41
Interface Hydrophobicity	-0.34	0.21
% of Ligand in Interface	0.27	0.13
Ligand Side-chain Change	-0.20	0.08
High-quality accuracy rate (R^2 -adjusted: 0.41)		
% of Ligand in Interface	0.50	0.24
Native Complex Close Contacts	0.39	0.21
Ligand Side-chain Change	-0.33	0.16
Complex Shape	0.36	0.14

Hydrophobicity, with a coefficient of partial determination of 0.21, is the second most important predictor in this regression model. Hydrophobicity was characterized for atoms in the interface that are completely buried upon binding using an atom-typing scheme (36) representing three categories: polar, hydrophobic, and charged. Unexpectedly, in this regression model, the relationship between ZDOCK performance and complexes with interfaces with a large amount of hydrophobic atoms buried is negative. Although the correlation is weak, simple linear regression of ACE score versus medium accuracy rate confirms this inverse relationship (see Fig. 8). Previous analysis (3) on an earlier test-case data set found a positive relationship between hydrophobicity and ZDOCK performance. However, the earlier data set included several homodimer test cases that were not included in the current benchmark. Homodimers are known to have strong hydrophobic interfaces, and their absence in the current benchmark explains the loss of a positive correlation between hydrophobicity and ZDOCK performance.

ZDOCK uses a 6-Å cutoff for defining the interface for calculating the desolvation energy of the prediction. In the multiple regression analysis, the relationship of hydrophobicity and ZDOCK performance is most significant when the interface is limited to the atoms that become buried upon complex formation. Although the results were surprising that hydrophobicity is negatively correlated with ZDOCK performance, it underscores a potential area for improvement in the ZDOCK algorithm. Calculating the desolvation energy of just the buried atoms instead of using a 6 Å contact radius may

better represent the role of the hydrophobic effect in protein–protein binding and consequently increase the accuracy of ZDOCK.

The size of the ligand interface relative to the size of the ligand is the third most important attribute in the highest correlated regression model for medium-quality accuracy rate. The relationship is positive and for ligands in which the interface represents a large proportion of the total size, ZDOCK performance increases for this regression model. From a probability standpoint, this certainly makes sense as the greater the ratio between ligand interface size and ligand size, the higher the probability any docking prediction can be considered near native.

The final attribute of the regression model is a measure of how much side-chain conformational change occurs in the ligand interface. Specifically, it is calculated by determining the percentage of ligand interface residues that differ in rotamer type between the unbound and bound states. Rotamers were defined using the Dunbrack rotamer libraries (37). Most of the conformational change that occurs in side chains does not result in large structural differences such as in

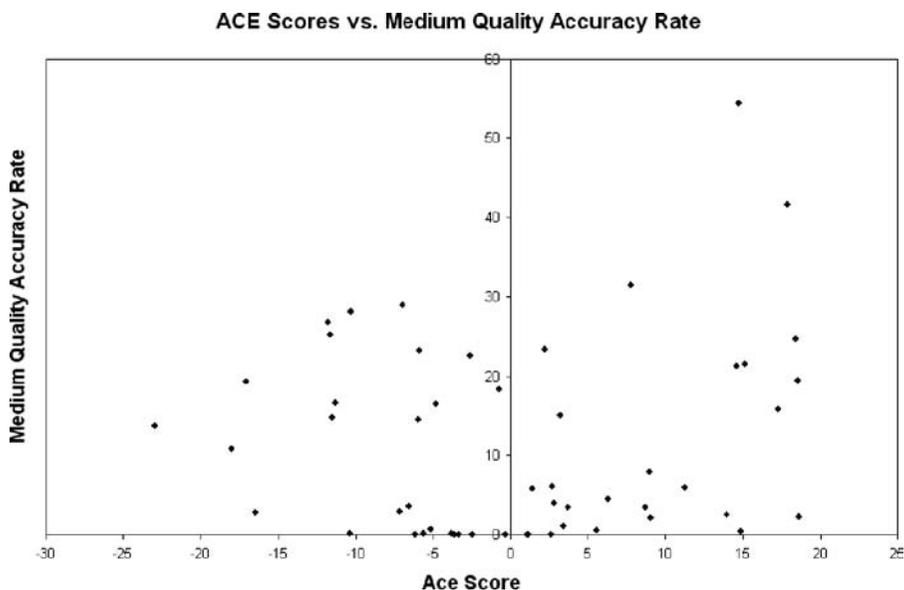


Fig. 8. Medium-quality accuracy rate shows a very weak but positive correlation with the atomic contact energy (ACE) score of the native complex interface. ACE scores decrease as interface hydrophobicity increases and medium-quality accuracy rate is therefore negatively correlated with interface hydrophobicity.

the complexes with backbone conformational change that were removed in the creation of the data set. However, even small differences in side-chain positions can cause large inaccuracies in the calculation of the scoring function especially within the vdW terms. Because ZDOCK does not attempt to move side chains during docking, interfaces with more side chains in different positions than in their unbound state will cause an inaccurate representation of the true bound interface and thus ZDOCK performance will suffer. Side-chain search is an actively pursued area in protein–protein docking research, and from the results of this regression analysis, it is understandable why accurate placement of side chains is a vital part of making successful docking predictions.

3.5.2. High-Quality Predictions

In comparison to the ability of ZDOCK to produce medium–quality predictions, there may exist a different set of characteristics of protein complexes that associate with ZDOCK’s ability to generate high-quality predictions.

To this end, all regression models with four predictors were run using the high-quality accuracy rate as the response variable. The highest correlated model (R^2 adjusted = 0.40) included the following four attributes: complex shape, size of the ligand’s interface relative to the size of the ligand, ligand side-chain movement, and number of close contacts in the native complex (see [Table 6](#)). Whereas two of these attributes are the same as in the medium-quality accuracy rate regression, two are different and will be explored further in this section.

The inclusion of native complex close contacts in the regression model was a surprising result, and even more unexpected was that the relationship between the number of close contacts and accuracy rate in the model was positive. Close contacts were calculated as all intermolecular atomic contacts less than 3 Å in the native complex structure. The positive relationship means that in the highest correlated model, ZDOCK performance is higher in complexes with many close contacts. It would seem that close contacts occur more often in larger interfaces and at least partly explain the positive relationship based on the aforementioned reasons why larger interfaces are preferred for better ZDOCK performance. However, there is no strong correlation between the two attributes of native complex close contacts and interface size ($R^2 = 0.25$). Thus, it may instead be that a complex with many close contacts represents a tightly packed interface. This would suggest once again the importance of the shape complementarity term in the ZDOCK energy function and in particular the necessity for a well struck balance between the vdW repulsion and attraction parameters.

Table 6
Intercorrelation of Attributes in the Regression Models for Medium- and High-Quality Accuracy Rates (R^2 values)

Medium-Quality Accuracy Rate				
Attribute	Interface Curvature	Interface Hydrophobicity	% of Ligand in Interface	Ligand Side-chain Change
Interface Curvature	1	–	–	–
Interface Hydrophobicity	0.03	1	–	–
% of Ligand in Interface	0.09	0.005	1	–
Ligand Side-chain Change	0.0001	0.009	0.02	1
High-Quality Accuracy Rate				
Attribute	% of Ligand in Interface	Native Complex Close Contacts	Ligand Side-chain Change	Complex Shape
% of Ligand in Interface	1	–	–	–
Native Complex Close Contacts	0.05	1	–	–
Ligand Side-chain Change	0.03	0.006	1	–
Complex Shape	0.28	0.01	0.02	1

Complex shape is the final attribute of the highest correlated regression model for high-quality accuracy rate. Complex shape is measured using the radius of gyration of the bound receptor and ligand. In this regression model, ZDOCK performance tends to increase with elongated complex shapes. The most commonly elongated complex shapes in the data set are the Antibody/Antigen cases, and removing these from the regression model reduces the coefficient of partial determination for this characteristic by more than half (0.14–0.06). The diminishing importance of complex shape when antibody–antigens are excluded suggests a relationship between ZDOCK’s high-quality accuracy rate and whether or not the complex is an antibody–antigen. Antibody–antigen complexes are known to be high-affinity binders and perhaps ZDOCK’s performance correlates well with binding affinity as such complexes would require very low energy conformations that simple scoring functions such as ZDOCK’s could find with greater success. Unfortunately, accurate binding affinity data

for each complex in the data set are not available to proceed further with such an analysis.

The coefficients of partial determination for the high-quality accuracy rate regression model for four predictors show more balance in the importance of the attributes than in the medium-quality accuracy rate model (see **Table 6**) Ligand interface size relative to ligand size and number of native complex clashes contribute almost equally to the reduction of regression error in the variation with coefficients of partial determination of 0.24 and 0.21, respectively. Ligand side-chain movement and complex shape were slightly less important with coefficients of 0.16 and 0.14, respectively.

3.6. Regression Analysis Conclusion

The relationships between complex characteristics and high-quality performance and medium-quality performance for ZDOCK are clearly similar especially with shape complementarity, side-chain conformational change, and the ratio of ligand interface size to ligand size. However, the difference in the two types of performance seems to be in how much each attribute contributes relative to the others. Shape complementarity, in the form of interface curvature, is ZDOCK's dominating discriminating force in medium-quality predictions. Yet, for high-quality predictions, it is clearly not as important and more attributes are equally as necessary. Understanding the differences in how ZDOCK performs with varying levels of prediction quality could allow for a future strategy of tweaking the parameters of the scoring function to fit a user's goals depending on what level of precision they require. Given the results of the regression analysis, it may be possible to target improvements to ZDOCK that would sacrifice high-quality performance for an increased amount of medium-quality predictions. Conversely, if only high-quality predictions are required, the quantity of medium level predictions could be sacrificed for a small amount of high-quality predictions.

Regression analysis is a good tool for finding the underlying relationships between characteristics of protein–protein complexes and ZDOCK performance. With this knowledge, it is possible to get a better idea of when and why ZDOCK makes successful predictions. Through this analysis, the shared importance of shape complementarity, side-chain conformational change, and interface size in ZDOCK's ability to predict high- and medium-quality protein complex structures is readily apparent.

In addition, understanding the relationships between each attribute in a comprehensive characterization of protein–protein complexes and how ZDOCK performs gives insight into where best to make future improvements to the

algorithm. Advancements in side-chain search and an approach for scoring only the buried interface atoms in the desolvation energy calculations are some possible avenues of pursuit for further ZDOCK development.

Acknowledgments

We are grateful to the Scientific Computing Facilities at Boston University and the Advanced Biomedical Computing Center at NCI, NIH for support in computing. This work was funded by NSF grants DBI-0133834 and DBI-0116574.

References

1. Betts, M.J. and M.J. Sternberg. An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Eng*, 1999, **12**(4): p. 271–83.
2. Katchalski-Katzir, E., et al. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci USA*, 1992, **89**(6): p. 2195–9.
3. Chen, R. and Z. Weng. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins*, 2002, **47**(3): p. 281–94.
4. Gabb, H.A., R.M. Jackson, and M.J. Sternberg. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol*, 1997, **272**(1): p. 106–20.
5. Vakser, I.A. Protein docking for low-resolution structures. *Protein Eng*, 1995, **8**(4): p. 371–7.
6. Ritchie, D.W. and G.J. Kemp. Protein docking using spherical polar Fourier correlations. *Proteins*, 2000, **39**(2): p. 178–94.
7. Palma, P.N., et al. BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins*, 2000, **39**(4): p. 372–84.
8. Abagyan, R., M. Totrov, and D. Kuznetsov. 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**(5): p. 488–506.
9. Gray, J.J., et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol*, 2003, **331**(1): p. 281–99.
10. Gardiner, E.J., P. Willett, and P.J. Artymiuk. Protein docking using a genetic algorithm. *Proteins*, 2001, **44**(1): p. 44–56.
11. Fischer, D., et al. A geometry-based suite of molecular docking processes. *J Mol Biol*, 1995, **248**(2): p. 459–77.
12. Morris, G.M., et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem*, 1998, **19**(14): p. 1639–62.

13. Comeau, S.R., et al. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 2004, **20**(1): p. 45–50.
14. Kuntz, I.D., et al. A geometric approach to macromolecule-ligand interactions. *J Mol Biol*, 1982, **161**(2): p. 269–88.
15. Mandell, J.G., et al. Protein docking using continuum electrostatics and geometric fit. *Protein Eng*, 2001, **14**(2): p. 105–13.
16. Dominguez, C., R. Boelens, and A.M. Bonvin. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc*, 2003, **125**(7): p. 1731–7.
17. Schneidman-Duhovny, D., et al. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*, 2005, **33**(Web Server issue): p. W363–7.
18. Chen, R., L. Li, and Z. Weng. ZDOCK: an initial-stage protein-docking algorithm. *Proteins*, 2003, **52**(1): p. 80–7.
19. Janin, J., et al. CAPRI: a critical assessment of predicted interactions. *Proteins*, 2003, **52**(1): p. 2–9.
20. Moulton, J., et al. Critical assessment of methods of protein structure prediction (CASP) –round 6. *Proteins*, 2005, **61 Suppl 7**: p. 3–7.
21. Vajda, S. Classification of protein complexes based on docking difficulty. *Proteins*, 2005, **60**(2): p. 176–80.
22. Delano, W.L. *The PyMOL Molecular Graphics System*, 2002.
23. Chen, R., et al. A protein-protein docking benchmark. *Proteins*, 2003, **52**(1): p. 88–91.
24. Kozakov, D., et al. Optimal clustering for detecting near-native conformations in protein docking. *Biophys J*, 2005, **89**(2): p. 867–75.
25. Duan, Y., B.V. Reddy, and Y.N. Kaznessis. Physicochemical and residue conservation calculations to improve the ranking of protein-protein docking solutions. *Protein Sci*, 2005, **14**(2): p. 316–28.
26. Tovchigrechko, A. and I.A. Vakser. Development and testing of an automated approach to protein docking. *Proteins*, 2005, **60**(2): p. 296–301.
27. Mintseris, J., et al. Protein-Protein Docking Benchmark 2.0: an update. *Proteins*, 2005, **60**(2): p. 214–6.
28. Murzin, A.G., et al. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol*, 1995, **247**(4): p. 536–40.
29. Berman, H.M., et al. The Protein Data Bank. *Nucleic Acids Res*, 2000, **28**(1): p. 235–42.
30. Zhang, C., et al. Determination of atomic desolvation energies from the structures of crystallized proteins. *J Mol Biol*, 1997, **267**(3): p. 707–26.
31. Chen, R. and Z. Weng. A novel shape complementarity scoring function for protein-protein docking. *Proteins*, 2003, **51**(3): p. 397–408.
32. Brooks, B.R., et al. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*, 1983, **4**: p. 187–217.

33. Li, L., R. Chen, and Z. Weng. RDOCK: refinement of rigid-body protein docking predictions. *Proteins*, 2003, **53**(3): p. 693–707.
34. Pierce, B., W. Tong, and Z. Weng. M-ZDOCK: a grid-based approach for Cn symmetric multimer docking. *Bioinformatics*, 2005, **21**(8): p. 1472–8.
35. Laskowski, R.A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph Model*, 1995, **13**(5): p. 323–30, 307–8.
36. Mintseris, J. and Z. Weng. Optimizing protein representations with information theory. *Genome Inform Ser Workshop Genome Inform*, 2004, **15**(1): p. 160–9.
37. Dunbrack, R.L., Jr. and M. Karplus. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *J Mol Biol*, 1993, **230**(2): p. 543–74.