PROTEINS: Structure, Function, and Genetics 52:000–000 (2003)

Combination of Scoring Functions Improves Discrimination in Protein–Protein Docking

AQ: 1

John Murphy, David W. Gatchell, Jahnavi C. Prasad, and Sandor Vajda*
Department of Biomedical Engineering, Boston University, Boston, Massachusetts

ABSTRACT **Two structure-based potentials are used for both filtering (i.e., selecting a subset of conformations generated by rigid-body docking), and rescoring and ranking the selected conformations. ACP (atomic contact potential) is an atom-level extension of the Miyazawa–Jernigan potential parameterized on protein structures, whereas RPScore (residue potential score) is a residue-level potential, based on interactions in protein–protein complexes. These potentials are combined with other energy terms and applied to 13 sets of protein decoys, as well as to the results of docking 10 pairs of unbound proteins. For both potentials, the ability to discriminate between near-native and non-native docked structures is substantially improved by refining the structures and by adding a van der Waals energy term. It is observed that ACP and RPScore complement each other in a number of ways (e.g., although RPScore yields more hits than ACP, mainly as a result of its better performance for charged complexes, ACP usually ranks the near-native complexes higher). As a general solution to the protein-docking problem, we have found that the best discrimination strategies combine either an RPScore filter with an ACP-based scoring function, or an ACP-based filter with an RPScore-based scoring function. Thus, ACP and RPScore capture complementary structural information, and combining them in a multistage postprocessing protocol provides substantially better discrimination than the use of the same potential for both filtering and ranking the docked conformations. Proteins 2003; 52:000–000.** © 2003 Wiley-Liss, Inc.

AQ: 2

Key words: fast Fourier transform (FFT); rigid-body docking; binding free energy; structure-based potential; ranking docked conformations

INTRODUCTION

The goal of protein–protein docking is to determine the structure of a complex in atomic detail, starting from the coordinates of the unbound component molecules. Most current docking methods use efficient rigid-body searches that generate large sets of docked conformations with favorable surface complementarity, and generally result in good (i.e., near-native, docked conformations) when starting from cocrystallized protein structures.[1,2] However, applications involving unbound (separately crystallized) proteins yield both near-native structures and many false positives that have good surface complementarity but are far from the native complex. Substantial efforts have been devoted to developing methods that can rank the docked conformations and select the ones close to the native, typically with use of a potential that accounts for the chemical activity between the molecules and possibly refines the interacting surfaces.[1] These procedures improve the discrimination such that conformations with less than 5 Å root-mean-square deviation (RMSD) are generally found within the top 10 to 100 structures. Nevertheless, no current method can eliminate all false positives, and the highest ranking docked conformations may be as far as 20 Å from the native complex.

We have previously described a combined filtering and scoring algorithm for finding near-native structures among the conformations generated by rigid-body protein docking.[3] The filtering was achieved by selecting a subset of conformations with low values of electrostatic interaction energy, and another set of conformations with low values of desolvation free energy,[4] the latter calculated with the structure-based atomic contact potential (ACP).[4] Because the formation of stable protein complexes requires favorable desolvation and/or electrostatic interactions,[5–9] filtering by these quantities tends to retain many of the near-native structures generated by the docking, resulting in a manageable number of conformations. In the rescoring step, the molecular mechanics energies of the retained structures were minimized, and the minimized structures reranked with a free-energy function combining electrostatic, desolvation, and van der Waals energy terms.[3] It is important to note that the addition of the last term substantially improved the discriminatory ability of the method.

The above algorithm has been tested with 5 sets of protein decoys, each comprising the 95 best scoring complexes generated by the docking program GRAMM.[10,11] Only a few of these 95 structures had RMSDs less than 20 Å from the native complex, and none less than 10 Å. Therefore, to each set we added a structure obtained by superimposing the unbound proteins over the complex,

Grant sponsor: National Institute of Health; Grant number: GM61867.

*Correspondence to: Sandor Vajda, Department of Biomedical Engineering, Boston University, Boston, MA 02215. E-mail: vajda@bu.edu

Received 20 November 2002; Accepted 15 March 2003

Published online 00 Month 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.10473

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

TABLE I. List of Receptor–Ligand Pairs Used in the Decoy Study[a]

	Cocrystallized structures					Unbound structures	
	PDB code	Receptor	Charge[b]	Ligand	Charge[c]	PDB code	PDB code
	1ahw	Inhibitory Immunoglobulin Fab 5G9	−3	Tissue factor (Thromboplastin)	2	1fgn	1boy
	1bgs	Barnase	2	Barstar	−6	1a2p	1a19
AQ: t1	1brc	Trypsin varaint (D189G, G226D)	−7	Amyloid β-precursor inhibitor	−3	1bra	1aap
	1bvk	Humanized antilysozyme Fv (Hulys11)	5	Lysozyme (muramidase)	8	1bvl	3lzt
	1cgi	α-chymotrysinogen	4	HPTI	−1	2cga	1hpt
	1dfj	Ribonuclease A	−22	Ribonuclease inhibitor	4	2bnh	7rsa
	1mlc	Monoclonal antibody Fab D44.1	−1	Hen egg lysozyme	8	1mlb	3lzt
	1ugh	Human uracil-DNA glycosylase	5	Uracil-DNA glycosylase inhibitor	−12	1akz	1ugi
	1wej	Igg1 Fab fragment of E8 antibody	−1	Horse cytochrome C	9	1qbl	1hrc
	1wq1	H-Ras-GTPase	−1	P120Gap	−8	1wer	5p21
	2kai	Porcine Kallikrein a	−17	Bovine pancreatic trypsin inhibitor	5	2pka	5pti
	2pcc	Yeast cytochrome C peroxidase	−11	Yeast Iso-1-cytochrome C	8	1cca	1ycc
	2sic	Subtilisin	−2	Subtilisin inhibitor	−4	1sup	3ssi

AQ: t2 [a]http://www.bmm.icnet.uk/docking/systems.html.
[b]Net charge of receptor at pH 7.0.
[c]Net charge of ligand at pH 7.0.

and four others created by superimposing the unbound proteins over good matches of the bound structures, resulting in five conformations with RMSDs less than 4 Å from the native. Although the combined potential of electrostatics, desolvation, and van der Waals energies clearly discriminated at least one of the near-native structures from the other 95 conformations, the large gap in RMSD between these two groups of structures left the generality of our results somewhat uncertain.

Our motivation in this article is twofold. First, Moont et al. have developed an empirical, residue potential score (RPScore) based on structural information from interactions in protein–protein complexes[12] that typically discriminates between near-native and non-native docked conformations better than the sum of electrostatic and desolvation
AQ: 3 terms we used in our previous work.[3] However, the discriminatory ability of the potential, ACP plus electrostatics (ACP+Elec), was substantially improved by the addition of a van der Waals term; therefore, we tested whether RPScore can also be improved in a similar manner (i.e., by accounting for the van der Waals energy). More generally, both the ACP and RPScore potentials combined with other energy terms to form an assortment of free-energy evaluation models[4,9,13,14] used for both filtering and scoring docked conformations and, as we show, provide substantially improved discrimination over use of either potential alone.

Second, the above potentials are tested on two sets of improved docking decoys. One set has been constructed by Sternberg and associates for 13 protein complexes[15,16] (see Table I and the website http://www.bmm.icnet.uk/docking/
T1 systems.html). It is similar to the one we used in an earlier study[3] (i.e., for each complex, it includes 96 docked conformations generated by the docking program FTDock[15] and 3 additional near-native structures, that is, complexes with RMSDs ≤10 Å). However, instead of retaining the 96 best scoring conformations, Sternberg and associates selected structures with relatively good scores that span an
F1 RMSD range from 5 to 60 Å [Fig. 1(a)]. Although the number of structures below 10 Å RMSD is still limited, the large gap in RMSD, which made it impossible to generalize conclusions from our previous results, is no longer present. In the second part of this article, we discriminate among 20,000 docked conformations generated by the rigid-body docking program DOT[17] for 10 protein complexes. As shown in Figure 1(b), only a very small fraction of these conformations are "hits" with RMSDs <10 Å; therefore, finding these complexes among the huge number of nonnative structures represents the full challenge of postprocessing in rigid-body docking. Although this problem is far from trivial, we are able to show that strategies combining an RPScore filter with an ACP-based scoring function, or an ACP-based filter with RPScore as the scoring function, perform much better than the use of either function in both roles. More generally, these results suggest that discrimination in docking can be substantially improved by multistep postprocessing based on two or more potentials.

METHODS

Structure-Based Potentials

The ACP,[4] an all-atom extension of the Miyazawa–Jernigan potential,[18,19] is calculated by the sum $\Delta G_{ACP} = \Sigma_{ij}\, e_{ij}$, where e_{ij} denotes the atomic contact energy of interacting atoms i and j, and the sum is taken over all atom pairs less than 6 Å apart.[4] Applying the quasichemical approximation,[18] e_{ij} is defined as the effective free-energy change when a solute–solute bond between two atoms of type i and j, respectively, is replaced by solute–solvent bonds. The atomic contact energies were derived for 18 atom types from a set of 89 nonhomologous, high-resolution protein structures. [We note that Zhang et al.[4] originally used the term atomic contact energy (ACE) for this contact potential. We changed this to ACP to distinguish the method from the analytic continuum electrostatics model of Schafer and Karplus,[20] also abbreviated ACE.]

The residue pairing score (RPScore), developed by Moont et al.[12] provides a score for each pair of residues across the

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

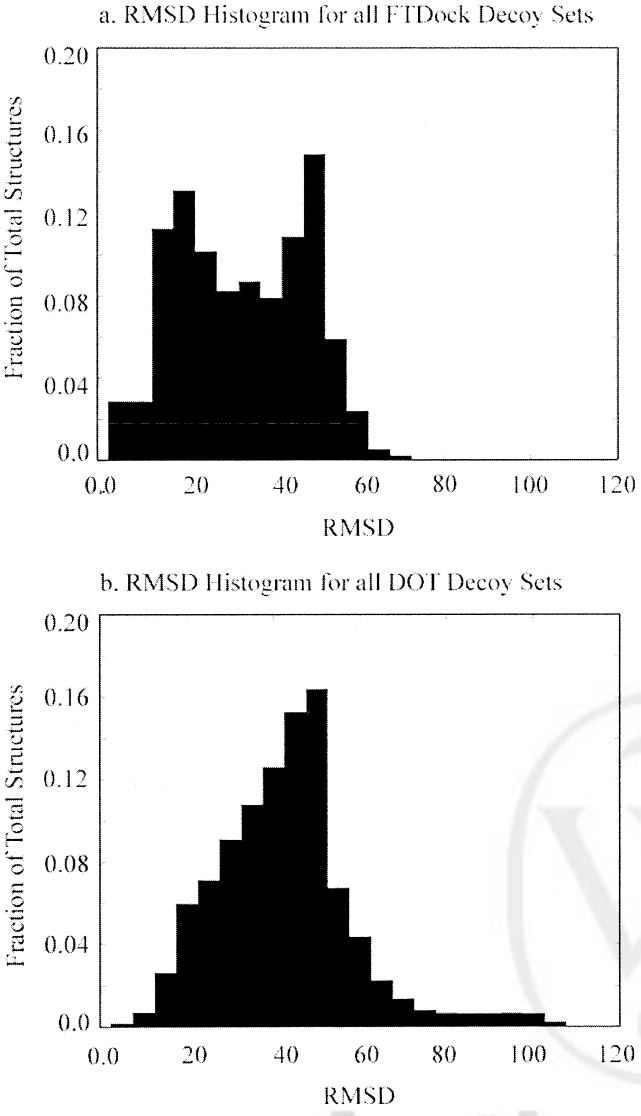

Fig. 1. Distribution of RMSD values. (**a**) Distribution of the 13 decoys sets generated by FTDock, each comprising 99 structures. (**b**) Distribution of the 10 decoy sets generated by DOT, each comprising 20,000 structures.

protein–protein interface and is defined as the log fraction of the actual frequency and the expected frequency of occurrence. Thus, the value of the score for each pair is a measure of the likelihood that the particular pair occurs, and the total score for a structure is the sum of all scores for residue pairs that satisfy a distance cutoff constraint of 6 Å.[12] The major difference between ACP and RPScore is that the latter has been constructed from statistics derived from interface data involving 90 nonhomologous protein complexes (http://www.bmm.icnet.uk/docking/downloads/3d-dock-manual.pdf).

Expressions for the Binding Free Energy

We combine molecular mechanics with empirical solvation and entropic terms.[14] In the most general case, the binding free energy, ΔG, is expressed as the sum,

$$\Delta G = \Delta E_{elec} + \Delta E_{vdW} + \Delta G_{des} + \Delta E_{int} - T\Delta S_{sc} + \Delta G_{other}, \quad (1)$$

where ΔE_{elec} is the change in electrostatic energy upon binding, $\Delta E_{vdW}$is the change in van der Waals energy, ΔG_{des} is the change in desolvation energy, and ΔE_{int} is the change in internal energy as a result of any flexing/straining of the backbone and sidechains.[21,22] The entropy term, $-T\Delta S_{SC}$, accounts for the decrease in entropy experienced by the interface sidechains upon binding. The term ΔG_{other} accounts for all other changes in the binding free energy that occur upon binding (e.g., due to changes in rotational, translational, vibrational, and cratic entropy). Because ΔG_{other} is considered to depend weakly on the size and shape of the conformation, it is treated as a constant.[21,22] In a free-energy expression of this form, $\Delta E_{vdW} + \Delta E_{int}$ is usually calculated by molecular mechanics, whereas the sum $\Delta E_{elec} + \Delta G_{des} - T\Delta S_{SC}$ is based on empirical models and/or continuum electrostatics calculations. In this work, the calculation of the latter sum involves the ACP and RPScore potentials. Because the change in the internal energy, ΔE_{int}, on docking is small, and we have shown that the correlation between internal energy and RMSD for docked complexes after minimization is poor,[3] this term is neglected.

Comparison of the ACP, ΔG_{ACP}, to experimental transfer free energies of amino acid sidechain analogs and to protein–protein-binding free energies suggests that the ΔG_{ACP} values include desolvation and entropic components[4] (i.e., $\Delta G_{ACP} = \Delta G_{des} - T\Delta S_{SC}$). Indeed, the atomic contact energy, e_{ij}, measures the free-energy change associated with moving the atoms i and j, exposed to the solvent (high entropy), to a position in which they are close to each other in the interior of the protein (low entropy). This assumption yields the free-energy expression,

$$\Delta G = \Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW} \quad (2a)$$

There is, however, still some uncertainty as to whether the empirical ACP accurately captures the sidechain entropy[4]; hence, we also consider the alternative free-energy expression,

$$\Delta G = \Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW} - T\Delta S_{sc} \quad (2b)$$

The van der Waals energy term is very sensitive to small structural perturbations; therefore, the potentials given by Eqs. (2a) and (2b) are meaningful only for minimized structures.[3] In addition to using the van der Waals energy, ΔE_{vdW}, in combination with other free-energy function components, we also use a modified van der Waals term that is zero for negative values of ΔE_{vdW} but otherwise identical. This "bump" energy, $\Delta E^{+}{}_{vdW}$, is introduced because we hypothesize that the structure-based potentials capture the attractive van der Waals energies but do not accurately represent the repulsive contributions. In fact, in measuring the likelihood of a solute–solvent pair becoming a solute–solute pair, only the attractive energies are accounted for, because overlaps are generally not present in a native data set of atom–atom contacts, as taken from the Protein Data Bank (PDB). Therefore, to

TABLE II. List of Receptor–Ligand Pairs Used in the Docking Study

Cocrystallized structures					Unbound structures	
PDB code	Receptor	Charge[a]	Ligand	Charge[b]	PDB code	PDB code
1brs	Barnase	+2	Barstar	−5	1a2p	1a19
1cgi	Chymotrypsinogen	+4	HPTI	−1	1chg	1hpt
1cho	α-chymotrypsin	+3	OMTKY	0	5cha	2ovo
1fss	Acetylcholinesterase	−12	Fasciculin II	+4	2ace	1fsc
1mlc	Fab D44.1 (antibody)	−3	Hen egg lysozyme	−3	1mlb	1lza
1ppf	Human leukocyte elastase	+11	OMTKY	0	1ppg	2ovo
2ptc	Trypsin	+6	BPTI	+6	2ptn	4pti
AQ: t3 2ptc	Trypsin	+6	BPTI	+6	2ptn	6pti
2sic	Subtilisin	−2	Subtilisin inhibitor	−4	2stl	3ssi
2sni	Subtilisin novo	−1	Chymotrypsin inhibitor 2	0	1sup	2ci2

[a]Net charge of receptor at pH 7.0.
[b]Net charge of ligand at pH 7.0.

capture the repulsive energy of a non-native overlapping interaction, a separate repulsive-only van der Waals energy term should be added, resulting in the free-energy expression,

$$\Delta G = \Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW} - T\Delta S_{sc} \quad (2c)$$

To incorporate RPScore into a thermodynamic potential, we rescale its values to an energy range that is physically meaningful. The scaling factor, α, is defined by

$$\alpha = -\sigma_{ACP+Elec}/\sigma_{RPScore} \quad (3)$$

where $\sigma_{ACP+Elec}$ and $\sigma_{RPScore}$ are the standard deviations of the $\Delta G_{ACP} + \Delta E_{elec}$ and RPScore values, respectively. We assume that the distribution range of RPScore values should be scaled to the range of $\Delta G_{ACP} + \Delta E_{elec}$ values, because RPScore is thought to account for both desolvation and electrostatic binding energies. A coefficient of −1 is assigned to α, because a more positive RPScore score corresponds to a more negative energy, or a more likely conformation. This scaling creates the free-energy expression, $\Delta G_{RP} = \alpha$ RPScore. Assuming that ΔG_{RP} captures all thermodynamic effects due to binding implies

$$\Delta G = \Delta G_{RP}. \quad (4a)$$

ΔG_{RP} is calculated both for unminimized and minimized structures, denoted by ΔG_{RP} and ΔG_{RPmin}, respectively. We combine ΔG_{RPmin} with the van der Waals energy to yield the expression,

$$\Delta G = \Delta G_{RPmin} + \Delta E_{vdW}, \quad (4b)$$

and examine the ability of ΔG to discriminate with and without the van der Waals energy term added. As shown for ACP, we also test the alternative free-energy expression,

$$\Delta G = \Delta G_{RPmin} + \Delta E^{+}_{vdW}, \quad (4c)$$

where ΔE^{+}_{vdW} represents the "bump" energy.

Docking, Filtering, and Scoring

In the first part of this article, we studied the decoys generated by Gabb et al. [because the 1BDJ complex is hypothesized not to be the biologically active conformation (see http://www.bmm.icnet.uk/docking/systems/1bdj.html), decoys are considered only for the other 13 complexes; see Table I].[15] Each set of decoys was minimized for 1000 adopted-basis Newton–Raphson (ABNR) steps in CHARMM, with use of the CHARMM 19 parameters.[23] Each set was then ranked according to the following potentials: ΔG_{ACP}, ΔE_{elec}, ΔE_{vdW}, ΔE^{+}_{vdW} ("bump" energy), $\Delta G_{ACP} + \Delta E_{elec}$, $\Delta G_{ACP} + \Delta E_{vdW}$, $\Delta G_{ACP} + \Delta E^{+}_{vdW}$, $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$, $\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$, $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW} - T\Delta S_{SC}$, $\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW} - T\Delta S_{SC}$, ΔG_{RP} (scaled RPScore), ΔG_{RPmin} (scaled RPScore, calculated for minimized structures), $\Delta G_{RPmin} + \Delta E_{vdW}$, and $\Delta G_{RPmin} + \Delta E^{+}_{vdW}$. AQ: 6

In the second part of this article, we discuss the analysis of structures generated with the use of Version 1.0 alpha of the software package DOT[17] to dock 10 pairs of monomeric, independently crystallized proteins for which the respective cocrystallized complexes are known (Table II). Using a 1-Å grid-step, a 12° Euler angle increment, and a 4-Å surface layer (the default value in DOT), we sampled approximately 10^{10} putative structures, from which 20,000 structures with the highest DOT scores, as determined by surface complementarity, were retained. Two alternative filters were applied to the 20,000 complexes. The ACP+Elec filter retains the 500 structures with the lowest values of ΔG_{ACP}, and an additional 1500 structures with the lowest values of ΔE_{elec}, with the electrostatic energy calculated with a Coulombic potential with a distance-dependent dielectric of 4r. The RPScore filter retains the 500 structures with the lowest values of ΔG_{RP}. The docked structures retained by the filters were minimized in CHARMM, as described above for the FTDock decoys. This created two sets of decoys for each complex studied: 2000 minimized complexes, as determined by the ACP+Elec filter, and 500 minimized complexes, as determined by the RPScore filter. Each set of decoys was then reranked according to the same potentials used for ranking the FTDock decoys (excluding $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW} - T\Delta S_{SC}$ and $\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW} - T\Delta S_{SC}$). T2

Measuring Deviations From the Native Complex

Recent articles on protein–protein docking use a variety of measures to describe the distance of a docked structure from the native complex. The most straightforward of these measures is the ligand RMSD, which involves superposing the receptor in the docked structure onto the receptor in the X-ray structure of the complex, with use of the derived transformation matrix to translate and rotate the docked ligand, and calculating the C_α RMSD only for the latter. Several groups use a different measure, termed the binding-site RMSD,[15,24,25] obtained by superposing both proteins of the predicted complex onto those of the crystal complex, and calculating the RMSD of only the C_α atoms belonging to residues with at least one atom within 10 Å of the other protein. Although it is a clear advantage that the binding site RMSD is not affected by parts of the molecule far from the interface, the definition of the interface residues requires an arbitrary parameter (i.e., the cutoff distance). A somewhat similar measure, describing the accuracy of the interface, is the number of correct contacts, which was recently used in the first CAPRI docking experiment.[26] However, because the number of correct contacts can be increased simply by enlarging the overlap between the two molecules (R. Abagyan, private communication), its value can be misleading; therefore, the CAPRI evaluators also included the ligand RMSD as a measure for ranking the submissions.[26] In view of the shortcomings of the other measures, we restrict consideration to the ligand RMSD as the single measure of quality. We note that this is the most appropriate measure in other applications (e.g., when docking methods are used for reconstructing electron microscopy images).[27,28] It is important to keep in mind that structures exhibiting RMSDs of 10 Å with this method typically would have binding-site RMSDs well below 5 Å. Indeed, calculating the RMSD for both the receptor and ligand with the latter method generally reduces its value by a factor of two, and this is further reduced by restricting consideration to the binding site.

AQ: 7

Throughout the discussion of results, we frequently use the term "hit," which is a structure that comes within some cutoff distance from the native complex. In the literature, the definition of a "hit" also varies. For example, in Gabb et al.,[15] a hit is considered to be a structure that has a binding site RMSD less than 2.5 Å from the native, where the binding site RMSD is defined as above. Gardiner et al.[25] defined a "hit" as a structure with less than a 4-Å binding-site RMSD from the native, plus the "standard error." The standard error is defined as the binding-site RMSD between the superposed cocrystallized and aporeceptor–ligand binding sites. In this work, we define a "hit" in terms of the ligand RMSD (as defined above) between the independently crystallized ligand, superposed onto the native complex, and the docked ligand, using 10 Å as the cutoff distance. Although 10 Å may appear to be too large to define a hit, we reiterate that this value corresponds to a binding-site RMSD of approximately 5 Å. In addition, we have recently developed a flexible docking method[29] that can refine structures from 10- to 4-Å ligand RMSD, thereby suggesting that the basin of attraction of the free-energy minimum around the native state extends at least to 10 Å. We also observed that both decoy sets have only a few conformations with RMSDs less than 5 Å, and further reducing the cutoff would yield very poor statistics in the near-native region.

Measuring the Quality of Scoring Functions

We determine the ability of different free-energy potentials to discriminate near-native and non-native docked conformations on the basis of the following four performance measures: (1) the number of structures within 10 Å RMSD of the native (number of "hits"), ranked within the top 10 energies; (2) the RMSD of the top-ranked structure; (3) the average RMSD of the 10 top-ranked structures; and (4) the rank of the first hit. Each measure yields different information about the discriminatory ability of the function. The ability of a function to select as many "hits" as possible is measured both by the number of hits within the top 10 structures and by the average RMSD of the top 10 structures, the first measure being particularly important for filtering. The rank of the first hit measures the ability of a function to bring an acceptable structure to the top and indicates how many structures must be retained to guarantee the presence of a single hit. The RMSD of the top-ranked structure is the most practical and, ultimately, useful of all quality measures, but we have not yet reached a proficiency level in protein docking to use this measure singularly.

RESULTS AND DISCUSSION FOR THE FTDOCK DECOYS

Because of the limited number of structures in each of the FTDock-derived decoy sets, it is computationally feasible to investigate the discriminatory abilities of various combinations of free-energy terms. In particular, we examined whether the additions of van der Waals and entropic terms to the structure-based potentials improve, or worsen, discrimination, which would suggest that ΔG_{ACP} and ΔG_{RP} already include the effect(s) of van der Waals and/or entropic contributions as suggested.[4,13]

Adding Van Der Waals Energy to ACP-Based Free-Energy Expressions

For the minimized FTDock decoys, adding ΔE_{vdW} to ΔG_{ACP} and $\Delta G_{ACP} + \Delta E_{elec}$ substantially improves discrimination by both potentials (Table III). This effect was also reported in our previous article.[3] Because of its sensitivity to small structural perturbations, the resulting potential can only be applied to structures refined by energy minimization. It has been hypothesized that the minimization creates an induced fit in the binding interface, leading to better shape complementarity, which is reflected in the van der Waals energy.[3] Following energy minimization, the addition of the van der Waals term does, in fact, improve discrimination by the ACP-based potentials across all quality measures. Therefore, we conclude that ΔG_{ACP} does not fully account for the van der Waals energy. When the bump energy, ΔE^{+}_{vdW}, is substituted for the van der Waals term, the improvement is typically equivalent, if

T3

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

TABLE III. Performance Means and Deviations for All FTDock Decoy Complexes[a]

Free-energy terms	RMSD of top 10 energies (Å)	Number of Hits[b] in the top 10 energies	Rank of 1st hit	Best energy RMSD (Å)
ΔG_{ACP}	33.7 ± 8.6	0.5 ± 1.1	29.9 ± 30.3	32.5 ± 17.2
ΔE_{elec}	30.8 ± 10.2	0.9 ± 1.3	22.4 ± 22.4	31.3 ± 18.2
ΔE_{vdW}	26.6 ± 7.2	1.3 ± 0.6	6.1 ± 7.5	33.4 ± 17.7
ΔE^{+}_{vdW}	25.1 ± 7.5	1.3 ± 1.0	6.4 ± 5.7	26.9 ± 21.4
$\Delta G_{ACP}+\Delta E_{elec}$[c]	37.0 ± 5.2	0.2 ± 0.6	37.5 ± 29.5	40.1 ± 15.8
$\Delta G_{ACP} + \Delta E_{elec}$	28.9 ± 5.3	0.8 ± 0.8	15.3 ± 14.8	28.2 ± 17.8
$\Delta G_{ACP} + \Delta E_{vdW}$	25.1 ± 6.1	1.6 ± 1.0	3.9 ± 5.4	18.1 ± 16.3
$\Delta G_{ACP} + \Delta E^{+}_{vdW}$	24.4 ± 6.5	1.5 ± 1.1	8.3 ± 12.2	16.2 ± 14.9
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$	24.9 ± 6.6	1.6 ± 0.9	3.2 ± 3.8	20.9 ± 18.1
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$	23.0 ± 6.5	1.9 ± 0.8	3.5 ± 3.1	13.6 ± 6.5
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW} - T\Delta S_{SC}$	24.6 ± 5.6	1.7 ± 1.0	3.7 ± 4.3	25.5 ± 19.0
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW} - T\Delta S_{SC}$	23.6 ± 6.2	1.6 ± 1.0	5.9 ± 7.1	22.1 ± 13.9
ΔG_{RP}[c]	19.2 ± 4.2	1.9 ± 1.3	8.9 ± 10.7	18.1 ± 8.8
ΔG_{RPmin}	18.5 ± 4.4	1.9 ± 1.4	9.2 ± 13.5	13.0 ± 7.0
$\Delta G_{RPmin} + \Delta E_{vdW}$	22.9 ± 7.0	1.7 ± 1.1	5.2 ± 6.8	14.7 ± 16.6
$\Delta G_{RPmin} + \Delta E^{+}_{vdW}$	22.2 ± 6.8	1.5 ± 1.0	5.4 ± 7.6	13.4 ± 9.7

[a]The best performance for each measure is shown in bold italics.
[b]Conformations with less than 10 Å RMSD (see text).
[c]Calculated for the unminimized decoys.

not better. This implies that whereas ΔG_{ACP} includes some of the attractive van der Waals interactions, and adding ΔE_{vdW} leads to partial double-counting, it most likely does not account for the repulsive contributions, allowing the addition of ΔE^{+}_{vdW} to improve discrimination for the ACP-based potentials.

Effect of Minimization on ΔG_{RP}

Because RPScore is a coarse-grained, residue-level potential, we expected that minimization would have little effect on ΔG_{RP}. As shown in Table III, however, minimization improves two of the four quality measures for ΔG_{RP} (RMSD of the top ranked structure and average RMSD of the top 10 structures), whereas values of the other two measures remain essentially unchanged.

Adding van der Waals Energy to RPScore-Based Free-Energy Expressions

Based on the results for ΔG_{ACP}, we anticipated that adding the van der Waals energy or the bump energy to ΔG_{RP} would improve discrimination. We found that the addition of ΔE_{vdW} to ΔG_{RPmin} leads to better discrimination than either ΔG_{RPmin} or ΔE_{vdW}, alone, as determined by the ranking of the first hit (see Table III). Although the other measures are slightly degraded, this makes $\Delta G_{RPmin} + \Delta E_{vdW}$ a useful potential for ranking near-native complexes in the top few energies. Adding ΔE^{+}_{vdW}, in place of ΔE_{vdW}, to ΔG_{RPmin} was observed to have a similar effect.

Accounting for Sidechain Entropy in ACP-Based Free-Energy Expressions

It has been shown that the loss of sidechain entropy substantially affects the binding free energy[21]; nevertheless, adding the sidechain entropy term ($-T\Delta S_{SC}$) to our free-energy expressions degrades nearly all performance measures. Thus, at least a fraction of the sidechain entropy loss is already represented in the ACP, as has been previously assumed.[4,13] However, the low accuracy of our sidechain entropy calculation[3,21,22] may also contribute to this observation.

Ranking of Potentials by Performance Measures

Table III shows the average performance measures for 16 scoring functions. The potentials that yield the highest number of hits are $\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$, ΔG_{RP}, and ΔG_{RPmin} (1.9 hits in the top 10 energies for each potential, respectively). ΔG_{RPmin} is also the best performer in terms of the lowest RMSD of the top-ranked structure and the mean RMSD of the top 10 structures. As previously mentioned, the sums $\Delta G_{RPmin} + \Delta E_{vdW}$ and $\Delta G_{RPmin} + \Delta E^{+}_{vdW}$ rank the first hit better than ΔG_{RPmin} alone. However, in terms of ranking the first hit, the best functions are $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ and $\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$. Thus, we conclude that RPScore generally places a higher number of near-native conformations in the top energies, whereas the ACP-based potentials provide better rankings of the top hit. This interesting complementarity is discussed further in the second part of this article.

Filtering

Because the FTDock decoy sets are small, all structures were minimized, allowing us to evaluate only the performances of the energy potentials in rescoring roles. However, dealing with larger sets of putative complexes requires filtering (i.e., reducing the number of conformations to be minimized). Therefore, we also applied the relatively smooth potentials, $\Delta G_{ACP} + \Delta E_{elec}$, and ΔG_{RP} to the unminimized FTDock structures. As shown in Table III, ΔG_{RP}, by all measures, performs significantly better in a filtering role than $\Delta G_{ACP} + \Delta E_{elec}$.

TABLE IV. Results From Rigid-Body Docking With the Use of DOT

Cocrystallized complex	Docked complex	Best RMSD	DOT ranking	Number of hits (RMSD <5.0 Å)	Number of hits (RMSD <10.0 Å)
1brs	1a2p-1a19	1.15	1705	22	226
1cgi	1chg-1hpt	1.27	12292	28	123
1cho	5cha-2ovo	3.57	348	20	148
1fss	2ace-1fsc	1.27	2861	25	88
1mlc	1mlb-1lza	2.50	3410	7	26
1ppf	1ppg-2ovo	4.4	6857	5	62
2ptc	2ptn-4pti	4.05	12575	1	47
2ptc	2ptn-6pti	2.74	8802	4	93
2sic	2st1-3ssi	1.02	7746	4	41
2sni	1sup-2ci2	1.38	3054	1	24

Potential Bias in RPScore

Because the RPScore was derived from residue–residue pairings across protein interfaces, it may be better suited for scoring docked structures than ACP, which was derived from atom–atom pairings within protein interiors. However, we have observed that RPScore is more likely to fail for a complex that is not represented (directly or by homology) in its training set. RPScore fails to find any hits within the 10 top-ranked structures for 4 of the 13 FTDock decoys. Three of these four were not represented in the training set (75%), whereas overall, only 7 of the 13 decoys (53%) were not represented. These two facts suggest that RPScore is less likely to select a near-native structure when either one or both components of the native protein complex are not represented in the training set. We do anticipate, however, that the performance of RPScore will improve as new complexes are added to the training set used in its derivation.

RESULTS AND DISCUSSION OF THE DOCKING STUDY

We have compared the ability of several scoring functions to discriminate near-native complexes among 10 sets of 20,000 docked conformations, generated from unbound protein structures using the program DOT, and ranked by a surface complementarity scoring function.[17] Retaining the top 20,000 complexes guaranteed, for all 10 complexes studied, at least 25 structures with RMSDs less than 10 Å (Table IV). T4

Filtering the Docked Conformations

In our earlier work,[3] we showed that complexes that comprise oppositely charged proteins tend to form in free-energy regions of low electrostatic potential, and that complexes with weak charge complementarity form in free-energy regions of low-desolvation free energy. From these considerations, we have developed a filter that retains a number of structures with the lowest ΔG_{ACP} values, as well as three times as many structures with the lowest values of ΔE_{elec}. We refer to this procedure as the ACP+Elec filter. The reason for retaining more conformations with low electrostatic energies is that the calculation of ΔE_{elec} is much more sensitive to the structural uncertainties associated with docking unbound protein conformations than the calculation of ΔG_{ACP}. For example, we have shown that the electrostatic energy of near-native docked structures can be substantially increased by changing the conformations of key sidechains, a frequent occurrence in the docking of unbound protein structures.[30]

According to the decoy study in this first part of this article, the RPScore (ΔG_{RP}) potential provides a better filter of docked complexes than $\Delta G_{ACP} + \Delta E_{elec}$. Figure 2 shows retention curves (i.e., the number of hits as a function of the number of retained complexes) for each filter: ACP+Elec, RPScore, as well as the individual energy terms, ΔG_{ACP} (ACP) and ΔE_{elec} (Elec). RPScore clearly performs the best of the four potentials over the range of filtered structures. In particular, the average number of hits in the 500 complexes with the best RPScore values nearly equals the average number of hits in the top 2000 complexes selected by the ACP. In this case, the ACP performs slightly better as a filter than the ACP+Elec filter. F2

As shown in Figure 3(a), the best filter for neutral complexes [i.e., at least one of the proteins comprising the complex has an absolute net charge ≤ 1 (see Table II)] is ACP alone. However, ACP is the worst performer for charged complexes [Fig. 3(b)]. Indeed, the ACP filter fails to find any hits among the top 2000 structures for the barnase–barstar complex, 1a2p-1a19, or for the trypsin-BPTI complex, 2ptn-6pti, and performs poorly for the other trypsin-BPTI complex, 2ptn-4pti (Table V). In contrast, the Elec filter finds near-native docked structures for all three complexes (Table V). These observations justify the usage of the combined ACP+Elec filter in our previous work.[3] As shown in Figure 3(b), however, RPScore is the best filter for charged complexes, and it is comparable to ACP for neutral complexes if less than 500 conformations are retained. However, as previously mentioned, it cannot be ignored that the performance results for RPScore are somewhat biased due to the overlap between its training set and the complexes considered here. Indeed, RPScore collects the fewest hits [i.e., 4, for 2ace-1fsc (Table V)], which is the only complex among the 10 that has no homolog for either of its constituent proteins in the RPScore training set (G. Smith, private communication). However, the only way to investigate further the possible bias of RPScore is to reparameterize the potential with F3 T5 AQ: 7

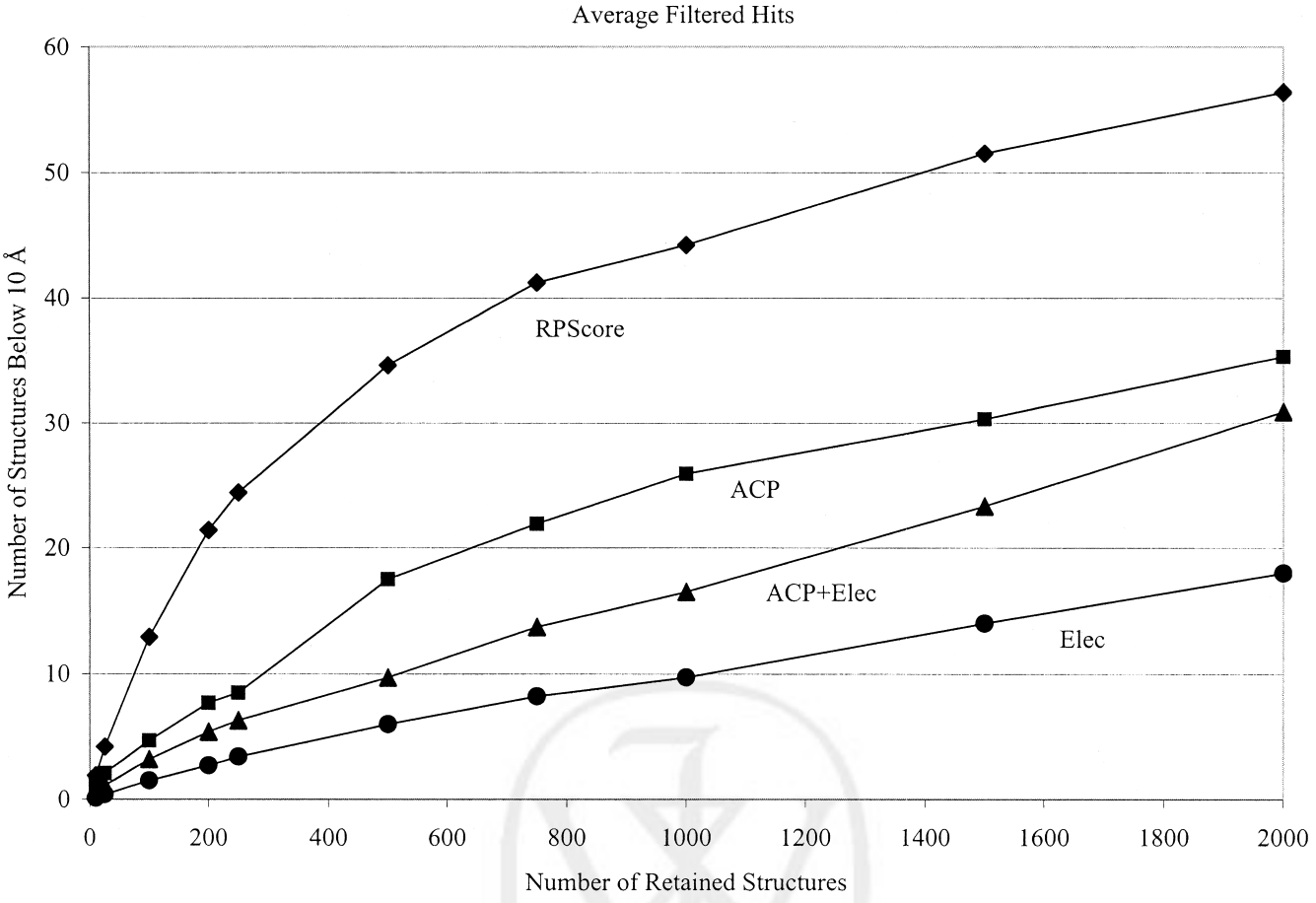

Fig. 2. The average number of hits retained by the various filters applied in the DOT docking study.

different training sets, which is obviously beyond the scope of this article.

Ranking the Structures Retained by the ACP+Elec Filter

For the docked decoy sets, two of the four quality measures (i.e., the average RMSD of the 20 lowest free-energy docked structures and the RMSD of the best energy structure) exhibit values so large that they are not useful for the comparison of scoring functions and are not discussed further (Table VI). The best performing potential is $\Delta G_{RPmin} + \Delta E_{vdW}$, which, on average, ranks 3.6 hits below 10 Å in its top 10 energies. For comparison, the second best potential is ΔG_{RPmin} alone, with 3.0 hits ranked in the top 10 energies (Table VI). According to the rank of the first hit ≤10 Å, the best performing functions are, again, $\Delta G_{RPmin} + \Delta E_{vdW}$ and ΔG_{RPmin}, with values of 4.3 and 9.7, respectively. These rankings indicate that both potentials, on average, find hits below 10 Å RMSD among their top 10 energies, whereas the improvement in ΔG_{RPmin} reaffirms the significance of the additional van der Waals term. In contrast to the FTDock decoys, however, the addition of the bump energy, ΔE^{+}_{vdW}, does not yield comparable improvement, possibly because of the different surface complementarity functions implemented in the FTDock and DOT programs (e.g., the bump limit established in DOT, which does not have an analogous counterpart in FTDock).

Table VII shows, for each decoy set, the number of hits in the top 10 energies and the rank of the first hit using the five best scoring functions. On average, $\Delta G_{RPmin} + \Delta E_{vdW}$ and ΔG_{RPmin} locate more hits in their top 10 energies, and rank the top hits much better than the other three potentials. In particular, $\Delta G_{RPmin} + \Delta E_{vdW}$ finds a hit among the top 10 docked structures for 9 of the 10 complexes. ΔG_{RPmin}, alone, is not as good, with 3 of the top hits below 10 Åranking worse than 10th, and 1, 1sup-2ci2, ranking 45th. We note that RPScore does not appear to depend on the charged or neutral character of the proteins. Of the three complexes with which it struggles, 1sup-2ci2, 2ace-1fsc, and 2ptn-4pti, one is neutral, and the other two are charged.

As expected [Fig. 3(a)], $\Delta G_{ACP} + \Delta E_{vdW}$ works very well for the complexes 1chg-1hpt, 1ppg-2ovo, 1sup-2ci2, and 5cha-2ovo, which are neutral or weakly charged [i.e., at least one of the proteins comprising the complex has an absolute net charge ≤1 (Table II)]. For the charged complexes, however, we observe much poorer discrimination, with 5 of the 6 complexes showing no hits in the top 10 $\Delta G_{ACP} + \Delta E_{vdW}$ energies. We hypothesized that the addition of the electrostatics (ΔE_{elec}) to the $\Delta G_{ACP} +$

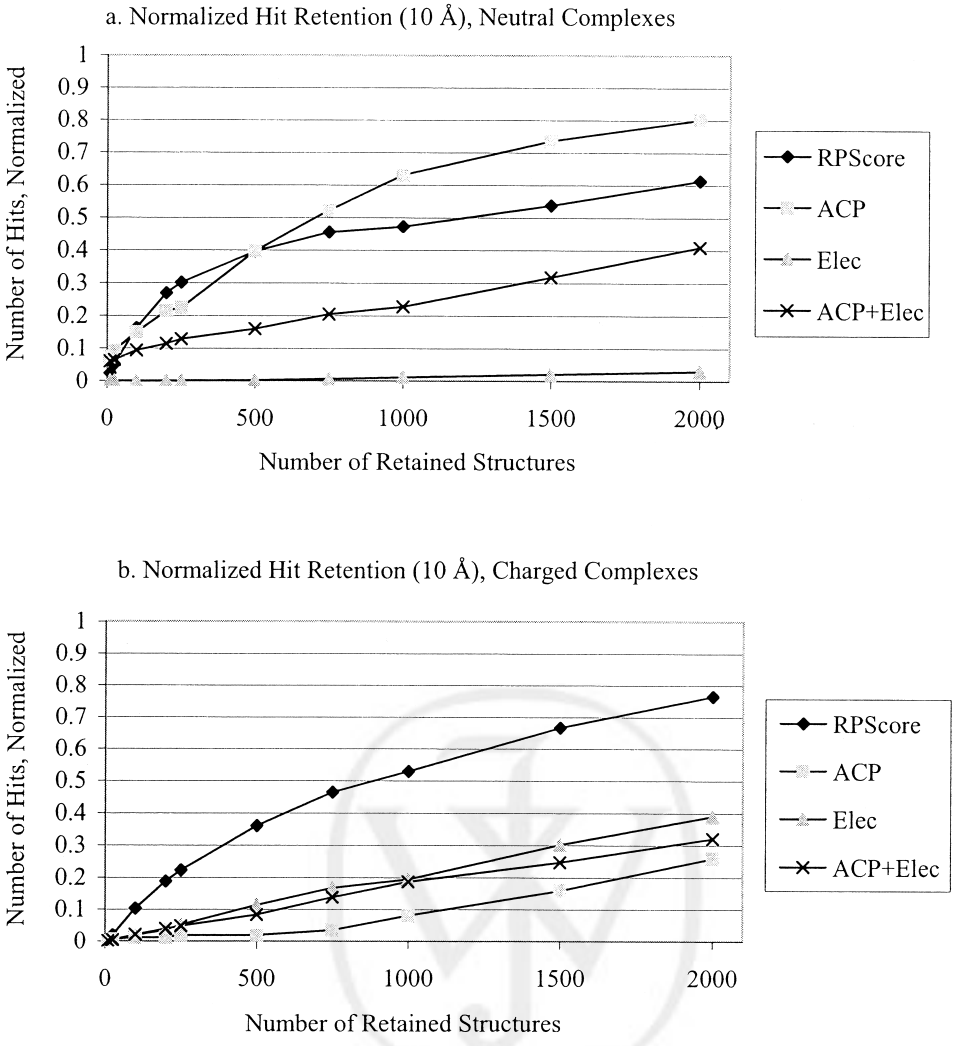

Fig. 3. Average normalized hit retention curves. (**a**) Neutral complexes. (**b**) Charged complexes. The term "normalized hits" refers to the number of hits retained by a filter divided by the total number of unique hits collected by all four filters (i.e., RPScore, ACP+Elec, ACP, and electrostatics). The normalization was calculated such that each complex was weighted equally in the averaging, regardless of the absolute number of hits collected for that complex.

TABLE V. Number of Hits Retained by the Various Filters in the Docking Study[a]

Complex	RPScore (500 Structures)	ACP + Elec (2000 Structures)	ACP (2000 Structures)	Elec (2000 Structures)
1a2p-1a19	63	84	—	102
1chg-1hpt	88	66	109	2
1mlb-1lza	11	3	17	—
1ppg-2ovo	7	24	45	—
1sup-2ci2	—	1	4	—
2ace-1fsc	4	13	20	17
2ptn-4pti	26	22	12	27
2ptn-6pti	29	10	—	17
2st1-3ssi	13	21	34	1
5cha-2ovo	108	65	111	14

[a]Dash indicates no hits found for this complex.

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

TABLE VI. Performance Means and Deviations Over All ACP + Elec-Filtered Complexes

Free-energy terms[a]	RMSD of top 10 energies (Å)	Number of hits[b] in the top 10 energies	Rank of 1st hit	Best energy RMSD(Å)
ΔG_{ACP}	29.5 ± 16.1	1.9 ± 2.5	242.3 ± 301.4	29.6 ± 25.9
ΔE_{elec}	36.4 ± 15.6	0.6 ± 1.6	211.0 ± 214.5	39.1 ± 15.4
ΔE_{vdW}	26.0 ± 11.4	1.6 ± 1.8	27.3 ± 45.5	24.2 ± 15.2
ΔE^{+}_{vdW}	37.0 ± 9.8	0.0 ± 0.0	140.6 ± 167.5	35.4 ± 19.4
$\Delta G_{ACP} + \Delta E_{elec}$	36.2 ± 13.5	0.5 ± 1.3	64.7 ± 75.8	38.1 ± 16.5
$\Delta G_{ACP} + \Delta E_{vdW}$	27.0 ± 11.4	1.4 ± 2.2	14.0 ± 16.3	21.1 ± 19.9
$\Delta G_{ACP} + \Delta E^{+}_{vdW}$	33.4 ± 13.5	1.0 ± 1.3	140.8 ± 217.2	38.6 ± 21.5
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$	30.0 ± 13.8	1.5 ± 1.9	20.5 ± 29.8	28.8 ± 19.5
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$	37.9 ± 13.9	0.4 ± 1.0	76.6 ± 88.4	36.5 ± 18.5
ΔG_{RPmin}	20.3 ± 13.4	3.0 ± 3.1	9.7 ± 14.4	19.3 ± 16.0
$\Delta G_{RPmin} + \Delta E_{vdW}$	17.1 ± 8.6	3.6 ± 2.9	4.3 ± 4.2	13.9 ± 7.5
$\Delta G_{RPmin} + \Delta E^{+}_{vdW}$	23.9 ± 12.6	1.6 ± 2.5	13.4 ± 22.3	28.7 ± 22.0

[a]See Table III.
[b]Conformations with less than 10 Å RMSD (see text).

TABLE VII. Ranking the Conformations Retained by the ACP+Elec Filter for Each Complex[a]

	ACP+vdW		ACP+Elec+vdW		RPScore		RPScore+vdW		RPScore+Bump	
Proteins	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit
1a2p-1a19	—	27	2	3	4	1	8	1	2	1
1chg-1hpt	7	1	3	2	9	1	8	1	8	1
1mlb-1lza	—	40	—	72	2	2	3	2	1	10
1ppg-2ovo	1	1	—	9	7	2	4	1	3	3
1sup-2ci2	1	1	—	13	—	45	1	7	—	75
2ace-1fsc	1	5	1	2	—	12	1	7	—	12
2ptn-4pti	1	10	2	4	—	24	3	2	—	16
2ptn-6pti	—	42	—	18	2	7	—	14	—	11
2stl-3ssi	—	11	—	80	4	1	2	6	1	2
5cha-2ovo	3	2	6	2	2	2	6	2	1	3

[a]Dash indicates no hits found for this complex.
[b]Number of hits in the top 10 energies.

ΔE_{vdW} potential would improve discrimination of the charged complexes, but we observed substantial improvement only for the highly charged barnase–barstar complex (1a2p–1a19).

Ranking the Structures Retained by the RPScore-Based Filter

For these filtered structures, we have applied the same discriminatory potentials as for the ACP+Elec filtered complexes. For comparison purposes, we have also included statistics on the unminimized RPScore (ΔG_{RP}) values for the 10 sets of 500 complexes (Table T8 VIII). Similar to the ACP+Elec-filtered decoys, the average RMSD of the 10 best energies and the RMSD of the lowest energy structure are high, and, again, these statistics are not very meaningful. In terms of the number of hits, the best performing function is $\Delta G_{ACP} + \Delta E_{vdW}$, although $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ is nearly as good (3.7 compared to 3.1). These two potentials also rank near-native structures higher than the other potentials, with average rankings of 11.5 and 11.2, respectively. In the decoy study, minimization did not lead to a clear improvement in RPScore performance. Here, however, minimization generally improves the discriminatory ability of the RPScore potential, with the largest improvement occurring within the rankings of the first hit (16.2 as opposed to 39.30; see Table VIII). The outcomes of the two studies may differ simply because the docking sets are much larger than the decoys sets; therefore, the significance of a change attributed to the minimization is easier to assess for the docking study.

The statistics for the two ACP-based potentials, $\Delta G_{ACP} + \Delta E_{vdW}$ and $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$, are significantly better for the RPScore-filtered complexes than for the ACP+Elec-filtered complexes (see Tables VII and IX). T9 Following the RPScore filter, $\Delta G_{ACP} + \Delta E_{vdW}$ and $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ find hits in the top 10 energies for 8 and 7 of the 10 decoy sets, respectively. $\Delta G_{RPmin} + \Delta E_{vdW}$, however, only finds hits in the top 10 energies for 6 decoy sets. It is interesting to note that the two ACP-based potentials do not favorably discriminate near-native complexes for the decoy set 2ace-1fsc, following the application of the RPScore filter, in stark contrast to the ACE+Elec-filtered complexes. These observations emphasize the importance of the free-energy filter, while reaffirming the plausible bias inherent in the RPScore potential. We also

TABLE VIII. Performance of Potentials in Ranking Conformations Retained by the RPScore Filter

Energy[a]	Average RMSDs of top 10 energies (Å)	Number of hits[b] in the top 10 energies	Rank of 1st hit	Best energy RMSD Å
ΔG_{ACP}	23.7 ± 12.1	1.8 ± 2.5	37.1 ± 42.8	20.2 ± 14.5
ΔE_{elec}	32.0 ± 13.3	0.8 ± 2.5	93.6 ± 142.3	39.1 ± 15.3
ΔE_{vdW}	22.6 ± 12.8	2.5 ± 3.2	21.1 ± 44.2	14.7 ± 10.3
ΔE^{+}_{vdW}	30.1 ± 11.5	0.7 ± 1.6	39.3 ± 44.6	28.1 ± 17.2
$\Delta G_{ACP} + \Delta E_{elec}$	28.3 ± 13.1	1.9 ± 2.7	34.4 ± 71.8	26.2 ± 18.9
$\Delta G_{ACP} + \Delta E_{vdW}$	19.1 ± 12.5	3.7 ± 3.4	11.5 ± 28.4	14.5 ± 15.1
$\Delta G_{ACP} + \Delta E^{+}_{vdW}$	25.6 ± 13.2	1.6 ± 2.9	26.9 ± 29.7	22.8 ± 13.0
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$	21.9 ± 12.7	3.1 ± 3.8	11.2 ± 16.2	20.1 ± 17.9
$\Delta G_{ACP} + \Delta E_{elec} + \Delta E^{+}_{vdW}$	27.3 ± 14.6	2.0 ± 2.9	37.5 ± 47.4	27.2 ± 18.4
ΔG_{RP}[c]	27.3 ± 17.2	2.2 ± 3.7	39.3 ± 46.8	20.7 ± 14.4
ΔG_{RPmin}	23.4 ± 16.2	2.4 ± 3.7	16.2 ± 21.6	25.9 ± 23.3
$\Delta G_{RPmin} + \Delta E_{vdW}$	21.6 ± 14.1	2.9 ± 3.4	17.3 ± 34.4	20.3 ± 16.2
$\Delta G_{RPmin} + \Delta E^{+}_{vdW}$	26.0 ± 15.4	2.2 ± 2.9	16.0 ± 18.7	23.5 ± 17.3

[a]See Table III.
[b]Conformations with less than 10 Å RMSD (see text).

TABLE IX. Ranking the Conformations Retained by the RPScore Filter for Each Complex[a]

	ACP+vdW		ACP+Elec+vdW		RPScore		RPScore+vdW		RPScore+Bump	
Proteins	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit	Number of hits[b]	Rank of top hit
1a2p-1a19	8	1	10	1	10	1	9	1	7	1
1chg-1hpt	9	1	7	1	8	1	7	1	6	1
1mlb-1lza	2	3	1	2	—	17	—	24	—	48
1ppg-2ovo	1	1	1	1	—	69	1	5	—	13
1sup-2ci2	—	—	—	—	—	—	—	—	—	—
2ace-1fsc	—	92	—	30	—	37	—	112	—	44
2ptn-4pti	3	2	2	4	2	3	2	1	1	5
2ptn-6pti	1	8	—	49	—	15	—	22	1	10
2st1-3ssi	4	1	2	7	1	3	1	4	—	34
5cha-2ovo	8	1	8	1	3	3	7	2	6	1

[a]Dash indicates no hits found for this complex.
[b]Number of hits in the top 10 energies.

found that the RPScore filter did not retain a single hit below 10 Å for the complex 1sup-2ci2 (Tables V and IX). This result is due to our use of the ligand RMSD as opposed to a binding-site RMSD. When a binding-site RMSD is applied to the complexes in this decoy set, many highly ranked near-native complexes are revealed (see Table X). T10

Comparison of Discrimination Protocols

According to data in Tables VI and VIII, the best discrimination strategies combine RPScore as a filter with an ACP-based potential for rescoring, or the ACP+Elec filter with an RPScore-based potential for rescoring. In particular, following an RPScore filter, $\Delta G_{ACP} + \Delta E_{vdW}$ and $\Delta G_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ yield 3.7 and 3.1 hits, respectively, in their top 10 energies. Because we retain only 500 conformations with this filter, these are stronger results than those obtained for the ACP+Elec-filtered complexes in which we retain the top 2000 complexes. The rankings of the first hits are less impressive, 11.5 and 11.2, respectively, for the two methods. However, these averages are heavily affected by the results for a single complex (i.e., 2ace-1fsc). In fact, retaining 2000 conformations rather than 500 for 2ace-1fsc from the RPScore-based filter improves these rankings to 3.9 and 7.8, respectively.

In terms of the number of hits, the next best strategy employs ACP+Elec as the filter and $\Delta G_{RPmin} + \Delta E_{vdW}$ as the scoring function, and yields 3.6 hits. Because we retained and minimized 2000 conformations rather than 500, these results are somewhat weaker than those provided by the combination of the RPScore-based filter and ACP-based scoring potentials. However, this combination ranks the conformations very well, yielding 4.3 as the average rank of the first hit. As already mentioned, the strategy finds hits within the top 10 structures for 9 of the 10 complexes (Table VII). The only exception is the complex 2ptn-6pti. Interestingly, the alternative strategy of using RPScore for filtering and $\Delta G_{ACP} + \Delta E_{vdW}$ for rescoring works relatively well for this complex.

In Figures 4 and 5, we plot the lowest 100 minimized $\Delta G_{ACP} + \Delta E_{vdW}$ and $\Delta G_{RPmin} + \Delta E_{vdW}$ energies versus RMSD for the ACP+Elec and RPScore-filtered structures of the antigen–antibody complex, 1mlb-1lza. Figures 4(d) (ACP+Elec-filtered, $\Delta G_{RPmin} + \Delta E_{vdW}$ rescored) and 5(d) (RPScore-filtered, $\Delta G_{ACP} + \Delta E_{vdW}$ rescored) show hits F4-F5

TABLE X. Comparison of Five Docking Protocols[a]

		DOT (unfiltered)		Gardiner et al.[25]		Chen and Weng[24]		This article[b]		This article[c]	
Complex	RMSD cutoff for hits (Å)[d]	Best rank	RMSD of the top hit (Å)	Best rank	RMSD of the top hit (Å)	Best rank	RMSD of the best hit (Å)	Best rank	RMSD of the top hit (Å)	Best rank	RMSD of the top hit (Å)
1brs	4.47	1705	1.15	1	3.14	—	—	1	2.41	1	1.15
1cgi	5.48	12292	1.27	34	5.04	6	2.61	1	2.56	1	4.84
1cho	4.33	348	3.57	1	1.79	72	1.53	1	1.79	2	3.56
2ptc	4.74	8802	2.74	1	4.38	6	1.62	8	2.9	5	4.63
2sic	4.38	7746	1.02	25	3.91	1	1.93	1	1.93	6	3.79
2sni	4.85	3054	1.38	51	2.05	2	1.20	1	4.69	7	0.86

[a]Dash indicates no hits reported.
[b]As determined by DOT, RPScore filtering, CHARMM minimization, and rescoring with the potential ACP+vdW (see text).
[c]As determined by DOT, ACP+Elec filtering, CHARMM minimization, and rescoring with the potential RPScore+vdW (see text).
[d]As shown in Chen and Weng[24] (see text).

ranked in the top 2 and 3 energies, respectively, and 2 or more hits ranked in the top 10 energies (see Tables VII and IX). In contrast, neither Figure 4(c) (ACP+Elec-filtered, $\Delta G_{ACP} + \Delta E_{vdW}$ rescored), nor Figure 5(c) (RPScore-filtered, $\Delta G_{RPmin} + \Delta E_{vdW}$ rescored) show even a single hit ranked in the top 10 energies. In Figures 4(b) and 5(b), we show the effectiveness of filtering the top 20,000 DOT complexes with the ACP+Elec and RPScore filters. In each case, we were able to reduce the number of putative complexes by at least a factor of 10, while retaining several near-native complexes. Comparisons among Figures 4(b, c, and d) and 5(b, c, and d) show the dramatic improvement that ΔE_{vdW} has on each potential following minimization.

These observations show that there is a beneficial complementarity between using RPScore as a filter and reranking with an ACP-dependent potential, and vice versa. The mixed strategies provide substantially better discrimination than our previous method[3] of selecting conformations with low values of ΔG_{ACP} and ΔE_{elec}, minimizing the structures, and scoring them by $\Delta_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ (see Tables VI and VIII). Indeed, the RPScore-based filtering is generally more effective than the filter combining Δ_{ACP} and ΔE_{elec} (Fig. 2). Because $\Delta_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$ ranks selected conformations better than ΔG_{RPmin}, primarily due to the inclusion of the van der Waals term, the strong performance observed when combining an RPScore filter with an ACP-dependent potential is understandable. However, it is interesting that the best performance is shown by the opposite combination, which uses ACP+Elec as the filter and $\Delta G_{RPmin} + \Delta E_{vdW}$ as the scoring function. The most likely explanation is that the RPScore and ACP potentials are based on different and complementary information.

As shown in Figure 3(b), for charged complexes, the RPScore filter performs much better than either the ACP or the ACP+Elec potentials. This is a major advantage of the RPScore potential, because the treatment of electrostatics is a central problem in rigid-body docking. We have shown that changing the conformations of a few sidechains in a near-native structure may lead to the loss of favorable electrostatic interactions, and the inability to favorably rank the structure with the free-energy function $\Delta_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$.[27] Indeed, the contribution of electrostatics to discrimination is very noisy and, by a number of measures, $\Delta_{ACP} + \Delta E_{vdW}$ performs better than $\Delta_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$. Therefore, reliable electrostatic interactions can be accurately calculated only if the interface sidechains are correctly oriented, which is rarely the case when docking separately crystallized proteins. Thus, it is an important observation that $\Delta G_{RPmin} + \Delta E_{vdW}$ is less sensitive to small errors in the coordinates than $\Delta_{ACP} + \Delta E_{elec} + \Delta E_{vdW}$.

Comparison with Earlier Results

Six of the 10 complexes considered in our docking study (Table III) were also studied by Gardiner et al.[25] and Chen and Weng.[24] Table X compares our results with these previously published docking calculations. The comparative rankings were based on a binding-site RMSD as detailed in Gardiner et al.[25]; the results were taken from Chen and Weng[24] (we have not included results for 2kai, 1udi, and TEM-BLIP, because these complexes were not part of our original data set in the docking study). Considering that neither of these earlier studies included any postprocessing, their results are extremely good compared to the results we have obtained using Version 1.0 alpha of the DOT program[15] without filtering and reranking (Table X). Indeed, although DOT did not find any hits among the top 1000 structures for 5 of the 6 complexes, Gardiner et al.[25] and Chen and Weng[24] find hits within the top 50 structures for almost all complexes.

In spite of our poor docking results, the postprocessing is clearly very effective, because the combinations RPScore filter, with $\Delta_{ACP} + \Delta E_{vdW}$ as the scoring function, and ACP+Elec filter, with $\Delta G_{RPmin} + \Delta E_{vdW}$as the scoring function, rank near-native hits in the top 10 energies for all 6 complexes. Although the average RMSDs of our best hits are larger than that of Chen et al.[24] (2.71 Å, and 3.14 Å compared to 1.78 Å), they are certainly comparable; both are also lower than that of Gardiner et al.[25] (3.39 Å). We note that this comparison does not include the complex, 1fss, with which we struggle, because Gardiner et al. do not include this complex in their analysis. A comparison with Chen and Weng, however, suggests that they struggle with this complex as well (the best reported hit is ranked 210th).[24]

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

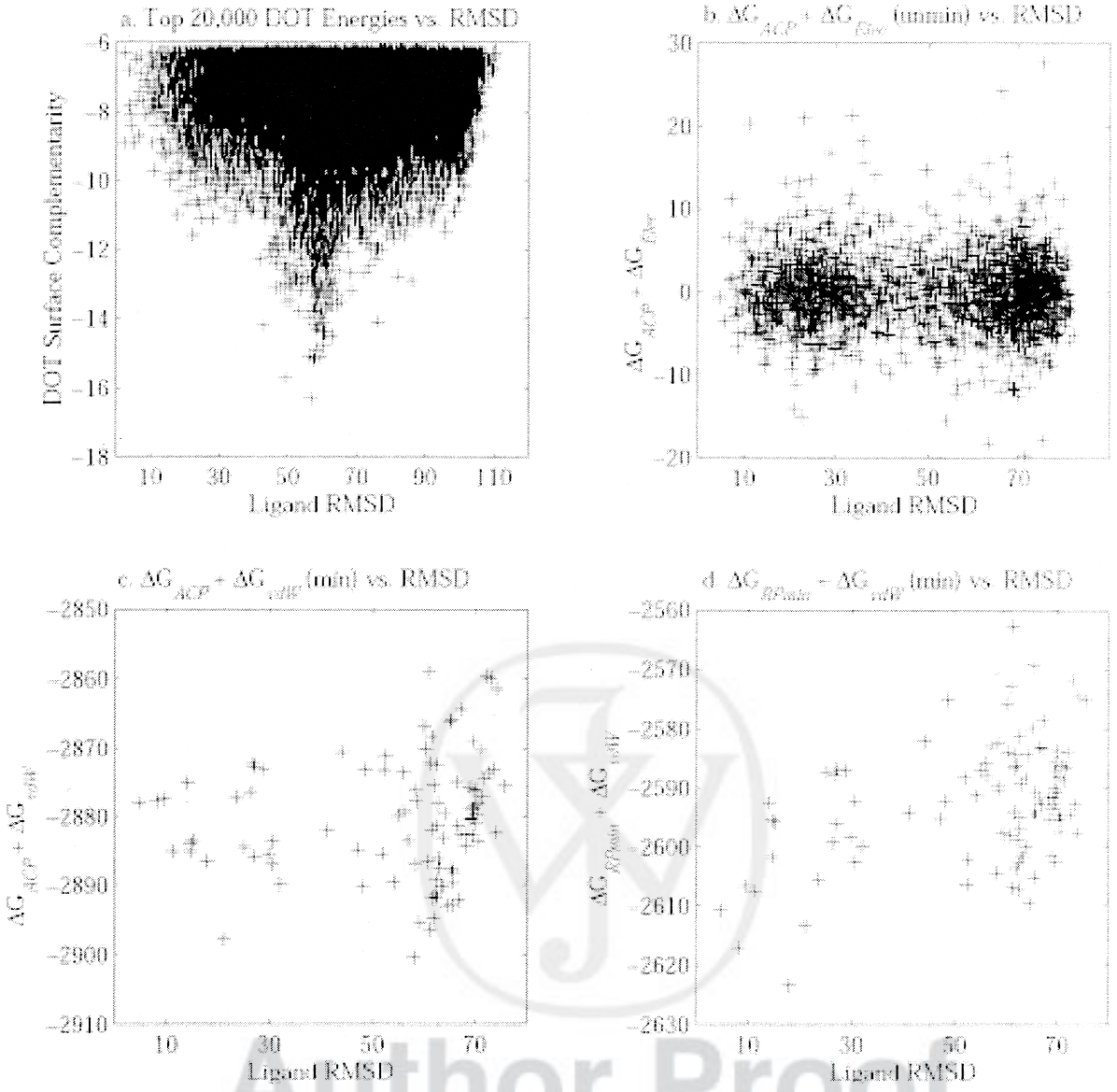

Fig. 4. Plots of four scoring functions versus RMSD for the antigen–antibody complex, 1mlb-1lza. (**a**) Top 20,000 DOT energies (as determined by surface complementarity only) versus RMSD, preminimization. (**b**) Top 2000 $\Delta G_{ACP} + \Delta E_{elec}$ energies (ACP+Elec filtered) versus RMSD, preminimization. (**c**) Top 100 $\Delta G_{ACP} + \Delta E_{vdW}$ energies (ACP+Elec filtered) versus RMSD, postminimization. (**d**) Top 100 $\Delta G_{RPmin} + \Delta E_{vdW}$ energies (ACP+Elec filtered) versus RMSD, postminimization.

CONCLUSIONS

We studied the abilities of two structure-based potentials to discriminate near-native from non-native structures in sets of conformations generated by rigid-body docking. The ACP (also denoted Δ_{ACP}) is an atom-level extension of the Miyazawa–Jernigan potential, and the RPScore (or ΔG_{RP}) is a residue-level pairwise potential derived from protein–protein interface data. Both potentials were combined with other energy terms, representing further contributions to the binding free energy, to improve discrimination. The resulting functions have been tested on two types of decoy sets containing conformations from rigid-body docking. The first decoys were obtained for 13 complexes from the output of the FTDock program.[15] For each complex, the set includes 99 structures with good surface complementarity, selected to cover the 5–40 Å RMSD range as evenly as possible. We minimized all conformations for 1000 steps in CHARMM and then rescored using 16 different scoring functions. We based the second part of this article on the docking of 10 unbound protein pairs, using the DOT program. For each complex, the 20,000 highest scoring docked structures were retained. Two of the scoring functions, RPScore and ACP+Elec, were applied before minimization to reduce the number of putative complexes. Following a filtering step in which we select either 500 or 2000 conformations, the retained structures were minimized and rescored, as in the first decoy study.

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | Art: 04132002

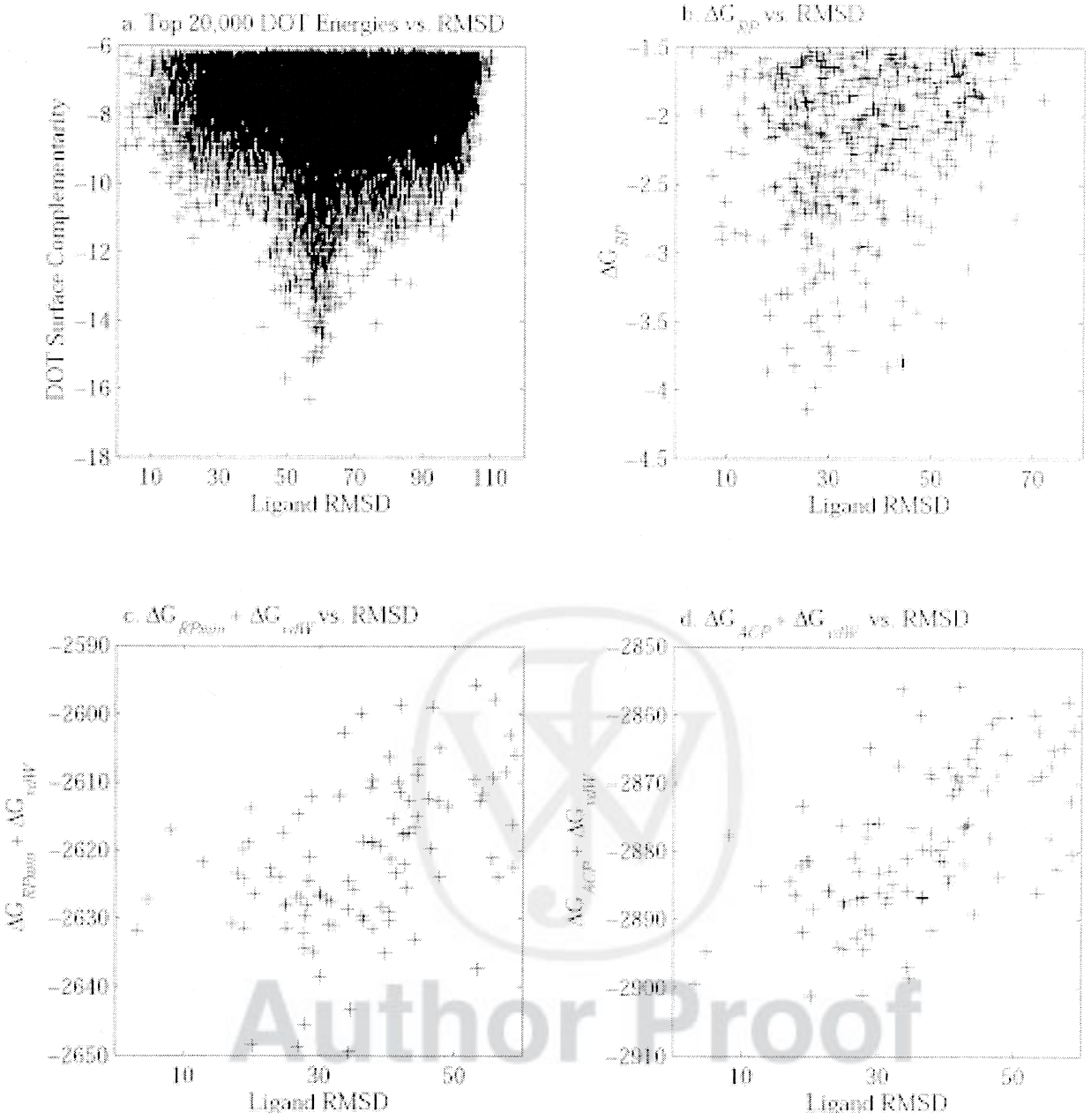

Fig. 5. Plots of four scoring functions versus RMSD for the antigen–antibody complex, 1mlb-1lza. (**a**) Top 20,000 DOT energies (as determined by surface complementarity only) versus RMSD, preminimization. (**b**) Top 500 RPScore values (RPScore filtered) versus RMSD, preminimization. (**c**) Top 100 $\Delta G_{RPmin} + \Delta E_{vdW}$ energies (RPScore filtered) versus RMSD, postminimization. (**d**) Top 100 $\Delta G_{ACP} + \Delta E_{vdW}$ energies (RPScore filtered) versus RMSD, postminimization.

We were able to substantially improve the discriminatory ability of both the ACP and RPscore potentials by refining the structures via energy minimization, and by adding terms representing further contributions to the binding free energy. In particular, the addition of the van der Waals energy term, ΔE_{vdW}, substantially improves the ranking of docked conformations; adding other terms to $\Delta G_{RPmin} + \Delta E_{vdW}$ does not yield further improvement. In contrast, the addition of the electrostatic term, ΔE_{elec}, to $\Delta G_{ACP} + \Delta E_{vdW}$ generally improves discrimination if the association is driven by electrostatic interactions (i.e., if the two proteins have large and opposite charges). However, adding ΔE_{elec} may slightly worsen discrimination if at least one of the proteins is neutral or weakly charged, and/or if any of the key sidechains sample incorrect rotamers. The addition of a term representing the sidechain entropy loss upon the formation of the complex does not yield any improvement.

Because the van der Waals energy is very sensitive to small structural perturbations, the combined potentials, $\Delta G_{ACP} + \Delta E_{vdW}$, and $\Delta G_{RPmin} + \Delta E_{vdW}$ can be calculated only for the minimized structures. ΔG_{RP} is a residue-level,

coarse-grained potential; thus, we expected that its behavior would not be heavily affected by minimization. However, we found that the minimization of docked structures leads to better discrimination even by RPScore alone, as opposed to the case in which the structures are not minimized. Although ΔG_{RP} is thought to be a smooth potential that will gloss over any smaller-scale defects, it may be specific enough to take advantage of the improved complementarity of the interface after minimization.

The analysis of docked conformations in the second part of this article involves both filtering and rescoring steps. Because filtering deals with unminimized structures, only ΔG_{ACP}, ΔE_{elec}, and ΔG_{RP} can be used. We have found that, on average, ΔG_{RP} is the best filter, resulting in the highest number of hits. Although ΔG_{ACP} works better than ΔG_{RP} for the neutral complexes, its performance deteriorates if the proteins are strongly charged. The ACP+Elec filter retains some conformations with good ACP values, and some with favorable electrostatics, but, on average, it is unable to match the number of hits produced by the RPScore filter for charged complexes.

The main result of this article is that the best discrimination strategies combine either the RPScore filter with an ACP-based scoring function, $\Delta G_{ACP} + \Delta E_{vdW}$, or the ACP+Elec filter with the RPScore-based scoring function $\Delta G_{RPmin} + \Delta E_{vdW}$. Both strategies work much better than using either the ACP or the RPScore potential for both filtering and scoring. This implies that ACP and RPScore, having been extracted from different sets of protein structures, capture complementary structural information. These observations also suggest that the ability to find near-native structures in large sets of structures produced by rigid-body docking can be improved by combining different potentials in a multistage postprocessing procedure.

ACKNOWLEDGMENTS

Our thanks to Prof. Michael J. E. Sternberg and Dr. Graham R. Smith for the set of FTDock-generated decoys and for helpful discussions, and to Dr. Lynn F. Ten Eyck for the DOT program.

REFERENCES

1. Camacho C, Vajda S. Protein–protein association kinetics and protein docking. Curr Opin Struct Biol 2002;12:36–40.
2. Smith GR, Sternberg MJE. Prediction of protein–protein interactions by docking methods. Curr Opin Struct Biol 2002;12:28–35.
3. Camacho CJ, Gatchell DW, Kimura SR, Vajda S. Scoring docked conformations generated by rigid-body protein–protein docking. Proteins 2000;40:525–537.
4. Zhang C, Vasmatzis G, Cornette JL. Determination of atomic desolvation energies from the structures of crystallized proteins. J Mol Biol 1997;267:707–726.
5. Janin J. Protein–protein recognition. Prog Biophys Mol Biol 1995;64:145–166.
6. Jones S, Thornton JM. Principles of protein–protein interactions. Proc Natl Acad Sci U S A 1996;93:13–20.
7. Lo Conte L, Chothia C, Janin J. The atomic structure of protein–protein recognition sites. J Mol Biol 1998;285:2177–2198.
8. Norel R, Petrey D, Wolfson HJ, Nussinov R. Examination of shape complementarity in docking of unbound proteins. Proteins 1999;36:307–317.
9. Camacho CJ, Weng Z, Vajda S, DeLisi C. Free energy landscapes of encounter complexes in protein–protein association. Biophys J 1999;76:1166–1178.
10. Vakser IA. Low-resolution docking: Prediction of complexes for undetermined structures. Biopolymers 1996;39:455–464.
11. Vakser IA, Matar OG, Lam CF. A systematic study of low-resolution recognition in protein–protein complexes. Proc Natl Acad Sci U S A 1999;96:8477–8482.
12. Moont G, Gabb HA, Sternberg MJE. Use of pair potentials across protein interfaces in screening predicted docked complexes. Proteins 1999;35:364–373.
13. Zhang C, Chan J, DeLisi C. Protein–protein recognition: Exploring the energy funnels near the binding sites. Proteins 1999;34:255–267.
14. Vajda S, Sippl M, Novotny J. Empirical potentials and functions for protein folding and binding. Curr Opin Struct Biol 1997;7:222–228.
15. Gabb HA, Jackson RM, Sternberg MJE. Modelling protein docking using shape complementarity, electrostatics, and biochemical information. J Mol Biol 1997;272:106–120.
16. Sternberg MJ, Gabb HA, Jackson, RM, Moont G. Protein–protein docking: Generation and filtering of complexes. Methods Mol Biol 2000;143:399–415.
17. Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. Protein docking using continuum electrostatics and geometric fit. Protein Eng 2001;14:105–113.
18. Miyazawa S, Jernigan RL. Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation. Macromolecules 1985;18:534–552. AQ: 8
19. Miyazawa S, Jernigan R. Residue–residue potentials with a favorable contact pair term and unfavorable high packing density terms, for simulation and threading. J Mol Biol 1996;256:623–644.
20. Schaefer M, Karplus M. A comprehensive analytical treatment of continuum electrostatics. J Phys Chem 1996;100:1578–1599.
21. Vajda S, Weng Z, Rosenfeld R, DeLisi C. The effect of conformational flexibility and solvation on receptor–ligand binding free energies. Biochemistry 1994;33:13977–13987.
22. Weng Z, Vajda S, DeLisi C. Prediction of protein complexes using empirical free energy functions. Protein Sci 1996;5:614–626.
23. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4:187–217.
24. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47:281–294.
25. Gardiner EJ, Willet P, Artymiuk PJ. Protein docking using a genetic algorithm. Proteins 2001;44:44–56.
26. Janin J. Welcome to CAPRI: A critical assessment of predicted interactions. Proteins 2002;47:257–257. AQ: 9
27. Roseman AM. Docking structures of domains into maps from cryo-electron microscopy using local correlation. Acta Crystallogr D 2000;56:1332–1340.
28. Wriggers W, Chacon P. Modeling tricks and fitting techniques for multi-resolution structures. Structure 2001;9:779–788.
29. Camacho CJ, Vajda S. Protein docking along smooth association pathways. Proc Natl Acad Sci U S A 2001;98:10636–10641.
30. Kimura SR, Brower R, Vajda S, Camacho CJ. Dynamical view of the positions of key side chains in protein–protein recognition. Biophys J 2001;80:635–642.

tapraid5/7e-protein/7e-protein/7e1103/7e1368d03g | royerl | S=11 | 5/21/03 | 10:32 | 4/Color Figure(s): | Art: 04132002

AQ1: If running head is not OK, please restrict new one to 45 or fewer characters.

AQ2: RPScore is spelled out two different ways in manuscript: residue potential score and residue pairing score. Choose one throughout for consistency?

AQ3: Sentence beginning “First,. . .” as meant? Moont et al. OK, as in references?

AQ4: Please make sure URL is current and active.

AQ5: Is “cratic” correct?

AQ6: Gabb et al. OK as in references?

AQ7: Exact date of private communication?

AQ8: Is 1985 correct?

AQ9: Please provide inclusive page nos.

AQT1: variant?

AQT2: Please make sure URL is current and active.

AQT3: meaning? Same as — for no data?