WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

The performance of ZDOCK and ZRANK in rounds 6–11 of CAPRI

Kevin Wiehe,[1] Brian Pierce,[1] Wei Wei Tong,[2] Howook Hwang,[1] Julian Mintseris,[1] and Zhiping Weng[1,2]*

[1] Bioinformatics Program, Boston University, Boston, Massachusetts 02215

[2] Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

ABSTRACT

We present an evaluation of our protein–protein docking approach using the ZDOCK and ZRANK algorithms, in combination with structural clustering and filtering, utilizing biological data in Rounds 6–11 of the CAPRI docking experiment. We achieved at least one prediction of acceptable accuracy for five of six targets submitted. In addition, two targets resulted in medium-accuracy predictions. In the new scoring portion of the CAPRI exercise, we were able to attain at least one acceptable prediction for the three targets submitted and achieved three medium-accuracy predictions for Target 26. Scoring was performed using ZRANK, a new algorithm for reranking initial-stage docking predictions using a weighted energy function and no structural refinement. Here we outline a practical and successful docking strategy, given limited prior biological knowledge of the complex to be predicted.

Proteins 2007; 69:719–725.

Key words: protein docking; CAPRI; ZDOCK; ZRANK.

INTRODUCTION

The CAPRI experiment gives the structural bioinformatics community an arena to practically test protein–protein docking algorithms. It also acts as a catalyst of innovation by presenting challenges that require new methods in protein–protein docking. Our own experience in CAPRI has led us to such innovations[1] and has also served to teach us how to tailor our docking approach to real-world docking problems. We participated in the Rounds 1–5 of CAPRI,[2,3] with the goal of attaining at least one "acceptable" prediction for each target (predictions are classified by the CAPRI assessor as high quality, medium quality, or acceptable quality).[3] We succeeded in doing so for 10 out of 16 targets with our fast Fourier transform (FFT)-based rigid-body protein docking algorithm, ZDOCK.[4–6] ZDOCK was trained and tested on two protein–protein docking benchmarks[7,8] that have been made freely available to the docking community. Recently we developed a new statistical potential called IFACE and implemented it in the ZDOCK program.[9] In addition, we have developed a new reranking algorithm called ZRANK to accurately score ZDOCK predictions.[10] In Rounds 6–11 of CAPRI, we applied these two new improvements to our docking strategy, and here we describe our performance. Our ability to rerank and refine predictions of various docking algorithms was also tested in three rounds of the new scoring section of CAPRI and is reported here.

MATERIALS AND METHODS

Our general docking approach throughout all rounds of CAPRI has been similar.[11,12] However, each target in CAPRI entails a specific strategy based on the biological data known prior to docking, as will be shown in the Results section. The general approach involves first mining the literature for biological data about the interaction. This information is applied in ZDOCK in two ways. During docking, ZDOCK can downweight predictions with interfaces that lack residues of interest, a method we refer to as "blocking." After docking, predictions can also be ranked and filtered based on distances between residues of interest. ZDOCK is typically run with 6° rotational sampling and generates 54,000 predictions, a subset of which is then reranked using one of our scoring algorithms. We can apply clustering to remove structural redundancy either before or after the scoring step. A portion of the remaining clusters is visually inspected to cull the list of predictions down to 10 submissions. We have recently developed a new scoring function and a new refinement algorithm, the details of which follow.

Conflict of interest: The authors state that the docking programs used in this article are licensed to Accelrys Inc. through Boston University.
Grant sponsor: NSF; Grant numbers: DBI-0133834, DBI-0116574.
*Correspondence to: Zhiping Weng, Bioinformatics Program, Boston University, Boston, MA 02215.
E-mail: zhiping@bu.edu
Received 31 May 2007; Revised 23 July 2007; Accepted 24 July 2007
Published online 5 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21747

ZDOCK 2.4 and 3.0

We developed two new versions of ZDOCK that utilize the IFACE statistical potential[13] to score protein docking predictions, versus the atomic contact energies[14] that were used in the previous version, ZDOCK 2.3. IFACE was developed as a pair potential specifically geared toward detecting transient protein–protein interfaces and was trained on a nonredundant dataset of 150 transient complexes. It is implemented in two different versions of ZDOCK.[9] In ZDOCK 2.4, IFACE energies are substituted for atomic contact energy values, and the scoring is implemented in the same method as in 2.3 by computing one additional FFT with the average value of all atom pairs for each atom type. In ZDOCK 3.0, atom pair energies are calculated explicitly by six additional FFTs. In practice ZDOCK 3.0 is about three times slower than previous versions of ZDOCK because of the computation time for the additional FFTs, but is more accurate than ZDOCK 2.4. Testing on the Benchmark 2.0 showed considerable improvement in overall success rate and number of hits, for both 2.4 and 3.0, over previous ZDOCK versions. ZDOCK 2.4 preceded 3.0 in development and we first applied it to targets 24 and 25 in CAPRI Round 9. ZDOCK 3.0 was used in all subsequent rounds.

ZRANK

ZRANK is a scoring method developed to quickly rerank initial-stage docking predictions using an energy-based potential.[10] In contrast to RDOCK, which was previously developed by our lab, ZRANK does not require structural minimization of the predictions; thus it is able to process all 54,000 predictions from a ZDOCK run using fine sampling. ZRANK was shown to lead to significant improvements in success rates for two versions of ZDOCK on Benchmark 2.0.[10]

RESULTS

Docking performance

Our overall performance in Rounds 6–11 is summarized in Table I. We were able to attain at least one acceptable classification for 5 out of the 6 targets. Two targets, 25 and 27, resulted in predictions of "medium" accuracy and a third target, 26, missed this classification by only 0.25-Å interface RMSD. Additionally, we were able to attain more than one acceptable prediction in Targets 25–27. Our improved success in the later rounds may be a reflection of the upgrading of our docking and scoring algorithms at that time. The details of the methods used for each target and their evaluation are described as follows.

T20 (RF-1/HemK)

Target 20 was an unbound homology-model docking case involving the prediction of the HemK/RF-1 complex. HemK is a methyl transferase that methylates RF-1, a prokaryotic ribosomal release factor, in a step that is critical in stop codon recognition.[15] From the literature, we were able to determine a putative interface for this complex. HemK recognizes the GGQ motif on a loop of RF-1 and methylates Gln 235.[16] *S*-Adenosyl-L-homocysteine of HemK is the methyl donor[17] and the motif NPPY is important for positioning the Gln substrate.[18] The unbound structure for RF-1 was homology modeled with the 3D-JIGSAW server,[19] using the X-ray structure of RF-2 as a template, which had a sequence identity to RF-1 of 40%. Our homology model lacked a loop conformation that exposed Gln 235. Therefore, we remodeled the loop by first using PSI-BLAST[20] to identify a homologous loop in the NMR structure of a peptidyl-tRNA hydrolase domain (PDB code 1J26), and then selecting a loop conformation from the 20 NMR models that maximally exposed Gln 235. We added this new loop conformation into the RF-1 model and used Insight II[21] to optimize side-chain positions. In addition, we removed all but 81 residues (211–292) of RF-1 to isolate the fold with the GGQ loop for docking. Blocking was used to give preference to binding at the putative interface hypothesized from the literature. Two independent runs of ZDOCK 2.1 and ZDOCK 2.3 resulted in 108,000 predictions at 6° rotational sampling. A distance filter was applied to this set of predictions. All predictions with a distance greater than 10 Å between the C_α carbon of the methylation target Gln 235 of RF-1 and the S_δ atom of *S*-adenosyl-L-homocysteine bound to HemK were eliminated. The full RF-1 model was added in place of the RF-1 fragment used in docking, and any predicted complexes with subsequent significant clashing were discarded. Structural clustering was used to aid in the manual inspection of structures, and the aforementioned distance was emphasized when ranking the final predictions. The top 10 predictions underwent a CHARMM[22] energy minimization step to relieve any remaining interface clashes.

Our top-ranked model made many of the appropriate contacts for methylation, as seen in the crystal structure of the complex.[15] However, because of the severe inaccuracy in the loop conformation of our homology model (13.13 Å RMSD), we did not achieve any acceptable-rated predictions. In addition, our docking of a fragment of RF-1 (residues 211–292) instead of the entire RF-1 may have affected ZDOCK's ability to identify near-native predictions. The RF-1 fragment we used did not include a region (residues 127–157) that is in contact with HemK in the solved complex.

Target 21 (Sir1/Orc1)

For Target 21 we were asked to perform unbound–unbound docking of the Orc1p subunit of the origin recognition complex of yeast with Sir1p, a silent informa-

DOI 10.1002/prot

Table I
Docking Performance

Target	Complex	Classification	Fnat[a] (%)	Interface RMSD (Å)[a]	Ligand RMSD (Å)[a]	At least acceptable[b]	ZDOCK[c]
20	HemK/RF-1[d]	Incorrect	19	13.38	28.98	0	2.1, 2.3
21	Sir1p/Orc1p	Acceptable	32	2.80	10.04	1	2.3
24	Arf1/ArfBD[d]	Acceptable	20	3.13	8.94	1	2.4
25	Arf1/ArfBD	Medium	81	1.51	3.80	3	2.4
26	TolB/Pal	Acceptable	50	2.25	5.64	2	3.0
27.2	Hip2/Ubc9	Medium	49	1.86	5.39	4	3.0

[a]Highest Fnat and lowest RMSD of any of the 10 submissions for each target.
[b]Number of predictions with at least an acceptable CAPRI classification.
[c]Version(s) of the ZDOCK program used for docking the target.
[d]Ligand was homology modeled.

tion regulator protein. Orc1p interacts with Sir1p to recruit Sir1p to a silencer.[23] Mutagenesis data from the literature suggested that the interaction involved the SRD region of Sir1p[24] and the H region of Orc1p.[25] Specifically, S491 and R493 in Sir1p[24] and K121 in Orc1p[25] were shown to be crucial for binding. Regions distal to these critical residues were blocked prior to docking, in order for preference to be applied to this putative interface.

For our best-rated prediction, we docked with ZDOCK 2.3 and filtered the results for predictions with close contacts between the SRD and H regions, in general, and the three aforementioned residues, in specific. The lowest energy structure from ZDOCK 2.3 that met the distance criteria was selected as our number one submission and achieved an "acceptable" classification. This prediction is shown in Figure 1(A).

Target 24 (Arf1/ArfBD homology model)

In Target 24, we predicted the complex formed between the ARF1 binding domain (ArfBD) of ARGHGAP21, a Rho family GTPase-activating protein, and Arf1, a small GTPase. This target required a homology model to be built for ArfBD, using the PH domain of β-spectrin (PDB: 1BTN), as a template with 30% sequence identity. We used the five homology models that were generously provided by Baker's lab to the CAPRI community, with the ROBETTA server.[26] Because of a lack of confidence in docking protein structures folded ab initio, we removed from all ArfBD models the 20-residue C-terminal extension that is missing in the template structure. Next, we used ZDOCK 2.4 to dock all five ArfBD models to Arf1 in five independent docking runs. The predictions were pooled and clustered to remove structural redundancy. The binding sites on Ras superfamily G proteins such as Arf1 are well conserved at the switch regions,[27] and we applied distance measures to ensure contacts of ArfBD with the switches. Three loops were also deemed important: the β3/β4 loop which is involved in binding in structurally homologous PH do-

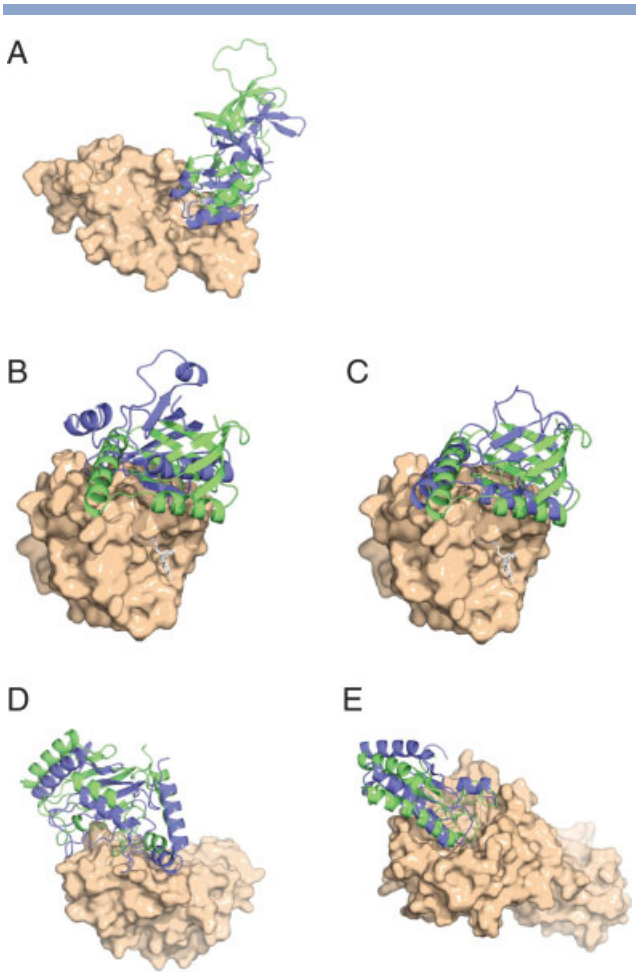

Figure 1
*Docking predictions for Targets 21–27 of CAPRI Rounds 7–11. Orientations of predicted (blue) and crystal structure (green) ligands after superposition of their receptors (beige, only crystal structure receptor shown). (**A**) Target 21: Orc1p (receptor) and Sir1p (ligand); (**B**) Target 24: Arf1 (receptor) and homology model of ArfBD (ligand); (**C**) Target 25: Arf1 (receptor) and bound ArfBD (ligand); (**D**) Target 26: TolB (receptor) and Pal (ligand); (**E**) Target 27: Hip2 (receptor) and Ubc9 (ligand). Figures were created using PyMOL.*

main/Ras complexes,[28,29] the β1/β2 loop which is structurally similar to β3/β4, and the β5/β6 loop which we found to be highly conserved in multiple sequence alignments and was therefore potentially important to binding. Clusters were filtered and ranked according to the distance between the switch regions of Arf1 and the loops of interest in ArfBD.

Our best prediction achieved an acceptable classification and came from a cluster in which the conserved loop of ArfBD contacted switch 2 of Arf1 [Fig. 1(B)]. The 20-residue C-terminal region of ArfBD was improperly folded in the homology model (16.40-Å RMSD for this region compared with 1.19-Å RMSD in the remaining regions) and was critical to binding as it contacted switch 1 of Arf1 in the complex. Removing this region prior to docking turned out to be a double-edged sword. On the one hand, we did not dock the incorrect fold; on the other hand, the contacts that it makes with Arf1 did not contribute to ZDOCK scoring and therefore prevented better predictions.

Target 25 (Arf1/bound ArfBD)

In Target 25, the structure of ArfBD was released from the complex and we were asked to perform unbound–bound docking of Arf1 and ArfBD. Because of missing atoms in the bound structure of ArfBD it was possible to eliminate some of the earlier identified regions of interest from the binding site. Of the ArfBD loops of interest we identified in Target 24, only the conserved β5/β6 loop was totally resolved in the electron density map and therefore we favored predictions with contacts to this loop. The C-terminal extension that we removed prior to docking in T24 was fully solved in the bound ligand structure as a hinge loop and long helix.[30] We preferred predictions involving this region in interactions, as the temperature factors of this region in the bound structure were very low, hinting at its participation in the binding site. For our best predictions, missing side-chain atoms were rebuilt using Insight II[21] and areas near missing residues were blocked prior to docking. We then docked with ZDOCK 2.4 and clustered to remove structural redundancy. Clusters were selected for further inspection based on distance for interactions involving the C-terminal extension and the conserved loop of ArfBD and the two switch regions of Arf1. Our best prediction achieved a medium-accuracy classification with 1.51-Å interface RMSD and included over 81% of correct contacts [Fig. 1(C)]. Because of our high confidence in our binding site model, we also had two predictions that achieved acceptable accuracy classifications (Table I).

Target 26 (TolB/Pal)

For Target 26, we predicted the complex of TolB and Pal, given both unbound components. The TolB-Pal complex is part of a supramolecular assembly of proteins that is important for the structural integrity of the outer membrane of Gram-negative bacteria.[31] TolB consists of two domains, a β-propeller domain and a secondary domain. Extensive mutagenesis data was available, showing that the face of the beta propeller domain of TolB distal to the secondary domain was involved in binding, and that, specifically, residues H246, A249, and T292 were critical.[32] Additional literature implicated contacts at residues 89–104 and 126–130 of Pal.[33] Given an abundance of relevant biological data, we removed the secondary domain of TolB and applied a liberal blocking scheme to place a heavy preference on predictions with contacts between the β-propeller domain and the Pal residues mentioned. We ran ZDOCK 3.0 and results were clustered to remove structural redundancy. Clusters were then ranked both by ZRANK and by the distance between the three critical residues of TolB and residues 89–130 of Pal. Predictions were selected on the basis of the best combination of these ranks. Two of our predictions received acceptable CAPRI classifications, one achieved an interface RMSD of 2.25 Å and the other had 50% of correct contacts predicted [Fig. 1(D)].

Target 27 (Hip2/Ubc9)

For Target 27, we were asked to predict the complex of Hip2 and Ubc9.[34] Ubc9 catalyzes a fusion between the C-terminal residue of the signaling protein Sumo-1 and a lysine of Hip2.[35] In addition to Hip2, Ubc9 also sumolyates RanGap1. The structure of the Ubc9-RanGap1 complex has been solved, and we assumed sumoylation of Hip2 would occur using the same catalytic groove on Ubc9.[36] The X-ray structure of unbound Ubc9 has two residues, Gln 126 and Asn 127, which partially block the putative catalytic groove.[37] We chose to dock with the Ubc9 structure from the Ubc9-RanGap1 complex in which this groove is more accessible[36] and this structure has an identical sequence as our target. In addition, we noticed that Lys 14 of Hip2, which lays in the putative catalytic groove and is fused to Sumo-1 via an isopeptide bond between its N_ζ atom and the terminal carbon atom, was not in an extended conformation as it is in the Ubc9/RanGap1/Sumo-1 complex.[38] We rotated Lys 14 of Hip2 into a nonnative rotamer, in order for it to be more accessible to Ubc9 during docking. Docking was accomplished with ZDOCK 3.0 followed by reranking with ZRANK. Cys 93 of Ubc9 is the catalytic residue involved in the conjugation at the active site and must be in close proximity to Lys 14 for the isopeptide bond to form.[39] We clustered all predictions with less than 5 Å distance between these two residues. Predictions were selected after manual inspection of the remaining clusters.

There are two possible interfaces proposed from the crystal structure of the Ubc9-Hip2 complex and our predictions were evaluated against both. For predictions

 DOI 10.1002/prot

Table II
Scoring Performance

Target	Complex	Classification	Fnat (%)	Interface RMSD (Å)[a]	Ligand RMSD (Å)[a]	At least acceptable[b]	Protocol[c]
25	Arf1/ArfBD	Acceptable	64	2.72	7.98	1	ZR
26	TolB/Pal	Medium	53	1.11 (0.86)	3.04 (3.76)	4	ZR + Ros
27.2	Hip2/Ubc9	Acceptable	32	2.39 (0.46)	6.42 (3.28)	7	ZR + Ros + ZR

[a]Ligand and interface RMSD of top submitted prediction from native. For Targets 26 and 27.2 (for which refinement was performed), the amount of RMSD improvement over the unrefined structure is given in parentheses.
[b]Number of predictions with at least an acceptable CAPRI classification.
[c]Scoring protocol employed; ZR, ZRANK; Ros, RosettaDock refinement.

evaluated using the second interface (T27.2), we achieved one "medium" classification and three acceptable classifications [Fig. 1(E)]. Our medium-classified prediction had an interface RMSD of 1.86 Å and our best acceptable classified prediction correctly identified 49% of the true contacts.

Scoring performance

We participated in the scoring rounds for Targets 25, 26, and 27. Our general strategy for these rounds was to test ZRANK performance on the sets of structures to be scored. In addition, for Targets 26 and 27 we utilized RosettaDock[40] to produce refined structures for the top 10 models. The results from the CAPRI evaluation of our submitted structures for these Targets are given in Table II; we obtained acceptable predictions for Targets 25 and 27, and medium-level predictions for Target 26. The Fnat, interface RMSD, and ligand RMSD of the top prediction for each target (as evaluated by interface RMSD) are also given in Table II. Details on the protocols and results for each target are provided later.

Target 25

For Target 25 (Arf1/Bound ArfBD), we were given 700 predicted complexes to score. We then used RosettaDock to add hydrogens to all structures and scored them using ZRANK. The top 10 models based on ZRANK score were submitted; of these, we had one acceptable prediction, which was ranked no. 7.

Target 26

The next target (Target 26; TolB/Pal) included 1567 predictions to rescore. In this case, we added hydrogens and scored all predictions as before, but also employed a distance filter based on the predicted contacts discussed in the docking portion of the Results section. In particular, we removed all predictions above a 15-Å cutoff, leaving 241 predictions. We found that several of the predictions that passed the filter were highly redundant (based on observations from the residue distances and ZRANK scores), and so we selected one representative structure with the highest ZRANK score and removed the rest.

From the remaining predictions, we took the top 10 based on ZRANK score and refined each of these predictions using RosettaDock (to produce 500 refined models per prediction). The refined model with the best RosettaDock score from the set of 500 was used for submission. Our submissions for this target included three medium predictions and one acceptable prediction, for a total of 4 out of the 10 submitted models that were at least acceptable. The top submitted model had ligand RMSD of 3.04 Å and interface RMSD of 1.11 Å, which is an improvement of 3.76 and 0.86 Å in RMSD over those models because of refinement (Table II; Marc Lensink, personal communication). This indicates the success of the RosettaDock refinement protocol in improving the structures.

Target 27

We employed a similar strategy for our scoring predictions for Target 27.2 (Hip2/UBC9). For this target, we utilized a residue distance measure based on the residues described in the docking portion of the Results section, and removed predictions above a 19-Å cutoff, leaving 350 out of the original 1489 predictions. We selected the top seven of the filtered predictions based on ZRANK score, and also the top three from the total set of predictions based on ZRANK score.

With these 10 predictions, we utilized a slightly different protocol than for Target 26. We refined them using RosettaDock as before (to produce 400 refined models per prediction), and the refined models were selected by ZRANK score (in contrast to Target 26 for which the RosettaDock scores were used to select the models from the refinement). This led to seven acceptable predictions for the second evaluated interface (models ranked no. 1–7), while the remaining three models submitted were those from the unfiltered set. Five of the seven acceptable models had improved structures after refinement, including the best-submitted model. This model, which had a ligand RMSD of 6.42 Å and an interface RMSD of 2.39 Å was improved because of refinement by 3.28 and 0.46 Å

for ligand and interface RMSDs, respectively (Table II; Marc Lensink, personal communication). These RMSD improvements are commensurate with those seen for Target 26, for which the RosettaDock score was used to choose the refined model. This, as well as the overall results for this target, indicate that using ZRANK in the context of scoring sets of RosettaDock refined predictions can lead to improved structures.

DISCUSSION

The goal of ZDOCK, as an initial-stage rigid-body docking algorithm, is to find at least one near-native structure within a set of predictions. In CAPRI, this is extended to finding at least one acceptable prediction within the 10 structures submitted for each target. Our performance in Rounds 6–11 demonstrates that this goal was very nearly reached, as we were able to find at least one acceptable in 5 out of 6 targets. Target 20 was the only target for which we did not attain a good prediction and serves as an example for the limitations of protein–protein docking when starting with inaccurate homology models. The challenges of homology docking can also be seen in the differing degrees of accuracy in Targets 24 and 25. When the bound ArfBD of Target 25 was substituted for its homology model in Target 24, we were able to increase our correct contact percentage from 20 to 81% and our interface RMSD was cut by more than half from 3.13 to 1.51 Å. Overcoming the challenges of homology docking will be an area of further research and may require shAring techniques from the similar problem of flexible docking.

Rounds 6–11 of CAPRI also served as another measure to evaluate the evolution of our ZDOCK algorithm. As seen in Table I, improvements in the ZDOCK algorithm coincided with increases in the accuracy and a higher number of good predictions we were able to achieve. ZRANK was also added to our approach in the later rounds and contributed to attaining successful predictions for Targets 26 and 27.2.

Our scoring performance (Table II) indicates that ZRANK can be successfully employed for rescoring many predictions from a variety of sources. This is encouraging, considering that ZRANK was initially developed and tested using predictions from one particular rigid-body docking algorithm (ZDOCK), whereas the predictions in the scoring round come from a variety of sources, and some of these are not necessarily rigid-body docking. It was possible that refined or energy-minimized false-positive structures would appear more favorable than near-native rigid-body predictions, particularly to an energy-based scoring function such as ZRANK. As was seen in particular for Target 27.2, the filtering step is quite useful for removing false positives and complements the ZRANK scoring well.

Our results from the CAPRI scoring rounds for Targets 26 and 27.2 indicate that it is helpful to refine the predictions prior to submission. We were able to successfully employ RosettaDock to generate the refined structural models for these targets. In addition, for Target 27.2, we rescored the refined models using ZRANK and found this to achieve improved structures. This protocol is currently being explored in more detail to determine its effectiveness and optimal usage. We are also considering optimizing the ZRANK scoring function to evaluate and compare refined docking models rather than just rigid-body docking models.

Overall our performance in the Rounds 6–11 of CAPRI demonstrates the progress of our approach in both protein–protein docking and scoring. We look forward to participating in future rounds of CAPRI in order to test the continued development of ZDOCK and ZRANK.

ACKNOWLEDGMENTS

We thank the CAPRI organizers and evaluation team for their tremendous efforts. In particular, we thank Marc Lensink for giving us the scoring evaluation data and answering many of our questions regarding the scoring evaluation process. We also thank our system administrator, Mary Ellen Fitzpatrick, for her continued technical support. We are grateful for the resources of the scientific computing facilities at Boston University.

REFERENCES

1. Pierce B, Tong W, Weng Z. M-ZDOCK: a superior grid-based approach for Cn symmetric multimer docking. Bioinformatics 2005;21:1472–1478.
2. Mendez R, Leplae R, De Maria L, Wodak SJ. Assessment of blind predictions of protein–protein interactions: current status of docking methods. Proteins 2003;52:51–67.
3. Mendez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in Rounds 3–5 shows progress in docking procedures. Proteins 2005;60:150–169.
4. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87.
5. Chen R, Weng Z. A novel shape complementarity scoring function for protein–protein docking. Proteins 2003;51:397–408.
6. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47:281–294.
7. Chen R, Mintseris J, Janin J, Weng Z. A protein–protein docking benchmark. Proteins 2003;52:88–91.
8. Mintseris J, Wiehe K, Pierce B, Anderson RJ, Chen R, Janin J, Weng Z. Protein–protein docking benchmark 2.0: an update. Proteins 2005;60:214–216.
9. Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, Weng Z. Integrating statistical pair potentials into protein complex prediction. Proteins 2007.
10. Pierce B, Weng Z. ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 2007;67:1078–1086.
11. Chen R, Tong W, Mintseris J, Li L, Weng Z. ZDOCK predictions for the CAPRI challenge. Proteins 2003;52:68–73.
12. Wiehe K, Pierce B, Mintseris J, Tong WW, Anderson R, Chen R, Weng Z. ZDOCK and RDOCK performance in CAPRI rounds 3, 4, and 5. Proteins 2005;60:207–213.
13. Mintseris J, Weng Z. Optimizing protein representations with information theory. Genome Inform 2004;15:160–169.

 DOI 10.1002/prot

14. Zhang C, Vasmatzis G, Cornette JL, DeLisi C. Determination of atomic desolvation energies from the structures of crystallized proteins. J Mol Biol 1997;267:707–726.
15. Graille M, Heurgue-Hamard V, Champ S, Mora L, Scrima N, Ulryck N, van Tilbeurgh H, Buckingham RH. Molecular basis for bacterial class I release factor methylation by PrmC. Mol Cell 2005;20:917–927.
16. Heurgue-Hamard V, Champ S, Engstrom A, Ehrenberg M, Buckingham RH. The *hemK* gene in *Escherichia coli* encodes the *N*(5)-glutamine methyltransferase that modifies peptide release factors. EMBO J 2002;21:769–778.
17. Yang Z, Shipman L, Zhang M, Anton BP, Roberts RJ, Cheng X. Structural characterization and comparative phylogenetic analysis of *Escherichia coli* HemK, a protein (*N*5)-glutamine methyltransferase. J Mol Biol 2004;340:695–706.
18. Schubert HL, Phillips JD, Hill CP. Structures along the catalytic pathway of PrmC/HemK, an *N*5-glutamine AdoMet-dependent methyltransferase. Biochemistry 2003;42:5592–5599.
19. Bates PA, Kelley LA, MacCallum RM, Sternberg MJ. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins 2001;Suppl 5:39–46.
20. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.
21. Accelyrs. Insight II, version 2005. San Diego, CA: Accelyrs; 2005.
22. 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.
23. Hou Z, Bernstein DA, Fox CA, Keck JL. Structural basis of the Sir1-origin recognition complex interaction in transcriptional silencing. Proc Natl Acad Sci USA 2005;102:8489–8494.
24. Gardner KA, Rine J, Fox CA. A region of the Sir1 protein dedicated to recognition of a silencer and required for interaction with the Orc1 protein in saccharomyces cerevisiae. Genetics 1999;151:31–44.
25. Zhang Z, Hayashi MK, Merkel O, Stillman B, Xu RM. Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing. EMBO J 2002;21:4600–4611.
26. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32:W526–W531.
27. Vetter IR, Wittinghofer A. The guanine nucleotide-binding switch in three dimensions. Science 2001;294:1299–1304.
28. Worthylake DK, Rossman KL, Sondek J. Crystal structure of the DH/PH fragment of Dbs without bound GTPase. Structure 2004;12:1078–1086.
29. Rossman KL, Worthylake DK, Snyder JT, Siderovski DP, Campbell SL, Sondek J. A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange. EMBO J 2002;21:1315–1326.
30. Menetrey J, Perderiset M, Cicolari J, Dubois T, Elkhatib N, El Khadali F, Franco M, Chavrier P, Houdusse A. Structural basis for ARF1-mediated recruitment of ARHGAP21 to Golgi membranes. EMBO J 2007;26:1953–1962.
31. Bonsor DA, Grishkovskaya I, Dodson EJ, Kleanthous C. Molecular mimicry enables competitive recruitment by a natively disordered protein. J Am Chem Soc 2007;129:4800–4807.
32. Ray MC, Germon P, Vianney A, Portalier R, Lazzaroni JC. Identification by genetic suppression of *Escherichia coli* TolB residues important for TolB-Pal interaction. J Bacteriol 2000;182:821–824.
33. Clavel T, Germon P, Vianney A, Portalier R, Lazzaroni JC. TolB protein of *Escherichia coli* K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. Mol Microbiol 1998;29:359–367.
34. Walker JR, Avvakumov GV, Xue S, Newman EM, Mackenzie F, Weigelt J, Sundstrom M, Arrowsmith CH, Edwards AM, Bochkarev A, Dhe-Paganon S. A novel and unexpected complex between the SUMO-1-conjugating enzyme UBC9 and the ubiquitin-conjugating enzyme E2-25 kDa, to be published.
35. Pichler A, Knipscheer P, Oberhofer E, van Dijk WJ, Korner R, Olsen JV, Jentsch S, Melchior F, Sixma TK. SUMO modification of the ubiquitin-conjugating enzyme E2-25K. Nat Struct Mol Biol 2005;12:264–269.
36. Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD. Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. Cell 2002;108:345–356.
37. Tong H, Hateboer G, Perrakis A, Bernards R, Sixma TK. Crystal structure of murine/human Ubc9 provides insight into the variability of the ubiquitin-conjugating system. J Biol Chem 1997;272:21381–21387.
38. Reverter D, Lima CD. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. Nature 2005;435:687–692.
39. Yunus AA, Lima CD. Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway. Nat Struct Mol Biol 2006;13:491–499.
40. Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D. Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. J Mol Biol 2003;331:281–299.