PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

How good is automated protein docking?

Dima Kozakov,[1]* Dmitri Beglov,[1] Tanggis Bohnuud,[1,2] Scott E. Mottarella,[1,2] Bing Xia,[1] David R. Hall,[3] and Sandor Vajda[1,4]*

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

[2] Program in Bioinformatics, Boston University, Boston, Massachusetts 02215

[3] Acpharis Inc., Holliston, Massachusetts 01746

[4] Department of Chemistry, Boston University, Boston, Massachusetts 02215

ABSTRACT

The protein docking server ClusPro has been participating in critical assessment of prediction of interactions (CAPRI) since its introduction in 2004. This article evaluates the performance of ClusPro 2.0 for targets 46–58 in Rounds 22–27 of CAPRI. The analysis leads to a number of important observations. First, ClusPro reliably yields acceptable or medium accuracy models for targets of moderate difficulty that have also been successfully predicted by other groups, and fails only for targets that have few acceptable models submitted. Second, the quality of automated docking by ClusPro is very close to that of the best human predictor groups, including our own submissions. This is very important, because servers have to submit results within 48 h and the predictions should be reproducible, whereas human predictors have several weeks and can use any type of information. Third, while we refined the ClusPro results for manual submission by running computationally costly Monte Carlo minimization simulations, we observed significant improvement in accuracy only for two of the six complexes correctly predicted by ClusPro. Fourth, new developments, not seen in previous rounds of CAPRI, are that the top ranked model provided by ClusPro was acceptable or better quality for all these six targets, and that the top ranked model was also the highest quality for five of the six, confirming that ranking models based on cluster size can reliably identify the best near-native conformations.

Proteins 2013; 81:2159–2166.

Key words: protein–protein docking; structure refinement; method development; CAPRI docking experiment; web-based server; user community.

INTRODUCTION

Our group has been working on protein-protein docking since 1996. During this period docking methods have substantially improved. Without even trying to be complete, we can mention the fast Fourier transform (FFT) correlation methods[1–5] based on the work of Katchalski-Katzir *et al.*,[6] the geometric methods by Wolfson and Nussinov,[7] Monte Carlo methods represented by RosettaDock,[8] and the high ambiguity driven biomolecular docking (HADDOCK).[9] Progress is continuously monitored by the critical assessment of prediction of interactions (CAPRI) experiment,[10–13] and the results show the emergence of new approaches.[14,15] However, the participants of CAPRI can use all available information, and hence it is particularly important that several methods are also implemented as automated servers, for example, ClusPro,[16] GRAMM-X,[17] ZDOCK,[3] RosettaDock,[18] HEX,[19] HADDOCK,[20] PatchDock and SymmDock,[21] and SwarmDock.[15] For each CAPRI target, servers should submit models within 48 h, and since the results should be reproducible, the possibilities of using *a priori* biological and structural information are more limited, resulting in a more unbiased measure of method performance. However, it should be noted that even just running different servers requires different amount of *a priori* knowledge. For example, while the FFT-based methods generally perform a global systematic search, Monte Carlo methods need initial conformations for the complex, and for HADDOCK the user should provide a list of interacting residues.[22]

Grant sponsor: NIH; Grant numbers: GM061867 and GM093147; Grant sponsor: National Science Foundation; Grant number: DBI1147082; Grant sponsor: Russian Ministry of Education and Science; Grant number: 14.A18.21.1973.
*Correspondence to: Sandor Vajda, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215. E-mail: vajda@bu.edu or Dima Kozakov, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215. E-mail: midas@bu.edu
Received 9 July 2013; Revised 17 August 2013; Accepted 21 August 2013
Published online 28 August 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24403

The protein docking server ClusPro has been participating in CAPRI since its introduction in 2004.[23] The server performs three computational steps as follows: (1) rigid-body docking using the FFT correlation approach; (2) root mean square deviation (RMSD)-based clustering of the structures generated to find the largest clusters that will represent the most likely models of the complex; and (3) refinement of selected structures. The first version of the ClusPro server used the docking programs DOT[24] and ZDOCK,[3] and employed an empirical energy function to select 2000 conformations for clustering. In 2006 we introduced PIPER, an FFT-based docking program that uses a scoring function including a pairwise potential,[4] and implemented it in the new server ClusPro 2.0, which clusters the top 1000 structures without any filtering.[25] Since ClusPro 2.0 was not properly tested when working on the targets in Rounds 13–19 of CAPRI, we have used both versions of the server. Thus, this article describes the first CAPRI submissions obtained solely by version 2.0.

ClusPro 2.0 is heavily used. By June 2013 we registered over 7000 unique user IPs, and the server completed over 46,000 docking jobs, currently about 1800 per month. Models built by ClusPro have been reported in over 200 publications. In many applications models generated by the server were validated by a variety of experimental techniques, including site-directed mutagenesis, cross-linking, and radiolytic protein footprinting with mass spectrometry. In view of this heavy usage and the availability of the new CAPRI results, it is timely to evaluate the performance of the server, exploring its strengths and weaknesses. While we focus on server performance, we also discuss our manual submissions that were obtained by further refinement of the ClusPro results.

METHODS

Docking using PIPER

PIPER is an FFT-based docking program that uses a pairwise interaction potential as part of its scoring function $E = E_{attr} + w_1E_{rep} + w_2E_{elec} + w_3E_{pair}$, where E_{attr} and E_{rep} denote the attractive and repulsive contributions to the van der Waals interaction energy E_{vdw}, E_{elec} is an electrostatic energy term, and E_{pair} represents the desolvation contributions.[4] E_{pair} has been parameterized on a set of complexes that included a substantial number of enzyme–inhibitor pairs and multisubunit proteins, and hence the resulting potential assumes good shape and electrostatic complementarity. The coefficients w_1, w_2, and w_3 specify the weights of the corresponding terms, and are optimally selected for different types of docking problems (see later). To evaluate the energy function E by FFT, it must be written in the form of correlation functions. The terms E_{vdw} and E_{elec} satisfy this condition, and E_{pair} can be expressed as a sum of a few correlation functions using the eigenvalue–eigenvector decomposition of the matrix of interaction energy coefficients.[4]

Unless specified otherwise in advanced options, the server generates four types of models using the scoring schemes called (1) balanced, (2) electrostatic favored, (3) hydrophobic favored, and (4) van der Waals + electrostatics. The balanced option works generally well for enzyme–inhibitor complexes, whereas Options (2) and (3) are suggested for complexes where the association is primarily driven by electrostatic and hydrophobic interactions, respectively. The fourth option, van der Waals + electrostatics, means that $w_3 = 0$, that is, the pairwise potential E_{pair} is not used. The need for this option occurs for proteins that are very different from the ones used for the parameterization of E_{pair}. Two specific cases can be selected as advanced options. In the "Antibody Mode," ClusPro uses a recently developed asymmetric potential for docking antibody and antigen pairs.[26] The "Other Mode" targets the so-called "other" type of complexes that primarily occur in signal transduction pathways,[27] and generally have substantially less perfect shape and electrostatic complementarity than the enzyme–inhibitor complexes. Due to the diverse nature implied by the "other" classification, this mode chooses 500 conformations from three diverse sets of weighting coefficients to give 1500 conformations. Our initial research using a diversity of coefficients is indicative that the "other" type of complexes can likely be further classified into subtypes for which a particular coefficient set works well.

While it is difficult to perform automated selection of the best scoring function, users generally have some information on the type of the particular complex considered. Accordingly, for the CAPRI targets we selected the scoring scheme we believed to be most appropriate for the proteins depending on their cellular function. If such information was not available, we selected the scoring function that yielded large clusters of docked structures with relatively low energies. In most cases two or three different scoring schemes provided similar results, facilitating the selection. Once the scoring function was selected for a particular target, we simply submitted the top 10 models generated by ClusPro. This simple rule provided very good models in the current rounds. As a matter of fact, we have found substantial robustness in terms of choosing the scoring function. In the process of writing this report, we rerun the docking calculations for the six targets we have successfully predicted. According to this analysis, using the balanced option would have given similarly good results in all cases for which this option was available, that is, apart from the heparin docking in target T57. We emphasize that beyond scoring function selection we have not used any other information such as *a priori* knowledge of some interface residues. Thus, while ClusPro has the advanced option of rewarding or penalizing the move of specified residues

into the interface, this option was not used for any of the CAPRI targets.

Scoring by clustering

The second step of the algorithm is clustering the top 1000 structures (1500 if the "Other Mode" is used), generated by PIPER, using pairwise RMSD as the distance measure.[16,28] The radius used in clustering is defined in terms of C_α interface RMSD. For each docked conformation we select the residues of the ligand that have any atom within 10 Å of any receptor atom, and calculate the C_α RMSD for these residues from the same residues in all other 999 ligands. Thus, clustering 1000 docked conformations involves computing a 1000 × 1000 matrix of pairwise C_α RMSD values. Based on the number of structures that a ligand has within a (default) cluster radius of 9 Å RMSD, we select the largest cluster and rank its cluster center number 1. Then, the members of this cluster are removed from the matrix, and we select the next largest cluster and rank its center number 2, and so on. After clustering with this hierarchical approach, the ranked complexes are subjected to a straightforward (300 step and fixed backbone) van der Waals minimization using the CHARMM potential[29] to remove potential side chain clashes. Unless requested otherwise by the user, ClusPro outputs the centers of the 10 largest clusters.

The biophysical meaning of clustering is isolating highly populated low energy basins of the energy landscape.[30] It is easy to show that large clusters are more likely to include native structures. The globally sampled conformational space can be considered as a canonical ensemble with the partition function $Z = \Sigma_j \exp(-E_j/\mathrm{RT})$, where E_j is the energy of the jth pose, and we sum over all poses. For the kth cluster the partition function is given by $Z_k = \Sigma_j \exp(E_j/\mathrm{RT})$, where the sum is restricted to poses within the cluster. Based on these values, the probability of the kth cluster is given by $P_k = Z_k/Z$. However, since the low energy structures are selected from a relatively narrow energy range, and the energy values are calculated with considerable error, it is reasonable to assume that these energies do not differ, that is, $E_j = E$ for all j in the low energy clusters. This simplification implies that $P_k = \exp(-E/\mathrm{RT}) \times N_k/Z$, and thus the probability P_k is proportional to N_k, where N_k is the number of structures in the kth cluster.

Selecting highly populated clusters rather than isolated low energy structures accounts for some of the entropic effects,[31] and makes our results less sensitive to the inherent inaccuracies in the structures generated by a rigid-body algorithm.[28]

Refinement of docked structures for manual submission

The only refinement currently used in ClusPro is minimizing the Charmm energy of the structures generated by the docking.[29] While the minimization generally removes potential steric clashes, it does not substantially change the conformation of the complexes, and thus the RMSD of our ClusPro submissions from the native complexes is fully determined by the rigid-body docking and clustering steps.

For manual submission selected clusters are refined using the method we have introduced as "stability analysis."[32] The method is based on the hypothesis that clusters of near-native structures are located in broad energy funnels.[32] This hypothesis is explored by starting short Monte Carlo minimization (MCM) simulations from randomly selected structures of the cluster. Each simulation step includes both repacking of the interface side chains and rotational/translational moves, and we use a higher accuracy scoring function than in the rigid-body step. Convergence for a substantial fraction of MCM trajectories to a region within the cluster indicates a broad funnel, and the point of convergence provides an improved estimate of the native structure. Conversely, diverging trajectories indicate that a substantive free energy funnel does not exist, and hence the cluster is not near native.[32]

RESULTS AND DISCUSSION

Table I shows the summary of results obtained by ClusPro, our manual submissions after refinement, as well as the quality of results by the entire CAPRI community, including all predictor groups and servers. Following the notation used by the evaluators,[11–13] results are shown in the form $x^{***} + y^{**} + z^{*}$, where x^{***}, y^{**}, and z^{*} denote the number of high, medium, and acceptable accuracy models, respectively. Table I also lists the models, ranked from M01 to M10, that were found at each level of accuracy. In addition, for the successfully predicted targets we show the accuracy of the top ranked ClusPro model, M01, which turned out to be acceptable or better for all targets for which the server provided correct predictions, emphasizing the progress we have made in ranking. Since the targets are well described in this volume, we provide only brief description of the difficulties we encountered for specific targets, and focus on the general properties of ClusPro and of our manual refinement in the Conclusions. We note that in our approach these two types of results are closely related, since our manual submissions were obtained by refining the ClusPro predictions, and thus also failed if ClusPro could not provide any model that was close to be acceptable.

Round 22

Round 22 included only target T46, predicting the complex between methyl transferase MTq2 and an activator protein (Trm112).[33] Homology modeling was required for both proteins. The difficulty of the target is shown by the fact that all 13 acceptable solutions

Table I
Submissions by ClusPro, by the Vajda/Kozakov Group Following Refinement, and by the Entire CAPRI Community

Target	Name	Type	Predictions by Cluspro/models[a]	Top-ranked ClusPro prediction[b]	Group predictions/models	Predictions by community
T46	Methyl transferase Mtq2/Trm112	Homology/homology	0	0	0	13*
T47	Colicin-E2/immunity protein 2 (IM2)	Homologous to Colicin-E9/IM9 known complex	2**/(M01 and M04)	M01**	10***/(M01–M10)	108*** + 95** + 13*
T48	T4moC/T4moH monooxygenase complex	Unbound/unbound	2**/(M06 and M07) + 5*/(M01 and M03 and M04 and M05 and M10)	M01*	2**/(M01 and M02) + 5*/(M03 and M05 and M06 and M08 and M10)	15** + 56*
T49	T4moC/T4moH monooxygenase complex	Unbound/unbound	3*/(M01 and M02 and M04)	M01*	5*/(M01 and M03 and M04 and M07 and M08)	2** + 46*
T50	HB36.3 designed protein/flu hemagglutinin	Homology/unbound	2**/(M01 and M03)	M01**	2**(M03 and M07) + 3*(M01 and M05 and M06)	17** + 35*
T51	Constructing a Xylanase protein of six domains	Part unbound, part homology	0	0	0	4*
T53	Designed Rep4/Rep2 a-repeat complex	Homology/unbound	1**/M01 + 1*/M03	M01**	1***/M02 + 1**/M03	1*** + 11** + 31*
T54	Designed neocarzinostatin/Rep16 a-repeat complex	Homology/unbound	0	0	0	6*
T57	BT4661/heparin complex	Unbound/homology	2*/(M01 and M02)	M01*	2**/(M01 and M04) + 1*/M05	5** + 26*
T58	PliG/SalG lysozyme complex with SAXS data	Unbound/unbound	0	0	0	15** + 18*

[a]As defined by CAPRI evaluators, ***, **, and * denote high, medium, and acceptable accuracy submissions, respectively.
[b]Accuracy of top-ranked ClusPro model (M01).

provided by the community came from two groups (Table I). As we will discuss, docking homology models is generally very difficult for rigid-body methods, particularly if there are backbone differences between target and template proteins, which was the case here. Although we attempted to build and dock models, none of the resulting complexes was acceptable accuracy.

Round 23

Round 23 included targets T47, T48, and T49. T47 was the prediction of the cognate complex formed by colicin E2 DNAse and the immunity protein 2 (Im2).[34] Since the resulting complex is highly homologous to the colicin E9–Im9 and to the E9–Im2 complexes with structures available [Protein Data Bank (PDB) entries 1EMV and 2WPT, respectively], human predictor groups could easily obtain medium or high accuracy models by template-based modeling, that is, by mutating a few side chains in the template complexes. However, ClusPro cannot use this type of information, and we docked the nuclear magnetic resonance structure of IM2 (PDB entry 2NO8) to the structure of colicin E2, resulting in two medium accuracy models, the better of which is shown in Figure 1(A). We admit that after refinement we used the information on the homologous complexes for model selection, which enabled us to submit 10 high accuracy models, similarly to many other groups. Target T47 also required predicting the positions of interface water molecules. The results of this task will be reported in a separate article, and hence are not discussed here. Nevertheless, we note that we added water as a probe in our solvent mapping program,[35] and obtained six good and two fair quality predictions.

Target T48 was to predict the complex formed by a ferredoxin molecule (T4moC) and a monooxygenase enzyme (T4moH). T4moH consists of two trimers forming a heterohexamer. According to the description of this target, participants were invited to submit only solutions that involved the first trimer. This was somewhat misleading, since the ferredoxin was in contact with subunits of both trimers, resulting in two ferredoxin binding sites on the T4moH hexamer. In view of this problem, the evaluators accepted as correct solutions the complexes of ferredoxin with the hexamer or any of the two trimers, and Table I shows the number of all of acceptable or better predictions by the CAPRI participants, submitted as either trimers or hexamers. We (correctly) used ClusPro to dock the unbound structure of ferredoxin to the T4moH hexamer, but submitted only models involving

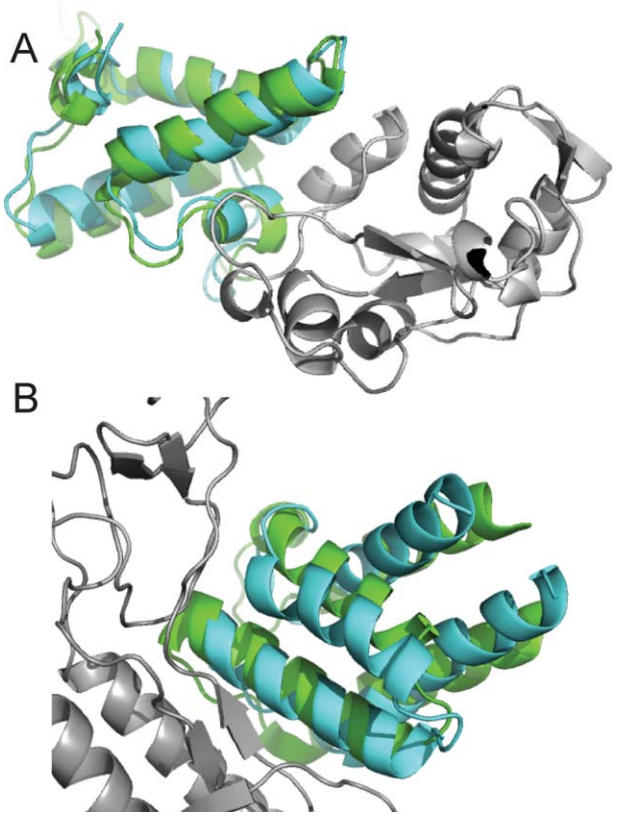

Figure 1

(**A**) Best ClusPro prediction for target T47, the complex formed by colicin E2 and Im2. The structure of colicin E2 is shown in gray, and the X-ray (PDB entry 3U43) and predicted structures of Im2 are shown in green and cyan, respectively, after superimposing the colicin structures. The interface backbone RMSD between the two IM2 structures is 1.45 Å. (**B**) Best ClusPro prediction for target T50, the complex formed by the designed protein HB36.3 and influenza hemagglutinin. The structure of hemagglutinin is shown in gray, and the X-ray (PDB entry 3R2X) and predicted structures of HB36.3 are shown in green and cyan, respectively, after superimposing the hemagglutinin structures. The interface backbone RMSD between the two HB36.3 structures is 1.94 Å.

trimers, resulting in two medium accuracy submissions, representing binding to the two sites. Although the refinement by stability analysis did not improve the quality of predictions, it helped to submit the two medium accuracy solutions as models M01 and M02 (see Table I). Target T49 was the same as T48 but with the monooxygenase in a different conformation, lacking the T4moD effector protein in the T4moH hexamer, again with two binding sites for ferredoxin. ClusPro provided three acceptable models, and although refinement did not substantially improve accuracy, it increased the number of acceptable models to five.

Round 24

Round 24 consisted of target T50, predicting the binding of the protein HB36.3, which was designed by the Baker group to replace an antibody binding to the stem region of influenza hemagglutinin,[36] and Target T51, involving a xylanase enzyme containing six modules. In T51 the design of HB36.3 was based on an all-helical protein of known structure (PDB entry 1U84) and involved introducing mutations at 15 positions. The backbone coordinates were also provided for HB36.3, attesting that the backbone conformation remains almost completely invariant. For the hemagglutinin, it was suggested to use chains A and B in an antibody-bound cocrystal structure (PDB entry 3GBN). We built the mutated side chains of HB36.3, and docked the resulting structure to the hemagglutinin. Among the top 10 models, ClusPro returned two medium quality predictions [Fig. 1(B)]. Refinement slightly improved accuracy, and brought three more models into acceptable range. For Target T51 the docking task was to predict three separate interactions among the six modules of the xylanase molecule, where one module had to built based on homology to a known structure. ClusPro did not yield any acceptable model, and the target was difficult for the entire CAPRI community, although four acceptable solutions were submitted (see Table I).

Round 26

Round 26 consisted of protein docking targets T53 and T54, both requiring the prediction of complexes formed by designed repeat proteins. T53 is a 1:1 complex between the artificial alpha-repeat Rep4 (PDB entry 3LTJ) and Rep2, a shorter alpha-repeat homologous to Rep4. Although constructing the structure of Rep2 required homology modeling, it was very easy due to high level of homology and the conserved backbone. Docking the model to 3LTJ by ClusPro yielded one medium quality and one acceptable submission (models M01 and M03, respectively). The two models were refined to high and medium accuracy, respectively, for manual submission. Although the problem was of moderate difficulty and several group obtained medium accuracy models, only our group had a high accuracy submission.

T54 is a 1:1 complex between the engineered neocazinostatin (2CBO) and Rep16, a short alpha-repeat selected to bind neocazinostatin. In spite of high level of homology between Rep16 and 3LTJ, ClusPro failed to yield any acceptable model. The fact that there were only six acceptable predictions from the entire community indicates that this was a difficult target. Although the structure of the complex is not yet published, we are convinced that the origin of difficulty is in the homology modeling. In fact, the Shen group also relied on ClusPro 2.0 to generate initial models for this target, and obtained an acceptable prediction using a homology model that was obviously better than ours (Dr. Yang Shen, personal communication).

Round 26 also included targets T55 and T56, which required predicting the effect of mutations on the interaction between two proteins rather than docking.[37] Results for these targets are available on the CAPRI website (http://www.ebi.ac.uk/msd-srv/capri/round26/round26.html), but the entire experiment will be published in a separate article and hence is not discussed here.

Round 27

T57 requires docking a six-sugar residue heparin to the unbound structure of fragment 423–700 of protein BT4661. A developmental version of ClusPro allows for docking molecules with heteroatoms, and we have tested cross-docking of heparin to several heparin-binding proteins in the PDB. We had no desolvation parameters for sugars, and hence used the van der Waals + electrostatics option of the server, but this should be a good choice, since heparin is negative charged and it usually binds to positively charged regions of the protein. In view of good test results we have decided to go forward with target T57, but used real heparin structures from the PDB instead of the model provided. ClusPro yielded two large clusters of structures, indicating two different binding modes, and our 10 submissions included two acceptable predictions (Table I). The structures were refined to medium accuracy by minimizing the Charmm energy of the complexes,[29] and we also obtained a new acceptable model. Based on this positive experience we have decided to add heparin docking as a new option to ClusPro, and are close to finishing validation of the method on structures of protein–heparin complexes in the PDB.

The second target (T58) of Round 27 required docking the unbound structure of an inhibitor (PDB entry 4G9S) to the unbound structure of a goose-type lysozyme (PDB entry 3MGW).[38] Small-angle X-ray scattering (SAXS) data were also provided to help the prediction. While ClusPro generally performs well when docking unbound structures, it failed to produce any acceptable prediction due to substantial backbone conformational change in the lysozyme structure upon inhibitor binding. The server did not have the ability to account for SAXS data, and although we now added this option in the developmental version, it was too late for improving the predictions for T58.

CONCLUSIONS

With the already introduced notation, our performance in Rounds 22–27 of CAPRI is described as 2*** + 3** + 1*, placing us as the fourth best predictor group. The six successful predictions were obtained for targets T47, T48, T49, T50, T53, and T57. With a performance of 4** + 2* and successful predictions for the same six targets, ClusPro was only the seventh best overall predictor, but the best in the automated server category. Although the number of targets is still too small for any significant conclusion, we believe that our results provide some information on the current state of automated protein docking, at least concerning methods that use rigid-body approximation in the first step. Our main observations are as follows.

1. ClusPro reliably yields correct predictions for the relatively "easy" targets with at most moderate conformational changes in the backbone. In addition to unbound proteins of known structure, such "easy" targets may include designed proteins obtained by mutating a few residues. Targets T50 and T53 were in this category, and ClusPro provided good results. The CAPRI community submitted many good predictions for targets T47, T48, T49, T50, T53, and T57, that is, exactly for the ones ClusPro also predicted well, confirming that these targets are relatively easy. Based on this logic we should have obtained an acceptable or better model for an additional target, T58, but the change in the backbone conformation of a lysozyme loop was too large for ClusPro, although other groups using rigid-body methods such as GRAMM were able to produce an acceptable model, but only for manual submission. The three other targets, T46, T51, and T54 that were difficult for ClusPro were also difficult for the entire CAPRI community, resulting in very few acceptable submissions. As will be further discussed, all these targets required homology modeling.
2. The quality of automated docking by ClusPro is very close to that of the best human predictor groups, including of our own. We consider this very important, because servers have to submit results within 48 h and the predictions should be reproducible by the server, whereas human predictors have several weeks and can use any type of information. In Rounds 22–27 three predictor groups (Bonvin, Bates, and Vakser) did extremely well, and submitted acceptable or better predictions for more than six targets. These three were followed by six groups that had good predictions for six targets: Vajda (2*** + 3** + 1*), Fernandez-Recio (1*** + 3** + 2*), Shen (1*** + 3** + 2*), Zou (1*** + 2** + 3*), Zacharias (1*** + 5*), and ClusPro (4** + 2*). The only difference between ClusPro and the other five groups is due to the ability of the human predictors obtaining high accuracy predictions for T47 by template-based modeling. Since ClusPro does not have this option, it had to use direct docking, and produced only a medium accuracy model. We emphasize that in the earlier rounds of CAPRI server predictions were substantially inferior to those of the human predictors—this is definitely not the case for ClusPro 2.0 in Rounds 22–27. However, ClusPro seems to be an exception, as for most other groups the manual submissions are generally much better than the submissions from their servers.

3. As mentioned, our manual submissions were obtained by refining the ClusPro results using "stability analysis," requiring a large number of relatively short MCM runs. In spite of substantial computational efforts, the improvements due to the refinement are moderate. Apart from T47, where obtaining high accuracy predictions were trivial, the refinement improved predictions only for two targets, T53 and T57. However, it appears that refining predictions to high accuracy was generally very difficult for all targets (again, not considering T47). In fact, the only high accuracy model submitted by any group for any target in Rounds 22–27 was our manual submission for target T53.
4. Fourth, a new development, not seen in previous rounds of CAPRI, is that the top ranked model M01 provided by ClusPro was acceptable or better quality for all the six targets that Cluspro was able to predict. M01 was also the highest quality model for five of these six targets. The only exception was T48, where models M06 and M07 were medium quality, while model M01 was only acceptable. Due to the very small number of targets the generality of this observation is not at all clear, but suggests that ranking predictions based on cluster size can reliably identify the highest accuracy models.[39]
5. The most difficult targets, T46, T51, and T54 required the construction of homology models based on templates with moderate sequence identity. The poor results for these targets, both by ClusPro and by the entire CAPRI community, show that the quality of homology models plays a critical role in docking. For example, while ClusPro did not produce any prediction for target T54 with the models we constructed, an acceptable submission was found by the Shen group, who also relied on the server for the initial docking, but used a better homology model. Thus, there is need for methods that are specifically designed for docking homology models, for example, by further reducing the sensitivity of the scoring function to steric clashes involving mutated side chains and predicted loop regions.

Apart from the difficulty of docking homology models, we believe that the predictions provided by ClusPro 2.0 in Rounds 22–27 of CAPRI show clear progress. Indeed, the server's performance was comparable to that of the best human predictors, and the top ranked model M01 was either medium or acceptable accuracy for all targets that were successfully predicted, which were also the targets that had good predictions submitted by several other groups.

REFERENCES

1. Gabb HA, Jackson RM, Sternberg MJE. Modelling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 1997;272:106–120.
2. Vakser IA. Low-resolution docking: prediction of complexes for underdetermined structures. Biopolymers 1996;39:455–464.
3. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87.
4. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 2006;65: 392–406.
5. Ritchie DW, Kemp GJL. Protein docking using spherical polar Fourier correlations. Proteins 2000;39:178–194.
6. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 1992;89:2195–2199.
7. Nussinov R, Wolfson HJ. Efficient computational algorithms for docking and for generating and matching a library of functional epitopes I. Rigid and flexible hinge-bending docking algorithms. Comb Chem High Throughput Screen 1999;2:249–259.
8. 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.
9. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. J Am Chem Soc 2003;125:1731–1737.
10. Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJ, Vajda S, Vakser I, Wodak SJ. CAPRI: a critical assessment of predicted interactions. Proteins 2003;52:2–9.
11. 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.
12. Lensink MF, Mendez R, Wodak SJ. Docking and scoring protein complexes: CAPRI 3rd Edition. Proteins 2007;69:704–718.
13. Lensink MF, Wodak SJ. Docking and scoring protein interactions: CAPRI 2009. Proteins 2010;78:3073–3084.
14. Zacharias M. Accounting for conformational changes during protein–protein docking. Curr Opin Struct Biol 2010;20:180–186.
15. Moal IH, Bates PA. SwarmDock and the use of normal modes in protein–protein docking. Int J Mol Sci 2010;11:3623–3648.
16. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 2004;20:45–50.
17. Tovchigrechko A, Vakser IA. GRAMM-X public web server for protein–protein docking. Nucleic Acids Res 2006;34:W310–314.
18. Lyskov S, Gray JJ. The RosettaDock server for local protein–protein docking. Nucleic Acids Res 2008;36:W233–238.
19. Macindoe G, Mavridis L, Venkatraman V, Devignes MD, Ritchie DW. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res 2010;38:W445–449.
20. de Vries SJ, van Dijk M, Bonvin AM. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 2010;5: 883–897.
21. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 2005;33:W363–367.
22. Vajda S, Kozakov D. Convergence and combination of methods in protein-protein docking. Curr Opin Struct Biol 2009;19:164–170.
23. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: a fully automated algorithm for protein–protein docking. Nucleic Acids Res 2004;32:W96–99.
24. 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.
25. Comeau SR, Kozakov D, Brenke R, Shen Y, Beglov D, Vajda S. ClusPro: performance in CAPRI rounds 6–11 and the new server. Proteins 2007;69:781–785.
26. Brenke R, Hall DR, Chuang GY, Comeau SR, Bohnuud T, Beglov D, Schueler-Furman O, Vajda S, Kozakov D. Application of asymmetric statistical potentials to antibody–protein docking. Bioinformatics 2012;28:2608–2614.

27. Chen R, Mintseris J, Janin J, Weng Z. A protein–protein docking benchmark. Proteins 2003;52:88–91.
28. Kozakov D, Clodfelter KH, Vajda S, Camacho CJ. Optimal clustering for detecting near-native conformations in protein docking. Biophys J 2005;89:867–875.
29. 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.
30. Lorenzen S, Zhang Y. Identification of near-native structures by clustering protein docking conformations. Proteins 2007;68:187–194.
31. Shortle D, Simons KT, Baker D. Clustering of low-energy conformations near the native structures of small proteins. Proc Natl Acad Sci USA 1998;95:11158–11162.
32. Kozakov D, Schueler-Furman O, Vajda S. Discrimination of near-native structures in protein–protein docking by testing the stability of local minima. Proteins 2008;72:993–1004.
33. Liger D, Mora L, Lazar N, Figaro S, Henri J, Scrima N, Buckingham RH, van Tilbeurgh H, Heurgue-Hamard V, Graille M. Mechanism of activation of methyltransferases involved in translation by the Trm112 'hub' protein. Nucleic Acids Res 2011;39:6249–6259.
34. Wojdyla JA, Fleishman SJ, Baker D, Kleanthous C. Structure of the ultra-high-affinity colicin E2 DNase–Im2 complex. J Mol Biol 2012;417:79–94.
35. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, Mattos C, Vajda S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. Bioinformatics 2009;25:621–627.
36. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch EM, Wilson IA, Baker D. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science 2011;332:816–821.
37. Whitehead TA, Chevalier A, Song Y, Dreyfus C, Fleishman SJ, De Mattos C, Myers CA, Kamisetty H, Blair P, Wilson IA, Baker D. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. Nat Biotechnol 2012;30:543–548.
38. Leysen S, Vanderkelen L, Weeks SD, Michiels CW, Strelkov SV. Structural basis of bacterial defense against g-type lysozyme-based innate immunity. Cell Mol Life Sci 2013;70:1113–1122.
39. Vajda S, Hall DR, Kozakov D. Sampling and scoring: a marriage made in heaven. Proteins 2013 Jun 15. doi: 10.1002/prot.24343. [Epub ahead of print]