WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

# RosettaDock in CAPRI rounds 6–12

Chu Wang,[1] Ora Schueler-Furman,[1,2] Ingemar Andre,[1] Nir London,[2] Sarel J. Fleishman,[1] Philip Bradley,[1] Bin Qian,[1] and David Baker[1,3]*

[1] Department of Biochemistry, University of Washington, Seattle, Washington

[2] Department of Molecular Genetics and Biotechnology, Hebrew University, Hadassah Medical School, Jerusalem, Israel

[3] Howard Hughes Medical Institute, University of Washington, Seattle, Washington

*ABSTRACT*

***A challenge in protein–protein docking is to account for the conformational changes in the monomers that occur upon binding. The RosettaDock method, which incorporates sidechain flexibility but keeps the backbone fixed, was found in previous CAPRI rounds (4 and 5) to generate docking models with atomic accuracy, provided that conformational changes were mainly restricted to protein sidechains. In the recent rounds of CAPRI (6–12), large backbone conformational changes occur upon binding for several target complexes. To address these challenges, we explicitly introduced backbone flexibility in our modeling procedures by combining rigid-body docking with protein structure prediction techniques such as modeling variable loops and building homology models. Encouragingly, using this approach we were able to correctly predict a significant backbone conformational change of an interface loop for Target 20 (12 Å rmsd between those in the unbound monomer and complex structures), but accounting for backbone flexibility in protein–protein docking is still very challenging because of the significantly larger conformational space, which must be surveyed. Motivated by these CAPRI challenges, we have made progress in reformulating RosettaDock using a "fold-tree" representation, which provides a general framework for treating a wide variety of flexible-backbone docking problems.***

Proteins 2007; 69:758–763.


***Key words: CAPRI; flexible-backbone docking; Monte Carlo minimization; loop modeling; homology docking.***

## INTRODUCTION

Protein–protein interactions play important roles in cellular processes and elucidating structures of these complexes will undoubtedly contribute to our understanding of their function. However, only a small fraction of the structures solved experimentally are of protein–protein complexes[1] and it is therefore important to develop computational docking methods that can reliably predict the structure of a complex starting from the structures of the uncomplexed monomers. Many docking methods have been developed over the years (see reviews[2–4]) and a community-wide double-blind experiment, CAPRI, has been organized since 2001 to evaluate the performance of different docking methods and provide a comprehensive review of the status of this field.[5–9]

RosettaDock,[10,11] a part of the Rosetta software suite, uses the Monte Carlo minimization[12] to sample the rigid-body degrees of freedom in 3D Cartesian space while treating sidechain flexibility explicitly, and the search is guided by a physically based free energy function dominated by an all-atom Lennard-Jones potential, solvation energy,[13] and an explicit hydrogen bonding potential.[14] The power of the method has been demonstrated in the previous CAPRI rounds (4 and 5) in which high-resolution models of atomic accuracy were created unambiguously when conformational changes were mainly restricted to protein sidechains.[15] However, the approach fails when there is significant backbone conformational change upon binding because the rigid-backbone approximation is no longer adequate. In this study, we present the results of our predictions in CAPRI rounds 6–12.

## MATERIALS AND METHODS

### Prediction protocol

A literature search was conducted prior to computation to retrieve relevant biological information, which could be potentially used as constraints. Special attention was paid to possible sources of information on the internal backbone flexibility within each partner. When possible, several solved

The authors state no conflict of interest.
Grant sponsors: Human Frontier Science Program, Knut and Alice Wallenberg Foundation, American Leukemia and Lymphoma Society, National Institutes of Health.
Chu Wang and Ora Schueler-Furman contributed equally to this work.
*Correspondence to: David Baker, Department of Biochemistry, University of Washington, Seattle, Box 357350, WA 98195. E-mail: dabaker@u.washington.edu
Received 31 May 2007; Accepted 19 June 2007
Published online 1 August 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.21684

structures of the protein partners (and their homologues) were compared. For each target, the standard fixed-backbone RosettaDock protocol was used to perform global docking followed by local refinement from each of the low-energy cluster centers as described previously.[15] Backbone flexibility was introduced in a target-dependent manner as described in the results section. Final model selection was based on the total energy of the entire system and if present, biological information was taken into account as well. When backbone flexibility was modeled explicitly, the interface energy (the difference between the total energy of the complex and the total energy when the two partners are separated) was also used to reduce noise introduced by optimizing regions far away from the interface.

### Loop modeling

A method for loop modeling was implemented to model backbone variable regions during docking if necessary. The details of the method have been described elsewhere (Wang C, Bradley P, Baker D. Protein–protein docking with backbone flexibility. Submitted for review). Briefly, the method incorporates the cyclic coordinate descent (CCD) algorithm[16] for loop closure into the Rosetta loop modeling method developed previously by Rohl *et al.*[17] The method starts from an initial loop conformation generated by randomly inserting local peptide segments ("fragments") into the loop region and performs a series of Monte Carlo CCD minimizations, first in the low-resolution stage and then in the high-resolution stage. Within each cycle of Monte Carlo CCD minimization, the backbone conformation of the loop region is perturbed first, and then the loop is closed by the CCD algorithm. After the gradient-based energy minimization is carried out with respect to the loop backbone and all sidechain torsional degrees of freedom, the resulting conformation is subjected to acceptance or rejection according to the Metropolis criterion.

### Fold-tree based docking

During CAPRI rounds 6–12, we reformulated the traditional RosettaDock using a "fold-tree" representation[18] to provide a general framework to incorporate backbone flexibility into protein–protein docking (Wang C, Bradley P, Baker D. Protein–protein docking with backbone flexibility. Submitted for review). The integration of backbone/sidechain torsional and rigid-body degrees of freedom allows optimization of all relevant degrees of freedom simultaneously. In the prediction of the last few targets, we introduced backbone minimization during docking using the fold-tree based approach in addition to the standard fixed-backbone docking.

**Table I**
*Summary of Performance in CAPRI Rounds 6–12*

| Target | Receptor-ligand | Type[a] | I_rmsd (Å)[b] | L_rmsd (Å)[b] | F_nat[b] | #h/m/a/i[c] |
|---|---|---|---|---|---|---|
| T20 | HemK-RF1 | U-H | **2.34** (4) | **7.74** (4) | **0.36** (8) | 0/0/3/7 |
| T21 | Orc1-Sir1 | U-U | 6.72 (4) | 32.151 (4) | 0.12 (4) | 0/0/0/0 |
| T24 | Arf1-ArfBD | U-H | 6.285 (2) | 18.86 (2) | 0.23 (2) | 0/0/0/0 |
| T25 | Arf1-ArfBD | U-B | 11.547 (8) | 30.36 (9) | 0.06 (4) | 0/0/0/0 |
| T26 | TolB-Pal | U-U | **1.06** (1) | **2.43** (1) | **0.54** (1) | 0/6/1/3 |
| T27 | Hip2-Ubc9 | U-U | 2.69 (9) | 5.98 (9) | 0.42 (9) | 0/0/8/2 |
| T28 | NEDD4L dimer | H-H | 10.01 (9) | 17.26 (9) | 0.00 | 0/0/0/0 |

Values are for the best of our 10 submitted models and those which are also best over all CAPRI submissions are shown in bold. The rank of the model is shown in parentheses.
[a]Starting structures: B, bound; U, unbound; H, homolog unbound.
[b]Measures according to CAPRI assessors (http://capri.ebi.ac.uk). I_rmsd: interface backbone rmsd; L_rmsd: ligand backbone rmsd; F_nat: Fraction of native contacts.
[c]Counts of models according to accuracy: h, high; m, medium; a, acceptable; i, incorrect.

### Rosetta@Home and distributed computing

A large amount of docking computation for the predictions of target 27 and 28 was carried out on Rosetta@Home (www.boinc.bakerlab.org), a distributed computing project running Rosetta software on personal computers of volunteers from all over the world using the Berkeley Open Infrastructure for Network Computing (BOINC) technology (www.boinc.berkeley.edu). The substantial computing resources allowed us to rapidly test multiple approaches for incorporating backbone flexibility, which would not have been possible with in-house computing resources.

## RESULTS

The variability of the targets led us to develop a target-tailored protocol for most of the targets in rounds 6–12. These protocols are described here in detail for each target separately and the results are summarized in Table I.

### Target 20: HemK and RF1

The docking task was to predict the complex structure of *E. coli* HemK and release factor 1 (RF1). HemK methylates a glutamine in the "Q-loop" of release factors *in vivo*.[19] The *E. coli* HemK structure[20] (pdb code: 1t43) is a binary complex with the methyl-donor product S-adenosyl-L-homocysteine. Another HemK orthologue structure from *T. Maritima*[21] (pdb code: 1nv8), contains in addition the methyl-acceptor product N5-methylglutamine (MEQ), which provides a strong constraint about the location of the methyl-acceptor glutamine when the release factor binds to HemK. Superimposing these two structures also reveals one loop region in HemK (from Glu189 to Phe201) with considerable structural variations and hence likely to be flexible. For the other docking

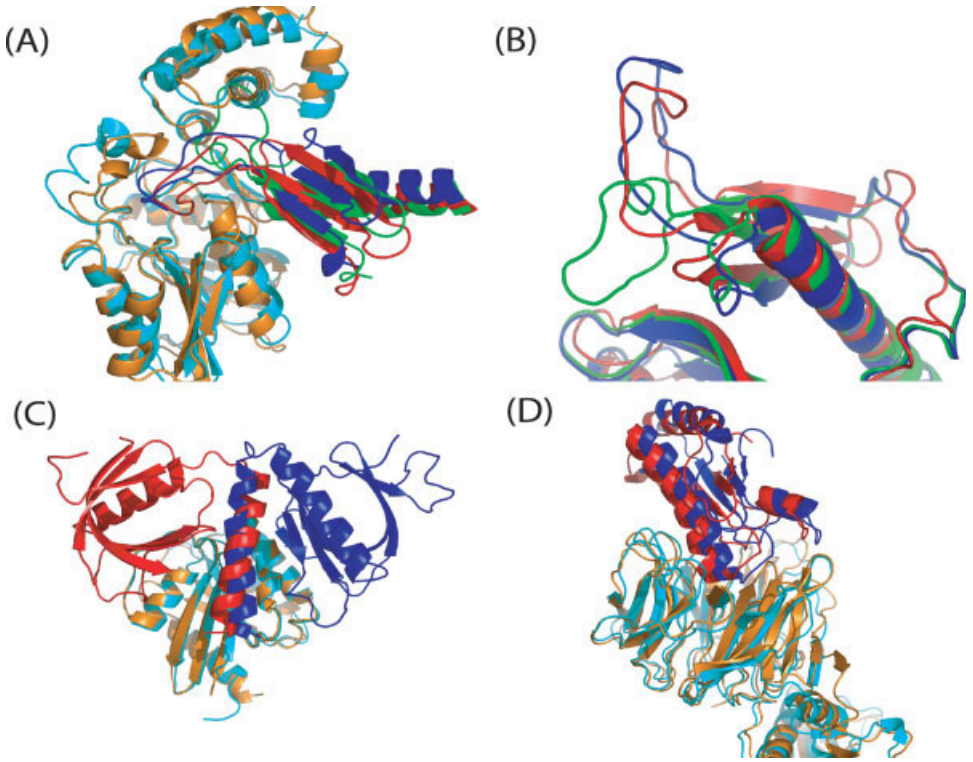


**Figure 1**

*Predictions for selected CAPRI targets. The receptor and ligand (Table I) from the native complex are shown in orange and red, respectively; the unbound ligand structure is shown in green; the receptor and ligand from the predicted model is shown in cyan and blue, respectively. (**A**) Target 20 HemK-RF1: superimposition of the native complex[23] and model 4, according to interface residues. The unbound ligand structure[22] is aligned to demonstrate that severe steric clashes would have prevented the native binding mode from being sampled in fixed-backbone docking. (**B**) Target 20 HemK-RF1: superimposition of the bound,[23] unbound[22] and predicted conformations of the ligand only, highlighting the substantial movement of the loop upon binding. (**C**) Target 24 Arf1-ArfBD: Superimposition of the native complex[24] and model 5 by the receptor. Homology modeling successfully predicted the C-terminal region of the ligand, which was completely missing in the template, to form an α-helix, and the helix was docked into a near-native rigid-body orientation. (**D**) Target 26 TolB-Pal: superimposition of the native complex[25] and model 1 by the receptor. The model has an interface backbone rmsd of 1.06 Å.*

partner, RF1, only the structure of its close homologue (40% sequence identity) release factor 2 (RF2) was available[22] (pdb code: 1gqe). However, the equivalent "Q-loop" region in this structure was completely disordered and was therefore computationally remodeled.

We first excluded both flexible loops in HemK and RF2, and carried out fixed-backbone global docking with the remaining template. Then, we rebuilt the missing loops onto the lowest-energy complex models (see Materials and Methods). The resulting docking models were filtered based on the "methylation" constraint, and the remaining full-length models underwent a second round of rigid-body docking refinement. The subsequently released structure of the "HemK-RF1" complex[23] (pdb code: 2bt3) shows that this protocol was successful: Our best model has an interface backbone Cα rmsd of 2.34 Å with respect to the native complex [Fig. 1(A)], and after superimposing the aligned template regions of RF1, the backbone Cα rmsd for the "Q-loop" in our model to that in the native complex is 4.8 Å, as compared to 11.8 Å for that in the unbound RF2 structure [Fig. 1(B)]. As illustrated in Figure 1(A), the dramatic conformational change of the "Q-loop" makes it impossible for fixed-backbone docking methods to identify (or even sample) the correct binding mode because it would clash badly with HemK.

### Target 21: Orc1 and Sir1

The unbound structures for Orc1[26] (pdb code: 1m4z) and Sir1[27] (pdb code: 1z1a) were provided as the starting structures for docking. We found some experimental information in the literature that supported that one domain within Sir1, the Orc interaction region (OIR), is necessary and sufficient for the Sir1-Orc interaction.[28] We therefore performed a fixed-backbone global docking using Orc1 and the OIR domain from Sir1 only. The evaluation shows that none of our predictions for this target was correct. Post-assessment of the native Orc1-Sir1 complex structure[28] (pdb code: 1zhi) indicates that

 DOI 10.1002/prot

the OIR domain in Sir1 only partially contributes to the Orc1-Sir1 binding interface and without the remaining part of Sir1 the native binding mode could not be selected based on energy criteria.

### Target 24 and 25: Arf1 and ARHGap10

In Target 24, the Arf1 binding domain (ArfBD) of ArhGap10 was modeled using the homologous structure, the PH domain of β-spectrin[29] (pdb code: 1btn). Notably, the 27 C-terminus residues of ArfBD do not align to any region of 1btn. PSIPRED[30] predicts high helical propensity for this region and when it was modeled by the Rosetta *ab initio* folding method,[31] the low-energy models indeed converged on a helical conformation although the orientation of the helix was varied. The resulting "helix-loop-helix" motif at the C-terminus of the ArfBD model is consistent with an analysis on the complex structures of GTPases (Arf and Rab family) with their effector proteins, which reveals that in most cases the GTPase-binding regions of the effectors are composed of two α helices that are aligned along the interswitch β strands between switch 1 and switch 2 of the GTPases.[32] We divided the ArfBD model into two parts: the core structure (modeled based on 1btn) and the C-terminus helix (modeled *ab initio*), and docked them individually onto the structure of Arf1. Low-energy models were screened to match the common interaction motif with the two helices binding close to the switch 1 and switch 2 regions. We then combined these two individual binding interfaces together by superimposing Arf1 and, if possible, connected the two parts of ArfBD by rebuilding the linker loop between them. Our approach did not yield any correct prediction for this target. Post-examination of the native complex structure[24] (pdb code: 2j59) shows that the C-terminus region of ArfBD does fold into an α-helix as we postulated and also binds to the interswitch β strands between switch 1 and switch 2 of Arf1 as seen in the common binding motif.[32] However, because the interface in the native complex structure is largely constituted by that between the C-terminus helix of ArfBD and Arf1, the sub-interface between the core region of ArfBD and Arf1 was not identified correctly when they were docked alone. Consequently, some of the submitted models have the C-terminus helix of ArfBD oriented in a near-native rigid-body position but the remaining core region in a completely wrong orientation [Fig. 1(C)].

Target 25 consisted of the same protein–protein interaction, except that this time the bound ArfBD structure (from 2j59) was given. Since in this case there were no indications of backbone flexibility, we performed a fixed-backbone global docking, followed by a local search starting from the lowest energy cluster centers. The released native complex structure showed that none of our predictions for this target was correct. The reason is probably insufficient sampling: although the native complex has very favorable energy, the minimal rmsd conformation found in the 200 lowest-energy structures in the global search was 10 Å away from the correct orientation, not near enough to yield the correct orientation upon minimization.

### Target 26: TolB and Pal

The unbound structures of TolB[33] (pdb code: 1c5k) and Pal[34] (pdb code: 1oap) were used to perform fixed-backbone global docking. In the lowest-energy model from the largest cluster, Pal binds to the propeller domain of TolB (residue 167–431) and this is consistent with the results from yeast two-hybrid[33] and mutagenesis[35] experiments suggesting that the propeller domain is important for binding of TolB and Pal. The interface on the Pal side found in this model also matches experimental data suggesting that residues 89–130 of Pal are sufficient to bind TolB,[36] and that the Pal mutants T93I, G101D, E102K, and E130K are deficient in formation of the TolB-Pal complex[37]: Both E102 and E130 are involved in salt bridges across the interface, and positions 93 and 101 cannot accommodate larger side chains due to tight packing with TolB. The model after the local refinement from this conformation was ranked first in our submitted predictions. Comparison with the native complex of TolB and Pal[25] (pdb code: 2hqs) reveals that, the model has, among all submissions, the highest fraction of native contacts (54.3%), the best ligand Cα rmsd (2.425 Å) and the best interface backbone rmsd (1.064 Å). In addition to the fixed-backbone local docking refinement, we also applied the "fold-tree" approach to perform flexible-backbone local docking refinement in which backbones of the short loops that connect the "propeller blades" in TolB are optimized together with the rigid-body degrees of freedom, but this does not produce better models (according to the accuracy measures in the CAPRI evaluation report).

### Target 27: Hip2 and Ubc9

The unbound structures of Hip2[38] (pdb code: 1yla) and Ubc9[39] (pdb code: 1a3s) were used to perform fixed-backbone global docking. The lowest-energy model obtained matched the expected biological constraint that Lys14 of Hip2 is close to the catalytic Cys93 of Ubc9 to allow sumoylation of Hip2 by Ubc9,[40] and this model was submitted as the top-ranking prediction after local refinement. However, the experimentally solved complex structure[41] reveals two novel and unexpected binding interfaces, which do not satisfy the sumoylation constraint. Model 9, which we selected due to its favorable electrostatic interaction across the interface, was measured to be close to the second binding mode with inter-

face backbone rmsd of 2.69 Å and 41.5% of native residue–residue contacts predicted correctly [Fig. 1(D)].

### Target 28: Homodimer of the catalytic domain of NEDD4L

This task requires the symmetrical docking of a homology model with potential internal hinge flexibility. The monomer structure of NEDD4L was built from its homologue, HECT domain of SMURF2[42] (pdb code: 1zvd) and a symmetry docking method (Andre I, Bradley P, Wang C, Baker D. Prediction of the Structure of Symmetrical Protein Assemblies. Submitted for review) was used to sample the rigid-body space with and without hinge flexibility to find the dimer interface. An alternative approach was tested where first the N-terminus domain of the first monomer was docked to the C-terminus domain of the second monomer, and the structure was completed by creating a full homodimer based on symmetry criteria, and evaluating whether the N- and C-terminus domains of each monomer could be connected. Unfortunately, none of these approaches was successful based on the evaluation report and no conclusion can be drawn as the native complex structure has not been made available to public yet.

## DISCUSSION

We predicted models with at least acceptable accuracy for three out of seven targets in CAPRI rounds 6–12. This success rate (as well as the prediction accuracy) is not as impressive as that achieved by RosettaDock in CAPRI rounds 4 and 5[15] (six out of eight). Among these seven targets, three (Target 20, 24 and likely 28) require accounting for backbone conformational changes in the docking protocol, and therefore are much more challenging to predict. However, these targets serve the aim of the CAPRI experiment well, as they motivate us to explore new approaches to incorporating backbone flexibility explicitly in our docking predictions. Encouragingly, our attempt of combining docking and loop modeling successfully predicted the significant interface loop conformational change in Target 20, and *ab initio* folding correctly predicted the conformation of the C-terminus region of ArfBD in Target 24. Inspired by these efforts, we have reformulated the RosettaDock method using a "fold-tree" representation, which provides a general framework to treat various types of backbone conformational change in protein–protein docking (Wang C, Bradley P, Baker D. Protein–protein docking with backbone flexibility. Submitted for review).

The strategies used in the predictions of Target 20 and Target 24 to handle backbone flexibility are similar in that flexible-backbone regions (defined based on the available experimental information) are dissected from the rest of docking partners first and then built back onto promising docking solutions (after the remaining docking partners are docked) to generate full-length models. The success of this approach depends heavily on whether the correct interface can be identified when the remaining portions of the docking partners are docked. For example, the interface between Arfl and ArfBD in Target 24 was not ranked favorably in docking when the C-terminal helix of ArfBD was not present, and therefore no prediction was correct when the C-terminal helix of ArfBD was added back.

Biological information played an important role in the predictions, especially for targets with backbone conformational change upon binding. The experimental data collected for each target helps to define flexible regions and types of movement on the one hand, and to significantly reduce the space to be sampled on the other hand. The successful predictions for Target 20 and Target 26 both benefit from the use of relevant experimental information in docking. However, the lesson learned from Target 21 and Target 27 suggests that such information can be misleading and therefore should be used with caution.

From a purely technical prospective, we have made substantial effort in developing flexible-backbone docking approaches in CAPRI rounds 6–12. However, it must be noted that the rigid-body Z-DOCK approach[43] was more successful for a number of the targets. Z-DOCK and other approaches employing softer potentials are less sensitive to backbone conformational changes. Our combination of explicit backbone optimization plus a strict potential should be optimal in the high sampling limit where the correct conformation is certain to be encountered, but with insufficient sampling the results can be inferior to fixed-backbone sampling with a soft potential.

## ACKNOWLEDGEMENT

We thank the many scientists who have participated in the development of the Rosetta software suite. In particular, David Kim developed and maintained Rosetta@Home; Keith Laidig and Chance Reschke built and maintained reliable, state-of-the-art computing resources. We thank all the Rosetta@Home users worldwide who have generously donated their computers to support our research, particularly users "StevenK," "Administrator," "QuickBeam," "devzero," "raptur," "cyclisttgb," "borekv," "yopjpeg," "orion2598," and "MK_I" who produced low-energy models for our submissions. We thank the CAPRI management and evaluation team for organizing the experiment.

## REFERENCES

1. Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The protein data bank and the challenge of structural genomics. Nat Struct Biol 2000;7(Suppl.):957–959.

 DOI 10.1002/prot

2. Vajda S, Camacho CJ. Protein–protein docking: is the glass half-full or half-empty? Trends Biotechnol 2004;22:110–116.
3. Smith GR, Sternberg MJ. Prediction of protein–protein interactions by docking methods. Current opinion in structural biology 2002;12: 28–35.
4. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins 2002;47:409–443.
5. 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.
6. Janin J. The targets of CAPRI rounds 3–5. Proteins 2005;60:170–175.
7. 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.
8. 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.
9. Wodak SJ, Mendez R. Prediction of protein–protein interactions: the CAPRI experiment, its evaluation and implications. Curr Opin Struct Biol 2004;14:242–249.
10. 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.
11. Wang C, Schueler-Furman O, Baker D. Improved side-chain modeling for protein–protein docking. Protein Sci 2005;14:1328–1339.
12. Li Z, Scheraga HA. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. Proc Natl Acad Sci USA 1987;84:6611–6615.
13. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133–152.
14. Kortemme T, Morozov AV, Baker D. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein–protein complexes. J Mol Biol 2003;326:1239–1259.
15. Schueler-Furman O, Wang C, Baker D. Progress in protein–protein docking: atomic resolution predictions in the CAPRI experiment using RosettaDock with an improved treatment of side-chain flexibility. Proteins 2005;60:187–194.
16. Canutescu AA, Dunbrack RL, Jr. Cyclic coordinate descent: a robotics algorithm for protein loop closure. Protein Sci 2003;12: 963–972.
17. Rohl CA, Strauss CE, Chivian D, Baker D. Modeling structurally variable regions in homologous proteins with rosetta. Proteins 2004;55:656–677.
18. Bradley P, Baker D. Improved beta-protein structure prediction by multilevel optimization of nonlocal strand pairings and local backbone conformation. Proteins 2006;65:922–929.
19. 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.
20. Yang Z, Shipman L, Zhang M, Anton BP, Roberts RJ, Cheng X. Structural characterization and comparative phylogenetic analysis of *Escherichia coli* HemK, a protein (N5)-glutamine methyltransferase. J Mol Biol 2004;340:695–706.
21. Schubert HL, Phillips JD, Hill CP. Structures along the catalytic pathway of PrmC/HemK, an N5-glutamine AdoMet-dependent methyltransferase. Biochemistry 2003;42:5592–5599.
22. Vestergaard B, Van LB, Andersen GR, Nyborg J, Buckingham RH, Kjeldgaard M. Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1. Molecular cell 2001;8:1375–1382.
23. 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. Molecular cell 2005;20:917–927.
24. 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.
25. 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.
26. 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.
27. 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.
28. Bose ME, McConnell KH, Gardner-Aukema KA, Muller U, Weinreich M, Keck JL, Fox CA. The origin recognition complex and Sir4 protein recruit Sir1p to yeast silent chromatin through independent interactions requiring a common Sir1p domain. Mol Cell Biol 2004;24:774–786.
29. Hyvonen M, Macias MJ, Nilges M, Oschkinat H, Saraste M, Wilmanns M. Structure of the binding site for inositol phosphates in a PH domain. The EMBO J 1995;14:4676–4685.
30. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999;292:195–202.
31. Rohl CA, Strauss CE, Misura KM, Baker D. Protein structure prediction using Rosetta. Methods Enzymol 2004;383:66–93.
32. Kawasaki M, Nakayama K, Wakatsuki S. Membrane recruitment of effector proteins by Arf and Rab GTPases. Curr Opin Struct Biol 2005;15:681–689.
33. Carr S, Penfold CN, Bamford V, James R, Hemmings AM. The structure of TolB, an essential component of the tol-dependent translocation system, and its protein–protein interaction with the translocation domain of colicin E9. Structure 2000;8:57–66.
34. Abergel C, Walburger A, Bouveret E, Claverie JM. MAD structure of the periplasmic domain of the *E. coli* pal protein, to be published.
35. 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.
36. Bouveret E, Benedetti H, Rigal A, Loret E, Lazdunski C. In vitro characterization of peptidoglycan-associated lipoprotein (PAL)-peptidoglycan and PAL-TolB interactions. J Bacteriol 1999;181:6306–6311.
37. 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.
38. Avvakumov GV, Choe J, Newman EM, Mackenzie F, Kozieradzki I, Bochkarev A, Sundstrom M, Arrowsmith C, Edwards A, Dhe-paganon S, Structural Genomics Consortium (SGC). Ubiquitin-conjugating enzyme E2-25 kDa (Huntington interacting protein 2), to be published.
39. 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.
40. 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.
41. 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.
42. Ogunjimi AA, Briant DJ, Pece-Barbara N, Le Roy C, Di Guglielmo GM, Kavsak P, Rasmussen RK, Seet BT, Sicheri F, Wrana JL. Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. Mol Cell 2005;19:297–308.
43. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47:281–294.