WILEY InterScience® DISCOVER SOMETHING GREAT

PROTEINS STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

# HADDOCK versus HADDOCK: New features and performance of HADDOCK2.0 on the CAPRI targets

**Sjoerd J. de Vries,[1] Aalt D. J. van Dijk,[1] Mickaël Krzeminski,[1] Mark van Dijk,[1] Aurelien Thureau,[1] Victor Hsu,[2] Tsjerk Wassenaar,[1] and Alexandre M. J. J. Bonvin*[1]**

[1] Bijvoet Center for Biomolecular Research, Science Faculty, Utrecht University, 3584CH, Utrecht, The Netherlands

[2] Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331-7305

*ABSTRACT*

***Here we present version 2.0 of HADDOCK, which incorporates considerable improvements and new features. HADDOCK is now able to model not only protein–protein complexes but also other kinds of biomolecular complexes and multi-component (N > 2) systems. In the absence of any experimental and/or predicted information to drive the docking, HADDOCK now offers two additional ab initio docking modes based on either random patch definition or center-of-mass restraints. The docking protocol has been considerably improved, supporting among other solvated docking, automatic definition of semi-flexible regions, and inclusion of a desolvation energy term in the scoring scheme. The performance of HADDOCK2.0 is evaluated on the targets of rounds 4-11, run in a semi-automated mode using the original information we used in our CAPRI submissions. This enables a direct assessment of the progress made since the previous versions. Although HADDOCK performed very well in CAPRI (65% and 71% success rates, overall and for unbound targets only, respectively), a substantial improvement was achieved with HADDOCK2.0.***

Proteins 2007; 69:726–733.


*Key words: HADDOCK; flexible docking; scoring; multibody docking.*

## INTRODUCTION

In recent years, docking has emerged as an important method to model biomolecular complexes, which is complementary to experimental structural methods. Despite the considerable progress shown in the previous rounds of CAPRI,[1–3] there still are major problems to be addressed, especially in dealing with flexibility and binding-induced conformational changes, and in scoring.[4–8] In the context of CAPRI rounds 6–11, which placed a heavy emphasis on unbound docking, these issues have become more pronounced. Exploiting biochemical data from literature to locate the interface has been shown to be of major importance.[4,6] In the absence of such data, interface predictions can be exploited as well.[9–12] HADDOCK is among the very few methods in which these data are used directly to drive the docking, rather than to filter pregenerated structures (for details see Dominguez *et al.*[13]).

Here we present version 2.0 of HADDOCK, which has been considerably extended, and evaluate its performance on the CAPRI targets of rounds 4–11, comparing it to our original submissions. This enables a direct assessment of the progress made since the previous versions.

## HADDOCK2.0 NEW FEATURES

HADDOCK2.0 has been modified to support docking of proteins, DNA,[14] RNA, oligosaccharides,[15] and small ligands, up to a total of six separate molecules (or domains) per docking. The new version allows the inclusion of anisotropy restraints from NMR (both residual dipolar couplings and relaxation data)[16,17] and supports solvated docking, that is, allowing the explicit inclusion of interfacial water molecules in the docking process.[18]

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat.

The authors state no conflict of interest.

Grant sponsor: Netherlands Organization for Scientific Research; Grant number: 700.56.442; Grant sponsor: European Community (FP6 STREP "ExtendNMR"); Grant number: LSHG-CT-2005-018988; Grant sponsor: FP6 STREP "BacAbs"; Grant number: LSHB-CT-2006-037325; Grant sponsor: Integrated Project SPINE2-COMPLEX; Grant number: 031220.

Aalt D. J. van Dijk's current address is Applied Bioinformatics, Plant Research International, Wageningen UR, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Aurelien Thureau's current address is CNRS, ICSN, Laboratoire de Chimie et Biologie Structurales, 1 avenue de la terrasse, 91190, Gif-sur-Yvette, France.

*Correspondence to: Alexandre M. J. J. Bonvin; Bijvoet Center for Biomolecular Research, Science Faculty, Utrecht University, 3584CH, Utrecht, The Netherlands. E-mail: a.m.j.j.bonvin@chem.uu.nl

Received 31 May 2007; Revised 6 July 2007; Accepted 10 July 2007

Published online 5 September 2007 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/prot.21723

HADDOCK defines ambiguous interaction restraints (AIRs) between residues (or chemical groups) of the different molecules. A residue can be defined as active or passive. For every AIR, the effective distance ($d^{\text{eff}}$) is computed as follows:

$$d_{\text{iAB}}^{\text{eff}} = \left( \sum_{m_{iA}=1}^{N_{\text{atoms}}} \sum_{k=1}^{N_{\text{res}}B} \sum_{n_{kB}=1}^{N_{\text{atoms}}} \frac{1}{d_{m_{iA}n_{kB}}^{6}} \right)^{-\frac{1}{6}}$$

where $N_{\text{atoms}}$ indicates all atoms of a given residue and $N_{\text{res}}$ the sum of active and passive residues for a given protein.

A restraining potential is defined to minimize all effective distances to a maximum value of 2 Å. This potential is harmonic up to 1 Å distance violation and then switches within a 1 Å window to a linear mode: the AIR forces become thus constant for violations larger than 2 Å.[19]

In the absence of experimental and/or predicted data, HADDOCK now offers two ab initio docking modes:

i. *random patch definition*: patches of active residues are randomly defined on the surface or part of the surface (e.g., CD-loop regions of an antibody) of each molecule
ii. *center-of-mass restraints*: one distance restraint is defined between the centers of mass of the molecules with the distance automatically defined based on the size and shape of the molecules

The docking protocol consists of three steps: rigid body docking by energy minimization driven by the interaction restraints (it0), semi-flexible refinement in torsion angle space in which side-chains and backbone atoms of the interface residues are allowed to move (it1), and finally, Cartesian dynamics refinement in explicit solvent, typically water, although DMSO is also supported. The last two steps are referred to as the refinement stage.

HADDOCK2.0 now offers the following improvements:

i. *Multibody docking:* the docking can be performed for up to six separate bodies representing different molecules or molecular domains.
ii. *Sampling of 180° rotated solutions*: To deal with the problem of symmetrical solutions (e.g., rotated conformations as was the case for targets 11 and 12) we now implemented a systematic sampling of 180° rotated solutions in it0 (and optionally in it1): each solution generated can automatically be compared to the same solution in which one protein is rotated by 180° around an axis perpendicular to the interface and energy minimized again. The rotation axis is automatically defined between the centers of masses of interface residues (within 5 Å of the partner molecule). The best solution is kept for further refinement.
iii. *Automatic definition of semi-flexible segments*: Previously, the flexible parts of the protein had to be defined manually. These can now be defined automatically, based on the residues that are in close proximity (within 5 Å) to a partner molecule.
iv. *Improved scoring*: An empirical desolvation energy term[20] is now implemented. Moreover, the scoring has been optimized for each stage of the docking protocol.

## MATERIALS AND METHODS

HADDOCK2.0 was run on all targets using default parameters, including random removal of 50% of the data for each separate docking trial, automatic definition of flexibility, and sampling of 180° rotated solutions in it0, but excluding solvated docking. Homology models were obtained using various programs and servers.[21–23] We usually presampled within HADDOCK to obtain an ensemble of starting structures. The protonation state of histidines residues was estimated using the WHATIF server.[24] During all unbound–bound docking, the bound coordinates were frozen. Cofactor and ions were explicitly included in the docking. Co-factor parameters were obtained from the PRODRG server.[25] As in our previous CAPRI submissions, HADDOCK was driven by data derived from literature and/or interface predictions.[9–12] The active and passive residues that were defined during the docking and target specific modifications to the protocols when applicable are given in Supplementary Tables S1 and S2. The HADDOCK parameter files and starting structures for individual targets are available upon request.

During the various stages of the docking protocols, solutions were scored as follows:

- it0: $0.01^{*}E_{\text{vdW}} + 1.0^{*}E_{\text{elec}} + 0.01^{*}E_{\text{AIR}} - 0.01^{*}\text{BSA} + 1.0^{*}E_{\text{desolv}}$
- it1: $1.0^{*}E_{\text{vdW}} + 1.0^{*}E_{\text{elec}} + 0.1^{*}E_{\text{AIR}} - 0.01^{*}\text{BSA} + 1.0^{*}E_{\text{desolv}}$
- water: $1.0^{*}E_{\text{vdW}} + 0.2^{*}E_{\text{elec}} + 0.1^{*}E_{\text{AIR}} + 1.0^{*}E_{\text{desolv}}$

$E_{\text{AIR}}$ is the ambiguous interaction restraint energy, BSA the buried surface area and $E_{\text{Desolv}}$ the desolvation energy.[20]

After water refinement, structures were clustered using a cut-off of 7.5 Å l-RMSD, only counting the interface backbone atoms of the ligand. The clusters were ranked according to the average energy of the four best structures in the cluster.

Assessment of the quality of the predictions was done according to the CAPRI evaluation criteria[26]:

- One-star (acceptable) predictions:
  [{(Fnat $\geq$ 0.1) && (Fnat $<$ 0.3)} && {(l-RMSD $\leq$ 10.0) || (i-RMSD $\leq$ 4.0)}]
- Two-star (medium quality) predictions:
  [{(Fnat $\geq$ 0.3) && (Fnat $<$ 0.5)} && {(l-RMSD $\leq$ 5.0) || (i-RMSD $\leq$ 2.0)}]
- Three-star (high quality predictions):
  [(Fnat $\geq$ 0.5) && {(l-RMSD $\leq$ 1.0) || (i-RMSD $\leq$ 1.0)}]

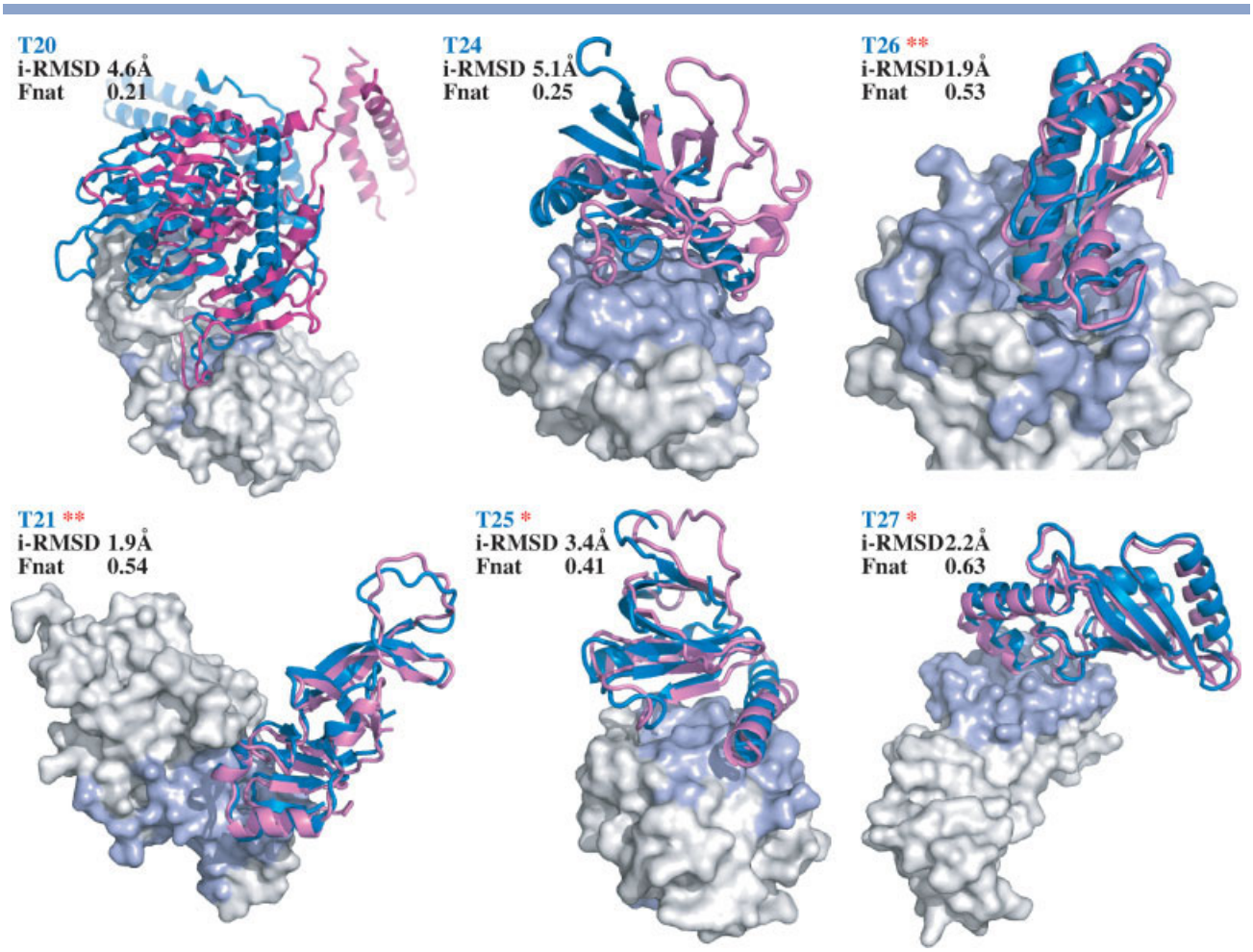


**Figure 1**

*Overlays of our best CAPRI submissions (those with the lowest i-RMSD values) and the reference crystal structures. Receptors are shown in surface representation with the area defined by the active and passive residues shown in light blue. The ligands are shown in cartoon representation (HADDOCK best: pink, reference target: blue). This figure was generated with PyMol.*[44]

Throughout the text, when the number of x-star predictions is mentioned, the number of x-star or better predictions is meant. For target 27, only the second interface (A–C′) was considered.

## RESULTS

Among the different groups participating in CAPRI rounds 6-11, HADDOCK performed well. Our best predictions (lowest i-RMSD; see Material and Methods) are shown in Figure 1. Of the six targets,[27–33] four were predicted successfully, with at least one one-star prediction among our submissions. For two of the targets, two-star predictions were submitted. Even for the two targets that were not successful, structures with high fractions of native contacts (Fnat) could be generated.

During our participation in these rounds of CAPRI, we used developmental versions of HADDOCK incorporating some of the features of HADDOCK2.0, but never all of them. In addition, the final selection of structures was performed manually, based on HADDOCK and other parameters such as DFIRE[34] and FastContact.[35] In contrast, HADDOCK2.0 was tested in a mostly automated manner (see Material and Methods) on the CAPRI targets of rounds 4–11, including the canceled targets for which data and starting structures had already been obtained. The same data (see Suppl. Table S1) and starting conformations were used as for our original submissions. The results of the HADDOCK2.0 redocking are shown in Table I. After water refinement, the structures

 DOI 10.1002/prot

**Table I**
*Performance of HADDOCK2.0 on the CAPRI Targets and Comparison with Our Original Submissions*

| | | All Structures | | | | | Best Cluster | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Target | (a) | Total (b) | *(c) | **(c) | ***(c) | <Fnat> (d) | Rank (e) | l-RMSD (Å) | i-RMSD (Å) | Fnat | Quality (f) | Subm.(g) |
| 10 | it0 | 1,000 | 31 | 0 | 0 | 0.19 | | | | | | |
| | selected | 200 | 18 | 0 | 0 | 0.22 | | | | | | |
| | it1 | 200 | 23 | 1 | 0 | 0.26 | | | | | | |
| | water | 200 | 24 | 0 | 0 | 0.25 | 1 | 2.08 | 1.77 | 0.27 | * | ** |
| 11 | it0 | 2,420 | 202 | 1 | 0 | 0.26 | | | | | | |
| | selected | 200 | 21 | 0 | 0 | 0.34 | | | | | | |
| | it1 | 200 | 30 | 0 | 0 | 0.32 | | | | | | |
| | water | 200 | 37 | 0 | 0 | 0.37 | 1 | 7.48 | 3.22 | 0.50 | * | ** |
| 12 | it0 | 1,100 | 151 | 26 | 3 | 0.21 | | | | | | |
| | selected | 200 | 22 | 7 | 3 | 0.29 | | | | | | |
| | it1 | 200 | 23 | 7 | 5 | 0.31 | | | | | | |
| | water | 200 | 26 | 7 | 5 | 0.34 | 1 | 1.41 | 0.64 | 0.91 | *** | 0 |
| 13 | it0 | 1,100 | 160 | 47 | 0 | 0.24 | | | | | | |
| | selected | 200 | 79 | 40 | 0 | 0.31 | | | | | | |
| | it1 | 200 | 83 | 41 | 3 | 0.34 | | | | | | |
| | water | 200 | 85 | 41 | 1 | 0.37 | 1 | 1.83 | 1.15 | 0.59 | ** | *** |
| 14 | it0 | 1,000 | 42 | 17 | 14 | 0.36 | | | | | | |
| | selected | 200 | 25 | 14 | 14 | 0.42 | | | | | | |
| | it1 | 200 | 27 | 14 | 11 | 0.43 | | | | | | |
| | water | 200 | 27 | 14 | 10 | 0.44 | 1 | 1.37 | 0.80 | 0.63 | *** | *** |
| 15 | it0 | 2,000 | 73 | 8 | 1 | 0.31 | | | | | | |
| | selected | 200 | 15 | 4 | 1 | 0.33 | | | | | | |
| | it1 | 200 | 13 | 4 | 1 | 0.32 | | | | | | |
| | water | 200 | 16 | 9 | 2 | 0.37 | 1 | 1.78 | 0.98 | 0.64 | *** | ** |
| 18 | it0 | 2,000 | 118 | 96 | 0 | 0.69 | | | | | | |
| | selected | 200 | 94 | 81 | 0 | 0.72 | | | | | | |
| | it1 | 200 | 94 | 10 | 0 | 0.29 | | | | | | |
| | water | 200 | 94 | 18 | 0 | 0.36 | 1 | 5.30 | 1.67 | 0.78 | ** | 0 |
| 19 | it0 | 10,000 | 80 | 9 | 0 | 0.23 | | | | | | |
| | selected | 200 | 16 | 1 | 0 | 0.25 | | | | | | |
| | it1 | 200 | 15 | 1 | 0 | 0.23 | | | | | | |
| | water | 200 | 17 | 1 | 0 | 0.24 | 1 | 5.82 | 1.13 | 0.56 | ** | 0 |
| 20 | it0 | 2,000 | 243 | 0 | 0 | 0.21 | | | | | | |
| | selected | 400 | 120 | 0 | 0 | 0.23 | | | | | | |
| | it1 | 400 | 136 | 1 | 0 | 0.32 | | | | | | |
| | water | 200 | 139 | 2 | 0 | 0.34 | 1 | 7.22 | 2.48 | 0.48 | * | 0 |
| 21 | it0 | 2,420 | 204 | 5 | 0 | 0.34 | | | | | | |
| | selected | 200 | 79 | 0 | 0 | 0.37 | | | | | | |
| | it1 | 200 | 76 | 3 | 0 | 0.37 | | | | | | |
| | water | 200 | 76 | 4 | 0 | 0.38 | 1 | 6.09 | 1.83 | 0.51 | ** | ** |
| 22 | it0 | 10,000 | 38 | 5 | 0 | 0.31 | | | | | | |
| | selected | 400 | 9 | 2 | 0 | 0.39 | | | | | | |
| | it1 | 400 | 9 | 2 | 0 | 0.36 | | | | | | |
| | water | 400 | 9 | 2 | 0 | 0.42 | 1 | 8.87 | 3.80 | 0.53 | * | NP |
| 24 | it0 | 3,300 | 20 | 0 | 0 | 0.18 | | | | | | |
| | selected | 200 | 2 | 0 | 0 | 0.19 | | | | | | |
| | it1 | 200 | 2 | 0 | 0 | 0.21 | | | | | | |
| | water | 200 | 2 | 0 | 0 | 0.21 | 7 | 11.62 | 3.40 | 0.22 | * | 0 |
| 25 | it0 | 2,200 | 129 | 43 | 0 | 0.34 | | | | | | |
| | selected | 200 | 32 | 18 | 0 | 0.38 | | | | | | |
| | it1 | 200 | 37 | 13 | 0 | 0.45 | | | | | | |
| | water | 200 | 39 | 15 | 0 | 0.44 | 2 | 5.49 | 1.85 | 0.57 | ** | * |
| 26 | it0 | 4,000 | 961 | 116 | 0 | 0.31 | | | | | | |
| | selected | 200 | 64 | 9 | 0 | 0.35 | | | | | | |
| | it1 | 200 | 65 | 17 | 0 | 0.42 | | | | | | |
| | water | 200 | 64 | 18 | 0 | 0.44 | 3 | 4.25 | 1.98 | 0.58 | ** | ** |
| 27 | it0 | 4,000 | 681 | 29 | 1 | 0.22 | | | | | | |
| | selected | 200 | 117 | 6 | 0 | 0.24 | | | | | | |
| | it1 | 200 | 136 | 4 | 0 | 0.29 | | | | | | |
| | water | 200 | 142 | 4 | 0 | 0.34 | 1 | 5.31 | 1.57 | 0.78 | ** | * |
| All (h) | it0 | 45,540 | 2973 | 289 | 5 | 0.26 | | | | | | |
| | selected | 3,000 | 594 | 87 | 4 | 0.29 | | | | | | |
| | it1 | 3,000 | 648 | 94 | 9 | 0.33 | | | | | | |
| | water | 2,800 | 676 | 103 | 8 | 0.36 | | | | | | |

Unbound targets are highlighted in gray. a – Various stages of the docking protocol: it0, initial rigid body docking; selected, the selection of structures after it0; it1, flexible refinement in torsion angle space; and water, refinement in explicit solvent (water). b – Total number of structures after each step in the protocol. c – Number of one-star or better, two-star or better, and three-star or better structures after each step in the protocol. d – Average fraction of native contacts, computed for all structures with i-RMSD $\leq$ 4 Å. e – Rank of the cluster with the highest number of stars among any structure in the top 4 structures of the cluster. If several clusters had the same highest number of stars, the cluster with the better rank is shown. f – Statistics for the structure with the highest number of stars among the top 4 of the cluster. If several structures had the same highest number of stars, the structure with the better rank is shown. g – The number of stars for our best submission in CAPRI. NP = not performed due to the target being canceled. h – Complex 14 and 18 were excluded from the pooled statistics due to the fact that intermolecular interactions were scaled down in it0 for those complexes to allow for more efficient inter penetration, resulting in a larger fraction of native (and non-native) contacts than for other complexes.

were clustered based on interface ligand-RMSDs and ranked according to the average score of the top 4 member of each cluster. The clusters with the highest-star prediction in their top 4 are shown in the table.

For target 10,[36] for which we had previously submitted the only two-star prediction of all submissions, redocking looked a bit disappointing: only one one-star prediction was generated. However, the RMSD values are well within the threshold of a two-star prediction and only a slight deterioration of Fnat (0.27) caused the prediction to drop to one-star. This may be due to the nature of the target (a trimeric spike). Still, the one-star cluster ranks at the top. Closer inspection showed that one two-star prediction was generated in it1, which was lost after the water refinement.

Target 11,[37] only yielded one-star clusters, but the top-scoring structure had an excellent Fnat of 0.5. This is worse than our original submission, which included a two-star prediction. A single two-star prediction was generated at it0, but not selected.

For target 12,[37] the results improved dramatically. Previously, we did not submit any correct solution. This time, the top-ranking structure of the top-ranking cluster is a highly accurate three-star prediction (i-RMSD = 0.64 Å; Fnat = 0.91).

HADDOCK2.0 came very close to generating a three-star prediction for target 13[38] as well (i-RMSD = 1.15 Å, Fnat = 0.59). Actually, three three-star predictions were generated during it1, but the best was ranked 51st; after water refinement, a single structure was left, ranking 25th, which was not high enough to reach the top 4 of the cluster. In CAPRI, we managed to submit a three-star prediction, although, in that case, visual inspection played an important role.

For target 14,[39] the default protocol was modified slightly. During it0, the non-bonded interaction energies were scaled by a factor 0.01 to allow deeper penetration. For this target, the largest number of three-star predictions (14) was generated at it0. Although some of them deteriorated during the refinement, the top-ranking structure of the top-ranking cluster is a three-star prediction. Note that our previous submission included a three-star prediction as well.

A three-star prediction was also obtained for target 15,[40] despite the fact that after water refinement only 2 out of 200 structures were three-star predictions, demonstrating the strength of the HADDOCK2.0 scoring function. Our previous submissions for that target (which was cancelled) were only two-star.

For target 18,[41] the it0 interaction energies were scaled down in the same way as for target 14. This was necessary, because the binding groove of xylanase is in a closed conformation, preventing the docking of the TAXI inhibitor. A large number of two-star predictions were generated; however, many of them contained clashes, with as result that, during refinement, the inhibitor was often ejected from the binding site. Still, the top-scoring structure of the top-scoring cluster was a two-star prediction with an excellent Fnat (0.78). We did not submit originally any correct solutions for this target. We note though, that essential literature data were overlooked, which were included in the redocking.

Target 19,[42] an antibody–antigen complex, proved a great challenge, as well as a great success of the HADDOCK2.0 scoring function. In the absence of data, HADDOCK2.0 was run using random definition of patches on the entire surface of the antigen and the CD-loop region of the antibody. Out of 10,000 structures, 9 two-star predictions were generated at it0. After water refinement, the single two-star prediction was the second-best ranking structure of the top-ranking cluster, ranking 16th among all structures. This is a great improvement over our original submission, which did not include any correct structures.

Target 20[28] was an unbound target for which our best submission was only at 4.6Å i-RMSD, despite an Fnat of 0.21. The prediction was outside the acceptable range due to a conformational change of the loop binding in the active site of HemK in the complex. This time we modified our protocol so that RF1 was cut into two molecules prior to docking; a three-body docking was then performed with additional restraints to restore the connectivity between the RF1 fragments. A large number of one-star structures were generated and selected, but only two of them resulted in refined two-star predictions. The top cluster has an excellent Fnat (0.48 for the top structure and 0.57 for the third-best), although the RMSDs are not good enough for two stars designation. The highest two-star prediction has a rank of 32 after water refinement. This is a large improvement over our submission, which did not yield any correct solutions. We note though, that correct solutions could not have been generated without the a-priori knowledge of the loop conformational change. Predicting conformational changes occurring upon binding remains an open challenge!

For target 21,[30] another unbound target, only four structures were refined into two-star predictions. However, in this case, one of these predictions is the top structure of the top-ranking cluster (i-RMSD = 1.83Å, Fnat = 0.51). This is similar in quality to the best model that we submitted.

Target 22,[32] also presented an unbound target, but was cancelled. This target, an antibody–antigen complex, was a challenge because we did not find any reliable experimental data to drive the docking. Therefore, center-of-mass restraints were used to drive the docking: 10,000 structures were generated, out of which 38 were one-star and 5 were two-star predictions. Strikingly, 9 one-star and 2 two-star predictions were among the top 200 structures. After water refinement, the third structure of the top-ranking cluster had an excellent Fnat of 0.57,

 DOI 10.1002/prot

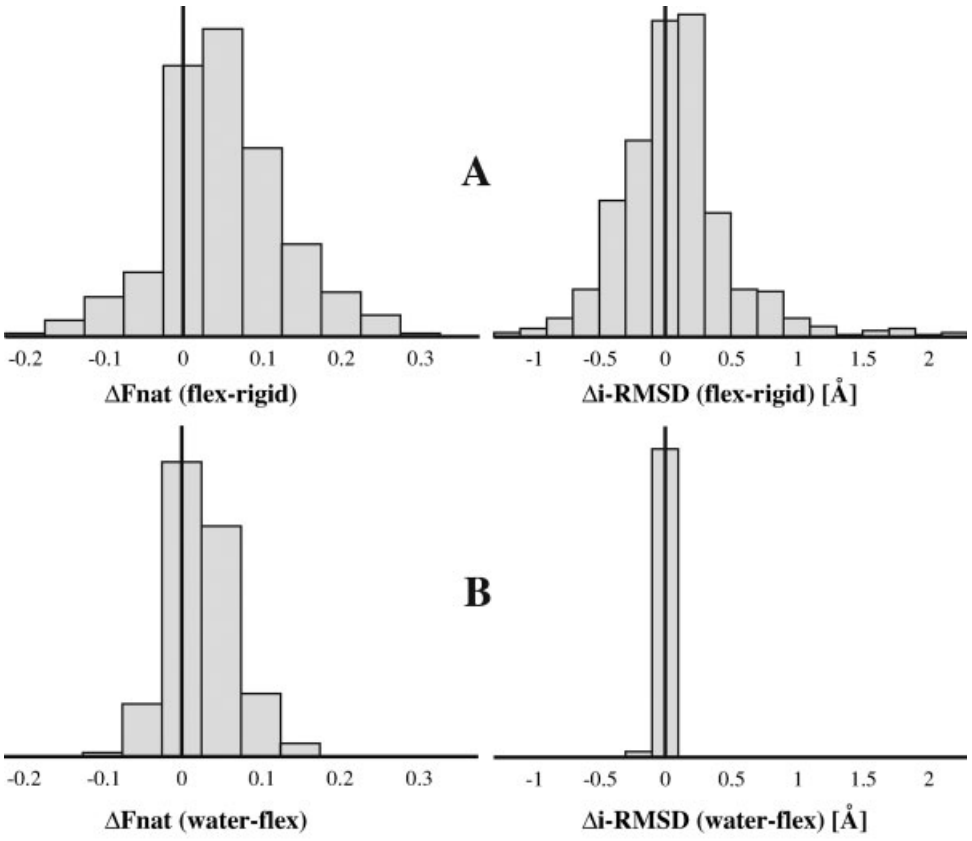


**Figure 2**

*Histograms of the differences in the fraction of native contacts (Fnat) (left panel) and the interface RMSD (i-RMSD) (right panel) between* ***(A)*** *the flexible refinement stage in vacuum (it1) and the rigid body docking (it0), and* ***(B)*** *between water refinement (water) and the flexible refinement stage in vacuum (it1). Statistics are shown only for dimeric targets for which data and unbound structures were available for both proteins (targets 21, 26, 27).*

although the RMSD statistics resulted in a one-star classification. The top-ranking two-star prediction had an overall rank of 16, which did not bring it to the top 4.

Target 24[27] consisted of a protein in the unbound form and a protein that had to be homology-modeled. The homology model proved inaccurate and only 20 one-star structures were generated in it0, of which two were selected after refinement. No two-star predictions were generated. Still, the seventh-ranking cluster contains a one-star prediction. This is an improvement over our submission, which did not contain any correct predictions.

Target 25[27] was the same complex as target 24, but with the homology model replaced by the bound form of the protein revealing a C-terminal α-helical region completely different from the homology model we used. This region was added to the interface definition. We were able to submit a one-star prediction. However, recalculation using HADDOCK 2.0 produced a two-star prediction, ranked first in the second-ranked cluster.

For target 26,[29] we submitted a two-star prediction, ranked second in l-RMSD among all predictions. Using HADDOCK2.0, the top structure from the third-ranking cluster was again a two-star prediction, with a Fnat of 0.58. While the RMSD values of this structure are worse than our best submission, the Fnat is better, and the fourth-best structure from the cluster has an overall better quality (l-RMSD = 2.63 Å, i-RMSD = 1.29 Å, Fnat = 0.61).

Finally, for target 27[33] experimental information was available on the residues that should be in contact for the sumoylation central to this target, namely Cys93 and Lys14.[43] However, in the crystal structure, it is not Lys14 but Lys10 which makes contacts. Despite this, we generated a large number of two-star predictions, although only one-star predictions were submitted. Recalculation using HADDOCK2.0 placed a two-star prediction as the second-ranking structure of the top-ranked cluster, even though few two-star predictions were generated. This again illustrates the strength of the HADDOCK2.0 scoring function; however, the scoring proved not to be perfect, since in it0, a single three-star prediction was generated, but it only ranked 1765 among the 4000 structures and therefore was further not refined.

Considering only the unbound targets, it is interesting to see the improvements that can be made in the refinement step of our protocol by explicit inclusion of both side-chain and backbone flexibility (Fig. 2): while a mixed result is obtained in i-RMSD values with the most improvement resulting from the semi-flexible refinement (it1), a consistent improvement is observed for Fnat, both in it1 and after water refinement. In many cases, this causes a one-star prediction to become two-star. The largest improvements achieved are 0.3 for the fraction of native contacts and 2.3 Å for the interface-RMSD. We also investigated how systematic sampling of 180° rotated solutions improved our results (Supplementary Table S3): on average over the CAPRI targets, the number of acceptable solutions after rigid-body docking increased by a factor 2.5.

## DISCUSSION

A direct comparison between the re-docking using HADDOCK2.0 and our submissions to CAPRI is unbalanced, since the former is semiautomated, whereas in the latter case we made extensive use of combining different docking runs, manual integration of different scoring functions, and visual inspection. Still, HADDOCK2.0 compares favorably with our CAPRI predictions. HADDOCK's submissions to CAPRI contained at least a one-star prediction in the top 10 for 65% of all targets, and 71% for unbound targets only. Now, with HADDOCK2.0, this is achieved for 100% of the targets. Note that in some cases the manual submissions were actually more accurate, and that in a few other cases the protocol had to be slightly modified. Nonetheless, a major improvement is achieved over the previous version of HADDOCK considering that no manual intervention was involved in the scoring. In particular, the scoring function performed much better in ranking correct solutions at the top, even in cases were the data to drive the docking were fuzzy or incomplete, and where only few correct solutions were obtained. We observed that in the large majority (13/15) of the cases, the top-scoring cluster is the one that is closest to the target structure. In most cases, the selection and clustering performed optimally: HADDOCK2.0 ranked at the top the cluster with the maximum number of stars that could be generated. In the other cases, the number of those structures was usually very small.

We could also assess the rank of the first structure with the maximum number of stars among all water-refined structures, disregarding cluster statistics: it is ranked first in 6/15 cases and in the top 10 in 11/15 cases. When the rank of the first structure with a number of stars equal to that in column 11 of Table I is assessed, it is rank 1 in 7/15 cases and in the top 10 in 14/15 cases. Usually, this is already the case after it1, however, if the first structure was outside of the top 10, the ranking generally improved after water refinement (results not shown). In particular, targets 19 and 22 illustrate the strength of our scoring function. Since the docking was done effectively ab initio, only a few one-star and two-star structures were generated. Even though only 2–4% of the structures were refined, those structures contained 20% or more of the one-star predictions. In both cases, the best cluster was ranked at the top and for target 19 this cluster was even of two-star quality.

Special attention should be paid to the case of target 27. In this case a single three-star prediction was generated at it0, due to the exceptional specificity of the available data. We correctly assumed that the target, a sumoylation complex, would transfer SUMO in a standard way. However, experimental data from the literature[43] indicated that K14 would be the SUMO-accepting/donating lysine, rather than K10, which was actually observed in the target crystal structure. In a new docking run we redefined the contacting lysine to be K10, which resulted in 16 three-star predictions at it0. Four of these were kept after water refinement, belonging to the top ranked cluster. For HADDOCK, this is unique for any complex that is docked from unbound structures alone.

In conclusion, we have shown that the combination of new features in HADDOCK2.0 has led to a significant improvement in both prediction and scoring capabilities. We want here to acknowledge CAPRI, which is an important driving force in recognizing problems and fostering new developments.

HADDOCK2.0 is freely available to academic institutions upon request (see http://www.nmr.chem.uu.nl/haddock).

## REFERENCES


1. Janin J, Henrick K, Moult J, Ten Eyck L, Sternberg MJE, Vajda S, Vasker I, Wodak SJ. CAPRI: a critical assessment of predicted interactions. Proteins: Struct Funct Genet 2003;52:2–9.
2. Méndez 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.
3. Janin J. Assessing predictions of protein–protein interaction: the CAPRI experiment. Protein Sci 2005;14:278–283.
4. van Dijk AD, Boelens R, Bonvin AM. Data-driven docking for the study of biomolecular complexes. FEBS J 2005;272:293–312.
5. Bonvin AM. Flexible protein–protein docking. Curr Opin Struct Biol 2006;16:194–200.
6. van Dijk ADJ, de Vries SJ, Dominguez C, Chen H, Zhou HX, Bonvin AMJJ. Data-driven docking: HADDOCK's adventures in CAPRI. Proteins: Struct Funct Bioinformatics 2005;60:232–238.
7. Vajda S, Camacho CJ. Protein–protein docking: is the glass half-full or half-empty? Trends Biotechnol 2004;22:110–116.
8. Gray JJ. High-resolution protein–protein docking. Curr Opin Struct Biol 2006;16:183–193.
9. Bradford JR, Westhead DR. Improved prediction of protein–protein binding sites using a support vector machines approach. Bioinformatics 2005;21:1487–1494.


 DOI 10.1002/prot

10. Chen H, Zhou HX. Prediction of interface residues in protein–protein complexes by a consensus neural network method: test against NMR data. Proteins 2005;61:21–35.
11. de Vries SJ, van Dijk AD, Bonvin AM. WHISCY: what information does surface conservation yield? Application to data-driven docking. Proteins 2006;63:479–489.
12. Neuvirth H, Raz R, Schreiber G. ProMate: a structure based prediction program to identify the location of protein–protein binding sites. J Mol Biol 2004;338:181–199.
13. 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.
14. van Dijk M, van Dijk AD, Hsu V, Boelens R, Bonvin AM. Information-driven protein-DNA docking using HADDOCK: it is a matter of flexibility. Nucleic Acids Res 2006;34:3317–3325.
15. Wu AM, Singh T, Liu JH, Krzeminski M, Russwurm R, Siebert HC, Bonvin AM, Andre S, Gabius HJ. Activity-structure correlations in divergent lectin evolution: fine specificity of chicken galectin CG-14 and computational analysis of flexible ligand docking for CG-14 and the closely related CG-16. Glycobiology 2007;17:165–184.
16. van Dijk AD, Fushman D, Bonvin AM. Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data. Proteins 2005;60:367–381.
17. van Dijk AD, Kaptein R, Boelens R, Bonvin AM. Combining NMR relaxation with chemical shift perturbation data to drive protein–protein docking. J Biomol NMR 2006;34:237–244.
18. van Dijk AD, Bonvin AM. Solvated docking: introducing water into the modelling of biomolecular complexes. Bioinformatics 2006;22: 2340–2347.
19. Nilges M, Gronenborn AM, Brunger AT, Clore GM. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. Protein Eng 1988;2:27–38.
20. Fernandez-Recio J, Totrov M, Abagyan R. Identification of protein–protein interaction sites from docking energy landscapes. J Mol Biol 2004;335:843–865.
21. Fiser A, Sali A. Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol 2003;374:461–491.
22. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32(Web Server issue):W526–W531.
23. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003; 31:3381–3385.
24. Vriend G. WHAT IF: a molecular modeling and drug design program. J Mol Graph 1990;8:52–56. 29.
25. Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr 2004;60(Part 8):1355–1363.
26. 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.
27. Ghosh A, Praefcke GJ, Renault L, Wittinghofer A, Herrmann C. How guanylate-binding proteins achieve assembly-stimulated processive cleavage of GTP to GMP. Nature 2006;440:101–104.
28. 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.
29. 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.
30. 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.
31. 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.
32. Nielsen TK, Liu S, Luhrmann R, Ficner R. Structural basis for the bifunctionality of the U5 snRNP 52K protein (CD2BP2). J Mol Biol 2007;369:902–908.
33. 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.
34. Zhou H, Zhou Y. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. Protein Sci 2002;11:2714–2726.
35. Camacho CJ, Zhang C. FastContact: rapid estimate of contact and binding free energies. Bioinformatics 2005;21:2534–2536.
36. Bressanelli S, Stiasny K, Allison SL, Stura EA, Duquerroy S, Lescar J, Heinz FX, Rey FA. Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. EMBO J 2004;23:728–738.
37. Carvalho AL, Dias FMV, Prates JAM, Nagy T, Gilbert HJ, Davies GJ, Ferreira LMA, Romao MJ, Fontes CMGA. Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. Proc Natl Acad Sci USA 2003;100:13809–13814.
38. Graille M, Stura EA, Bossus M, Muller BH, Letourneur O, Battail-Poirot N, Sibai G, Gauthier M, Rolland D, Le Du MH, Ducancel F. Crystal structure of the complex between the monomeric form of *Toxoplasma gondii* surface antigen 1 (SAG1) and a monoclonal antibody that mimics the human immune response. J Mol Biol 2005; 354:447–458.
39. Terrak M, Kerff F, Langsetmo K, Tao T, Dominguez R. Structural basis of protein phosphatase 1 regulation. Nature 2004;429:780–784.
40. Graille M, Mora L, Buckingham RH, van Tilbeurgh H, de Zamaroczy M. Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity protein. EMBO J 2004;23:1474–1482.
41. Sansen S, De Ranter CJ, Gebruers K, Brijs K, Courtin CM, Delcour JA, Rabijns A. Structural basis for inhibition of *Aspergillus niger* xylanase by Triticum aestivum xylanase inhibitor-I. J Biol Chem 2004;279:36022–36028.
42. Eghiaian F, Grosclaude J, Lesceu S, Debey P, Doublet B, Treguer E, Rezaei H, Knossow M. Insight into the PrPC -> PrPSc conversion from the structures of antibody-bound ovine prion scrapie-susceptibility variants. Proc Natl Acad Sci USA 2004;101:10254–10259.
43. 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.
44. DeLano WL. The PyMOL Molecular Graphics System. Palo Alto, CA, USA: DeLano Scientific; 2002.