

An overview of data-driven HADDOCK strategies in CAPRI rounds 38-45

Panagiotis I. Koukos¹ | Jorge Roel-Touris¹ | Francesco Ambrosetti^{1,2} |
Cunliang Geng¹  | Jörg Schaarschmidt^{1,3} | Mikael E. Trellet¹ |
Adrien S. J. Melquiond¹ | Li C. Xue¹ | Rodrigo V. Honorato¹ | Irina Moreira^{1,4} |
Zeynep Kurkcuoglu¹ | Anna Vangone¹ | Alexandre M. J. J. Bonvin¹ 

¹Faculty of Science, Department of Chemistry, Bijvoet Center for Biomolecular Research, Computational Structural Biology Group, Utrecht University, Utrecht, The Netherlands

²Department of Physics, Sapienza University, Rome, Italy

³Multiscale Materials Modelling and Virtual Design, Institute of Nanotechnology, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

⁴CNC—Center for Neuroscience and Cell Biology, Rua Larga, FMUC, Polo I, 1° andar, Universidade de Coimbra, Coimbra, Portugal

Correspondence

Alexandre M. J. J. Bonvin, Faculty of Science, Department of Chemistry, Bijvoet Center for Biomolecular Research, Computational Structural Biology Group, Utrecht University, Utrecht, The Netherlands.
Email: a.m.j.j.bonvin@uu.nl

Present address

Anna Vangone, PharmaResearch and Early Development, Large Molecule Research, Roche Innovation Center Munich, Nonnenwald 2. 82377 Penzberg, Germany.

Funding information

China Scholarship Council, Grant/Award Number: 201406220132; Fundação para a Ciência e a Tecnologia, Grant/Award Number: IF/00578/2014; H2020 Marie Skłodowska-Curie Actions, Grant/Award Numbers: BAP 659025, 659826; Horizon 2020 Framework Programme, Grant/Award Numbers: 675728, 823830, 777536, 653549; Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Numbers: 718.015.001, 722.014.005; Netherlands eScience Center, Grant/Award Number: 50 027016G04

Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1002/prot.25869>.

Abstract

Our information-driven docking approach HADDOCK has demonstrated a sustained performance since the start of its participation to CAPRI. This is due, in part, to its ability to integrate data into the modeling process, and to the robustness of its scoring function. We participated in CAPRI both as server and manual predictors. In CAPRI rounds 38-45, we have used various strategies depending on the available information. These ranged from imposing restraints to a few residues identified from literature as being important for the interaction, to binding pockets identified from homologous complexes or template-based refinement/CA-CA restraint-guided docking from identified templates. When relevant, symmetry restraints were used to limit the conformational sampling. We also tested for a large decamer target a new implementation of the MARTINI coarse-grained force field in HADDOCK. Overall, we obtained acceptable or better predictions for 13 and 11 server and manual submissions, respectively, out of the 22 interfaces. Our server performance (acceptable or higher-quality models when considering the top 10) was better (59%) than the manual (50%) one, in which we typically experiment with various combinations of protocols and data sources. Again, our simple scoring function based on a linear combination of intermolecular van der Waals and electrostatic energies and an empirical desolvation term demonstrated a good performance in the scoring experiment with a 63% success rate across all 22 interfaces. An analysis of model quality indicates that, while we are consistently performing well in generating acceptable models, there is room for improvement for generating/identifying higher quality models.

KEY WORDS

biomolecular interactions, complexes, integrative modeling, prediction, scoring

1 | INTRODUCTION

In recent years, molecular simulation techniques have been gaining traction as alternative methods for the elucidation of structural details underpinning cellular mechanisms.¹ Methods like molecular docking can complement experimental techniques such as X-RAY crystallography, NMR and cryo-electron microscopy and allow us to gain insight into the structural machinery of the cell. Understanding the structural elements at the core of interacting biomolecules in atomic detail is the first step toward understanding the nature of those interactions and being able to mechanistically explain what goes awry when they are disturbed (eg, in disease phenotypes).²⁻⁴

Initiatives like those of the D3R consortium^{5,6} test the ability of various modeling software to predict 3D structure and relative binding affinities of pharmaceutically interesting protein receptors bound to drug-like molecules. For protein-protein complexes, and to a lesser extent protein-peptide ones, the performance of various docking codes has been continuously evaluated over a period spanning almost 20 years in the worldwide CAPRI experiment.⁷⁻¹³ We participated in all three experiments (server, manual, and scoring) for all targets of rounds 38-45 with HADDOCK—our data-driven integrative modeling platform.^{14,15} HADDOCK (High Ambiguity Driven DOCKing) makes use of biochemical/biophysical experimental information which is translated into distance restraints that drive the docking toward conformations that satisfy the experimentally available data. This cuts down on the need to exhaustively sample the conformational space and instead allows focusing on flexibly refining a subset of the models generated. Data-driven approaches have, of course, downsides as well, most prominently the fact that if the information that is provided to HADDOCK is incorrect, it will likely not sample the region of the conformational landscape close to the native state.

Our strategies for rounds 38-45 of CAPRI can be grouped in four categories depending on the nature of the target and the availability of information we could use to drive the docking. The first category concerns targets 121, 134, and 135 which featured protein-peptide complexes. Our approach for these targets boils down to variations of our protein-peptide protocol¹⁶ or threading on available templates and refining. The second category concerns targets 122 and 136 for which we could identify good template structures in the PDB.¹⁷ Our main strategy here was to either superimpose the models we generated on the template structures and refine them or extract C α -C α interface distance restraints¹⁸ from the available templates, map them to the target sequence numbering and use those to drive the docking. The third—and most populated category—concerns targets for which no good templates were available but some experimental information was, mainly in the form of evolutionarily conserved residues or mutagenesis data. Targets 123-125 and 131-133 were modeled in this way. The last category concerns the protein-saccharide complexes

(targets 126-130). For these targets, we identified structures of the receptors of interest bound to oligosaccharide molecules of varying lengths and extracted binding site information in the form of the hydrogen bonds between the sugars and the residues of the pocket. We then used those restraints to dock conformers of the sugars we had generated from their SMILES strings.

In the following, we describe the various strategies in more details, and present and discuss our results in light of the official CAPRI evaluation results, since not all reference structures of the complexes are yet available.

2 | METHODS

Here, we summarize the approaches we followed for the various targets of CAPRI rounds 38-45. Due to the diverse nature of the targets, it is impossible to describe details of our protocols that hold true for all of them. Instead, we will first describe our high-level strategy and then list the details on a per-target basis. Additional details about the nature of the targets and the modeling process for all can be found in the Supporting Information.

Owing to the data-driven nature of HADDOCK, our first step after receiving the target sequences was to search structural databases like the PDB as well as make use of tools like HHpred¹⁹ to identify close and remote structural homologues, respectively. Whenever high-quality templates that necessitated no mutations relative to our target sequences were available in the PDB we used those for the docking. Often though, differences in the sequences between the identified template structures and our targets meant that we needed to model the target sequence onto the template structure and to that end we used MODELLER²⁰ and ROSETTA.²¹ When a high-quality template of the full complex was available, we followed a template-based strategy such as modeling of the components, superimposition and refinement or extraction of interface C α -C α carbon distance restraints from the template followed by distance-restrained docking. If we could not identify any usable templates then we turned to protein structure prediction servers such as RAPTOR-X^{22,23} or iTasser.²⁴⁻²⁶

After identifying or generating input models for the docking we searched the literature for putative interface information that might be available from (among others) bioinformatics predictions, mutagenesis or NMR titration experiments or even homologous complexes. For the targets where we could not identify any experimental information to assist in the modeling process, we made use of ab initio docking mode of HADDOCK, which defines restraints between the centers of mass of the various molecules and identified from an analysis of all generated models the most contacted residues.

After assembling all input structures and identifying all relevant interface information, we docked the individual molecules using

HADDOCK. A typical HADDOCK run consists of the following three stages:

1. Random orientation of all molecules followed by energy minimization-driven rigid body docking—it0.
2. Semi-flexible refinement in torsion angle space during which flexibility is introduced to the system starting with the interface (defined at a 5 Å cut-off) sidechains before expanding to the interface backbone atoms as well—it1.
3. Short flexible refinement in explicit solvent—itw.

After every stage, the generated models are scored with the simple yet robust HADDOCK scoring function (HS)^{27,28} which is a linear weighted sum of energetic and structural terms:

$$HS_{it0} = 0.01 E_{vdw} + 1.0 E_{elec} + 1.0 E_{desolv} - 0.01 BSA + 0.01 E_{AIR}$$

$$HS_{it1} = 1.0 E_{vdw} + 1.0 E_{elec} + 1.0 E_{desolv} - 0.01 BSA + 0.1 E_{AIR}$$

$$HS_{itw} = 1.0 E_{vdw} + 0.2 E_{elec} + 1.0 E_{desolv} + 0.1 E_{AIR}$$

where E_{vdw} , E_{elec} , E_{desolv} , and E_{AIR} stand for van der Waals, Coulomb electrostatics, desolvation and restraint energies, respectively, and BSA for Buried Surface Area upon complex formation (in Å²). The non-bonded components of the score (E_{vdw} , E_{elec}) were calculated with the OPLS forcefield,²⁹ the desolvation energy is a solvent accessible surface area-dependent empirical term³⁰ which estimates the energetic gain or penalty of burying specific sidechains upon complex formation.

For our participation in the server category, all models came from a single run that was submitted to the HADDOCK2.2 webserver, whereas for the manual category our submission usually consisted of a combined analysis of the various strategies that we tried out for each individual target. That combined analysis consists of clustering of all generated models using the Fraction of Common Contacts (FCC) metric with a cut-off of 0.75, meaning all models for which 75% of their interface residue contacts are shared would end up in the same cluster. The models are then scored according to the HS_{itw} function shown above (minus the restraint term). The top clusters and models are then visually inspected, and the top 10 models that make up the submission are selected, typically selecting more models from the top ranked cluster and then spreading the remaining models over other good ranking clusters when more than one cluster are generated. Specifically, when the best scoring cluster stands out from the remaining clusters (in a statistically significant way, according to its HADDOCK score), we selected the top four models of that cluster and selected the top two models from each remaining cluster for the prediction experiments. If the cluster-based scores were overlapping, we selected the top two models from each cluster. For the remaining 90 models that are only considered for the scoring experiment, we simply submitted the best-scoring models of a single run or combination of runs that were not already present in the selected top 10 models. Figure 1 provides a visual summary of the strategies we followed for some of our successful predictions.

For the scoring experiment, we first identify the largest meaningful common subset of residue contacts across all models as a way of standardizing the set on which we score. That sometimes involves excluding models which are not docked or have very small/peculiar interfaces. Then, all missing atoms are rebuilt according to OPLS topologies. Finally, all models underwent a short energy minimization (50 steps of steepest descent EM) and are scored according to the HS_{itw} function shown above (minus the restraint term) when dealing with protein-protein systems. The scoring term weights are adjusted accordingly for different systems (see below for details).

2.1 | Targets involving peptides

2.1.1 | Target 121 (round 38)

We modeled the protein (TolA) on PDB entry 1lr0, renumbering the residues as required to match the target sequence. For the peptide, we created an ensemble of three models: the first and second models were created in PyMOL³¹ using the known sequence of the peptide and ideal backbone angles that corresponded to beta and polyproline-II conformations; the third was obtained from the PEPFOLD webserver.³²

For the server submission, we performed ensemble docking with increased sampling using restraints we identified in the literature³³ to drive the docking. For the manual submission, we performed the joint analysis mentioned previously (see second to last paragraph of the Methods section) over a plethora of runs, which centered around the concept of specifying restraints that would mimic hydrogen bonds between the peptide and the protein.

2.1.2 | Targets 134 and 135 (round 44)

For the server submission of target 134, we modeled the protein on PDB entry 4d07 and the peptide on 4d07 and 4qh8, threading the target sequence on the peptide structure of the templates. For the manual submission, we defined restraints to dock instead of threading and refining.

For target 135, we used the same templates for the peptide as target 134 (4d07 & 4qh8) and added 5el0 and followed similar strategies as for target 134.

2.2 | Template-based targets

2.2.1 | Target 122 (round 39)

We modeled the IL23R receptor using MODELLER and the RAPTOR-X webserver creating an ensemble of two models, IL12B using FATCAT^{34,35} in addition to the model that was provided by the organizers. For IL23A, we used the unbound structure that was provided by the organizers.

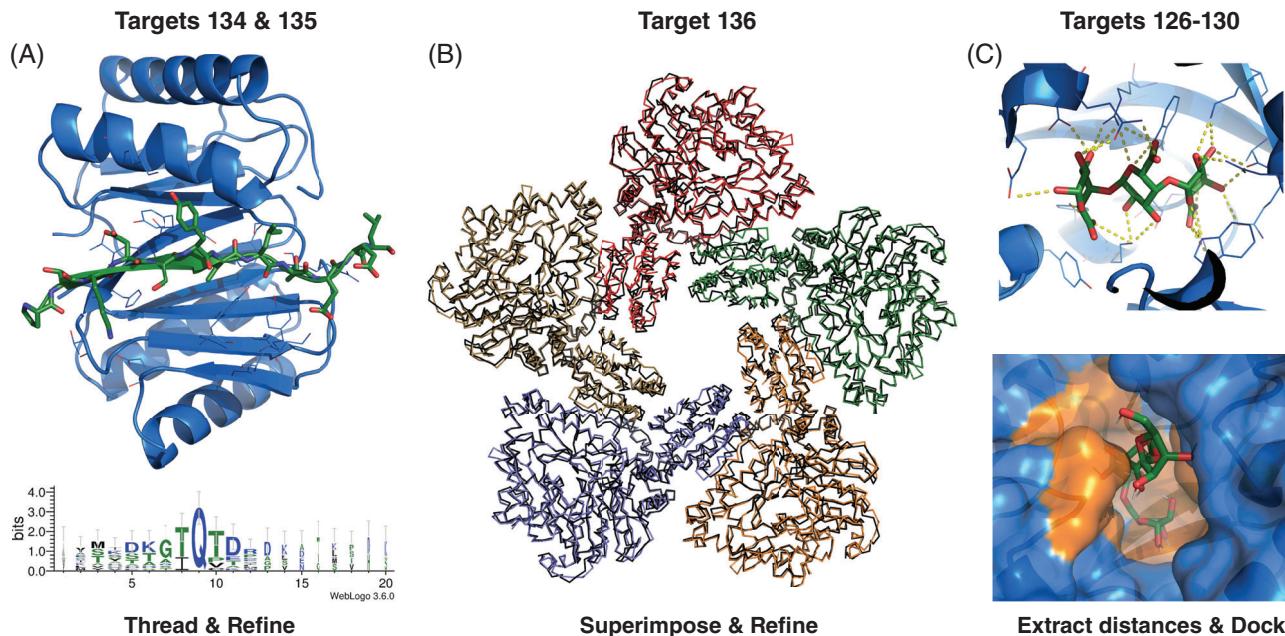


FIGURE 1 Visual summary for some of our successful predictions. A, Concerns the strategy for the peptide targets, and specifically targets 134 and 135, which centered on threading and refinement. Our top-ranking model for the manual submission (medium quality) is shown. The receptor is depicted as cartoon and colored blue with the binding site residues shown as lines, whereas the peptide is depicted as cartoon and sticks and colored dark green. The bottom figure shows a sequence logo of all the peptides we identified in the literature which highlights a conserved TQT motif. B, Concerns target 136 for which we generated models based on available templates, superimposed the models on respective templates and refined them. It depicts the dimer models we generated based on template 2vyc (only the top half of the decameric ring is shown to simplify the figure), superimposed on the template structure chains and refined in HADDOCK. The template structure is shown as black ribbon with the superimposed models colored red, green, orange, blue, and pale gold and represented as ribbons. C, Depicts the two main strategies for the sugar targets. The top figure concerns target 130 and shows the restraints extracted from structure (among others) 2uvj. The receptor is shown as blue cartoons with the binding site residues depicted as lines too. The sugar is shown as sticks with the carbon and oxygen atoms being colored dark green and red, respectively. The bottom figure concerns the remaining sugar targets and shows our top-ranking model for target 128 (medium quality). The consensus binding site that was defined based on the available templates is colored orange, whereas the rest of the receptor is shown as transparent surface and cartoon and colored blue

For the server submission, we superimposed the models of IL12B and IL23A onto PDB structure 5e4e and then, using the IL23A sub-unit, which is shared between PDB entries 5e4e and 2d9q, we superimposed all models onto 2d9q to create a model of the entire complex which was then refined in HADDOCK. For the manual submission, we extracted interface C α -C α restraints¹⁸ from multiple templates and used those to drive the docking instead of superimposing and refining.

2.2.2 | Target 136 (round 36)

Our main strategy for this target was to model the monomers based on available templates and then recreate the full decamer using symmetry restraints. For the manual round, we extracted interface information from the top model (according to the HADDOCK score) of the server submission and turned them into restraints to drive the docking. These runs were performed using our new MARTINI^{36,37} coarse grain implementation.³⁸

2.3 | Information-driven targets

2.3.1 | Targets 123 and 124(round 39)

For the server submission of target 123, we generated an ensemble of models for the PorM protein using the Raptor-X and Robetta³⁹ web servers and an ensemble of models for the nanobody using MOD-ELLER and PDB entries 5imo and 4qlr as templates. We docked using the CDR residues of the nanobody as active residues and the surface residues of PorM as passive residues with increased sampling (10 000/400/400 instead of the default 1000/200/200 models). For the manual submission, we repeated the same run, but we replaced the nanobody ensemble with one kindly provided to us by Dr. Jeff Gray (Johns Hopkins Medical School, Baltimore) (based on PDB entry 5lmw). We submitted the top model of the top 10 clusters for the server and the top 10 models (irrespective of clustering) for the manual prediction.

We could not identify any usable information for target 124, either in the form of template structures or interaction data. Instead,

we defined custom restraints in order to generate models with the global shape of a H (for Haddock, or Hopeless) in order to provide some unrealistic but well optimized interfaces for scoring.

2.3.2 | Target 125 (round 40)

For the server round, we used available structures for the LLT1 dimer (PDB entries 4qkh, 4qki, 4qkj, and 4wco) after renumbering them so they matched the target numbering to create a dimer ensemble. We modeled the NKR-P1 dimer on PDB entries 3m9z and 3t3a using MODELLER to cover all possible combinations of types of interfaces and interacting loops that were available in the literature.^{40,41} We drove the docking using interface information extracted from the available complexes⁴²⁻⁴⁴ and the literature.^{40,41,45} For the manual submission, we repeated the same runs as for the server submission but with a filtered list of residues.

2.3.3 | Targets 131 and 132 (round 42)

A structure of CEACAM1 available in the PDB (entry 2gk2) was used for the docking of both targets. For the modeling of HopQ1, we used template chain A of PDB entry 5lp2 which has an almost perfect sequence identity with our target sequence but does not cover the entire target sequence. To model those gaps, we used MODELLER and selected the top 10 models based on the DOPE score. For HopQ2, we created an ensemble of three models, in which the first came from the RAPTOR-X webserver, the second from iTasser and the last was created in MODELLER after manually curating the alignment between our target sequence and chain A of 5lp2. For the docking, we made use of information available in the literature,⁴⁶ which suggested that a beta strand of HopQ is crucial for the CEACAM1-HopQ interaction and two mutations of CEACAM1 almost knocked out the interaction altogether.

2.3.4 | Target 133 (round 43)

For the server round, to drive the docking, we mapped the sequence differences between target and template proteins to the sequence of the target and defined those restraints as active. For the manual round, we performed a plethora of runs, trying different ways of specifying the restraints (with or without random removal), using single models instead of ensembles, using ROSETTA and molecular dynamics (with GROMACS^{47,48}) to sample alternative conformations of the proteins.

2.4 | Targets involving glycans

2.4.1 | Targets 126-130 (round 41)

We modeled the receptors with ROSETTA-CM protocol⁴⁹ using the templates 2xd3, 2xd3, 3k00, 2uvj for targets 126-129, respectively, and template 3cu9 for target 130. In addition to the already

mentioned templates, we identified many more with HHPRED using the sequence of the target and filtering for the presence of saccharides in the binding pocket. We used those templates to define both interfaces and more specific ambiguous restraints between the sugar molecules and the receptors. For the manual round, we further refined the list of templates that we were using as well as expand the selection of residues specified as active. No other settings were changed compared to the server submission.

For the ligands, we generated up to 500 conformers using OpenEye OMEGA,^{50,51} clustered them and selected representative structures that resulted in the creation of ensembles with 6, 9, 6, 4, and 4 conformations for targets 126-130, respectively.

3 | RESULTS AND DISCUSSION

Table 1 summarizes the results we obtained for the server and manual predictions and for the scoring experiments. In total, we generated acceptable- or higher-quality models for 13 and 11 interfaces for the server and manual submissions, respectively, which corresponds to success rates of 59 and 50% when considering the top 10 models submitted for evaluation. If we consider targets instead of interfaces, and consider successful any target for which at least one interface was correctly predicted, then our success rates jump to 63% and 56%, respectively. Our performance in the scoring experiment is even better, with 14 (63%) interfaces and 11 (69%) targets successfully predicted.

Breaking down our performance by target category, we generated medium- and acceptable quality models for two (targets 134 and 135) of the three peptide targets (targets 121, 134, and 135), with target 121 being a challenging target for which only six acceptable or higher quality models could be found among all submissions. For target 134, we correctly predicted the binding 12mer sequence out of the 50-residue long peptide for both manual and server submissions making good use of the information available in the literature, but failed to generate higher quality models for target 135 for which the binding residues were known, although we did generate more models of the same quality in the top 10. Targets 134 and 135 serve as a good example of the difficulty we are facing in scoring good models in the top 10 as can be seen in Table S1, where for the manual submission of target 135 we only generated four medium- and one acceptable-quality model but when considering 43 models instead of 10 we are generating 11/23/4 high-/medium-/acceptable-quality models instead. This pattern holds true for almost all targets for which more than the top 10 models were evaluated indicating there is plenty of room for improvement in the scoring of our models.

Despite the fact targets 122 and 136 were template-based, we only achieved good performance for the latter, generating acceptable models for all three interfaces and medium quality ones for two of the three for the server submission. Target 136 is particularly interesting as its structure was determined with cryo-EM and it is the largest complex featured so far in CAPRI. Therefore, it was an excellent test case for our implementation of the MARTINI coarse grain forcefield into HADDOCK.³⁸ Using less than ideal information (interface restraints extracted from the

TABLE 1 Summary of the prediction and scoring results obtained by the HADDOCK server and manual team

Target	Target Category	Modeling strategy	Server ***/**/*	Manual ***/**/*	Scoring ***/**/*
121	Protein-Peptide	Information-driven	0/0/0	0/0/0	0/0/0
122	Protein-Protein	Template-based	0/0/0	0/0/0	0/1/0
123	Protein-Protein	Information-driven	0/0/0	0/0/0	0/0/0
124.1	Protein-Protein	Information-driven	0/0/0	0/0/0	0/0/0
124.2			0/0/0	0/0/0	0/0/0
125.1	Protein-Protein	Information-driven	0/0/4	0/0/0	0/4/0
125.2			10/0/0	7/1/2	7/3/0
125.3			0/0/0	0/0/0	0/0/0
125.4			0/0/0	0/0/0	0/0/0
126	Protein-Glycans	Information-driven	0/0/1	0/0/0	0/0/3
127			0/1/1	0/0/4	0/0/5
128			0/0/10	0/2/7	0/2/8
129			0/0/5	0/0/6	0/0/7
130			1/5/4	2/4/3	1/5/2
131	Protein-Protein	Information-driven	0/0/0	0/0/0	0/0/0
132	Protein-Protein	Information-driven	0/0/0	0/0/0	0/0/1
133	Protein-Protein	Information-driven	0/3/3	0/3/3	0/1/5
134	Protein-Peptide	Information-driven/Threading	0/0/1	0/1/1	0/2/0
135			0/0/5	0/4/1	0/0/0
136.1	Protein-Protein	Template-based	0/5/5	0/9/1	0/4/5
136.2			0/4/1	0/0/8	0/1/3
136.3			0/0/8	0/0/1	0/1/5

Note: The first column is the target and interface number as reported by CAPRI and the fourth through sixth columns refer to the number of high-(***)/medium-(**)acceptable-(*)quality models generated during the server, manual and scoring experiments, respectively, when considering the top 10 models submitted for evaluation.

top model of the all-atom, server submission), we were able to accurately recreate the full decamer using five dimers with interface restraints specified only between the first two adjacent ones (ie, simultaneous five-body docking) and symmetry restraints to recreate the rest of the complex, while at the same time massively reducing the computational cost from approximately 10-12 to 2 hours per model (5-6-fold speedup).

For the majority of the targets, we followed an information-driven strategy as we were able to identify some information about the putative interfaces of the target complexes that could aid the modeling process. Unfortunately, the information we uncovered was not enough to overcome the challenging nature of most of these targets (considering submissions from all participants, one acceptable-quality model for target 123, no acceptable-quality models for 124, one medium-quality model for 131, and one medium- and four acceptable-quality models for 132 when considering the top 10 models). We successfully modeled two of the four interfaces of target 125 (1 and 2), generating acceptable quality models for the first as well as high and medium quality for the second. We were not able to generate acceptable models for the other two interfaces which proved too challenging for all the participants as only one acceptable model was generated for the third interface and none for the fourth. We were able to generate good models for the last of the information-driven targets, 133, achieving the same

performance for server and manual predictions. Preliminary analysis of these targets (the crystal structures were not yet available at the time of writing) indicates that the limiting factor for these targets was poor modeling of the individual components of the complex rather than the docking itself.

The last targets were those of round 41 which featured glycans. These targets represent the first protein-small molecule complexes to be featured in CAPRI and HADDOCK performed well across all of them, generating acceptable models for all but the manual submission of target 126, and even generating highly accurate models for both the manual and server submission of target 130. The topologies and parameters for describing the glycans were automatically obtained from PRODRG,⁵² after modifying their numbering and naming to make them look as one single heteroatom residue.

Our scoring performance ranks us as one of the best performers of the scoring experiment when considering the number of targets for which acceptable models were identified with acceptable- or higher quality models in 14 of 22 interfaces. The robustness of our scoring protocol and function is also demonstrated in targets such as 132 for which we were able to pick out one of the 27 acceptable models in a pool of over 2000 models as well as the fact we could identify near-native models in all target categories. One limitation of our scoring protocol, which also

affects our CAPRI ranking, is the difficulty in distinguishing medium- or higher quality models from acceptable-quality ones.

4 | CONCLUSIONS

CAPRI rounds 38-45 featured many firsts such as the inclusion of a cryo-EM determined target in target 136 of round 45 and protein-small molecule (glycans) complexes in targets 126-130 of round 41. HADDOCK was able to generate near-native models for these systems as well as most of the traditional protein-protein systems that featured in the remaining targets, including the peptide ones. While it is evidently clear the entire docking community still has difficulties in dealing with systems for which no good templates are available, this is an issue that seems to be affecting our efforts in CAPRI for the remaining targets as well: as it appears poor initial modeling of the components rather than the docking itself is affecting our performance. The second limiting factor is our scoring performance during the docking experiments (manual and server). As can be seen in Table S1, for every target for which more than the top 10 models have been evaluated, there are many higher quality models that we are failing to rank in the top 10. This creates a contradiction with our performance in the scoring experiment for which we can reliably identify near-native models in a pool of thousands and indicates that there is room to re-optimize our scoring function. Development and testing of the next major version of HADDOCK (v2.4) is currently ongoing and we expect that many new features and small improvements will allow us to both generate higher quality models as well as rank them more accurately.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Cristina Paissoni for contributions she made for target 133.

CONFLICT OF INTERESTS

Dr. A. M. J. J. Bonvin is a member of the CAPRI management committee but has no access to target information.

ORCID

Cunliang Geng  <https://orcid.org/0000-0002-1409-8358>

Alexandre M. J. J. Bonvin  <https://orcid.org/0000-0001-7369-1322>

REFERENCES

- Vakser IA, Deeds EJ. Computational approaches to macromolecular interactions in the cell. *Curr Opin Struct Biol.* 2019;55:59-65. <https://doi.org/10.1016/j.sbi.2019.03.012>.
- Ganesan A, Coote ML, Barakat K. Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discov Today.* 2017; 22(2):249-269. <https://doi.org/10.1016/j.drudis.2016.11.001>.
- Śledź P, Caflisch A. Protein structure-based drug design: from docking to molecular dynamics. *Curr Opin Struct Biol.* 2018;48:93-102. <https://doi.org/10.1016/j.sbi.2017.10.010>.
- Cournia Z, Allen B, Sherman W. Relative binding free energy calculations in drug discovery: recent advances and practical considerations. *J Chem Inf Model.* 2017;57(12):2911-2937. <https://doi.org/10.1021/acs.jcim.7b00564>.
- Gathiaka S, Liu S, Chiu M, et al. D3R grand challenge 2015: evaluation of protein-ligand pose and affinity predictions. *J Comput Aided Mol Des.* 2016;30(9):651-668. <https://doi.org/10.1007/s10822-016-9946-8>.
- Gaieb Z, Liu S, Gathiaka S, et al. D3R grand challenge 2: blind prediction of protein-ligand poses, affinity rankings, and relative binding free energies. *J Comput Aided Mol Des.* 2018;32(1):1-20. <https://doi.org/10.1007/s10822-017-0088-4>.
- Janin J, Henrick K, Moult J, et al. CAPRI: a critical assessment of predicted interactions. *Proteins Struct Funct Genet.* 2003;52(1):2-9. <https://doi.org/10.1002/prot.10381>.
- Méndez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures. *Proteins.* 2005;60(2):150-169. <https://doi.org/10.1002/prot.20551>.
- Janin J. Assessing predictions of protein-protein interaction: the CAPRI experiment. *Protein Sci.* 2005;14(2):278-283. <https://doi.org/10.1110/ps.041081905>.
- Lensink MF, Méndez R, Wodak SJ. Docking and scoring protein complexes: CAPRI 3rd edition. *Proteins.* 2007;69(4):704-718. <https://doi.org/10.1002/prot.21804>.
- Lensink MF, Wodak SJ. Docking and scoring protein interactions: CAPRI 2009. *Proteins.* 2010;78(15):3073-3084. <https://doi.org/10.1002/prot.22818>.
- Lensink MF, Wodak SJ. Docking, scoring, and affinity prediction in CAPRI. *Proteins.* 2013;81(12):2082-2095. <https://doi.org/10.1002/prot.24428>.
- Lensink MF, Velankar S, Wodak SJ. Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition. *Proteins.* 2017;85(3): 359-377. <https://doi.org/10.1002/prot.25215>.
- Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J Am Chem Soc.* 2003;125(7):1731-1737. <https://doi.org/10.1021/ja026939x>.
- van Zundert GCP, Rodrigues JPGLM, Trellet M, et al. The HADDOCK2.2 web server: user-friendly integrative Modeling of biomolecular complexes. *J Mol Biol.* 2016;428(4):720-725. <https://doi.org/10.1016/j.jmb.2015.09.014>.
- Trellet M, Melquiond ASJ, Bonvin AMJJ. A unified conformational selection and induced fit approach to protein-peptide docking. *PLoS One.* 2013;8(3):e58769. <https://doi.org/10.1371/journal.pone.0058769>.
- Berman H, Henrick K, & Nakamura H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology.* 2003;10 (12):980. <https://doi.org/10.1038/nsb1203-980>.
- Xue LC, Rodrigues JPGLM, Dobbs D, Honavar V, Bonvin AMJJ. Template-based protein–protein docking exploiting pairwise interfacial residue restraints. *Brief Bioinform.* 2016;18(3):458-466, bbw027. doi: <https://doi.org/10.1093/bib/bbw027>
- Zimmermann L, Stephens A, Nam S-Z, et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol.* 2018;430(15):2237-2243. <https://doi.org/10.1016/j.jmb.2017.12.007>.
- Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993;234(3):779-815. <https://doi.org/10.1006/jmbi.1993.1626>.
- Leaver-Fay A, Tyka M, Lewis SM, et al. Rosetta3. In: Johnson ML, Brand L, eds. *Methods in Enzymology.* Vol 487. Methods in Enzymology. Academic Press; New York 2011:545-574. doi:<https://doi.org/10.1016/B978-0-12-381270-4.00019-6>
- Källberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012;7(8):1511-1522. <https://doi.org/10.1038/nprot.2012.085>.
- Wang S, Li W, Liu S, Xu J. RaptorX-property: a web server for protein structure property prediction. *Nucleic Acids Res.* 2016;44(W1):W430-W435. <https://doi.org/10.1093/nar/gkw306>.

24. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc.* 2010;5(4):725-738. <https://doi.org/10.1038/nprot.2010.5>.
25. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: protein structure and function prediction. *Nat Methods.* 2015;12(1):7-8. <https://doi.org/10.1038/nmeth.3213>.
26. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* 2015;43(W1):W174-W181. <https://doi.org/10.1093/nar/gkv342>.
27. Vangone A, Rodrigues JPGLM, Xue LC, et al. Sense and simplicity in HADDOCK scoring: lessons from CASP-CAPRI round 1. *Proteins.* 2017;85(3):417-423. <https://doi.org/10.1002/prot.25198>.
28. Vangone A, Rodrigues JPGLM, Xue LC, et al. Sense and simplicity in HADDOCK scoring: lessons from CASP-CAPRI (page 418). *Proteins.* 2017;85(8):1589-1590. <https://doi.org/10.1002/prot.25339>.
29. Jorgensen WL, Tirado-Rives J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc.* 1988;110(6):1657-1666. <https://doi.org/10.1021/ja00214a001>.
30. Fernández-Recio J, Totrov M, Abagyan R. Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol.* 2004;335(3):843-865. <https://doi.org/10.1016/j.jmb.2003.10.069>.
31. Schrödinger, LLC. *The {PyMOL} Molecular Graphics System, Version~1.8.;* 2015.
32. Beaufays J, Lins L, Thomas A, Brasseur R. In silico predictions of 3D structures of linear and cyclic peptides with natural and non-proteinogenic residues. *J Pept Sci.* 2012;18(1):17-24. <https://doi.org/10.1002/psc.1410>.
33. Cassel Eoin. *Pseudomonas aeruginosa TolA-TolB interaction* Eoin Cassels, Doctor of Philosophy, University of York, Department of Biology December 2012. 2012; (December).
34. Ye Y, Godzik A. Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics.* 2003;19(suppl 2):ii246-ii255. <https://doi.org/10.1093/bioinformatics/btg1086>.
35. Ye Y, Godzik A. FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Res.* 2004;32 (Web Server):W582-W585. <https://doi.org/10.1093/nar/gkh430>.
36. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tielemans DP, Marrink S-J. The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput.* 2008;4(5):819-834. <https://doi.org/10.1021/ct700324x>.
37. de Jong DH, Singh G, Bennett WFD, et al. Improved parameters for the Martini coarse-grained protein force field. *J Chem Theory Comput.* 2013;9(1):687-697. <https://doi.org/10.1021/ct300646g>.
38. Roel-Touris J, Don CG, V. Honorato RV, Rodrigues JPGLM, Bonvin AMJJ. Less is more: coarse-grained integrative modeling of large biomolecular assemblies with HADDOCK. *J Chem Theory Comput.* 2019;15(11):6358-6367. doi:<https://doi.org/10.1021/acs.jctc.9b00310>
39. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 2004;32(Web Server):W526-W531. <https://doi.org/10.1093/nar/gkh468>.
40. Kolenko P, Rozbeský D, Vaněk O, et al. Molecular architecture of mouse activating NKR-P1 receptors. *J Struct Biol.* 2011;175(3):434-441. <https://doi.org/10.1016/j.jsb.2011.05.001>.
41. Kolenko P, Rozbeský D, Vaněk O, Bezouška K, Hašek J, Dohnálek J. Structure of the H107R variant of the extracellular domain of mouse NKR-P1A at 2.3 Å resolution. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2011;67(12):1519-1523. <https://doi.org/10.1107/S1744309111046203>.
42. Skálová T, Bláha J, Harlos K, et al. Four crystal structures of human LLT1, a ligand of human NKR-P1, in varied glycosylation and oligomerization states. *Acta Crystallogr Sect D Biol Crystallogr.* 2015;71(3):578-591. <https://doi.org/10.1107/S1399004714027928>.
43. Kamishikiryō J, Fukuhara H, Okabe Y, Kuroki K, Maenaka K. Molecular basis for LLT1 protein recognition by human CD161 protein (NKR-P1A/KLRB1). *J Biol Chem.* 2011;286(27):23823-23830. <https://doi.org/10.1074/jbc.M110.214254>.
44. Li Y, Wang Q, Chen S, Brown PH, Mariuzza RA. Structure of NKp65 bound to its keratinocyte ligand reveals basis for genetically linked recognition in natural killer gene complex. *Proc Natl Acad Sci.* 2013;110(28):11505-11510. <https://doi.org/10.1073/pnas.1303300110>.
45. Kita S, Matsubara H, Kasai Y, et al. Crystal structure of extracellular domain of human lectin-like transcript 1 (LLT1), the ligand for natural killer receptor-P1A. *Eur J Immunol.* 2015;45(6):1605-1613. <https://doi.org/10.1002/eji.201545509>.
46. Javaheri A, Kruse T, Moonens K, et al. *Helicobacter pylori* adhesin HopQ engages in a virulence-enhancing interaction with human CEACAMs. *Nat Microbiol.* 2017;2(1):16189. <https://doi.org/10.1038/nmicrobiol.2016.189>.
47. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. *J Comput Chem.* 2005;26(16):1701-1718. <https://doi.org/10.1002/jcc.20291>.
48. Pronk S, Páll S, Schulz R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 2013;29(7):845-854. <https://doi.org/10.1093/bioinformatics/btt055>.
49. Song Y, DiMaio F, Wang RY-R, et al. High-resolution comparative modeling with RosettaCM. *Structure.* 2013;21(10):1735-1742. <https://doi.org/10.1016/j.str.2013.08.005>.
50. Omega Toolkit 2.6.4 OpenEye Scientific Software, Santa Fe, NM, USA. 2017.
51. Hawkins PCD, Skillman AG, Warren GL, Ellingson BA, Stahl MT. Conformer generation with OMEGA: algorithm and validation using high quality structures from the protein databank and Cambridge structural database. *J Chem Inf Model.* 2010;50(4):572-584. <https://doi.org/10.1021/ci100031x>.
52. Schüttelkopf AW, van Aalten DMF. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr Sect D Biol Crystallogr.* 2004;60(8):1355-1363. <https://doi.org/10.1107/S0907444904011679>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Koukos PI, Roel-Touris J, Ambrosetti F, et al. An overview of data-driven HADDOCK strategies in CAPRI rounds 38-45. *Proteins.* 2020;1-8. <https://doi.org/10.1002/prot.25869>