RESEARCH ARTICLE

Proteins: Structure, Function and Bioinforr
DOI 10.1002/prot.25870

Modeling protein-protein, protein-peptide and protein-oligosaccharide complexes: CAPRI 7th edition

Marc F. Lensink[1*], Nurul Nadzirin[2], Sameer Velankar[2] and Shoshana J. Wodak[3]

[1] *University of Lille, CNRS UMR8576 UGSF, Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France*

[2] *European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom*

[3] *VIB-VUB Structural Biology Research Center, Pleinlaan 2, 1050 Brussels, Belgium*

---

* Corresponding author; E-mail: marc.lensink@univ-lille.fr

Keywords: CAPRI, protein docking, protein-protein interaction, protein-peptide interaction, protein-polysaccharide interaction, protein complexes, protein assemblies, blind prediction, docking

Received: Nov 13, 2019;Revised: Dec 19, 2019;Accepted: Dec 26, 2019

ABSTRACT

We present the seventh report on the performance of methods for predicting the atomic resolution structures of protein complexes offered as targets to the community-wide initiative on the Critical Assessment of Predicted Interactions (CAPRI). Performance was evaluated on the basis of 36114 models of protein complexes submitted by 57 groups – including 13 automatic servers – in prediction Rounds held during the years 2016-2019 for 8 protein-protein, 3 protein-peptide, and 5 protein-oligosaccharide targets with different length ligands. Six of the protein-protein targets represented challenging hetero-complexes, due to factors such as availability of distantly related templates for the individual subunits, or for the full complex, inter-domain flexibility, conformational adjustments at the binding region, or the multi-component nature of the complex. The main challenge for the protein-peptide and protein-oligosaccharide complexes was to accurately model the ligand conformation and its interactions at the interface. Encouragingly, models of acceptable quality, or better, were obtained for a total of 6 protein-protein complexes, which included 4 of the challenging hetero-complexes and a homo-decamer. But fewer of these targets were predicted with medium or higher accuracy. High accuracy models were obtained for 2 of the 3 protein-peptide targets, and for one of the protein-oligosaccharide targets. The remaining protein-sugar targets were predicted with medium accuracy. Our analysis indicates that progress in predicting increasingly challenging and diverse types of targets is due to closer integration of

template-based modeling techniques with docking, scoring and model refinement procedures, and to significant incremental improvements in the underlying methodologies.

INTRODUCTION

Proteins carry out their function by interacting with other proteins, and with macromolecular components such as DNA or RNA [1]. The disruption or deregulation of these interactions often leads to disease [2,3]. Charting these interactions and characterizing their biochemical, biophysical, and structural properties therefore remains an important goal in molecular biology and medicine.

Experimental and computational techniques have played a decisive role in pursuing this goal. Proteomics techniques have been providing increasingly more comprehensive descriptions of the protein interaction landscape of specific cellular processes, of cell types, and entire organisms [4,5]. These descriptions yield information on the identities of the interacting proteins, and the compositions of multi-protein assemblies, or complexes, but offer little information on the stoichiometry of the interactions, or their molecular details.

Over the last 5 decades, structural biology techniques such as X-ray crystallography and Nuclear Magnetic Resonance (NMR) have been very effective in charting the 3D structure repertoire of individual proteins, with the data deposited in the PDB (Protein DataBank) [6]. But so far these techniques have been slow in providing information on the 3D structures of macromolecular complexes, in particular those of larger assemblies that are active in the cell and can be detected by modern proteomics and other methods [7,8]. But this is changing rapidly

thanks to recent spectacular advances in single molecule cryo-EM techniques, specifically geared at determining the structure of large macromolecular assemblies to atomic resolution [9,10].

Computational methods have been playing an important role in efforts to populate the uncharted landscape of protein assemblies, by exploiting the rapidly increasing body of data on protein sequences and three-dimensional structures of individual proteins and their complexes stored in public databases [11]. In recent years these methods are often cast into a wider framework of the so-called 'hybrid' or 'integrative' modeling techniques [12,13], whereby larger macromolecular assemblies are modeled by integrating sequence information and structural data at various levels of resolution, with various other types of data.

An important role in these recent developments is played by methods for modeling the atomic structure of macromolecular assemblies, starting from the structures of the individual components of these assemblies. These are the so-called ''docking'' algorithms and the associated scoring schemes, the energetic criteria for singling out stable binding modes [14-16]. A key factor fueling progress in docking algorithm and more generally in methods for modeling protein assemblies has been the Critical Assessment of PRedicted Interactions (CAPRI) (http://pdbe.org/capri/; http://www.capri-docking.org/). CAPRI is a community-wide initiative established in 2001, offering computational biologists the opportunity to test the

performance of their computational procedures in blind predictions of experimentally determined 3D structures of protein complexes, the 'targets', provided to CAPRI prior to publication. CAPRI initially focused on targets representing protein-protein complexes. Over the years its focus broadened by including targets representing complexes of proteins with other molecules (including peptides, nucleic acids and sugars) [17], tackling the problem of predicting protein binding affinities [18,19] and modeling the positions of water molecules in protein-protein interfaces [20].

Since CAPRI's inception the task of modeling the atomic structure of protein complexes has likewise evolved. Initially, this task involved mainly classical docking procedures, which sample and score putative binding poses of two or more proteins using as input the known unbound structures of the individual interacting components [see ref [14]]. With the growing ease with which structural templates can be found in the PDB, docking calculations now routinely accept as input homology-built models of individual protein components of an assembly, with a growing rate of success [21,22]. Furthermore, it is not uncommon to find templates for the full protein assembly. This is most often the case for assemblies of identical subunits (homo-dimers, or higher order homo-oligomers), because closely related proteins tend to adopt the same assembly mode (or oligomeric state) [23,24]. In such instances, classical docking calculations may no longer be required because the protein assembly can be modeled directly from the template, a task also called 'template-based docking' [11,25,26].

Methods for predicting the 3D structures of proteins *ab-initio* – from sequence information alone and in absence of available templates – have also come a long way. By leveraging the information on protein sequences, it is now possible to predict residues that make contacts in the protein 3D structure and therefore facilitate its prediction [27-29]. Also, information on structure and sequence features of proteins (see for examples [30-32]) is being exploited much more efficiently, thanks to new developments in Artificial Intelligence Deep Learning techniques [33,34], enabling the prediction of the 3D structure of proteins from sequence information alone, as demonstrated recently in the *ab-initio* structure prediction challenge in CASP (Critical Assessment of protein Structure Prediction) [35]. As these methods progress, one would hope that they will ultimately be successfully integrated with assembly prediction procedures, and efforts in this direction have been undertaken by a close collaboration between CAPRI and CASP [22,36-38].

Here, we present the prediction results for the 16 targets of CAPRI Rounds 38-45 that took place during the years 2016-2018. These targets comprised protein-protein, protein-peptide and protein-oligosaccharide complexes. Protein-peptide complexes are a growing focus of CAPRI given the important role the recognition of short peptide motifs plays notably in regulatory processes [39,40], whereas protein-sugar complexes represent an important category of complexes where method developments are needed. Prediction results for these targets

were presented at the 7th CAPRI evaluation meeting held in April 2019 at the EBI (the European Bioinformatics Institute) (see https://bioexcel.eu/7th-capri-evaluation-meeting/). An additional prediction Round (Round 46) was run in collaboration with CASP during the CASP13 prediction challenge in the summer of 2018, with results described in a separate publication [38].

The results reported here were analyzed on the basis of a total of 36114 evaluated models, submitted by 57 participating groups including 13 automatic servers. The challenges posed by the different targets are discussed and the extent to which current modeling procedures were able to address them is evaluated. This evaluation is furthermore used to assess emerging trends in modeling protein complexes, and to evaluate progress in the field.

THE CAPRI TARGETS OF ROUNDS 38-45

Information on the 16 CAPRI targets of Rounds 38-45 for which predictions are evaluated here is summarized in **Table I.** These targets comprise 8 protein-protein complexes, 3 protein-peptide complexes and 5 protein-oligosaccharide complexes. For all these complexes the prediction task involved modeling the 3D conformation of the different components (protein subunits, bound peptides or oligosaccharides) as well as the detailed binding mode between the components. The level of difficulty of these targets hence depended on the availability of templates for the individual subunits, or in some cases, for the hetero-complex as a whole, and on the ability to identify the correct binding mode and optimize the corresponding interface.

Protein-protein complexes (T122-125, T131-T133, T136)

This category of targets comprised 7 hetero-complexes, including 1 heterotrimer, 5 heterodimers and one homo-decamer (depicted in **Figure 1**). In the following they will be described in the order they were offered in CAPRI Rounds.

The first 3 targets **T122, T123, T124** were offered in Round 39. **T122** was the x-ray structure of the human cytokine-receptor heterotrimer complex (IL23/IL23R) determined at 2.8 Å resolution, contributed by *Savvas Savvides, Ghent University, Belgium* [41]. The stoichiometry of the complex was experimentally verified to consist of 1 receptor monomer (IL23R) comprising 330 residues bound to 1 cytokine heterodimer (IL23A/IL12B) (comprising respectively 198

and 328 residues). In the 3D structure of the complex the receptor binds both chains of IL-23, forming interfaces that bury a total of 2330 Å^2. This was in principle a medium difficulty target. Templates adopting somewhat different conformations were available for the receptor, (e.g. PDB code 2d9q, displaying 3.78 Å backbone rmsd relative to the bound receptor), and an unbound structure for the IL23 cytokine heterodimer was provided by the authors (later released as PDB code 5mxa; with a 3.59 Å backbone rmsd relative to the bound structure). Nevertheless, the inter domain flexibility of the IL12B moiety and the IL23 receptor represented a challenge for the docking procedures.

T123 and T124, offered by *Alain Roussel, Université Aix Marseille, France*, were X-ray structures of the complexes between, respectively, the N-terminal (174 residues) and C-terminal (202 residues) domains of PorM, a component of the type 9 secretion system (T9SS), from bacteria (*Porphyromonas gingivalis*), and Camelid nanobodies: nb-02 (121 residues) and nb-130 (141 residues). These complexes, determined at the resolution of 2.5 Å (PorM-Nt/nb-02) and 2.15 Å (PorM-Ct/nb-130) [42], featured reasonable size interfaces of respectively 1240 Å^2 and 1120 Å^2. These were difficult targets because although several good templates were available for the nanobodies, none were available for either of the PorM domains. The PorM-Nt domain is helical and adopts an up-down fold, but was missing the N-terminal trans-membrane helix, whereas the PorM-Ct domain is a domain-swapped dimer, containing predominantly β-strands [42]. **T124** was offered as two challenges, separated into the PorM-Ct

dimeric structure, and the complex between a PorM-Ct monomer with the nanobody molecule. To predict the structure of these complexes the PorM domains had to be modeled *ab-initio*, a challenging task with which CAPRI participant have very limited expertise. The PorM-Ct dimer was offered as target T0907 in the CASP12 structure prediction Round [43,44], but no acceptable structures were produced then.

T125, offered in Round 40 was the multi-protein assembly of the extracellular domain of the Lectin-like transcript 1 (LLT1) (protein A, 135 residues) with the extracellular domain of its inhibitor NKR-P1 (protein B, 146 residues). The structure was determined at 1.9 Å resolution by *Ondřej Vaněk, Charles University, Prague, Czech Republic* (PDB code 5mgt). According to the authors, the stoichiometry of the complex is A2B4, e.g. a homodimer of protein A, bound to 2 homodimers of protein B, forming a total of 5 distinct interfaces sized between 440 and 775 $Å^2$ of buried area (**Table I**). Two interfaces are those of the A2 and B2 homodimers, respectively, two other interfaces are formed between LLT1 and NKR-P1 proteins and a fifth small contact is formed between the NKR-P1 proteins (see **Figure 1**). This stoichiometry was however questioned by predictor groups, based on an analysis of various interfaces of the target structure (PDB code 5mgt) (see Dapkunas *et al.* this issue). A good quality template was available for the LLT1 homodimer, and a more distantly related template was available for the LLT1-NKR-P1 hetero-dimer. The main difficulty of this target likely resided in building the full assembly, given the uncertainty concerning the announced stoichiometry of this target.

The main interfaces of this target, i.e. the A2, B2, and AB interfaces, were therefore assessed separately.

T131 and **T132,** the next two protein-protein complexes offered in Round 42, were x-ray structures of host-pathogen complexes of the cell adhesion protein CEACAM1 from human (hCEACAM1 N-terminal domain; 108 residues), bound to the cell adhesion proteins of *Helicobacter pylori*, HopQ Type I (HopQ-I; 404 residues), and HopQ Type II (HopQ-II; 418 residues), respectively. These structures were solved to 2.8 Å and 2.59 Å resolution by *Han Remaut, Free University of Brussels, and the Flemish Institute for Biotechnology, Brussels Belgium* [45]. The HopQ-I and HopQ-II are rather closely related (sequence ID ~60%), and engage in similar binding modes with the CEACAM1 protein. However HopQ-I features four loops, which contribute to the binding interface with the host protein, without altering the binding location. Excellent templates were available for the HopQ components (displaying backbone rmsd values as low as 1.0 Å, relative to HopQ-I). But none include the extra loop of HopQ-I. Excellent templates were also available for the host protein domain (rmsd below 1 Å). **T131** (HopQ-I/CEACAM complex) was a more challenging target than **T132**, since the missing loops had to be modeled.

T133, offered in Round 43, was a particularly interesting case. It represents a redesigned version of the *wt* Colicin E2/DNase-Im2 complex from *E.coli* (PDB code 3u43), which was

previously a CAPRI target (T47) [17]. The redesigned version, contributed by *Sarel Fleishman and Ravit Netzer from the Weizmann Institute, Israel, is* denoted as E^{des3}/Im^{des3} (referring to a specific version of the designed complex, where both components have been redesigned). Its structure was determined at high resolution [46] and shown to contain two complexes per asymmetric unit, which displayed only marginal differences. However, residues at the interface, and residue-residue contacts of both versions differed more relative to those of the *wt* complex (possibly as a result of a small rigid body rotation). The authors also indicated that the affinity of the redesigned complex was 73 nM (K_d), and that the latter complex displays high specificity relative to 'chimeric' versions, which contain one *wt* component, e.g. E^{des3}/Im2 and E2/Im^{des3}, which display 3 orders of magnitude lower affinity, and could therefore be used as controls for 'non-binders'.

Considering the relatively small differences between the designed and *wt* complexes, this was not a classical CAPRI prediction problem, because the challenge here was not to identify the correct association mode, which was well known, but to correctly identify the structural perturbation relative to the *wt* complex, which resulted in the gain in affinity.

T136, offered in Round 45, was the lysine decarboxylase LdcA (751 residues), from *Pseudomonas aeruginosa*, a homo-decamer adopting D5 symmetry, determined by cryo-EM to 4.5 Å resolution. The target was contributed by *Irina Gutsche, Institut de Biologie Structurale,*

Grenoble, France (personal communication), and was the first CAPRI target representing a cryo-EM multi-protein complex. The homo-decamer is composed of 2 rings, comprising 5 subunits each. These rings are packed on top of each other, such that subunits interact between rings, as well as within rings. This arrangement displays 3 distinct interfaces: a very large interface (5600 Å^2 of buried area) involving interactions between the subunits within and between rings, and two smaller interfaces of 1100 Å^2 and 800 Å, respectively, between subunits mainly within rings. This was in principle as easy target, since a 2.4 Å resolution x-ray structure of a related enzyme, the arginine decarboxylase AdiA from *E. coli* (PDB code 2vyc; backbone rmsd 2.33 Å; sequence identity, 40%) was available, and could therefore be used as template for building the structure of the full target complex.

Protein-peptide complexes (T121, T134, T135)

The protein-peptide complexes of targets **T134** and **T135** are depicted in **Figure 2**; **T121** cannot be displayed at this time due to confidentiality reasons. The amino acid sequences of the peptides ligands of all 3 targets are provided in **Figures 2** and **7.**

With 8 protein-peptide complexes offered as targets in previous CAPRI Rounds [21], and the important role these complexes play in regulating many cellular processes, the CAPRI community has been working hard to meet the particular challenge of accurately predicting the 3D structure of such complexes. Protein-peptide complexes are indeed significantly different from protein-protein complexes. Their binding interface is usually much smaller, not exceeding 1000 Å^2, reflecting the lower affinity of the corresponding associations. In addition,

the bound peptides tend to retain a significant degree of conformational flexibility, such that often only a portion of the peptide becomes structured upon binding and the remaining residues remain too flexible to be visible in the X-ray structure. Since the protein moieties in these complexes generally undergo small conformational changes upon peptide binding [47], the main challenge for predictors is to accurately model the bound conformation of the peptide and the corresponding protein-peptide interactions, including the side-chain conformations at the interface. This in turn required tightening the parameters used by the assessors to define model accuracy levels, as described below.

Target T121, offered in Round 38, was the complex between the TolAIII domain, a small C-terminal domain from the periplasmic protein TolA of *P. aeruginosa* (115 residues), and the 13 residue N-terminal peptide of *P. aeruginosa* TolB. The solution structure, solved by NMR, was contributed by *Colin Kleanthous, University of Oxford UK* (personal communication). The peptide used for the structure determination had an unblocked N-terminus and the C-terminus blocked with an amide group. This was the most difficult target of this category. While an unbound x-ray structure was available for the TolAIII domain (PDB code 1lr0, backbone rmsd 1.7 Å relative to the target NMR average conformation), the buried interface in this solution structure was quite extensive for this type of targets (1400 $Å^2$). Correctly predicting the conformation of the otherwise flexible peptide and its interactions with the globular domain was therefore challenging. Regrettably, not much else can be revealed about

this target at this point, since the target structure is still on-hold at the PDB, and cannot be described in any detail at this time.

Targets T134, T135, the two other protein-peptide complexes, were offered as targets in Round 44. These targets are complexes of the Dynein Light Chain subunit 8 (DLC8) from human (a homodimer) with segments from the cognate L-MAG protein (**Figure 2**), and were contributed by *Petri Kursula, University of Bergen Norway* [48]. The two complexes were crystallized starting from, respectively, a 50-residue segment of the cognate protein, and a 12-residue peptide from the same segment, with the resulting complexes adopting very similar structures, burying respectively 1540 Å^2 and 1640 Å^2 in the protein-peptide interface. For **T134** the challenge was two-fold. Predictors were given the sequence of the full 50-residue segment and asked to predict the 12-residue peptide of this segment that actually binds DLC8, and to model the resulting complex. For **Target 135**, predictors were given the sequence of the 12-residue peptide whose DLC8-bound structure was determined. The main challenge for the latter target was to produce an accurate model for the protein-peptide complex. Overall these were easy targets, given the conserved binding mode of diverse peptides to the DLC8 homodimer [49] in structures available in the PDB. Furthermore, the DLC8 component of the target was the same as that of the PDB entry 1f3c (100% sequence identity).

Protein-oligosaccharide complexes (T126-130)

CAPRI Round 41 offered 5 targets representing protein-oligosaccharide complexes, which were contributed by *Shifra Lansky and Gil Shoham, Hebrew University of Jerusalem*. Four of these targets were complexes of the same protein the arabino-oligosaccharide-binding protein AbnE (an ABC transporter-bound periplasmic protein from *Geobacillus stearothermophilus*, involved in arabinose import), bound to arabino-oligosaccharide chains of decreasing lengths. These chains were A6 (arabinohexaose, 6 sugar units) for **T126** through A3 (arabinotriose, 3 sugar units) for **T129**. The fifth target, **T130**, was the complex of a mutant (E201A) of the intracellular arabinanase AbnB, an enzyme involved in the degradation of intra-cellular arabino-oligosaccharides, bound to A5 (arabinopentaose, 5 sugar units). This was the second time protein-oligosaccharide complexes were offered as targets in CAPRI. The first such target was T57, a complex between the bacterial surface protein Bt4661 and a fragment of heparin comprising 6 sugar units, offered in CAPRI Round 27, over 6 years ago [17].

A number of templates were available for all 5 targets. Templates representing both the apo (3CU9) and holo forms of the AbnB enzyme were available for **T130**. The holo forms included two mutant enzymes, one (D147A) bound to arabinobiose (PDB code 3d61) and another (E201A) mutant bound to arabinotriose, but none were found to bind longer oligosaccharides, such as the arabinopentaose ligand of **T130**. For **T126-T129**, structures of AbnE homologs bound to various ligands with 6-membered sugar rings were available. These ligands included

trigalacturonid acid, galactose, mannitol, tagatopyranose, glucose, but none featured the 5-membered sugar rings of arabinose.

These templates were clearly helpful in guiding the challenging task of docking and modeling the association modes of these flexible sugar ligands, but only to a point, given that they bind different sugar ring moieties. On the other hand, valuable insights into the binding modes of arabinose ligands of different sizes were ultimately obtained from analyzing the experimental structures of **T126-T129**, which cannot be displayed at this time due to confidentiality reasons. Examining these structures revealed that arabinohexaose occupies binding sites 1-6 of the enzyme, arabinopentaose occupies sites 1-5 and arabinotetraose binds to sites 1-4. Furthermore, the A4-A6 sugars in **T126-T128** were seen to occupy the same 4 binding pockets (sites 1-4), whereas the A5 and A6 sugars occupied in addition the same pocket at site 5. This in turn suggests that the binding of any additional arabinose ring can be readily modeled by attaching it to the bound structure of the shorter sugar, or vice versa, namely, that the binding mode of a shorter sugar can be modeled by deleting the extra sugar ring (s) from the bound structure of the longer version. But the latter didn't seems to apply to **T129**, as shortening the sugar ligand below A4 seems to brake this 'rule' by leading to a shift within the enzyme binding pocket. Indeed the arabinotriose (A3) ligand of **T129** occupies binding sites 2-4 and not 1-3. Interestingly, the sugar rings bound to sites 2-4 feature lower atomic fluctuations (lower crystallographic B-factors) than the arabinose rings bound to the other

sugar binding sites on the enzyme, which may be indicating tighter binding at this site. Details of the 5 protein–oligosaccharide complexes can unfortunately not be displayed, since the structures of the 5 complexes have not been released at this time.

TARGET AND PARTICIPATION STATISTICS

A total of 57 groups participated in the 8 CAPRI Rounds evaluated here (Rounds 38-45).

The overall participation statistics for the 8 prediction rounds, including scorer and server participation are summarized in **Table II**. The target participation statistics and the number of models submitted for the 8 prediction rounds are summarized in **Table S1** of the **Supplementary Material**. It is noteworthy that 13 automatic servers, a larger number than in previous assessments [21,22], participated in Rounds 38-45. Nearly half of these servers submitted predictions for all targets – including the protein-peptide and protein sugar complexes – and 5 servers also submitting scoring predictions for certain targets (see **Table S2** of the **Supplementary Material**).

Starting in Round 39 (target **T122** onwards), CAPRI participants were requested to submit up to 100 models by the prediction deadline directly to the CAPRI-EBI website. This eliminated the extra step of uploading 100 models to the Lille site, as done previously. Predictors and servers were nevertheless asked to rank the first 10 models of their set, which were assessed

as usual, whereas all the 100 models were pooled with those submitted by other participants for each target and offered for the scoring experiment, after anonymizing their origins, as usual [17].

Overall, a total of 36114 models were evaluated for the 8 CAPRI Rounds. Depending on the target, 28-35 predictor groups submitted models. As previously, a smaller number of groups (16-20) submitted models for the scoring predictions. Overall 4835 and 2830 models were assessed for predictors and scorers, respectively. Scoring Rounds were offered for all targets, including the 3 protein-peptide and 5 protein-oligosaccharide complexes. This is the first time scoring Rounds have been offered for these categories of targets.

ASSESSMENT CRITERIA AND PERFORMANCE RANKING

Distinct assessment criteria were applied for the different categories of targets.

The classical CAPRI evaluation and ranking protocol [50,51] was applied to models submitted for targets of hetero- and homo- protein-protein complexes. This protocol was complemented with the DockQ score [52], a continuous quality metric that integrates the main quality measures of the standard CAPRI protocol, as detailed below. Modified protocols were used to evaluate the protein-peptide and protein-sugar complexes.

The quality of the modeled 3D structure of individual protein subunits was also evaluated by computing the 'molecular' root mean square deviation, M-rms, of backbone atoms of the model versus the target. It was used mainly to gauge the influence of the quality of subunit models on the predicted structure of the assembly.

Assessment of protein-protein complexes

The classical CAPRI evaluation and ranking protocol uses three main parameters, *L_rms*, *i_rms* and *f(nat)*, to derived the quality score of a predicted model [50,51]. *f(nat)* represents the fraction of native contacts in the target that is recalled in the model. Atomic contacts below 3 Å are considered clashes and predictions with too many clashes are disqualified (for the definition of native contacts, and the threshold for clashes see ref [50]. *L_rms* represents the backbone rmsd (root means square deviation) over the common set of ligand residues after the receptor proteins have been superimposed, and *i_rms* represents the backbone rmsd calculated over the common set of interface residues after the structural superposition of these residues. An interface residue is defined as such, when any of its atoms (hydrogen atoms excluded) are located within 10 Å of any of the atoms of the binding partner. On the basis of the values of these 3 parameters models are ranked into 4 categories of quality scores: high quality, medium quality, acceptable quality and incorrect (as previously described [36] and listed in **Table III**).

Considering the uncertainty associated with the stoichiometry of the hetero-hexamer target **T125**, prediction results for the three main interfaces of this target, the homodimeric LLT1 and NKR-P1, T125.2 and T125.4, respectively, and the two heterodimeric LLT1/NKR-P1 interfaces, T125.1 and T125.3, were evaluated independently. For **T124**, the PorM-Ct homodimer interface and the binding interface of the PorM-Ct monomer with the nanobody were considered as representing separate modeling problems and were therefore evaluated separately. For the homo-decamer target (**T136**) with 3 distinct interfaces, the quality score (or CAPRI category) for the assembly as a whole was taken as the score of the best-predicted individual interface for the assembly. This criterion was used in evaluating models in the recent CASP13-CAPRI challenge [38] and is more lenient than criteria applied in earlier Rounds, where the score for the entire assembly was taken as that of the worst predicted interface. Schemes of intermediate leniency, representing linear combinations of weighted scores for individual interfaces of the assembly were also tested, but found not to affect the performance ranking for these 3 targets.

Assessment of protein-peptide complexes

In protein-peptide complexes the interfaces are usually small. Only a portion of the bound peptide makes specific interactions with the protein, whereas the remainder of the peptide often remains very flexible. Therefore, for the predicted interface to be sufficiently informative, its accuracy needs to be higher than that of a protein-protein interface. To meet

this requirement we used the modified evaluation and ranking criteria approved by the CAPRI Management Committee and successfully applied to models of protein-peptide targets in previous CAPRI Rounds [21]. The modifications were as follows: 1) an 8 Å distance cut-off on Cβ atoms is used to define interface residues; 2) a 4 Å distance atom-atom cut-off is used to define native contacts; 3) at least 90% of the sequence of the peptide has to be recovered in the model (compared to at least 70% for the protein). Moreover, parameter values for classifying model quality as 'incorrect', 'acceptable', 'medium' and 'high', have been tightened (see **Table III** for details).

For target **T134**, the ability of predictors to correctly identify the correct 12-peptide that actually binds the DLC8 domain from the full 50-residue segment of the MAG protein was also evaluated. For this target, models lacking the 10 central residues of the segment that forms the protein-peptide interface were considered as incorrect. This is slightly more lenient than the 90% sequence coverage for the peptides, mentioned above and applied to **T135**.

Assessment of protein-oligosaccharide complexes

The protein-oligosaccharide complexes, although not an entirely new category of targets in CAPRI, remain a challenging modeling problem for predictors and assessors alike, owing to the particularities of the sugar moieties. In an attempt to standardize the format of the submitted models, participants were invited to use the following conventions: 1) use the PDB HETATM

records for the atomic coordinates of the arabinose oligosaccharide moieties; 2) the chain identifier for the sugar moieties should be the same as for the protein chain; 3) sugar residue numbering should start at 1. Participants were also provided with examples of oligosaccharide structures in the PDB, which listed the correct atomic numbering conventions.

While the requested format conventions were more or less followed, departures from the correct residue and atom number conventions were all too common. These included reversing the sugar residue numbering; wrong atom names, e.g. using O1' instead of O5' and attaching this oxygen to the wrong (previous) residue; using the same residue numbering for all sugar residues; using two different residue numbers for a single monoarabinose; first listing all carbons, then all oxygens, and numbering them sequentially; and finally, submitting models to the wrong target (e.g. AbnE/A6 submitted to **T129**, the AbnE/A3 complex, instead of to **T126**).

With a few exceptions, these problems could fortunately all be corrected through visual inspection followed by rule-based customized scripts, applied to all the models submitted by individual predictor groups. However models submitted by scorer groups required a special treatment. Since these groups used pooled models from multiple predictors, and adopted either the source (predictor) format or their own, two transformations were applied to each Scorer submission in order to carry out the assessment. One transformation was applied to

convert the scorer submission to the correct standard representation. In addition, thanks to the CAPRI tracking code, the provenance (source) of models in the uploaded set was identified and a second transformation was applied to models contributed by the same predictor group to the set. Both transformations were applied separately and the resulting models were evaluated. The more favorable evaluation was then assigned to the scorer submission. This convoluted protocol finally led to the rejection of only a handful of models over all five sugar targets.

We assessed only models where all the required sugar residues were present. The sugar atoms C1', O4', C4' and C5' were defined as backbone atoms. These atoms were used to perform the structure superimpositions. In addition, light of the small size of the protein-sugar interaction interface, which resembles the situation in protein-peptide interfaces, the following changes with respect to the assessment of protein-protein complexes were implemented: the threshold distance to determine interface residues was reduced to 6 Å, and the threshold distance for ligand-receptor contacts was reduced to 4 Å (**Table III**). However, the standard protocol (e.g. the standard definitions and thresholds for *f(nat), L-rms* and *I-rms*) was used for classification into the by now well-known categories of 'incorrect', 'acceptable', 'medium' and 'high' quality, models. **Figure 2b** illustrates model versus target fits of high, medium, and acceptable quality arabinopentaose models submitted for target T127 (AbnE/A5 complex)**.**

Additional assessment measures

To enable a higher-level analysis of the performance across targets, we used a continuous quality metric, as formulated by the DockQ score, to evaluate each modeled interface [52] :

$$DockQ = [f(nat) + rms_{scaled}(L_{rms}, d_1) + rms_{scaled}(i_{rms}, d_2)]/3$$

with $$rms_{scaled} = 1/[1 + (\frac{rms}{d_i})^2]$$

where *f*(nat), *i_rms*, and *L_rms* are as defined above. The rms_{scaled} represents the scaled rms deviations corresponding to either *L_rms* or *i_rm, s* and d_i is a scaling factor, d_1 for *L_rms* and d_2 for *i_rms,* which were optimized to fit the CAPRI model quality criteria, yielding d_1 = 8.5 Å and d_2 = 1.5 Å (see ref [52]).

Performance ranking

The performance of predictors scorers and servers for individual targets was ranked on the basis of a per-target score (*Score_T*). This score was computed as the weighted sum of the number of models in each of the 3 CAPRI categories (acceptable, medium and high) submitted by a given group as part of their top 5 ranking models, as follows:

$$Score\_T = \omega_1 N_{ACC} + \omega_2 N_{MED} + \omega_3 N_{HIGH}$$

where N_{ACC}, N_{MED} and N_{HIGH} are the number of models of acceptable-, medium- and high-quality, respectively, submitted for a given target (among the top 5 models); the values of the weights 'ω' were taken as ω_1=1, ω_2=2 and ω_3=3.

The performance of predictors, scorers and servers across targets was ranked on the basis of their best quality model in the 5-model submission for each target. The global score assigned to a group or a server, denoted as *Score_G,* has the same form as the score for individual targets shown above. Only this time, N_{ACC}, N_{MED} and N_{HIGH} are the number of targets for which the best model (among the top 5) was of acceptable, medium and high quality, respectively.
For both the per-target and across target ranking we also discuss the rankings for the top 1 and top 10 submitted models.

This ranking protocol, introduced recently for the CASP13-CAPRI challenge [38], differs from previous protocols, where priority was given to the number of targets for which medium or high quality models were submitted, and then to the number of targets with acceptable models. The new protocol leads to fairer ranking of groups, as it also takes into account the number of acceptable models that they submitted, especially for the more challenging targets.

PREDICTION RESULTS

In the following we summarize the prediction results for the different categories of targets, and present the global performance ranking across groups and automatic servers. The performance of predictors, scorers and servers was ranked on the basis of the top 5 submissions for each target, but results for the top 1 and top 10 submitted models are also discussed, whenever appropriate. The ranked group performance for individual interfaces of the targets of Round 38-45 are provided in the supplementary **Tables S3-S18**, and the full account of all the results for these targets are available at the official CAPRI website (https://www.ebi.ac.uk/msd-srv/capri/).

Results per target category

Protein-protein complexes (Targets T122-T125; T131-T133, T136)

T122, the x-ray structure of the human cytokine-receptor complex, comprised the IL23 hetero dimers (IL23A/IL12B) bound to its receptor protein IL23R. This turned out to be a medium difficulty target, since templates were available for both the IL23 hetero dimer and the IL23R receptor protein. But the flexibility of the inter domain arrangements in IL12B and the IL23R receptor differed sufficiently to affect the prediction performance. Probably as a result only a small number of groups were able to submit medium quality models for this target. A total of three predictor groups (Venclovas, Seok, Kozakov/Vajda) submitted such models among their top 5 ranking ones. The best of these models is illustrated in **Figure 3**. An additional 7 groups (including the CLUSPRO server) submitted only acceptable models. The scorer performance

was lower, with some scorer groups either ranking lower (e.g. among their top 10 submission) or missing altogether the medium quality models they themselves ranked higher (top 5) as predictors. The scorer performance improves when considering the top 10 ranking submissions.

The prediction results for **T123 and T124,** the PorM-Nt/nb-130 and PorM-Ct/nb-02 complex respectively, were very poor, due to the absence of templates for the two PorM domains. With predictors attempting and failing to model the PorM protein *ab-initio* – as during the CASP12 round for T0907 – only incorrect models were submitted for these targets, by the ~30 participating predictor groups and servers. However, one acceptable model was produced for **T123** by the group of Andreani/Guerois. This model was ranked too low to be considered a predictor submission, but was available in the scoring set, where it was unfortunately not recognized by any of the scorer groups.

Mitigated results were obtained for **T125**, the LLT1/NKR-P1 multi-protein hetero assembly, where the announced stoichiometry consisted of 1 homodimer of LLT1 binding 2 homodimers of NKR-P1, forming a hexamer with a total of 5 interfaces. Prediction results were evaluated only for 4 interfaces, which buried areas between 540 and 775 $Å^2$. These interfaces, numbered from 1-4, include the LLT1-NKR-P1 main hetero-dimer interface (interface 1), the LLT1 homodimer interface (interface 2), the NKR-P1 homodimer interface (Interface 4) and an

alternate binding mode of the hetero-dimer (Interface 3), see **Table I** and **Figure 4**. Models of acceptable quality or higher were submitted only for the 2 largest interfaces, for which templates were available. The very small size of the remaining interfaces, and the fact that consistent models featuring these interfaces were difficult to generate, are reasons for doubting the complex stoichiometry defined by the author, as mentioned above.

Eleven groups including 2 servers submitted at least 1 models of medium quality or better among their 5 top ranking models for interface 1. The two best performing groups for this interface were those of Pierce and Venclovas, with respectively 5 and 3 high quality models among their top 5 submissions. Moreover, the Pierce group had 8 high quality models among their 10 best predictions, whereas Venclovas had 5 high quality and 4 medium quality models among their top 10 predictions. It is also noteworthy that CLUSPRO and MDOCKPP, the 2 servers with correct predictions for this interface, each submitted 5 medium quality models.

Superb performance was achieved for the LLT1 homodimer (interface 2), with the majority of the predictor groups obtaining at least 1 high quality model among their top 5 submissions, and as many as 12 groups generating 5 high quality models among the top 5 submissions. The latter set of best performing groups comprised those of Bates, Fernandez-Recio, Gray, Moal, Shen, Weng, Zacharias, Huang and Grudinin, and included in addition the HADDOCK, PYDOCKWEB and SWARMDOCK servers. The vast majority of predictor groups also had a high

quality model as their single top prediction for this homodimer, and 10 groups including the 3 servers had high quality models for all their top 10 submissions.

Echoing the prediction performance for the different interfaces of **T125**, scorer groups submitted models of acceptable quality or higher only for interfaces 1 and 2. The scorer performance was good for interface 1, and excellent for interface 2, but overall lower than that of predictors (human & servers), for both interfaces. Whereas 14 scorer groups submitted at least 2 medium quality models for interface 1, only the group of Gray submitted high quality models for this interface (1 among their top 5 predictions and 2 among their top 10). The best scoring performance for this interface was by the LZERD server and the Bates group, each submitting 5 medium quality models, followed by the groups of Kihara (the authors of LZERD), Weng and Bonvin, with 4 medium quality models.

For the well predicted interface 2 of T125 as many as 16 scorer groups submitted at least 1 high quality model among their top 5 predictions. Among these groups, those of Laine, Moal and Carbone submitted 4 such models, whereas Zou, Gray, and Bonvin submitted 3 high quality models among their top 5 predictions.

The prediction performance was rather weak for targets **T131** and **T132**, the hetero-complexes of human CEACAM1 N-terminal domain with, respectively, the related bacterial cell

adhesion proteins HopQ-I and HopQ-II, despite the availability of templates for both the bacterial HopQ components and the human CEACAM1 domain. For **T131** (CEACAM1/HopQ-I complex), the lower performance may be explained by the presence of extra loops in the pylori HopQ-I protein, which contributes to the binding interface, but were absent in the available HopQ templates. Of the 30 predictors and servers groups submitting models for this target, only the group of Kozakov/Vajda submitted 1 medium quality model among their 5 top (and top 10) ranking ones. The results were only marginally better for **T132**, the complex with HopQ-II that did not contain the extra loops, and therefore should have been easier to model. Here, only the LZERD server and the group of Moal submitted 1 medium quality, and 1 acceptable quality model, respectively. The best model submitted for **T131** is shown in **Figure 5a**)

The scorer performance for **T131** and **T132** was only somewhat better than that of predictors, with the group of Carbone submitting a medium quality model as their top submission, in addition to one acceptable model among their top 5 submissions.

T133, the redesigned version of the Colicin E2/DNase-Im2 complex (E^{des3}/Im^{des3}), was a non-classical CAPRI target, where predictors had to accurately model the structural perturbations produced by redesigned residues in the *wt* complex. In order to gauge the effect of these perturbations on the CAPRI model quality criteria, we assessed the *wt* complex (PDB 3u43)

used as input for the design calculations, against **T133**. This yielded a medium-quality model (*f(nat)* 0.698; *L-rms* 2.043; *i-rms* 1.207), just like those submitted here by a good number of groups. Therefore, the fact that no high quality models were submitted for **T133** indicates that predictors, servers and scorer groups were all unable to reproduce the finer details of interface of the redesigned complex. Ten predictor groups, those of Venclovas, Kozakov/Vajda, Birini, Bonvin, Seok, Takeda-Shitaka, and Kihara, and 3 servers, INTERPRED, HADDOCK, LZERD, each submitted at least 2 medium quality models among their top 5 submissions. Whereas, 3 groups (Venclovas, Kozakov/Vajda and Brini, a recent CAPRI participant, submitted 5 such models. The best model obtained for this target, and the corresponding CAPRI quality measures, is illustrated in **Figure 5b**. Overall however, as many as 28 groups, of which 9 were servers, submitted at least 1 acceptable, or medium quality model among their top 5 predictions.

The scorer performance for this target was likewise good, with 7 scorer groups submitting at least 2 models of medium quality of better among their top 5 predictions, and 2 groups, the LZERD server and Bates, submitting 5 acceptable models of which 3 were of medium quality.

The 8th protein-protein target, **T136**, was the 4.5 Å resolution cryo-EM structure of the lysine decarboxylase (LdcA) homo-decamer, determined by cryo-EM to 4.5 Å resolution. This decamer comprises 3 distinct interfaces. A very large interface, involving interactions between

the subunits within and between rings (Interface 1), and 2 smaller, yet sizable, interfaces: interface 2 between subunits within the ring, and interface 3, again between subunits of different rings. This was in principle an easy target that could be predicted using 'template-based modeling' using as template the high-resolution x-ray structure of a related arginine decarboxylase (PDB code 2vyc) as template. It is therefore not surprising that good prediction results were obtained for all three interfaces, with the best results obtained for interface 1, followed by that for interfaces 2 and 3. Indeed, 18 groups, including 5 servers, submitted 5 medium quality models as their top prediction for interface 1, with another 6 groups (including 1 server) submitted at least 2 medium quality models for this interfaces. For interface 2, only 6 groups (including 1 server) submitted medium quality models for all their top 5 predictions, whereas for interface 3, the best performance was by the GALAXYPPDOCK server and the Vakser group, each submitting 4 medium quality models among their top 5 predictions. The best overall model obtained for this target is illustrated in **Figure 6.**

Not too surprisingly, the performance of scorer groups (and servers) reflected that of predictors but was in general lower, with a very good performance by groups such as Venclovas, Seok, and the LZERD server.

The low resolution of the target cryo-EM structure, or the genuine differences between the target structure and the available template(s), or both factors, may have contributed to the

absence of high-quality models for any of these interfaces, and therefore for the homo decamer as a whole.

Protein-peptide complexes (T121, T134, T135)

Target T121 was the solution structure of the complex between the TolAIII domain and the 13-residue N-terminal peptide of *P. aeruginosa* TolB. This turned out to be a difficult target, although, as already mentioned, a 1.94 Å resolution x-ray unbound structure was available for the TolAIII domain. The prediction performance for this target was very poor for predictor and server groups, and consequently also in the scoring challenge. In total, only 3 groups submitted 4 acceptable models among their 5 top ranking predictions: Seok submitted 2 such models, with Bates and Zou submitting 1 acceptable model each. The groups of Schueler-Furman and Kozakov/Vajda did however submit a medium quality model among their 10 top predictions.

As to the scorer performance, the LZERD server and 5 additional scorer groups submitted acceptable models for this complex, with LZERD, Kihara and Seok each submitting 1 acceptable model as their top prediction. However, 1 medium quality model and an additional 15 acceptable models were among the 10 best models of as many as 8 scorer groups.

The only 3 medium quality models were submitted for **T121** by the groups of Schueler-Furman, Kozakov/Vajda (predictors) and Moal (scorer) among their top 10 best models. These models displayed i-rms values of 2.5 Å, 2.8 Å, and 3.7 Å, respectively. Interestingly, the uploaded models (100 models submitted per participating group), displayed larger *i-rms* values than the predictor set (only from T122 onward were all predictor submissions automatically included in the uploaded set), with none below 3 Å, and the single medium quality model submitted by Moal scorer group represented an improvement upon its source model from the uploaded set.

Among the possible reasons for the generally poor performance may have been the flexibility of the peptide, of which only a small segment was predicted to adopt a β-strand conformation. The latter possibility may have been compounded by a conformational change of the TolAIII domain relative to the available template. With the target structures still on-hold, as this report is being written, we cannot comment on which of these two possibilities played any role.

Not too surprisingly, very good results were obtained for Targets **T134** and **T135**, the complexes of the DLC8 homodimer with a 12-residue peptide from the cognate L-MAG protein. The additional task of having to identify the 12-residue segment that binds to the DLC8 dimer in **T134** led to a generally lower prediction performance for this target, compared to T135

where the sequence of the 12-residue peptide was provided. In order to be evaluated, models submitted for **T134** had to feature a sufficient overlap (of at least 10 residues, with only one residue out of phase) between the bound peptide segment and the 12-residue segment of the target complex. Only 10 out of the 31 predictors groups (including servers) achieved this goal, including 6 groups that were able to identify exactly the 12-residues segment of the target complex (**Figure 7**). Among the predictor and server groups whose models for **T134** were evaluated, 7 groups (but no servers) had at least 1 model of medium or higher quality among their top 5 submissions. Of these groups the best performance was obtained by the group of Schueler-Furman (3 high quality and two medium quality model) followed by 4 other groups (Andreani-Guerois, Venclovas and Moal) each submitting 1 high quality model. For **T135**, as many as 14 predictor groups including 3 servers (HDOCK, CLUSPRO, and GALAXYPPDOCK) produced at least 1 model of medium or higher quality for this target among their top-5 submission, with several of these groups submitting 3 or 4 such models. The 5 best performers for this target were those of Schueler-Furman, with high quality models for all their top-5 submission, and the groups of Huang, Zacharias, Andreani/Guerois and the HDOCK sever, with only high and medium quality models among their top-5 submissions.

The scorer performance for these targets was quite good as well. As many as 11 groups, including the HDOCK and LZERD servers, submitted at least 1 medium quality model among their top-5 submissions for **T134**, with the groups of Venclovas, Oliva and Kihara submitting in addition between 1 and 5 high quality models. For **T135**, at least 1 medium quality model

or better was submitted by as many as 14 groups, with the 3 best performing groups also submitting higher quality models: 4 such models were submitted by the group of Venclovas, 2 by Fernandez-Recio and 1 by the HDOCK server.

The most accurate models submitted for the DLC8-peptide complex (by both predictor and scorer groups) is depicted in **Figure 8.**

Protein-oligosaccharide complexes (T126-T130)

The prediction results for the 5 protein-oligosaccharide complexes were quite interesting because they provide insight into the challenges that CAPRI groups faced in modeling these complexes. For **Targets 126** and **127**, the 2 complexes of AnbE with the longer-chain arabinohexaose (A6) and arabinopentaose (A5) ligands, respectively, only the HDOCK server submitted a medium quality model (for **T127**), whereas all other predictor and scorer groups submitted only acceptable models among their top 5 submissions. The total number of groups submitting such models was however quite high (15 for **T126** and 23 for **T127**; see **Figure 9**).

On the other hand, higher quality models were obtained for **T128** and **T129**, the AbnE complexes with the shorter sugars, arabinotetraose (A4) and arabinotriose (A3). Indeed, a number of groups (7 for **T128** and 6 for **T129**) submitted at least one medium quality model for these targets, with many more groups submitting acceptable models among their top 5 predictions. The number of groups with this performance increased to 9 or 10, depending on

the target, when the top 10 submissions were tallied (**Figure 9**). Five groups (Bonvin, Andreani/Guerois, Zacharias, Kihara and the LZERD server) performed best for **T128** and 6 groups (Gray, Seok, Chang, Huang, Andreani/Guerois , and the HDOCK server) performed best for **T129**.

These results indicate that predictors performed better in modeling the AbnE complexes with the shorter chain arabinose sugars than the longer versions. Interestingly furthermore, they were marginally more successful in modeling the complex with arabinotriose (A3), which binds to sites 2-4 of the AbnE binding pocket, and not sites 1-3, as might have been expected from the binding modes adopted by the longer chain arabinose sugars (A4-A6). The likely reason for this trend is that the available templates consisted of complexes of AbnE homologs bound only to 6-membered sugar ligands. These complexes were helpful in more or less correctly positioning the flexible longer arabinose oligosaccharides in the AbnE binding pocket, but they provided limited guidance for optimizing the bound sugar conformation and protein-sugar contacts. To perform such optimizations predictors had to rely on various modeling and refinement procedures. These procedures may not have been very effective in sampling the conformational landscape of the larger bound sugars moieties, but produced better results for the smaller sugars, owing to the fewer degrees of freedom that need to be sampled. This seems to have allowed predictor groups, including several servers, to produce medium quality models also for AbnE/A3, where the arabinotriose occupies sites 2-4.

The 5th sugar target, **T130**, the complex of AbnB with arabinopentaose, was the easiest of the lot, because templates representing the holo and apo forms of the enzymes were available, although the holo forms only bound shorter arabinose oligosaccharides. This was the only target of this category where high quality models were submitted. Nine predictor groups, including 2 servers (HADDOCK and LZERD), submitted such models among their top 5 predictions, and 5 of these groups (Venclovas, Chang, Grudinin, Seok, and Carbone) had a high quality model as their top ranking submission. Over 20 groups submitted at least one medium quality model for T130, either among the top 5 or top 10 submissions.

Lastly, as expected, the scorer performance mirrored that of predictors. It was relatively poor and roughly on par with those of predictors for **T126** and **127**; it was better and again essentially on par for **T128** and **129**, and excellent for **T130**.

Performance across predictor groups scorers and servers.

Groups (predictors, servers and scorers) were ranked according to their prediction performance for the 16 targets evaluated here. This ranking represents the official performance ranking for Rounds 38-45. We also ranked the performance of predictor groups and servers for each of the 3 categories of targets separately, as these categories represent distinct modeling problems. But these rankings are only presented for analysis purposes, as

the number of targets per category was too small for such ranking to be sufficiently informative.

All the rankings presented followed the revised ranking protocol, which uses a more balanced weighting scheme for models of different accuracy levels, as detailed in the section on the assessment and ranking procedures.

For a full account of the results obtained by each group the reader is referred to the CAPRI web site (http://pdbe.org/capri).

Performance of predictor groups

The consolidated ranking of predictor groups (including servers) across all 16 targets assessed in this report (**Table IV.a**), shows the 5 best performing groups to be those of Andreani/Guerois, Seok, Venclovas, Kozakov/Vajda, and Zacharias. The first 5 groups submitted correct models for 11-13 targets, which included high quality models (4 by Andreani/Guerois and Venclovas; 3 by Zacharias, 2 by Seok and 1 by Kozakov/Vajda), as well as medium quality models (8, 7, 5, 4 and 3, by Kozakov/Vajda, Seok, Andreani/Guerois, Venclovas and Zacharias, respectively). The first server in the ranked list is CLUSPRO, with 10 correctly predicted targets, including 6 medium quality ones.

It is furthermore noteworthy that the three top ranking groups (Andrani/Guerois, Seok, and Venclovas) also rank highest when considering their single best ranking models for each target (top 1); and that the same 5 groups remain the top performers, when considering their top 10 ranking models for each target (**Table IV.a**), as was the practice in earlier CAPRI evaluations. This supports the current more stringent approach of ranking the performance on the basis of the top 5 submitted models per target.

Performance of scorer groups

As already observed for the prediction results for individual targets, the across target performance of scorer groups is also somewhat weaker than that of predictors. The consolidated ranking of scorer across all targets (**Table IV.c**) lists 7 top performing scorer groups, as those submitting correct models for at least 11 of the 16 evaluated targets. In order of their rank, these groups are: Kihara, Zou, Bonvin and Seok, followed by the servers MDOCKPP and LZERD, and the group of Bates. The models submitted by these groups included a sizable fraction of targets for which medium quality models were submitted (7 targets by Zou, 6 by Bonvin, Seok, Huang and the LZERD server, and 5 by Kihara). On the other hand, the 5 best performing groups obtained high quality models for only 1 or 2 targets, whereas the group of Venclovas, which ranks lower in the list, did obtain high quality models for 3 targets, but overall submitted correct predictions for only 10 targets.

On the whole scorer rankings were less consistent, when considering also the top 1 or top 10 submitted models for each target (**Table IV.c**).

Performance of prediction servers

Like in the last CAPRI assessment [21], the overall performance of prediction servers in Rounds 38-45 (**Table IV.b**) is lower than that of human predictor groups, most likely due to their lower performance on the protein-peptide and protein-sugar complexes. These complexes – more particularly the protein sugar complexes – represent relatively recent target categories in CAPRI, for which automatic servers still need to be optimized.

Nonetheless, the CLUSPRO server maintains its high-ranking performance of previous CAPRI assessments, and is closely followed by those of HDOCK, a server that also performed very well in the recent CASP13-CAPRI challenge [38], MDOCKPP, LZERD, and the veteran HADDOCK server. CLUSPRO loses its high ranking position when considering only the single best predicted model (top 1), but ranks second (after HADDOCK) when the top 10 submitted models for each target are considered.

It is satisfying to see that as many as 20 predictor groups (and servers) submitted correct models for 8 to 12 targets (out of the 16 total), including models of medium quality or better, for between 4 to 9 targets (**Table IV.a**). The performance of predictor and scorer groups did

display some differences between the 3 target categories. While the top ranking predictor groups across all 16 targets were also among the best performers in each of the 3 target categories (see **Supplementary Tables S19-S21**), depending on the target category, a few additional groups percolated to the top of the rank. For example, the group of Pierce moved to the 3rd position for the protein complexes, and the groups of Zacharias and Schueler-Furman, rose to the 1st and 2nd position in the ranking for the protein-peptide complexes. These differences are clearly not statistically significant, but they seem to correlate with the proficiency of the group or server in tackling the modeling problem of the corresponding target category.

General trends and factors influencing prediction performance

The 16 targets of Round 38-45 comprised 3 very different categories of complexes representing distinct modeling challenges. Furthermore, with the exception of LdcA homo-decamer (**T136**), all the targets were hetero-complexes, where the partners are different proteins, a protein and a peptide or a protein and an oligosaccharide.

The prediction performance for these targets critically depended on several factors. For the 7 protein hetero-complexes, and considering the increased reliance of predictor groups on template-based modeling rather than on classical docking procedures, an important role was

played by their ability to leverage information on the binding modes and interfaces in structures of related hetero-complexes, and use this information to model the structure of the complex directly or to guide docking calculations. For all three target categories – but more particularly for the protein-peptide and protein-oligosaccharide complexes – a decisive factor was the successful optimization of the interactions at the binding interfaces by sampling both internal and rigid-body degrees of freedom of the interacting partners.

A convenient overview of how these challenges were met by predictor groups and servers across all three categories of targets is presented in **Figure 10**. This Figure plots the DockQ quality scores, color-coded by the CAPRI model quality categories for all the interfaces in individual models submitted by predictors (human and severs) for all the 16 targets of Rounds 38-45. These scores are contrasted with those obtained for the best of the 5 models submitted by CLUSPRO, the top performing automatic server in this evaluation. These models are used here as the baseline performance, analogous to that produced by the 'naïve' predictions considered previously [22], which employ 'off the shelf' standard modeling tools, although, given the consistently good performance of this server, this may put the bar for baseline performance rather high.

Examining the DockQ scores for the 7 protein-only hetero-complexes (**T122-T125** and **T131-T133**), most of which are also deemed to be difficult targets, confirms that the prediction

performance, measured as the fraction of models of acceptable quality or higher submitted across the ~40 human predictor and server groups, was uneven, and generally on the weak side. This global trend reflects the inherent difficulty of these targets. For predictors employing *ab-initio* docking approaches of various flavors, the availability of mostly distantly related templates for the individual subunits was often an important limitation, which allowed groups relying on template based modeling of the entire complex to gain some advantage. An exception was the well-predicted LLT1/NKR-P1 interface of **T125**, with several high accuracy models submitted for this complex, despite its relatively small size interface (775 $Å^2$). For this set of hetero-complex targets the baseline performance of the CLUSPRO server was often comparable to those of the best performing manual predictors. Other times however, the base line performance was also significantly lower, like for **T131**, **T132**, and for the targets or interfaces for which only incorrect models were submitted. The better performance for **T133** (the redesigned DNAse/Im2 complex) and the homo-decamer target **T136**, the easier targets for which adequate templates were available for the entire complex, illustrates the advantage that such targets represent for the modeling task. This advantage is likewise witnessed by the relatively good baseline performance of the CLUSPRO server. Nonetheless the failure to produce high quality models for these targets suggest that the modeling procedures remain in general limited in their ability to further optimize the initial near-correct models that are being generated.

Not much can be concluded from the performance trends for the 3 protein-peptide targets. The poor results obtained for **T121** cannot be analyzed here due to confidentiality requirement at this time, whereas the good performance for **T134-T135**, the DLC8-peptide complexes, is not particularly revealing, given the accurate template available for the DLC8 dimer, and the well conserved peptide binding site of this protein. Interestingly nevertheless, for **T134**, the poor performance of CLUSPRO and a good fraction of predictor groups, is mainly due the failure to identify the correct 12-residue that binds DLC8 from the sequence of a 50-residue peptide, a task which most groups performed using template based methods. These methods involved leveraging information on known structures of complexes of DLC8 with other peptides, aligning the sequences of these peptides and looking for conserved motif(s) that match the target sequence. On the other hand, the reason that only medium quality models where obtained for T135, where the exact sequence of the MAG peptide was provided, may be explained by the fact that the C-terminal leucine residue of the peptide is stabilized by crystal contacts rather than by interactions with the DLC8 domain (Khramushin *et al.*, this issue). The effects of such contacts on peptide conformation may need to be accounted for in future assessments.

Lastly, the performance trend across the 5 protein-oligosaccharide targets (**T126-T130**) confirms the conclusions reached by examining the per-target results, namely that the main bottleneck for both predictors and servers was the limited ability to optimize the binding

modes of the larger sugar ligands, the arabinopentaose (A5) and arabinohexaose (A6), due to their high degree of conformational flexibility.

To gain insight into the factors contributing to the quality of predicted models we examined the relationships between the quality of these models and 2 important parameters: the accuracy of the 3D structures of individual subunits in the predicted complexes and the accuracy of side chain conformations of interface residues in these complexes. **Figure 11b** displays the distributions of the M-rms values (the backbone rms values of the individual subunits of the submitted models versus those of the target), color-coded by the CAPRI model quality categories for all the interfaces in individual models submitted by predictors (human and severs) for the evaluated targets. For each target/interface, distributions are displayed for the ligand (L) and receptor (R) components of each complex (see legend of Figure 11 for definition). These distributions reveal interesting trends. The M-rms values of models submitted for individual targets span a wide range, even within models of the same CAPRI quality category. As expected, the lower M-rms values (< 1 Å) are generally observed for the high quality models (interfaces 1 and 2 of **T125**, the protein-peptide complexes **T134** and **T135,** and the protein sugar complex **T130**), confirming that accurately modeling the 3D structures of the individual subunits contributes to the accuracy of the predicted complex. For some targets however, submitted models ranked at different CAPRI quality levels comprise components that feature very similar M-rms values. For example, models submitted for **T122**

and ranked as medium-quality, acceptable, or incorrect, all feature very similar M-rms values for the component proteins. This indicates in turn that accurately modeling the 3D structure of the individual subunits, while helpful, is not sufficient for accurately, or even correctly, predicting the resulting complex. A similar conclusion can be drawn for the more poorly predicted interfaces of **T125** (interfaces 3, the alternate LLT1/NKR-P1 interface). Interestingly, the low M-rms values for interface 2 of **T125** are not reproduced in the LLT1 homodimer (interface 1 of **T125**), most likely because submitted models were derived from better templates for this homodimer.

A less equivocal relationship is observed between the accuracy of the modeled side chains of interface residues in predicted complexes and the quality level of these complexes, as illustrated in **Figure 11a**. This Figure displays the distributions of the root mean square deviation of side chain atoms (S-rms) of interface residues, color-coded by the CAPRI model quality categories for all the interfaces in individual models submitted by predictors (human and servers) for the evaluated targets. These distributions reveal that for the vast majority of the target or interfaces, the S-rms values decrease as the model quality level improves, highlighting the close relationship between the CAPRI model quality measures, which notably evaluate inter-residue contacts at the binding interface, and the accuracy with which the side chains of residues at this interface are modeled. On the other hand, S-rms values remain in general rather high (up to 2.4 Å) even for high quality models, and may span a wide range among models with similar quality level. This indicates that even in models of high quality,

according to CAPRI criteria, side chain conformations of interface residues may not be modeled at high enough accuracy for quantitative downstream applications such as rational protein or drug design.

Lastly, having noticed that some predicted complexes evaluated in previous Rounds, notably in the recent CASP13-CAPRI challenge, tended to feature a significant fraction of non-native residues at the interface [38] and hence to over-predict the interface regions, we examine if this also occurs in predicted complexes for the 16 targets evaluated in this study. To this end we plotted the quantity *f(non-nat)* in models submitted by predictor and scorer groups for individual targets, as a function of the DockQ score of the corresponding models, which are color coded according to the CAPRI model quality category (see **Figure 12**). *f(non-nat)* represents the fraction of contacts formed between residues the interface of predicted complexes, which are not found in the interface of the target. This quantity is routinely computed in CAPRI assessments, but has so far not been used for model ranking.

As one would hope, the scatter plots of **Figure 12** reveal a clear linear inverse correlation between *f(non-nat)* and the DockQ score of models submitted by predictor groups (including servers) (**Figure 12a;** Pearson Correlation –0.915), and by scorers (**Figure 12b**; Pearson Correlation –0.912). These plots also show quite a good correspondence of the plotted quantities with the 4 CAPRI model quality categories, each of which can be characterized by a

pair of *f(non-nat)* and *DockQ* values. But values for individual models display a significant spread, with those of *f(non-nat)* spreading systematically toward higher values across models of all 4 CAPRI quality categories. Most strikingly, a majority of the medium quality models and a large fraction of the high quality models in **Figure 12** feature *f(non-nat)* values of 40% or higher. Thus, even though interfaces in this fraction of high quality models recall between 60-95% of the native contacts of the target complexes, nearly half of all the predicted contacts in some of these interfaces are incorrect. As a result, the interfaces are of limited predictive value for any downstream application, as the probability of picking a native contact in such interfaces is only slightly above random.

Upon closer inspection we could verify however, that the high quality models displaying this behavior, numbering 60 in total, were all submitted for interface 2 of **T125** (that of the LLT1 homodimer) by different scorer and predictor groups (including servers). With several templates available for this homodimer, it is reasonable to assume that most of these high quality models were generated by template-based modeling protocols, but that these protocols probably involved limited subsequent model refinement. Interestingly, the HADDOCK prediction server, and the scorer group of Bonvin, the developer of this server, submitted 2 high quality models each, for this complex. These models featured, in contrast, a very small fraction (< 10%) of non-native contacts, demonstrating that it was possible to reach a high predictive value for this complex.

NEW TRENDS AND FUTURE DIRECTIONS

This 7th CAPRI report presented the evaluation of 36114 models of predicted protein complexes and their interfaces, submitted by about 60 groups, for 16 targets in CAPRI Round 38-45. These targets belonged to 3 different categories: protein-protein, protein-peptide and protein-oligosaccharide complexes, each representing distinct modeling challenges.

Of the 8 protein-protein complexes, most of which were hetero-complexes, at least 5 turned out to be non-trivial modeling problems. For **T122**, the IL23/IL23R complex, a bottleneck was the inter-domain flexibility of 2 of the subunits of this hetero-trimer and their homologs, requiring additional information and human expertise to guide the modeling. For **T123** and **T124**, the C-ter and N-ter PorM domain/nb complexes, the complete absence of templates for the PorM moieties was the reason that no correct models were submitted. For **T131** and **T132**, the CEACAM1/HopQ-Type I and II complexes, differences between the interface region in the target and templates, and more particularly the presence of several loops contributing to the interface of **T131**, that were absent in the templates, represented the major challenge for these targets.

Not too surprisingly, the prediction performance was overall poor for these targets, with only a relatively small number of groups submitting models of acceptable quality or better for most of these targets. As can be seen from the reports of individual groups in this issue, the groups

that performed best for these difficult targets used increasingly sophisticated 'integrative' template-based modeling techniques. These techniques leveraged information not only on templates for the individual subunits of the complex, but also on templates for the complex as a whole, whenever available, and on the interfaces of these templates. Information from different templates and interfaces was often also combined. Furthermore, in cases where the templates suggested several competing solutions to the modeling problem, these solutions were prioritized using additional approaches such as re-docking of the modeled subunits, sometimes guided by biochemical information (Padhorny *et al.,* Burman *et al.,* this issue), or computing various model quality scores (Dapkunas *et al.,* this issue) including approximate binding affinities, (Park *et al.,* this issue) to evaluate and rank the corresponding interfaces.

So far, similar but in general more straightforward template-based modeling of the entire complex has been successfully applied to predict the structure of homo-oligomer targets, such as those commonly offered in the CASP-CAPRI challenges [22,38]. It is therefore quite encouraging to see that this type of template-based modeling in its augmented and more 'integrative' forms is making important headway in the prediction of hetero-complexes, even for difficult targets. It hence comes as no surprise that these methods performed particularly well for 2 of the interfaces of **T125**, and for **T133**, where excellent templates were available for the binary complexes of the corresponding assemblies.

Nevertheless, the integrative template-based modeling of the full complex still falls short of generating accurate models in cases where more consequential conformational adjustments or remodeling part of the interface are necessary (like for **T131**, CEACAM1/HopQ-Type I, and **T133** the redesigned DNase/Im2 complex). Other limitations reside in the fact that crystal structures of hetero-complexes of related proteins are more prone to displaying different binding modes than homomeric protein assemblies, owing to their transient nature and to the influence of other protein components (e.g. nanobodies as well as additional native interactions partners), with whom they may be co-crystallized. A related issue played out in the prediction of the LLT1/NKR-P1 hetero complex (**T125**), where inconsistencies between the announced stoichiometry of the target and the interfaces formed in available templates, adversely affected the prediction performance of the complex as a whole.

Template-based modeling also played an important role in the successful predictions for **T134** and **T135**, the DLC8-peptide complex, and for the 5 protein-oligosaccharide complexes (**T126-T130**). Nevertheless, for the AbnE/arabinose oligosaccharide complexes (**T126-T129**), the available templates featured different sugar ligands, and could therefore serve mainly to guide the modeling procedures. The latter involved various strategies, including specialized flexible docking and model refinement methods, which alternate conformational perturbations of the sugar ligand with side chain adjustments of the protein binding pocket (see e.g. Burman *et al.* this issue).

Highlighting the progress in template-based modeling of the full protein-protein of protein-ligand complex should not be taken to mean that docking algorithms, which use as input homology built models of individual subunits of the complex, have not been holding their own in terms of prediction performance. Indeed, these algorithms have significantly progressed in their ability to speed up costly computations by using for example GPU (Graphic Processing Unit) processors. A number of algorithms are now able to more efficiently incorporate conformation flexibility during the docking calculations (Glaschagen *et al.*; Torchala *et al.;* Burman *et al.,* this issue), and to employ more accurate scoring functions to prioritize docking solutions, allowing their authors to rank prominently among the top performing groups. Docking procedures, and the criteria they use to score docking poses are also particularly helpful in augmenting the performance of template-based modeling (Burman *et al.*; Padhorny *et al*.; Park *et al.,* this issue), and they clearly remain the only viable approach when templates are available only for the individual subunits of the complex.

In cases where such templates cannot be found, even if only for one of the interacting components the prediction task cannot be tackled, as was the case for **T123** and **T124**, the C-ter and N-ter PorM/nb complexes. Hopefully this may change in the not too distant future, if Deep Learning techniques confirm their recent successful performance in predicting the 3D

structure of individual protein subunits [35], and these *ab-initio* predicted subunits are taken as input for docking procedures to generate useful models for the full assembly.

Acknowledgements

We gratefully acknowledge the EU grant WestLife (H2020-EINFRA-2015-1-675858) for support to Sameer Velankar and Nurul Nadzirin. We express gratitude to all the structural biologists who provided the targets of Rounds 38-45. Our thanks also go to the CAPRI management team and all predictor groups for stimulating discussion, valuable input and cooperation.

REFERENCES

1. Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 1998;92(3):291-294.
2. Ideker T, Sharan R. Protein networks in disease. Genome research 2008;18(4):644-652.
3. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nature reviews Genetics 2011;12(1):56-68.
4. Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buchholz F, Hyman AA, Mann M. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell 2015;163(3):712-723.
5. Babu M, Bundalovic-Torma C, Calmettes C, Phanse S, Zhang Q, Jiang Y, Minic Z, Kim S, Mehla J, Gagarinova A, Rodionova I, Kumar A, Guo H, Kagan O, Pogoutse O, Aoki H, Deineko V, Caufield JH, Holtzapple E, Zhang Z, Vastermark A, Pandya Y, Lai CC, El Bakkouri M, Hooda Y, Shah M, Burnside D, Hooshyar M, Vlasblom J, Rajagopala SV, Golshani A, Wuchty S, J FG, Saier M, Uetz P, T FM, Parkinson J, Emili A. Global landscape of cell envelope protein complexes in Escherichia coli. Nature biotechnology 2018;36(1):103-112.
6. ww PDBc. Protein Data Bank: the single global archive for 3D macromolecular structure data. Nucleic acids research 2019;47(D1):D520-D528.
7. Berman HM, Coimbatore Narayanan B, Di Costanzo L, Dutta S, Ghosh S, Hudson BP, Lawson CL, Peisach E, Prlic A, Rose PW, Shao C, Yang H, Young J, Zardecki C. Trendspotting in the Protein Data Bank. FEBS letters 2013;587(8):1036-1045.
8. Dutta S, Berman HM. Large macromolecular complexes in the Protein Data Bank: a status report. Structure 2005;13(3):381-388.
9. Bai XC, McMullan G, Scheres SH. How cryo-EM is revolutionizing structural biology. Trends in biochemical sciences 2015;40(1):49-57.
10. Cheng Y, Glaeser RM, Nogales E. How Cryo-EM Became so Hot. Cell 2017;171(6):1229-1231.
11. Kundrotas PJ, Zhu Z, Janin J, Vakser IA. Templates are available to model nearly all complexes of structurally characterized proteins. Proceedings of the National Academy of Sciences of the United States of America 2012;109(24):9438-9441.
12. Karaca E, Bonvin AM. Advances in integrative modeling of biomolecular complexes. Methods 2013;59(3):372-381.
13. Smith MT, Rubinstein JL. Structural biology. Beyond blob-ology. Science 2014;345(6197):617-619.
14. Wodak SJ, Janin J. Structural basis of macromolecular recognition. Advances in protein chemistry 2002;61:9-73.
15. Ritchie DW. Recent progress and future directions in protein-protein docking. Current protein & peptide science 2008;9(1):1-15.

16. Vajda S, Kozakov D. Convergence and combination of methods in protein-protein docking. Current opinion in structural biology 2009;19(2):164-170.
17. Lensink MF, Wodak SJ. Docking, scoring, and affinity prediction in CAPRI. Proteins 2013;81(12):2082-2095.
18. Fleishman SJ, Whitehead TA, Strauch EM, Corn JE, Qin S, Zhou HX, Mitchell JC, Demerdash ON, Takeda-Shitaka M, Terashi G, Moal IH, Li X, Bates PA, Zacharias M, Park H, Ko JS, Lee H, Seok C, Bourquard T, Bernauer J, Poupon A, Aze J, Soner S, Ovali SK, Ozbek P, Tal NB, Haliloglu T, Hwang H, Vreven T, Pierce BG, Weng Z, Perez-Cano L, Pons C, Fernandez-Recio J, Jiang F, Yang F, Gong X, Cao L, Xu X, Liu B, Wang P, Li C, Wang C, Robert CH, Guharoy M, Liu S, Huang Y, Li L, Guo D, Chen Y, Xiao Y, London N, Itzhaki Z, Schueler-Furman O, Inbar Y, Potapov V, Cohen M, Schreiber G, Tsuchiya Y, Kanamori E, Standley DM, Nakamura H, Kinoshita K, Driggers CM, Hall RG, Morgan JL, Hsu VL, Zhan J, Yang Y, Zhou Y, Kastritis PL, Bonvin AM, Zhang W, Camacho CJ, Kilambi KP, Sircar A, Gray JJ, Ohue M, Uchikoga N, Matsuzaki Y, Ishida T, Akiyama Y, Khashan R, Bush S, Fouches D, Tropsha A, Esquivel-Rodriguez J, Kihara D, Stranges PB, Jacak R, Kuhlman B, Huang SY, Zou X, Wodak SJ, Janin J, Baker D. Community-wide assessment of protein-interface modeling suggests improvements to design methodology. Journal of molecular biology 2011;414(2):289-302.
19. Moretti R, Fleishman SJ, Agius R, Torchala M, Bates PA, Kastritis PL, Rodrigues JP, Trellet M, Bonvin AM, Cui M, Rooman M, Gillis D, Dehouck Y, Moal I, Romero-Durana M, Perez-Cano L, Pallara C, Jimenez B, Fernandez-Recio J, Flores S, Pacella M, Praneeth Kilambi K, Gray JJ, Popov P, Grudinin S, Esquivel-Rodriguez J, Kihara D, Zhao N, Korkin D, Zhu X, Demerdash ON, Mitchell JC, Kanamori E, Tsuchiya Y, Nakamura H, Lee H, Park H, Seok C, Sarmiento J, Liang S, Teraguchi S, Standley DM, Shimoyama H, Terashi G, Takeda-Shitaka M, Iwadate M, Umeyama H, Beglov D, Hall DR, Kozakov D, Vajda S, Pierce BG, Hwang H, Vreven T, Weng Z, Huang Y, Li H, Yang X, Ji X, Liu S, Xiao Y, Zacharias M, Qin S, Zhou HX, Huang SY, Zou X, Velankar S, Janin J, Wodak SJ, Baker D. Community-wide evaluation of methods for predicting the effect of mutations on protein-protein interactions. Proteins 2013;81(11):1980-1987.
20. Lensink MF, Moal IH, Bates PA, Kastritis PL, Melquiond AS, Karaca E, Schmitz C, van Dijk M, Bonvin AM, Eisenstein M, Jimenez-Garcia B, Grosdidier S, Solernou A, Perez-Cano L, Pallara C, Fernandez-Recio J, Xu J, Muthu P, Praneeth Kilambi K, Gray JJ, Grudinin S, Derevyanko G, Mitchell JC, Wieting J, Kanamori E, Tsuchiya Y, Murakami Y, Sarmiento J, Standley DM, Shirota M, Kinoshita K, Nakamura H, Chavent M, Ritchie DW, Park H, Ko J, Lee H, Seok C, Shen Y, Kozakov D, Vajda S, Kundrotas PJ, Vakser IA, Pierce BG, Hwang H, Vreven T, Weng Z, Buch I, Farkash E, Wolfson HJ, Zacharias M, Qin S, Zhou HX, Huang SY, Zou X, Wojdyla JA, Kleanthous C, Wodak SJ. Blind prediction of interfacial water positions in CAPRI. Proteins 2014;82(4):620-632.
21. Lensink MF, Velankar S, Wodak SJ. Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition. Proteins 2017;85(3):359-377.

22. Lensink MF, Velankar S, Baek M, Heo L, Seok C, Wodak SJ. The challenge of modeling protein assemblies: the CASP12-CAPRI experiment. Proteins 2018;86 Suppl 1:257-273.
23. Levy ED, Teichmann S. Structural, evolutionary, and assembly principles of protein oligomerization. Progress in molecular biology and translational science 2013;117:25-51.
24. Xu Q, Canutescu AA, Wang G, Shapovalov M, Obradovic Z, Dunbrack RL, Jr. Statistical analysis of interface similarity in crystals of homologous proteins. Journal of molecular biology 2008;381(2):487-507.
25. Negroni J, Mosca R, Aloy P. Assessing the applicability of template-based protein docking in the twilight zone. Structure 2014;22(9):1356-1362.
26. Szilagyi A, Zhang Y. Template-based structure modeling of protein-protein interactions. Current opinion in structural biology 2014;24:10-23.
27. Marks DS, Hopf TA, Sander C. Protein structure prediction from sequence variation. Nature biotechnology 2012;30(11):1072-1080.
28. Ovchinnikov S, Kim DE, Wang RY, Liu Y, DiMaio F, Baker D. Improved de novo structure prediction in CASP11 by incorporating coevolution information into Rosetta. Proteins 2016;84 Suppl 1:67-75.
29. Jones DT, Singh T, Kosciolek T, Tetchner S. MetaPSICOV: combining coevolution methods for accurate prediction of contacts and long range hydrogen bonding in proteins. Bioinformatics 2015;31(7):999-1006.
30. Kocher JP, Rooman MJ, Wodak SJ. Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches. Journal of molecular biology 1994;235(5):1598-1613.
31. Rooman MJ, Kocher JP, Wodak SJ. Prediction of protein backbone conformation based on seven structure assignments. Influence of local interactions. Journal of molecular biology 1991;221(3):961-979.
32. Rooman MJ, Kocher JP, Wodak SJ. Extracting information on folding from the amino acid sequence: accurate predictions for protein regions with preferred conformation in the absence of tertiary interactions. Biochemistry 1992;31(42):10226-10238.
33. Min S, Lee B, Yoon S. Deep learning in bioinformatics. Briefings in bioinformatics 2017;18(5):851-869.
34. Wang S, Sun S, Li Z, Zhang R, Xu J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS computational biology 2017;13(1):e1005324.
35. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Zidek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D. Protein structure prediction using multiple deep neural networks in CASP13. Proteins 2019.
36. Lensink MF, Velankar S, Kryshtafovych A, Huang SY, Schneidman-Duhovny D, Sali A, Segura J, Fernandez-Fuentes N, Viswanath S, Elber R, Grudinin S, Popov P, Neveu E, Lee H, Baek M, Park S, Heo L, Rie Lee G, Seok C, Qin S, Zhou HX, Ritchie DW, Maigret B,

Devignes MD, Ghoorah A, Torchala M, Chaleil RA, Bates PA, Ben-Zeev E, Eisenstein M, Negi SS, Weng Z, Vreven T, Pierce BG, Borrman TM, Yu J, Ochsenbein F, Guerois R, Vangone A, Rodrigues JP, van Zundert G, Nellen M, Xue L, Karaca E, Melquiond AS, Visscher K, Kastritis PL, Bonvin AM, Xu X, Qiu L, Yan C, Li J, Ma Z, Cheng J, Zou X, Shen Y, Peterson LX, Kim HR, Roy A, Han X, Esquivel-Rodriguez J, Kihara D, Yu X, Bruce NJ, Fuller JC, Wade RC, Anishchenko I, Kundrotas PJ, Vakser IA, Imai K, Yamada K, Oda T, Nakamura T, Tomii K, Pallara C, Romero-Durana M, Jimenez-Garcia B, Moal IH, Fernandez-Recio J, Joung JY, Kim JY, Joo K, Lee J, Kozakov D, Vajda S, Mottarella S, Hall DR, Beglov D, Mamonov A, Xia B, Bohnuud T, Del Carpio CA, Ichiishi E, Marze N, Kuroda D, Roy Burman SS, Gray JJ, Chermak E, Cavallo L, Oliva R, Tovchigrechko A, Wodak SJ. Prediction of homoprotein and heteroprotein complexes by protein docking and template-based modeling: A CASP-CAPRI experiment. Proteins 2016;84 Suppl 1:323-348.

37. Lafita A, Bliven S, Kryshtafovych A, Bertoni M, Monastyrskyy B, Duarte JM, Schwede T, Capitani G. Assessment of protein assembly prediction in CASP12. Proteins 2018;86 Suppl 1:247-256.

38. Lensink MF, Brysbaert G, Nadzirin N, Velankar S, Chaleil RAG, Gerguri T, Bates PA, Laine E, Carbone A, Grudinin S, Kong R, Liu RR, Xu XM, Shi H, Chang S, Eisenstein M, Karczynska A, Czaplewski C, Lubecka E, Lipska A, Krupa P, Mozolewska M, Golon L, Samsonov S, Liwo A, Crivelli S, Pages G, Karasikov M, Kadukova M, Yan Y, Huang SY, Rosell M, Rodriguez-Lumbreras LA, Romero-Durana M, Diaz-Bueno L, Fernandez-Recio J, Christoffer C, Terashi G, Shin WH, Aderinwale T, Raghavendra Maddhuri Venkata Subraman S, Kihara D, Kozakov D, Vajda S, Porter K, Padhorny D, Desta I, Beglov D, Ignatov M, Kotelnikov S, Moal IH, Ritchie DW, Chauvot de Beauchene I, Maigret B, Devignes MD, Echartea MER, Barradas-Bautista D, Cao Z, Cavallo L, Oliva R, Cao Y, Shen Y, Baek M, Park T, Woo H, Seok C, Braitbard M, Bitton L, Scheidman-Duhovny D, DapkUnas J, Olechnovic K, Venclovas C, Kundrotas PJ, Belkin S, Chakravarty D, Badal VD, Vakser IA, Vreven T, Vangaveti S, Borrman T, Weng Z, Guest JD, Gowthaman R, Pierce BG, Xu X, Duan R, Qiu L, Hou J, Ryan Merideth B, Ma Z, Cheng J, Zou X, Koukos PI, Roel-Touris J, Ambrosetti F, Geng C, Schaarschmidt J, Trellet ME, Melquiond ASJ, Xue L, Jimenez-Garcia B, van Noort CW, Honorato RV, Bonvin A, Wodak SJ. Blind prediction of homo- and hetero- protein complexes: The CASP13-CAPRI experiment. Proteins 2019.

39. Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. Chemical reviews 2014;114(13):6733-6778.

40. Tompa P, Davey NE, Gibson TJ, Babu MM. A million peptide motifs for the molecular biologist. Molecular cell 2014;55(2):161-169.

41. Bloch Y, Bouchareychas L, Merceron R, Skladanowska K, Van den Bossche L, Detry S, Govindarajan S, Elewaut D, Haerynck F, Dullaers M, Adamopoulos IE, Savvides SN. Structural Activation of Pro-inflammatory Human Cytokine IL-23 by Cognate IL-23

Receptor Enables Recruitment of the Shared Receptor IL-12Rbeta1. Immunity 2018;48(1):45-58 e46.
42. Leone P, Roche J, Vincent MS, Tran QH, Desmyter A, Cascales E, Kellenberger C, Cambillau C, Roussel A. Type IX secretion system PorM and gliding machinery GldM form arches spanning the periplasmic space. Nature communications 2018;9(1):429.
43. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. Critical assessment of methods of protein structure prediction (CASP)-Round XIII. Proteins 2019.
44. Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP)-Round XII. Proteins 2018;86 Suppl 1:7-15.
45. Moonens K, Hamway Y, Neddermann M, Reschke M, Tegtmeyer N, Kruse T, Kammerer R, Mejias-Luque R, Singer BB, Backert S, Gerhard M, Remaut H. Helicobacter pylori adhesin HopQ disrupts trans dimerization in human CEACAMs. The EMBO journal 2018;37(13).
46. Netzer R, Listov D, Lipsh R, Dym O, Albeck S, Knop O, Kleanthous C, Fleishman SJ. Ultrahigh specificity in a network of computationally designed protein-interaction pairs. Nature communications 2018;9(1):5286.
47. Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins 2010;78(9):2029-2040.
48. Myllykoski M, Eichel MA, Jung RB, Kelm S, Werner HB, Kursula P. High-affinity heterotetramer formation between the large myelin-associated glycoprotein and the dynein light chain DYNLL1. Journal of neurochemistry 2018;147(6):764-783.
49. Fan J, Zhang Q, Tochio H, Li M, Zhang M. Structural basis of diverse sequence-dependent target recognition by the 8 kDa dynein light chain. Journal of molecular biology 2001;306(1):97-108.
50. Mendez 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.
51. Lensink MF, Mendez R, Wodak SJ. Docking and scoring protein complexes: CAPRI 3rd Edition. Proteins 2007;69(4):704-718.
52. Basu S, Wallner B. DockQ: A Quality Measure for Protein-Protein Docking Models. PloS one 2016;11(8):e0161879.

FIGURE CAPTIONS

Figure1: The 8 protein-protein targets of CAPRI Rounds 38-45.

The eight protein-protein complexes (see Table 1) are labeled by their CAPRI target numbers.

Different protein chains in each target are displayed using ribbon diagrams of different colors;

redundant subunit copies in T136 are depicted in grey. For the multi-protein complexes of T125 and T136, dashed lines indicate the inter-subunit interfaces that were evaluated (see text and Table 1 for target descriptions and further details).

Figure 2: Pictorial tips on protein-peptide and protein-oligosaccharide targets of CAPRI Rounds 38-45.

(a) Components of T121, the complex between TolAIII domain (represented by the crystal structure of the unbound form (PDB code 1lr0), and the amino acid sequence of the TolB N-terminal peptide. The structure of the target against which models were evaluated could not be shown. (b) T134/T135 the DLC8-peptide complex (see Table 1 and text for details). (c-e) Superimposition of the arabinopentaose ligands of AbnE in predicted versus target complexes for T127, illustrating the match obtained for different CAPRI model quality categories: acceptable (e) medium (d) and high (c). The corresponding CAPRI quality measures are listed alongside each model. The full complex of T127 could not be shown for confidentiality reasons.

Figure 3: Overview of the prediction results for T122, the Il23/Il23R complex.

(a) Shown is T122, the complex of IL23 (Il23A/IL12B) bound to the IL23R receptor, with individual subunits depicted in ribbon representations of different colors. The dots surrounding the complex are the positions of the geometric centers of the 'ligand' (here the

IL23R subunit) in the models submitted by predictors, relative to each 'receptor' subunit (here the IL23 hetero complex). The geometric centers of incorrect models (the majority) are colored yellow; acceptable models among the top 10 submissions, including the one by the group of Venclovas, detailed in (b), are colored light blue. (b) The acceptable model for T122 submitted by the group of Venclovas, superimposed onto that of the target. Shown are the positions of the IL23R subunit in the submitted mode (blue), versus that of the target (green), relative to the 'receptor' (e.g. the IL23 subunits) used as reference.

Figure 4: T125, the LLT1/NKR-P1 complex and its best predicted model.

(a) The author-announced A2B4 stoichiometry for the LLT1/NKR-P1 complex, with individual subunits shown in ribbon representations. Dashed lines illustrate the 4 evaluated inter-subunit interfaces (see Table 1, and the text for details). (b) The best, high quality, model submitted for this target by the group of Pierce, illustrating the correspondence between the LLT1 homodimers in the model versus the target (interface 2 in (a)) and between the LLT1 homodimer and the NKR-P1 subunit (interface 1 in (a)); the other, incorrectly placed, subunits are not shown.

Figure 5: Prediction results for T131, and T133.

(a) The best predicted model for T131, the CEACAM1/HopQ-Type-I complex submitted by Kozakov/Vajda, superimposed onto the target structure. The correspondence between the

predicted and target interfaces is illustrated in the inset (the target backbone, with the extra loop of HopQ-Type_I subunit, is colored bright red). (b)Close up of the best medium quality predicted model for T133, the redesigned E2 DNase/Im2 complex, submitted by the group of Venclovas.

Figure 6: Prediction results for T136, the LdcA homo-decamer.

(a) The best medium quality model submitted for the full LdcA homo-decamer, by the group of Andreani/Guerois, superimposed over the target complex. (b) Close-up of the subunit of the same model, superimposed over its equivalent in the target.

Figure 7: Overview of the results on identifying the 12-residue peptide of the MAG protein that binds DLC8 in the protein-peptide complex of targets T134/T135.

The amino acid sequence of full 50-residue segment from the MAG protein that was used in the co-crystallization experiment of T134 is listed above the top line, with the 12-residue segment that binds the DLC8 dimer highlighted in red fonts. The sequences of various putative peptide segments, predicted to bind DLC8 in the submitted models for T134, are listed in subsequent rows. The last 3 columns of the Table, list the number of models containing the listed sequence, the number of predictor groups, and the number of scorer groups that submitted models with the listed sequence, respectively. Sequence segments depicted in red fonts correspond to the segment that forms the protein-peptide interface in the complex with

the DLC8 dimer. Only models featuring at least the center 10 residues of this segment were evaluated for T134 (see text).

Figure 8: Best model(s) for T134-135, superimposed on the target complex.
The DLC8 dimer is depicted in ribbon representation (red). The best high quality model of the bound peptide submitted by the group of Schueler-Furman (light blue) is shown superimposed onto the 12-residue bound peptide in the target structure.

Figure 9: Overview of the results obtained by predictor and scorer groups (including servers) for the protein-oligosaccharide targets T126-T130.
Plots showing the cumulative number of predictor (including servers) and scorer groups that submitted correct models of acceptable quality or better (color coded according to the CAPRI model quality, as indicated in the legends), for each of the 5 protein/arabinose oligosaccharide complexes of T126-T130. The numbers of such models are tallied separately for the top 1, top 5 and top 10, submissions respectively. Performance ranking was performed only on the basis of the top 5 submitted models.

Figure 10: DockQ scores for models submitted for targets/interfaces of Rounds 38-45.
Shown are the distributions of the DockQ scores computed for the top 5 models submitted by all predictor groups for individual targets or interfaces of Rounds 38-45. Targets are labeled

by their CAPRI target number, with the additional digit referring to individual interfaces (see Table 1 and text for details). The target category (protein-protein, protein-peptide, and protein-oligosacharide) is listed at the top. Individual points are color-coded according to the CAPRI model quality category; yellow: incorrect; blue: acceptable; green: medium; red: high. For each target, a baseline-level prediction, represented by the best model of the top-performing automatic server (CLUSPRO, see **Table 4b**), is represented by black triangles. The box-plot distributions (with whiskers at 9th and 91st percentiles) of each target and prediction category are shown on the lower panel; color-coding is as for the upper panel, but with a lighter shade of blue for better visibility.

Figure 11: Model quality of individual subunits, and of sidechain conformations of interface residues in models submitted for targets of Rounds 38-45.

(a) Distributions of the S-rms values (the sidechain rms values of interface residues of the submitted models versus those of the target) for all the targets and association modes in individual models submitted by predictors (human and severs) for the evaluated targets. (b) Distributions of the M-rms values (the backbone rms values of the individual subunits of the submitted models versus those of the target) for all the targets and association modes in individual models submitted by predictors (human and servers) for the evaluated targets. For each target/interface, distributions are displayed for the ligand (L) (left column) and receptor (R) (right column) components of each complex. For target T122, the IL23 hetero-complex is

considered as the 'receptor' moiety, where the IL23R subunit is taken as the 'ligand'. Individual points in (a) and (b) are color-coded according to the CAPRI model quality category; yellow: incorrect; blue: acceptable; green: medium; red: high.

Figure 12: *f(non-nat)* versus *DockQ* scores for models submitted for targets of Rounds 38-45. Scatter plots of the quantity *f(non-nat)* in models submitted by Predictor (left plot) and Scorer (right plot) groups for individual targets, as a function of the DockQ score of the corresponding models, color coded according to the CAPRI model quality category; yellow: incorrect; blue: acceptable; green: medium; red: high. *f(non-nat)* represents the fraction of contacts formed between residues of the interface of predicted complexes, which are not found in the interface of the target, and represent the extent to which an interface may have been over-predicted; the DockQ score is the consolidated quality score of submitted models as detailed in the text. The plots show that some models ranked as high quality by the CAPRI criteria (and DockQ scores >0.7), feature *f(non-nat)* values between 50-60%, indicating in turn that while the predicted interfaces contain the majority (between 60-95%) of the native contacts, half of the predicted contacts are non-native. Interestingly, this occurs mainly in high quality models submitted for interface 2 (that of the LLT1 homodimer) of target 125 (see text for details).

Table I

Target	Stoich.	#Int.	Area (Å^2)	#Res.	PDB	Difficulty level	Description
Protein-Protein complexes							
T122	A1B1C1	1	2330	198/328/330	5mzv	Difficult	Human cytokine hetero-dimer/receptor complex IL23/IL23R
T123	A1B1	1	1240	174/121	6ey0	Difficult	PorM-Nt/nb(02)
T124	A1B1	1	1120	202/141	6ey6	Difficult	PorM-Ct/nb(130)
T125	A2B4	5	775 (a) 630 (b) 540 (c) 540 (d) 440 (e)	135/146	5mgt	Difficult	Hetero-hexamer of LLT1/NKR-P1 (extra-cellular domains)
T131	A1B1	1	2030	108/404	6gbg	Difficult	Human CEACAM1/HopQ-Type-I *H. pylori*
T132	A1B1	1	1860	108/418	6gbh	Medium	Human CEACAM1/hopQ-Type-II *H. pylori*
T133	A1B1	1	1600	69/95	6ere	Easy	Redesigned Colicin E2 DNase/Im2 complex
T136	A10	3	5600 (a) 1100 (b) 800 (c)	751	N/A	Easy	LdcA *P. aeroginosa*
Protein-Peptide complexes							
T121	A1B1	1	1440	115/13	N/A	Difficult	*P.aeroginosa* TolAIII domain/N-terminus *P. aeruginosa* TolB
T134	A2B1	1	1540	88/50	6gzj	Easy	DLC8 dimer/MAG 50-residue fragment
T135	A2B1	1	1640	88/12	6gzl	Easy	DLC8 dimer (Rat)/MAG 12-residue fragment
Protein Oligosaccharide complexes							

T126	A1B1	1	1340	415/6	N/A	Difficult	Arabino-oligosaccharide binding protein,*G stearothermophilus, with* AbnE/A6
T127	A1B1	1	1220	415/5	N/A	Difficult	Arabino-oligosaccharide binding protein,*G stearothermophilus, with* AbnE/A5
T128	A1B1	1	810	415/4	N/A	Medium	Arabino-oligosaccharide binding protein,*G stearothermophilus, with* AbnE/A4
T129	A1B1	1	820	415/3	N/A	Medium	Arabino-oligosaccharide binding protein,*G stearothermophilus, with* AbnE/A3
T130	A1B1	1	1160	315/5	6f1g	Easy	Arabino-oligosaccharide binding protein,*G stearothermophilus, with* AbnB/A5

Table I – The targets of CAPRI Rounds 38 – 45.

The target list is subdivided into categories: protein-protein (R39:T122-T124; R40:T125; R42:T131,T132; R43:T133; R45:T136), protein-peptide (R38:T121; T44:T134,T135) and protein-polysaccharide (R41:T126-T130). Columns 1 – 3 list the CAPRI target ID, its stoichiometry and the number of assessed interfaces; columns 4 – 6 the buried surface area per interface, the number of residues and the PDB ID (if available); column 7 lists the target difficulty and column 8 contains a textual description of the target.

Table II

	Total	Human	Server
Predictors	51	40	11
Scorers	29	24	5
All	57	44	13

Table II – CAPRI Rounds 38 – 45 participation statistics.
The table shows the number of registered Predictor and Scorer groups, both human and server.

Table III.a – CAPRI criteria for ranking models of protein-peptide complexes

f(nat)	incorrect	$< 0.2 \leq$	acceptable	$< 0.5 \leq$	medium	$< 0.8 \leq$	high
L-rms	incorrect	$> 5.0 \geq$	acceptable	$> 2.0 \geq$	medium	$> 1.0 \geq$	high
i-rms	incorrect	$> 2.0 \geq$	acceptable	$> 1.0 \geq$	medium	$> 0.5 \geq$	high

Table III.b – CAPRI criteria for ranking models of protein-protein and protein-polysaccharide complexes

f(nat)	incorrect	$< 0.1 \leq$	acceptable	$< 0.3 \leq$	medium	$< 0.5 \leq$	high
L-rms	incorrect	$> 10.0 \geq$	acceptable	$> 5.0 \geq$	medium	$> 1.0 \geq$	high
i-rms	incorrect	$> 4.0 \geq$	acceptable	$> 2.0 \geq$	medium	$> 1.0 \geq$	high

Table III.c – CAPRI distance and sequence conservation criteria for complexes

	protein-protein	protein-peptide	protein-polysaccharide
d(interface)	10 Å	8 Å	6 Å
d(contact)	5 Å	4 Å	4 Å
Sequence identity	70%	90%	90%

Table III – CAPRI criteria used to rank submitted models for protein-protein, protein-peptide and protein-polysaccharide complexes. Ranking is performed using three parameters: two for interface accuracy, f(nat) and i-rms, and one for ligand positioning, L-rms. The distance threshold for f(nat) contacts is d(contact), the distance threshold for i-rms calculation is d(interface). For further details we refer to the text and references therein.

Table IV.a – Predictor group ranking (human and servers)

			top 1		top 5		top 10	
Rank	Group name	#targets	rank	performance	rank	performance	rank	performance
1	Andreani/Guerois	19	3	10/1***/5**	1	12/4***/5**	1	12/4***/6**
2	Seok	19	1	11/1***/5**	2	13/2***/7**	2	13/2***/8**
3	Venclovas	16	1	9/3***/3**	3	11/4***/4**	3	11/5***/3**
4	Kozakov/Vajda	18	5	9/1***/4**	4	12/1***/8**	3	13/1***/9**
5	Zacharias	19	4	11/2***/1**	5	12/3***/3**	5	12/3***/3**
6	Zou	18	8	8/1***/2**	6	11/1***/4**	11	11/1***/4**
6	Moal	16	19	4/1***/3**	6	10/2***/3**	7	11/2***/3**
6	Huang	17	6	8/2***/2**	6	9/2***/4**	7	10/2***/4**
9	Kihara	19	8	9/1***/1**	9	10/1***/4**	6	11/1***/6**
9	CLUSPRO	18	22	5/2**	9	10/6**	13	10/6**
11	Shen	19	15	6/1***/2**	11	10/1***/3**	11	11/1***/4**
11	HDOCK	15	22	4/1***/1**	11	8/1***/5**	13	9/1***/5**
11	Gray	16	15	6/1***/2**	11	9/1***/4**	16	9/1***/4**
11	Fernandez-Recio	19	15	6/1***/2**	11	10/1***/3**	13	10/1***/4**
11	Bonvin	19	8	6/1***/4**	11	7/2***/4**	7	9/2***/5**
16	Pierce	11	8	6/2***/2**	16	7/2***/3**	19	7/2***/3**
16	MDOCKPP	18	12	7/1***/2**	16	9/1***/3**	19	9/1***/3**
16	LZERD	19	12	8/3**	16	8/1***/4**	17	8/1***/5**
16	HADDOCK	19	15	5/1***/3**	16	8/2***/2**	7	11/2***/3**
16	Chang	19	12	6/2***/1**	16	8/2***/2**	13	9/2***/3**
21	Grudinin	19	7	7/2***/2**	21	7/2***/2**	19	8/2***/2**
22	Takeda-Shitaka	14	19	6/1***/1**	22	7/1***/2**	24	7/1***/2**
23	GALAXYPPDOCK	16	22	5/2**	23	6/4**	23	7/1***/3**
23	Bates	19	19	6/1***/1**	23	7/1***/1**	19	10/1***/2**
25	Vakser	17	26	4/2**	25	6/1***/1**	26	6/1***/1**
25	SWARMDOCK	19	22	4/1***/1**	25	6/1***/1**	25	7/1***/1**
27	Carbone	10	28	3/1***	27	4/1***/1**	27	5/1***/1**
28	Weng	11	26	3/1***/1**	28	3/1***/1**	29	3/1***/1**
28	Schueler-Furman	3	28	2/1***/1**	28	2***	27	3/2***/1**
30	PYDOCKWEB	14	33	1***	30	3/1***	29	4/1***
31	Wolfson	7	30	2/1***	31	2/1***	32	2/1***
31	Del Carpio	18	30	3/1**	31	3/1**	29	5/1**
31	Brini	3	30	2**	31	2**	32	2**
34	Ritchie	7	33	2/1**	34	2/1**	32	2**
34	Laine	3	33	1***	34	1***	35	1***
36	Negi	9	36	2	36	2	36	2
36	Iwadate	1	36	1**	36	1**	36	1**
36	INTERPRED	1	36	1**	36	1**	36	1**
39	UUcourse	1	39	1	39	1	36	1**
39	Czaplewski	3	39	1	39	1	36	2
41	ZDOCK	1	41	0	41	0	41	0
41	Wang	1	41	0	41	0	41	0
41	Wallner	5	41	0	41	0	41	0
41	Tuffery	1	41	0	41	0	41	0
41	Schneidman	1	41	0	41	0	41	0
41	Sanner	1	41	0	41	0	41	0
41	Niv	1	41	0	41	0	41	0
41	GRAMM-X	3	41	0	41	0	41	0
41	Gong	1	41	0	41	0	41	0
41	Carazo	1	41	0	41	0	41	0

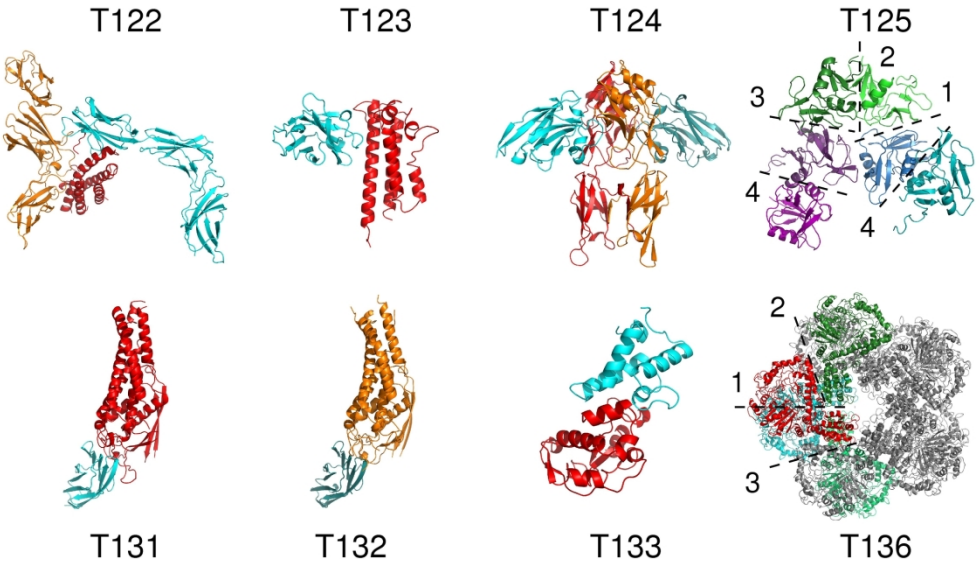
Table IV.b – Server group ranking

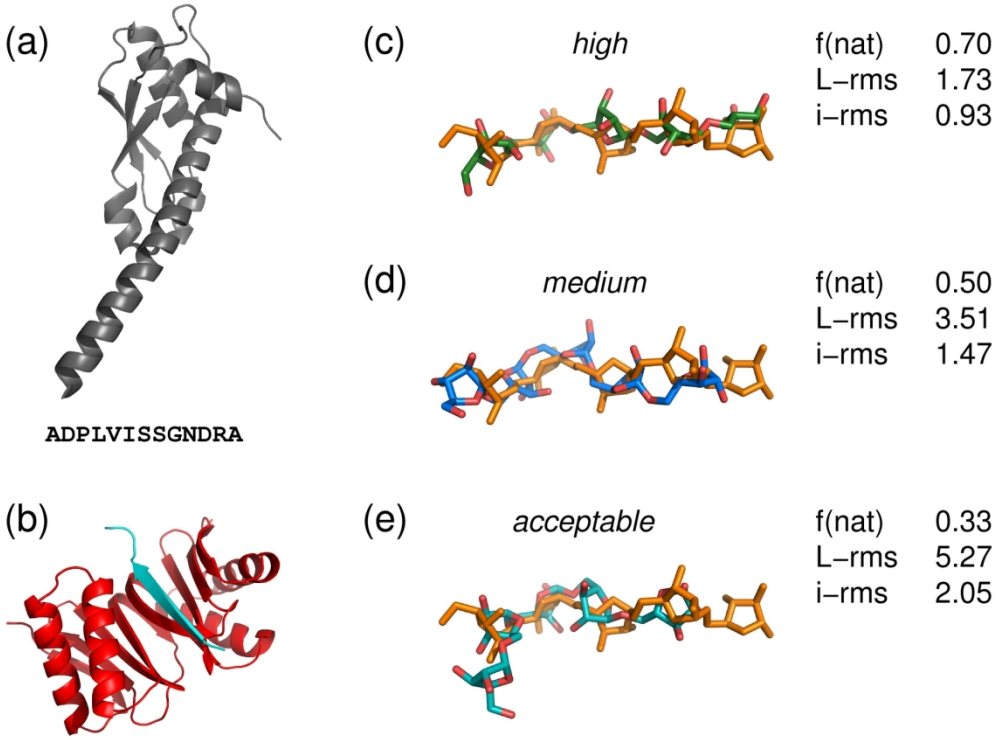
			top 1		top 5		top 10	
Rank	Group name	#targets	rank	performance	rank	performance	rank	performance
1	CLUSPRO	18	4	5/2**	1	10/6**	2	10/6**
2	HDOCK	15	4	4/1***/1**	2	8/1***/5**	2	9/1***/5**
3	HADDOCK	19	3	5/1***/3**	3	8/2***/2**	1	11/2***/3**
3	LZERD	19	1	8/3**	3	8/1***/4**	4	8/1***/5**
3	MDOCKPP	18	1	7/1***/2**	3	9/1***/3**	5	9/1***/3**
6	GALAXYPPDOCK	16	4	5/2**	6	6/4**	6	7/1***/3**
7	SWARMDOCK	19	4	4/1***/1**	7	6/1***/1**	7	7/1***/1**
8	PYDOCKWEB	13	8	1***	8	3/1***	8	4/1***
9	INTERPRED	1	9	1**	9	1**	9	1**
10	ZDOCK	1	10	0	10	0	10	0

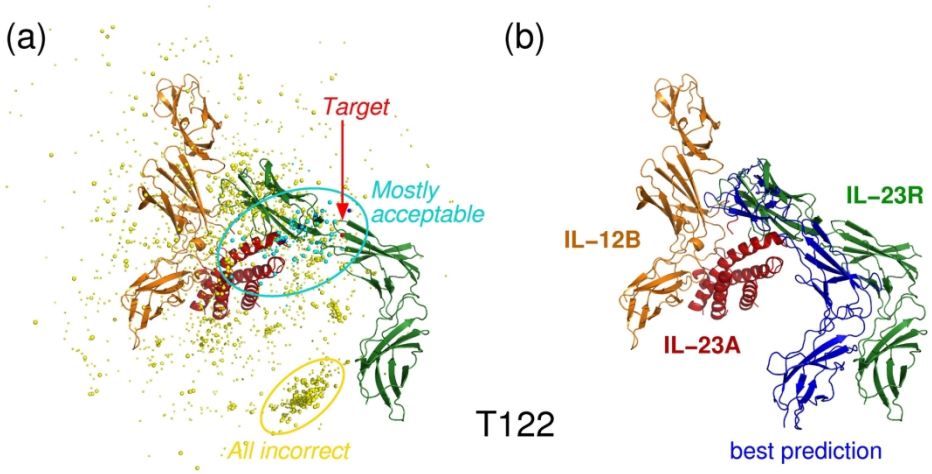
Table IV.c – Scorer group ranking (human and servers)

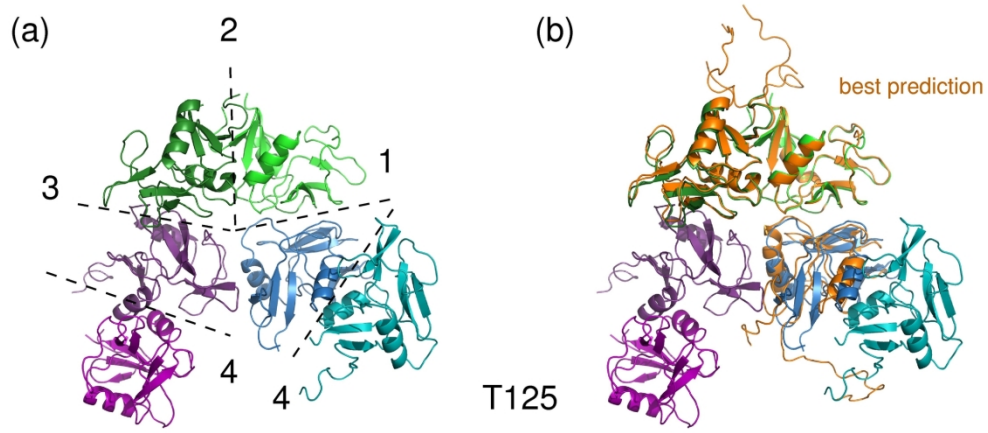
			top 1		top 5		top 10	
Rank	Group name	#targets	rank	performance	rank	performance	rank	performance
1	Kihara	19	1	11/6**	1	13/2***/5**	4	13/2***/6**
2	Zou	18	8	8/1***/2**	2	12/1***/7**	8	12/1***/7**
2	Bonvin	19	3	8/5**	2	11/2***/6**	2	13/2***/7**
4	Seok	18	3	8/5**	4	12/1***/6**	4	13/1***/8**
4	Huang	18	3	8/5**	4	10/2***/6**	4	12/2***/7**
6	Venclovas	19	3	6/2***/3**	6	10/3***/3**	2	12/4***/4**
6	MDOCKPP	15	9	7/1***/2**	6	11/2***/4**	9	12/2***/4**
6	LZERD	19	2	10/5**	6	13/6**	7	13/1***/7**
9	Gray	15	3	8/1***/3**	9	9/2***/3**	12	11/2***/3**
9	Chang	17	11	6/4**	9	9/1***/5**	9	10/2***/6**
11	HDOCK	12	12	6/3**	11	8/2***/3**	13	9/2***/4**
11	Fernandez-Recio	18	15	4/1***/1**	11	8/2***/3**	1	13/2***/10**
11	Bates	19	13	6/2**	11	11/4**	11	12/1***/5**
14	Grudinin	17	18	4/2**	14	8/2***/2**	15	8/3***/1**
15	Carbone	11	9	6/2***/1**	15	7/2***/2**	14	9/2***/3**
16	Oliva	14	15	5/1***	16	6/1***/2**	15	9/2***/2**
16	Moal	11	13	4/1***/2**	16	6/1***/2**	17	6/1***/4**
18	Weng	10	18	3/1***/1**	18	4/1***/3**	18	4/1***/3**
18	QASDOM	8	15	4/1***/1**	18	6/1***/1**	18	6/1***/1**
20	Takeda-Shitaka	10	20	3/2**	20	3/1***/1**	18	4/1***/3**
21	Yan	6	21	2**	21	2**	22	2**
21	Xue	2	21	2**	21	2**	22	2**
23	Laine	3	23	1***	23	1***	21	2/1***/1**
24	ISCORE	1	24	1	24	1**	24	1**
25	Wallner	5	25	0	25	1	25	1
26	Wolfson	2	25	0	26	0	26	0
26	Wang	1	25	0	26	0	26	0
26	Anashkina	1	25	0	26	0	26	0

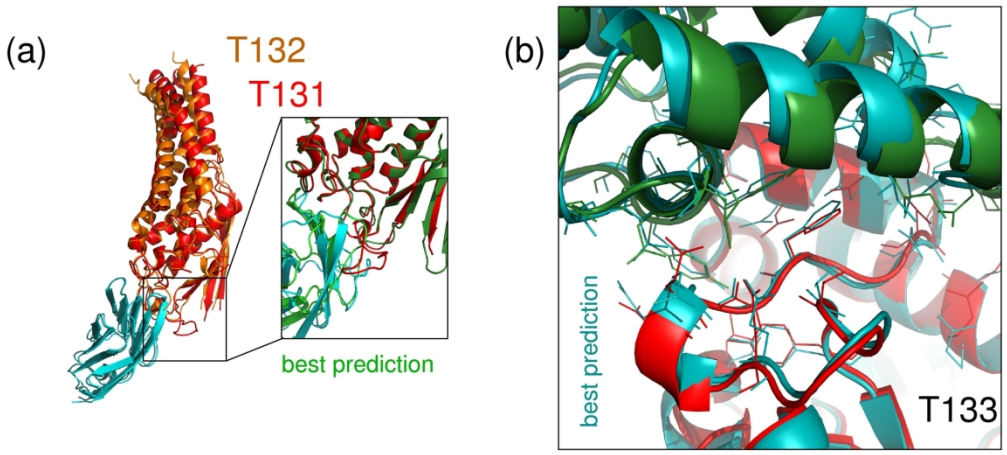
Table IV – Predictor and Scorer performance rankings in Rounds 38 – 45. **(a)** lists all Predictor groups, human and server combined; **(b)** lists only the automatic servers; **(c)** lists all Scorer groups, human and server combined. For all groups, the model of highest quality in their top 5 submission for every target is considered, and ranking is performed following the formula for $Score\_G$ as outlined in the text. Groups are listed by the name of the corresponding PI. Servers are listed by their acronym. The performance is listed by the number of targets for which a model of acceptable quality or better was produced, followed by the number of these models that were of high (***) and/or medium (**) quality. The total number of "#targets" to score was 19, three more than the actual 16 targets, because the PorM-Ct dimeric interface (T124.2) and the interfaces 1, 2 and 4 of T125 were considered independent and are therefore counted separately.

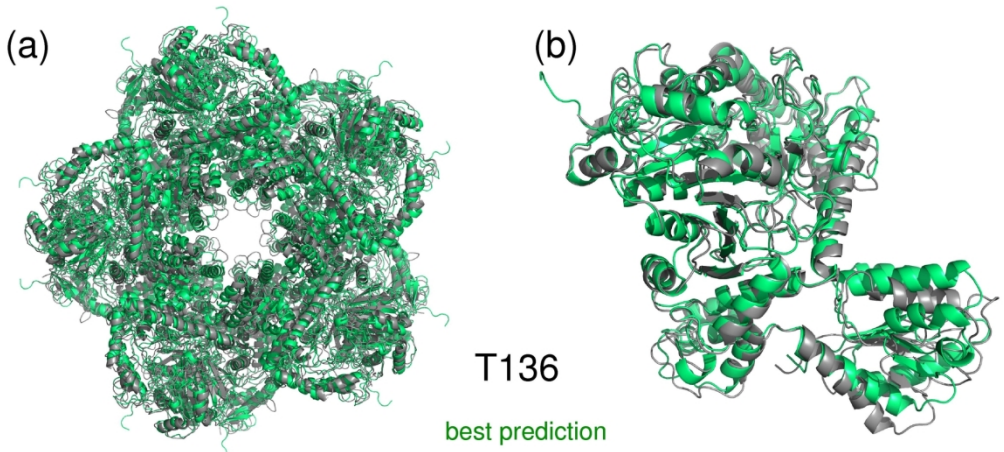












KYESEKRLGSERRLLGLRGESPELDLSYSHSDLGKRPTKDSYTLTEELAEYAEIRVK	#model[a]	#P[b]	#S[c]
		(31)	(18)
RPTKDSYTLTEEL	67	11	14
KYESEKRLGSERRLLGLRGESPELDLSYSHSDLGKRPTKDSYTLTEELAEYAEIRVK	51	4	4
EELAEYAEIRVK	39	4	4
SPELDLSYSHSD	28	10	10
LGSERRLLGLRG	23	4	11
KRPTKDSYTLTE	20	6	5
PELDLSYSHSDL	18	9	4
GESPELDLSYSH	15	5	9
ELDLSYSHSDLG	12	7	4
ESPELDLSYSHS	12	7	2
PTKDSYTLTEEL	12	6	5
SYTLTEELAEYA	11	4	1
SDLGKRPTKDSY	11	4	1
GSERRLLGLRGE	10	6	3
ESEKRLGSERRLLGLRGESPELDLSYSHSDLGKRPTKDSYTLTEELAEYAE	10	1	0
RRLLGLRGESPE	9	5	4
104 submissions skipped			
TKDSYTLTEELA	3	3	0
35 submissions skipped			

[a]number of submitted models; [b]number of Predictor groups; [c]number of Scorer groups