Protein-protein and peptide-protein docking and refinement using ATTRACT in CAPRI

Christina Schindler, Isaure Chauvot de Beauchêne, Sjoerd de Vries, Martin Zacharias

HAL Id: hal-01505867

https://hal.archives-ouvertes.fr/hal-01505867

Submitted on 30 Nov 2020

Protein-protein and peptide-protein docking and refinement using ATTRACT in CAPRI

Running title: protein-protein docking and refinement
Keywords: protein-protein interaction; induced fit; elastic network model; docking minimization; protein-protein complex formation

Christina EM Schindler [1,2], Isaure Chauvot de Beauchêne [1],
Sjoerd de Vries[1] and Martin Zacharias [1,2]*

1 Physics Department T38, Technische Universität München, 85748 Garching,
Germany
2 Center for Integrated Protein Science Munich, 81377 München, Germany

The ATTRACT coarse-grained docking approach in combination with various types of atomistic, flexible refinement methods has been applied to predict protein-protein and peptide-protein complexes in CAPRI rounds 28–36. For a large fraction of CAPRI targets (12 out of 18), at least one model of acceptable or better quality was generated corresponding to a success rate of 67%. In particular for several peptide-protein complexes excellent predictions were achieved. In several cases, a combination of template-based modeling and extensive molecular dynamics-based refinement yielded medium and even high quality solutions. In one particularly challenging case, the structure of an ubiquitylation enzyme bound to the nucleosome was correctly predicted as a set of acceptable quality solutions. Based on the experience with the CAPRI targets, new interface refinement approaches and methods for ab-initio peptide-protein docking have been developed. Failures and possible improvements of the docking method with respect to scoring and protein flexibility will also be discussed.

*Correspondence should be addressed to: Martin Zacharias, Physik Department T38, Technische Universität München, James-Franck-Strasse 1, 85748 Garching, Germany, martin.zacharias@ph.tum.de

1 Introduction

Protein-protein and peptide-protein interactions are abundant in the cell and are involved in virtually all important biological processes. However, so far only a small fraction of complex structures has been characterized experimentally. Especially, structure determination of transient interactions between proteins is experimentally challenging and probably in many cases impossible to obtain. Since atomic structural knowledge is vital for understanding the biological roles of these interactions, protein-protein docking methods that predict the 3D structure of complexes and efficient refinement approaches have become increasingly important in structural biology. The Critical Assessment of PRediction of Interactions (CAPRI) experiment [22, 23, 24] has provided a framework for blind testing and comparative assessment of protein-protein docking and refinement methods. Apart from evaluating the current state of the field, CAPRI's diverse, challenging targets have stimulated the development of new and more sophisticated protocols and pushed the limits of what is achievable in protein-protein docking. Our protein–protein docking approach ATTRACT [48, 29, 31, 47, 30, 17, 42, 38, 12, 13] can predict protein-protein and protein-nucleic acid interactions and has already been used successfullly in various rounds of CAPRI [47, 31, 17, 14]. ATTRACT's main characteristics are its coarse-grained (CG) force field, the ability to incorporate conformational flexibility already during the initial large-scale search [29, 4, 30] and the possibility to dock any number of (protein) partners [42]. The CG model was derived from statistical analysis of protein-protein and protein-nucleic acid interfaces and is intermediate between a residue/base-level and full atomistic description. It represents each amino acid by up to four pseudoatoms (two for the backbone and one or two for the side chains) [17]. A systematic docking search consists of several potential energy minimizations starting from hundreds of thousands of initial configurations. To speed up the docking calculations, the potential energy can be precalculated on a grid [**deVries2016attractgrid**]. Global flexibility (e.g., domain–domain motion) can be included explicitly during docking by energy minimization along the directions of precalculated soft normal modes [30]. Side chain and loop conformational changes can be accomodated by a multi-copy [4] or an ensemble docking approach. The flexible interface refinement method iATTRACT [38] can be used to further optimize the rigid body docking solutions. We have participated in CAPRI rounds 28–36 and in the following report on our predictions for targets 59–107 and related new methodological developments. This also includes our efforts to design a fully blind peptide-protein protocol and molecular dynamics-based refinement approaches.

2 Methods

2.1 ATTRACT rigid body docking

The protein and peptide structures were converted to the ATTRACT atom type representation [48] with the ATTRACT tool reduce. The empirical coarse grained force field in ATTRACT represents the amino acid side chains by one or two pseudoatoms and

takes all polar backbone atoms without hydrogens into account. Interactions are based on LJ type potentials and can be either attractive or repulsive (using a saddle point instead of an energy minimum in the LJ-potential) [17]. In addition, interactions between charged residues are calculated via a Coulomb term with a distance-dependent dielectric constant ($\varepsilon = 15r$). Starting points were generated by choosing random positions and orientations for the association partners with an appropriate center-of-mass distance to prevent steric overlap in the initial configuration. The starting structures are subjected to rigid body optimizations in a potential energy minimization of 1,000 minimization steps with the ATTRACT metric minimizer [29, 30]. Energy calculation was accelerated using a precalculated grid [**deVries2016attractgrid**]. In cases where distance restraints were employed, the optimization of the six rigid body degrees of freedom was preceded by a minimization in which the center of mass positions were fixed and the docking partners could orient towards each other. During this stage, only the restraint potentials were applied ("ghost" mode). If multiple conformations were available for the association partners, these conformations were docked separately (ensemble docking). Finally, the docking models were ranked by their ATTRACT energy evaluated within a squared cutoff of 50 Å^2 and highly similar models were removed with the deredundant tool.

2.2 iATTRACT flexible interface refinement

The protein and the peptide structures were converted into the OPLS atom type description with the ATTRACT tool aareduce. Missing hydrogens were built with PDB2PQR [15, 16] and protonation states were determined by PropKa[27]. Peptide termini were charged (unless the peptides were part of a larger protein), protein termini left uncharged. The atomistic refinement uses a physical force field based on the OPLS parameters to calculate non-bonded and electrostatic interactions between the protein partners. Contacts from the input structure are treated as flexible during a simultaneous potential energy minimization in rigid body degrees of freedom and interface flexibility [38]. A structure-based force field is determined on-the-fly to evaluate intra-protein interactions for the flexible interface residues. Depending on the size of the target, a few hundreds to thousand models from rigid body docking with ATTRACT were selected for iATTRACT refinement. The refinement parameters were chosen as specified in [38]. Structures were in general not rescored after iATTRACT refinement.

2.3 Molecular dynamics refinement

Atomistic refinement simulations were performed sing the AMBER14 molecular dynamics package [9] in combination with an implicit solvent description. The structures were converted to the AMBER atom type description using the pdb4amber tool. A Generalized-Born (GB) implicit solvent model (igb=8) was used with the newest version of the AMBER force field ff14SB [9]. The structures were first minimized with the sander program (500 steepest descend steps) with a short cutoff and a small initial step size to relax possible atom overlap and deformations resulting from the structure-based force

field used in iATTRACT refinement. Then two short molecular dynamics (MD) simulations were run with the pmemd.cuda program of the AMBER package for 5000 and 2500 time steps (2 fs) at temperatures $T = 400K$ and $T = 350K$ respectively. During the MD simulations, intra-molecular distances and inter-molecular distances between backbone atoms were restrained to prevent large deformations and dissociation of the binding partners. The intra-molecular distances were restrained with a harmonic potential to the distance found in the unbound structure with force constant of 2 kcal/(mol $Å^2$). The inter-molecular distances were allowed to change by up to 10 Å with respect to the distance found in the initial structure. Larger deviations beyond 10 Å were penalized by a harmonic potential with force constant 0.25 kcal/(mol $Å^2$). Finally, the structures were minimized for 5000 steps with a large cutoff using the pmemd.cuda program without restraints. The force field energy was evaluated for the complex and the individual protein partners by the sander program. The binding interaction energy score was calculated by subtracting the energy of the free protein partners from the energy of the complex. The final models were ranked by their binding interaction energy score without using any reweighting scheme of the energy contributions. Note, that the pmemd.cuda program employs graphical processing units (GPUs) and results in refinement simulations within seconds to minutes for averaged-sized complexes (100-200 aminos acids per partner protein). This translates to refinement of hundreds to thousands of complexes per day on a single GPU node.

For refinement in explicit solvent with GROMACS version 4.6 (www.gromacs.org) [6, 35, 1], structures were converted into the gro format with the tool pdb2gmx. Simulations were run with explicit solvent using the TIP3P water model and the AMBER99SB-ILDN force field at a temperature of 300 K. The positions of the backbone atoms were restrained with force constant 1000 kJ/mol/$Å^2$ in x, y and z direction.

3 Results and Discussion

In CAPRI rounds 28–36, we submitted predictions for targets 59-67, 95-97 and 102-107 (we did not participate in the CASP-CAPRI experiment in round 30 and in CAPRI round 32; round 36 was canceled before the submission deadline). Since the templates for the targets in round 32 (targets 98–101) were all of very low sequence identity we did not attempt to model these targets and to predict possible complex structures. Note, that indeed none of the predictors achieved any successful prediction for round 32. Table 1 shows a summary of the results. The targets vary strongly in terms of interaction type and docking difficulty. Especially, many targets of the most recent CAPRI rounds involved a high degree of flexibility and consequently, results for all predictor groups were rather poor. In the following, we discuss our predictions and the challenges we faced for some of the targets.

3.1 Round 28 (Targets 59–64)

Round 28 consisted of one protein-protein complex (target 59) and five peptide-protein complexes (targets 60–64). Target 59 corresponded to the Edc3 LSm domain, an ac-

Table 1 Results for ATTRACT predictions in CAPRI rounds 28-36.

Target	Type	Best model	fnat	IRMSD [Å]	Classification
59	protein-protein	5	0.18	3.8	acceptable (2⋆)
60	peptide-protein	5/6	1.0/0.94	0.44	high (2⋆ ⋆ ⋆,4⋆⋆)
61	peptide-protein	1	0.76	0.5	medium (5⋆⋆)
62	peptide-protein	1	0.92	0.39	high (2⋆ ⋆ ⋆, 3⋆⋆)
63	peptide-protein	1	0.81	0.49	high (2⋆ ⋆ ⋆, 3⋆)
64	peptide-protein	2	0.87	0.42	high (3⋆ ⋆ ⋆, 2⋆⋆)
65	peptide-protein	2	0.27	3.7	incorrect
66	peptide-protein	1/1	0.5/0.75	1.3/2.1	acceptable (1⋆)
67	peptide-protein	5	0.88	0.8	medium (2⋆⋆, 8⋆)
68–94	protein-protein	-	-	-	no submission
95	protein-protein	3	0.5	3.3	acceptable (5⋆)
96	protein-protein	3	0.15	2.96	acceptable (1⋆)
97	protein-protein	8	0.05	12.2	incorrect
98–101	protein-protein	-	-	-	no submission
102	protein-protein	-	-	-	not yet available
103	protein-protein	10	0.27	12.2	incorrect
104	interfacial water	4	0.68	0.92	high (1⋆ ⋆ ⋆,9⋆⋆), water (5++)
105	interfacial water	3	0.66	1.2	medium (10⋆⋆), water (7+++)
106	protein-protein	-	-	-	not yet available
107	protein-protein	7	0.05	17.42	incorrect
108-109		-	-	-	canceled

tivator of the mRNA decapping complex, with a motif from Rps28B. NMR ensembles of structures for the LSm domain were available both in the apo form and bound to another motif and we used all NMR models removing the highly flexible regions prior to docking [40, 19]. For Rps28B, we used an ensemble of homology models created with MODELLER [] based on available structures [46, 2]. We performed ATTRACT ab-initio rigid body docking and subsequent molecular dynamics refinement. For this target, we obtained two acceptable structures. The best submitted model deviated from the native structure at the interface by 3.8 Å and retrieved 18 % of the native contacts. To date, a more detailed analysis for this target is not possible, since the experimental structure has not yet been published. Target 60–64 were structures of importin-α binding to different peptides derived from nuclear localization signals (NLSs). NLSs contain one or two clusters of basic residues and are recognized by the import receptor importin-α [10]. Several structures of importin-α in complex with different peptides were available at the time of round 28 which showed two binding sites: a major and a minor site. The peptides in Round 28 were derived from nuclear localization signals and had been optimized towards binding to the minor site [10]. In the crystal structure, the major site was also occupied and so evaluation of these targets was carried out for both the major and the minor site. We generated models for the peptide based on the available crystal structures and used this in an ATTRACT rigid body docking with ambiguous distance restraints twoards the major and the minor sites. The best models were then refined by an energy minimization with AMBER. For all but one target, we achieved
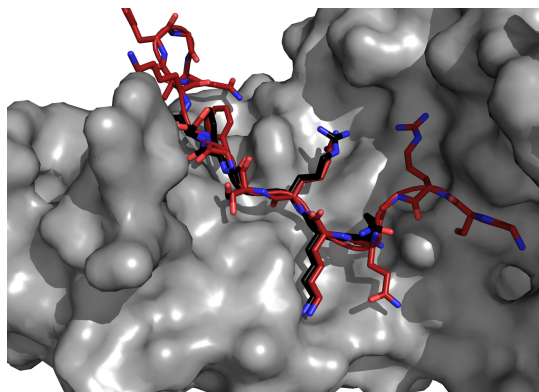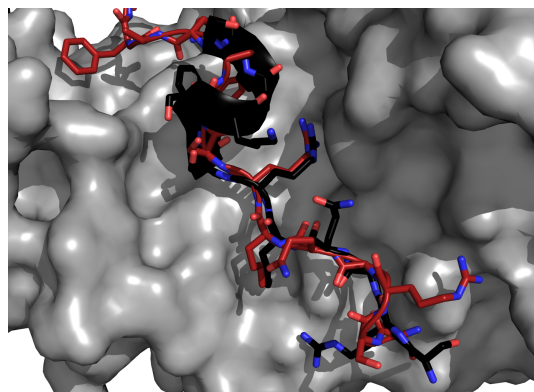
A Major site.   **B** Minor site.

Figure 1 Peptide-protein docking model for target 60 (docked peptide indicated as red sticks, importin-α receptor shown as grey surface) superimposed on the native structure (PDB 3ZIN, bound peptide in black). Peptide binding to importin-α was modeled using a combination of homology modeling and molecular dynamics refinement. For this target, and related targets 62-64, several three-star quality models were submitted

high quality predictions (Figure 1). However, we failed to predict the $\alpha-$helical turn in the peptide when binding to the minor site. Here, possibly more extensive peptide structure modeling [39] prior to complex prediction could have improved the results.

3.2 Round 29 (Targets 65–67)

In Round 29, three distinct peptide-protein complexes were proposed as targets. Target 65 and 66 were complexes of proteins with a SSB C-terminal peptide. For both targets, we analyzed available structures of other proteins in complex with a SSB C-terminal peptide (PDB 3UF7, 3Q8D, 3C94) [36, 28] and identified a conserved binding mode with an hydrophobic anchor at the C-terminus and a solvent exposed part of the peptide, although in general the peptide appeared relatively flexible. We extracted the peptide conformations from the available complexes, recombined parts of the peptides from different crystal structures to generate additional conformations and used this peptide ensemble in a fully blind coarse-grained search of the entire protein surface combined with two stages of flexible refinement [37]. Before refinement, the rigid body models were filtered to detect models with a buried phenylalanine at the C-terminus.

Target 65 was a complex with RNase Hi[34]. For this target we did not generate any near-native prediction, since we did not sufficiently model the protein's flexibility at the peptide binding site (we only used a single crystal structure for the protein during docking [20]) and failed to detect the correct binding pocket. Target 66 corresponded to a PriA helicase in complex with a SSB C-terminal peptide [7]. We obtained one model of acceptable quality (top 1) and one model (top 2) which had a very similar orientation of the peptide but an incorrect peptide conformation (α-helical turn) (Figure 2). In our top-ranked model, the whole peptide was in contact with the protein surface, whereas in the
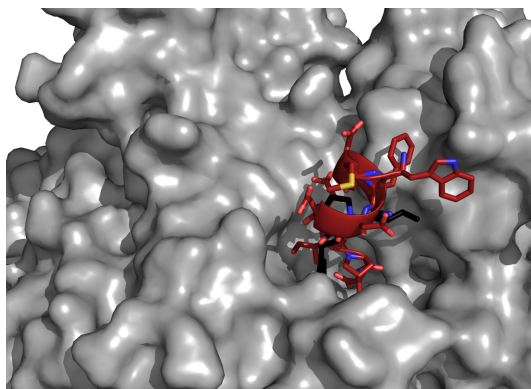
A top 1
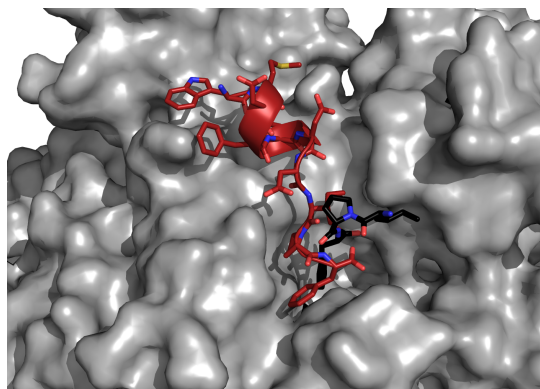
B top 2

Figure 2 Predicted structural models (top1 and top2) for target 66 (docked peptide in red, receptor as grey surface) superimposed on the native structure (PDB 4NL8, bound peptide in black).

crystal structure the N-terminal part of the peptide is solvent-exposed. The tendency to maximize the interface between a flexible peptide and a binding region can be attributed to the use of an implicit solvent model during final refinement (see Methods). In future cases, we will consider explicit solvation during refinement of peptide-protein complexes which may help to improve the accuracy of the predictions.

Target 67 corresponded to a WW domain in complex with an extended peptide containing proline (P) residues. Several structures of PPXY motifs bound to WW domains were available in the PDB [8] and we used these to build initial peptide-protein complexes. Since only a limited number of complexes were generated it was possible in this case to use restrained molecular dynamics simulations in explicit solvent with GRO-MACS for refinement (1 ns simulations at room temperature and normal pressure, see Methods). Finally, the models were energy-minimized with sander of the AMBER package. All our 10 submitted models were at least of acceptable quality and the best model had an IRMSD of 0.8 Å and retrieved 88 % of the native contacts (medium quality).

3.3 Round 31 (Targets 95–97)

Round 31 comprised three very challenging protein-protein complexes as targets. For target 95, binding of an ubiquitylation enzyme PRC1 (Bmi1/Ring1b ubiquitin ligase) to the nucleosome had to be predicted [32]. PRC1 ubiquitylates the histone H2A tail at residue LYS 119 [33, 44]. Furthermore, several residues on PRC1 (ASP 56, LYS 92, LYS 93, LYS 97 and ARG 98 on Ring1b and LYS 62 and ARG 64 on Bmi1) had been identified as important for binding by mutational experiments [5]. There was also experimental evidence for PRC1 binding to DNA (although only for isolated DNA, not for the nucleosome) [5] and data pointing to an important role for the acidic patch in PRC1 function [26]. We used the unbound protein structures for PRC1 [5] and the nucleosome [41] and performed a large scale rigid body docking search with a 10 Å upper harmonic distance restraint between the C_α atoms of residue LYS 118 on the H2A tail
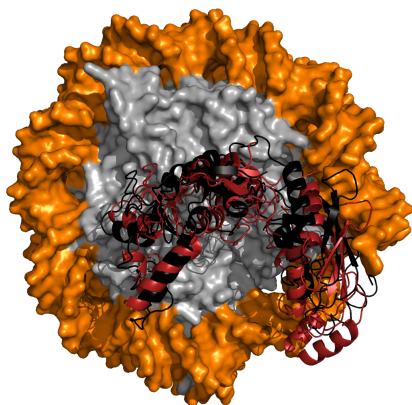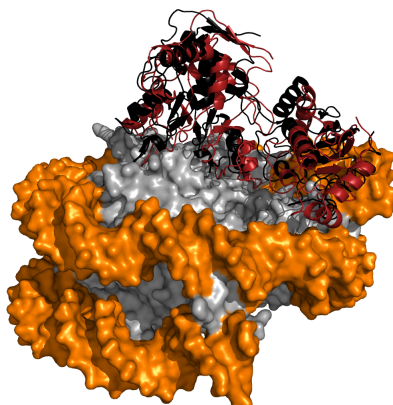
A Top view.   **B** Side view.

Figure 3   Best model (top 3) for target 95 superimposed on the native structure (PDB 4R8P, black). PRC1 binds to the nucleosome at the histone acidic patch.

(LYS 119 was not resolved in the structure) and residue CYS 85 in the active site of the ubiquitin ligase. Subsequently, we refined the best clusters with iATTRACT and short molecular dynamics refinement with AMBER. We found two types of solutions among the top-ranked models: one with PRCC1 contacting the acidic patch and one with PRCC1 binding to nucleosomal DNA similar to an earlier model [5]. We hence submitted 5 models for each type of solution and achieved 5 acceptable model with the best model being very close to medium quality (LRMSD 5.03 Å, fnat 0.5 Å, Figure 3). A posterior comparison of our models with the experimental crystal structure of the complex indicated a conformational adjustment of a histone tail upon PRCC1 binding not accounted for in our docking efforts. Hence, the sterical clash of the docked PRCC1 with the histone tail in the unbound nucleosome structure limited the accuracy of our models to the level of acceptable CAPRI solutions. Similar to other CAPRI cases it appears to be necessary in the future to carefully check for highly flexible segments (termini or loops) in partner structures prior to docking and possibly eliminate these elements during docking followed by rebuilding after docking.

Targets 96 and 97 were complexes of designed α-repeat proteins binding to GFP (PDB 4XL5 and 4XVP) [11]. We used our standard ab-initio docking protocol in combination with additional scoring with the Rosetta force field which has been used in the past for protein design [43, 21]. We achieved acceptable quality for target 96. However, we failed to accurately predict the smaller target 97. In this case docking sampling gave solutions where the ligand protein (GFP) in our models was rotated by 180° with respect to the native structure. This failure can be attributed to deficiencies of the scoring of the docking models selected for further refinement.

3.4 Round 33 and 34 (Targets 102–105)

In Round 33, prediction of two protein-protein complexes was offered (targets 102 and 103). Results for target 102 were not available at the time of writing. For target 103, we did not achieve any successful predictions although for one submitted model an overlap with the correct binding interface was achieved for model 10 (fnat=0.27, Table 1). In Round 34, the challenge was to model the placement of interfacial waters. The targets were pyocin DNase domains in complex with immunity proteins which were structurally similar to previously solved structures in colicin. We built an initial model of the complex by homology modeling using a colicin complex [45] and refined this with backbone restrained molecular dynamics simulations in explicit solvent. We achieved a high quality model for target 104 and medium quality models for target 105. For the prediction of the interfacial water placements we followed our molecular dynamics based protocol described previously [14, 25]. Briefly, in this protocol water molecules are placed using the AMBER leap module allowing for partial overlap of waters with solute atoms at the interface (resulting in a slight over-hydration of the interface). The hydration structure is then allowed to relax during short MD simulation including positional restraints on the protein backbone. Finally, all waters outside the interface are removed followed by energy minimization to optimize the position and orientation of each water molecule at the interface (see details in reference [25]). Interestingly, the prediction of interfacial waters was more accurate for target 105 (medium quality model), probably because there were less water molecules to predict for this target. Interestingly, in the overall CAPRI evaluation it was found that high accuracy water predictions were not necessarily generated only for high quality models of the protein-protein complex (Lensinck, unpublished data). This is in contrast to previous results for interfacial water predictions in CAPRI [25] and indicates that probably only certain key groups at the interface have to be in near-native position to allow correct placement of waters.

3.5 Round 35 (Targets 102–103 and 106–107)

In these two rounds, the targets were a complex of haemopexin with the haemopexin utilization protein (huxA). The complex was solved in the apo state and with haem bound. It was previously speculated that the C-terminal domain of huxA interacts with haemopexin [18]. However, at the time of submission for Round 33, structural data was only available for the N-terminal secretion domain [3]. Hence, the target was proposed again as a target in Round 35 offering an unbound structure of full-length huxA. Still, we were not able to generate any near-native predictions, since binding of huxA to haemopexin involves a conformational change in a loop on HuxA which is partly disordered in the unbound form. Unfortunately, this loop yields steric clashes with haemopexin when superimposing the unbound huxA structure on the complex. A strategy involving detailed flexibility analysis prior to docking, removal of flexible loops and a-posteriori loop rebuilding would have certainly yielded improved results for this target.

4 Conclusion and Outlook

In CAPRI Rounds 28-36, we submitted predictions for 18 targets (16 distinct targets) and achieved at least acceptable predictions in 12 cases which can be considered as highly successful given the difficulty of many of the targets. The coarse-grained ATTRACT docking approach in combination with different refinement schemes proved to be versatile in dealing with a variety of different targets that ranged from peptide-protein interactions to docking of large proteins to the nucleosome. The CAPRI challenge has also triggered the development of several extensions of the original ATTRACT approach in the area of peptide-protein docking (pepATTRACT) [37] and of refinement (iATTRACT) [38] within the last rounds. For several CAPRI targets, we found that homology modeling in combination with restrained molecular dynamics refinement yielded high quality predictions. In some cases, significant conformational change or inaccuracies in the homology modeling have contributed to the failure of the docking search. More detailed flexibility analysis prior to docking is needed to select appropriate conformational ensembles and to identify flexible loops. It might be beneficial in many cases to eliminate such highly flexible loops prior to docking if the chances to correctly model the bound forms are small and to tackle the generation of a bound loop structure after docking. Such approaches may also be useful for improving protein-protein complexes predicted based on sequence similarity to a known template complex. Furthermore, improvements in scoring towards distinguishing structures of medium or higher quality from just acceptable solutions is highly desirable and at the focus of our future research.

Acknowledgements

References

[1] Mark James Abraham et al. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1 (2015), pp. 19–25.

[2] James M Aramini et al. "Solution NMR structure of the 30S ribosomal protein S28E from Pyrococcus horikoshii". In: *Protein science* 12.12 (2003), pp. 2823–2830.

[3] Stéphanie Baelen et al. "Structure of the secretion domain of HxuA from Haemophilus influenzae". In: *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 69.12 (2013), pp. 1322–1327.

[4]   Karine Bastard, Chantal Prévost, and Martin Zacharias. "Accounting for loop flexibility during protein–protein docking". In: *Proteins: Structure, Function, and Bioinformatics* 62.4 (2006), pp. 956–969.

[5]   Matthew L Bentley et al. "Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex". In: *The EMBO journal* 30.16 (2011), pp. 3285–3297.

[6]   Herman JC Berendsen, David van der Spoel, and Rudi van Drunen. "GROMACS: a message-passing parallel molecular dynamics implementation". In: *Computer Physics Communications* 91.1 (1995), pp. 43–56.

[7]   Basudeb Bhattacharyya et al. "Structural mechanisms of PriA-mediated DNA replication restart". In: *Proceedings of the National Academy of Sciences* 111.4 (2014), pp. 1373–1378.

[8]   Romel Bobby et al. "Structure and dynamics of human Nedd4-1 WW3 in complex with the αENaC PY motif". In: *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1834.8 (2013), pp. 1632–1641.

[9]   DA Case et al. "AMBER 14". In: *University of California, San Francisco* (2014).

[10]   Chiung-Wen Chang et al. "Distinctive Conformation of Minor Site-Specific Nuclear Localization Signals Bound to Importin-α". In: *Traffic* 14.11 (2013), pp. 1144–1154.

[11]   Anne Chevrel et al. "Specific GFP-binding artificial proteins (αRep): a new tool for in vitro to live cell applications". In: *Bioscience reports* 35.4 (2015), e00223.

[12]   Sjoerd J de Vries et al. "A web interface for easy flexible protein–protein docking with ATTRACT". In: *Biophys J* 108.3 (2015), pp. 462–465.

[13]   Sjoerd J de Vries et al. "Cryo-EM data is superior to contact or interface information in integrative modeling". In: *Biophys. J.* 110.4 (2016), pp. 462–465.

[14]   Sjoerd de Vries and Martin Zacharias. "Flexible docking and refinement with a coarse-grained protein model using ATTRACT". In: *Proteins: Struct, Funct, Bioinf* 81.12 (2013), pp. 2167–2174.

[15]   Todd J. Dolinsky et al. "PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations". In: *Nucleic Acids Res* 32.suppl 2 (2004), W665–W667.

[16]   Todd J. Dolinsky et al. "PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations". In: *Nucleic Acids Res.* 35.suppl 2 (2007), W522–W525.

[17]   Sébastien Fiorucci and Martin Zacharias. "Binding site prediction and improved scoring during flexible protein–protein docking with ATTRACT". In: *Proteins: Struct., Funct., Bioinf.* 78.15 (2010), pp. 3131–3139.

[18]   Clémence Fournier, Ann Smith, and Philippe Delepelaire. "Haem release from haemopexin by HxuA allows Haemophilus influenzae to escape host nutritional immunity". In: *Molecular microbiology* 80.1 (2011), pp. 133–148.

[19] Simon A Fromm et al. "The structural basis of Edc3-and Scd6-mediated activation of the Dcp1: Dcp2 mRNA decapping complex". In: *The EMBO journal* 31.2 (2012), pp. 279–290.

[20] Kohki Ishikawa et al. "Stabilization of Escherichia coli ribonuclease HI by cavity-filling mutations within a hydrophobic core". In: *Biochemistry* 32.24 (1993), pp. 6171–6178.

[21] Kristian W Kaufmann et al. "Practically useful: what the Rosetta protein modeling suite can do for you". In: *Biochemistry* 49.14 (2010), pp. 2987–2998.

[22] Marc F. Lensink, Raúl Méndez, and Shoshana J. Wodak. "Docking and scoring protein complexes: CAPRI 3rd Edition". In: *Proteins: Struct., Funct., Bioinf.* 69.4 (2007), pp. 704–718.

[23] Marc F. Lensink and Shoshana J. Wodak. "Docking and scoring protein interactions: CAPRI 2009". In: *Proteins: Struct., Funct., Bioinf.* 78.15 (2010), pp. 3073–3084.

[24] Marc F. Lensink and Shoshana J. Wodak. "Docking, scoring, and affinity prediction in CAPRI". In: *Proteins: Struct., Funct., Bioinf.* 81.12 (2013), pp. 2082–2095.

[25] Marc F Lensink et al. "Blind prediction of interfacial water positions in CAPRI". In: *Proteins: Structure, Function, and Bioinformatics* 82.4 (2014), pp. 620–632.

[26] Justin W Leung et al. "Nucleosome acidic patch promotes RNF168-and RING1B/BMI1-dependent H2AX and H2A ubiquitination and DNA damage signaling". In: *PLoS Genet* 10.3 (2014), e1004178.

[27] Hui Li, Andrew D. Robertson, and Jan H. Jensen. "Very fast empirical prediction and rationalization of protein pKa values". In: *Proteins: Struct, Funct, Bioinf* 61.4 (2005), pp. 704–721.

[28] Duo Lu and James L Keck. "Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I". In: *Proceedings of the National Academy of Sciences* 105.27 (2008), pp. 9169–9174.

[29] Andreas May and Martin Zacharias. "Accounting for global protein deformability during protein–protein and protein–ligand docking". In: *Biochim. Biophys. Acta, Proteins Proteomics* 1754.1–2 (2005), pp. 225–231.

[30] Andreas May and Martin Zacharias. "Energy minimization in low-frequency normal modes to efficiently allow for global flexibility during systematic protein–protein docking". In: *Proteins: Struct, Funct, Bioinf* 70.3 (2008), pp. 794–809.

[31] Andreas May and Martin Zacharias. "Protein–protein docking in CAPRI using ATTRACT to account for global and local flexibility". In: *Proteins: Struct, Funct, Bioinf* 69.4 (2007), pp. 774–780.

[32] Robert K McGinty, Ryan C Henrici, and Song Tan. "Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome". In: *Nature* 514.7524 (2014), pp. 591–596.

[33] Mariana de Napoles et al. "Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation". In: *Developmental cell* 7.5 (2004), pp. 663–676.

[34] Christine Petzold et al. "Interaction with single-stranded DNA-binding protein stimulates Escherichia coli Ribonuclease HI enzymatic activity". In: *Journal of Biological Chemistry* 290.23 (2015), pp. 14626–14636.

[35] Sander Pronk et al. "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit". In: *Bioinformatics* (2013), btt055.

[36] Mikhail Ryzhikov et al. "Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein". In: *Nucleic acids research* 39.14 (2011), pp. 6305–6314.

[37] Christina EM Schindler, Sjoerd J de Vries, and Martin Zacharias. "Fully blind peptide-protein docking with pepATTRACT". In: *Structure* 23.8 (2015), pp. 1507–1515.

[38] Christina EM Schindler, Sjoerd J de Vries, and Martin Zacharias. "iATTRACT: Simultaneous global and local interface optimization for protein–protein docking refinement". In: *Proteins: Struct., Funct., Bioinf.* 83.2 (2015), pp. 248–258.

[39] Pierre Thévenet et al. "PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides". In: *Nucleic Acids Res.* 40.W1 (2012), W288–W293.

[40] Felix Tritschler et al. "A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting". In: *Molecular and cellular biology* 27.24 (2007), pp. 8600–8611.

[41] Dileep Vasudevan, Eugene YD Chua, and Curt A Davey. "Crystal structures of nucleosome core particles containing the 601strong positioning sequence". In: *Journal of molecular biology* 403.1 (2010), pp. 1–10.

[42] SJ de Vries and M Zacharias. "ATTRACT-EM: a new method for the computational assembly of large molecular machines using cryo-EM maps." In: *PLoS One* 7.12 (2012), e49733–e49733.

[43] Chu Wang, Philip Bradley, and David Baker. "Protein–protein docking with backbone flexibility". In: *Journal of molecular biology* 373.2 (2007), pp. 503–519.

[44] Hengbin Wang et al. "Role of histone H2A ubiquitination in Polycomb silencing". In: *Nature* 431.7010 (2004), pp. 873–878.

[45] Justyna Aleksandra Wojdyla et al. "Structure of the ultra-high-affinity colicin E2 DNase–Im2 complex". In: *Journal of molecular biology* 417.1 (2012), pp. 79–94.

[46] Bin Wu et al. "Solution structure of ribosomal protein S28E from Methanobacterium thermoautotrophicum". In: *Protein science* 12.12 (2003), pp. 2831–2837.

[47] Martin Zacharias. "ATTRACT: protein–protein docking in CAPRI using a reduced protein model". In: *Proteins: Structure, Function, and Bioinformatics* 60.2 (2005), pp. 252–256.

[48] Martin Zacharias. "Protein–protein docking with a reduced protein model accounting for side-chain flexibility". In: *Protein Sci.* 12.6 (2003), pp. 1271–1282.