Coarse-grain modelling of protein-protein interactions

ARTICLE *in* CURRENT OPINION IN STRUCTURAL BIOLOGY · OCTOBER 2013

Impact Factor: 7.2 · DOI: 10.1016/j.sbi.2013.09.004 · Source: PubMed

2 AUTHORS:

Marc Baaden
Institute of Physical and Chemical Biology
106 PUBLICATIONS **1,863** CITATIONS
SEE PROFILE

Siewert J Marrink
University of Groningen
202 PUBLICATIONS **13,694** CITATIONS
SEE PROFILE

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Coarse-grain modelling of protein–protein interactions

Marc Baaden[1,3] and Siewert J Marrink[2,3]

Here, we review recent advances towards the modelling of protein–protein interactions (PPI) at the coarse-grained (CG) level, a technique that is now widely used to understand protein affinity, aggregation and self-assembly behaviour. PPI models of soluble proteins and membrane proteins are separately described, but we note the parallel development that is present in both research fields with three important themes: firstly, combining CG modelling with knowledge-based approaches to predict and refine protein–protein complexes; secondly, using physics-based CG models for *de novo* prediction of protein–protein complexes; and thirdly modelling of large scale protein aggregates.

Addresses
[1] Laboratoire de Biochimie Théorique, CNRS, UPR9080, Univ Paris Diderot Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, Paris 75005, France
[2] Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, AG Groningen 9747, The Netherlands

Corresponding author: Baaden, Marc (baaden@smplinux.de)
[3] MB and SJM are co-corresponding authors.

Current Opinion in Structural Biology 2013, **23**:878–886

This review comes from a themed issue on **Protein-protein interactions**

Edited by **Joël Janin** and **Alexandre M J J Bonvin**

For a complete overview see the Issue and the Editorial

Available online 27th October 2013

0959-440X/$ – see front matter,

http://dx.doi.org/10.1016/j.sbi.2013.09.004

Introduction

The cytoplasm contains thousands of different proteins, at an estimated concentration of 200–300 mg/ml. Together with other biomolecular constituents, the total macromolecular concentration (proteins, lipids, nucleic acids, and sugars) of the cytoplasm can be as high as 400 mg/ml and occupy 20–30% of its volume [1]. In such a crowded environment, protein–protein interactions are essential, and in fact regulatory of almost all cellular processes. Likewise, membrane proteins function at high concentration in the cell membrane with a protein area coverage of 15–30% [2] and many membrane-related processes require formation of membrane protein (super) complexes. Therefore it is of fundamental importance to understand and predict protein–protein binding modes, association–dissociation equilibria, as well as mechanisms by which proteins avoid sticking to each other under such crowded conditions [3].

To understand the driving forces governing the organization of cellular components, computer simulation has become an indispensible tool. To deal with the range of spatio-temporal scales involved in cellular processes, modelling at a coarse-grained (CG) level of description is required. The use of reduced representations enables us to consider large and complex systems, often beyond binary complexes, towards multi-protein assemblies [4,5]. Multiple levels of coarse-graining are now available and CG models of proteins can be combined with other biomolecules to simulate the structure and dynamics of complex biosystems [6,7]. Increasingly the aim is to derive quantitative rather than qualitative data. The accurate prediction of protein–protein binding affinities is seen as a major target for future developments [8], for instance. The ultimate challenge is to predict the structure and dynamics of protein–protein complexes and multi-protein assemblies *de novo*, that is without prior knowledge of the bound state. These endeavours are limited by several challenges for CG computational methods [9], which form the topic of this review. We restrict ourselves to CG PPI models that retain chemical specificity, focusing on progress over the past few years. We first describe progress in CG modelling of soluble proteins, followed by membrane proteins, and we end with a short outlook section.

Soluble proteins

Coarse-graining using knowledge-based approaches

For many protein assemblies, experimental data is available that may be used as effective restraints to guide CG models. Such prior knowledge about protein–protein assemblies can be integrated in CG models at various levels. The spatial symmetry of an assembly, for instance, may be used as an original way to coarse-grain spatial degrees of freedom as has been demonstrated for the rational computational design and optimization of self-assembled protein nanomaterial interfaces [10]. Symmetry can furthermore be combined with experimental data such as cryo-EM maps to drive the computational assembly of protein complexes [11]. Another level of integration of prior knowledge is to directly embed reference data within knowledge-based CG potentials, for example by refining the description of PPI in docking scores [12,13]. A promising novel route to guide protein–protein assembly is to include multi-body interactions and evolutionary information in scoring functions for the prediction of protein interactions. This idea has been implemented in a CG statistical potential, showing significant improvement on a common protein docking benchmark [14•]. Knowledge-based CG representations may be designed based on publicly available databases

and resources, many of which are referenced at http://www.vls3d.com/links.html (section 12). It may be very profitable to combine such knowledge-based potentials with structure modelling approaches using low-resolution biochemical or biophysical data to devise integrative computational tools [15,16]. In this context it is important to incorporate inherent uncertainties into CG or multi-scale models, for example relating to the shapes and positions of proteins in an assembly. This can be achieved in a quantitative manner with so-called 'toleranced' models [17].

Physics-based simplified representations of protein–protein interfaces

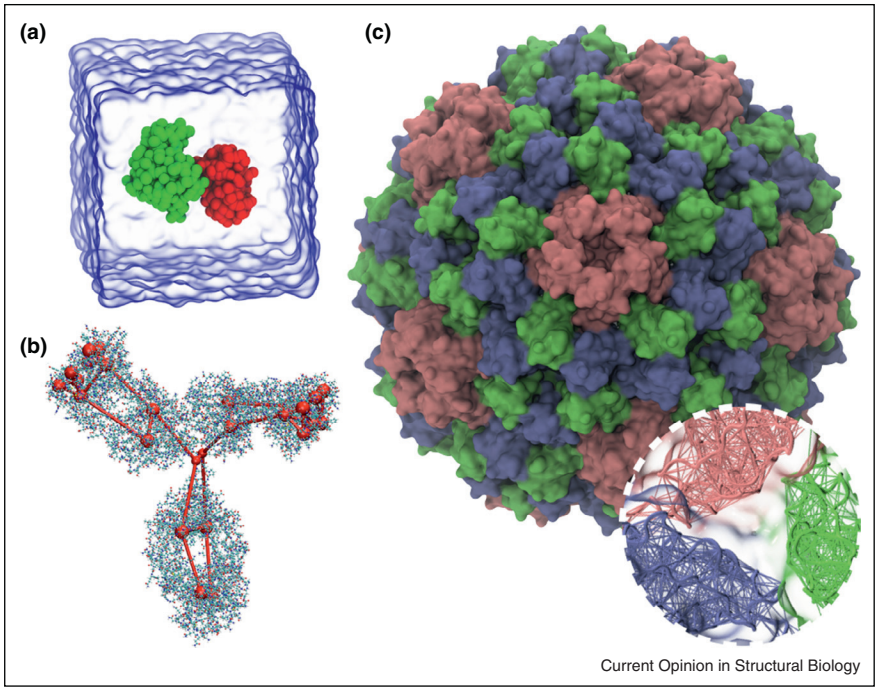
A major challenge in computational modelling of PPI is the *de novo* prediction of protein–protein interfaces. The large sampling space naturally invokes the use of CG representations. A recurrent issue with CG models, however, is the implementation of realistic structural flexibility. An original approach has been suggested for the coarse-grained ATTRACT representation, where a concurrent hybrid all-atom/CG representation was used [18] allowing full flexibility of the protein. In such a model, however, computational efficiency is a major challenge. The SCORPION CG force field was specifically designed to describe the dynamics of protein–protein recognition in water, focusing on the parameterization of non-bonded potentials that are compatible with a polarizable particle-based solvent model. The internal structure and flexibility of the protein is maintained via an elastic network model [19]. Starting from different initial configurations, the model reproduces the association of the Barnase/Barstar complex (Figure 1a) in Molecular Dynamics (MD) simulations. Such a dynamic description including surface flexibility may be very complementary to coarser static models used in Brownian Dynamics (BD) approaches [20]. Solernou and Fernandez-Recio introduced the pyDockCG CG potential for computational protein–protein docking [21], an extension of the UNRES model with additional terms accounting for a solvation energy contribution and explicit electrostatic interactions. Predictions are comparable to the all-atom pyDock scoring function but at a significantly lower computational cost, which enables the future introduction of flexibility at the interface. Concerning the importance of interface solvation, Ceres *et al.* showed that a solvation energy deficit term calculated at a CG level discriminates the nature of a protein interface patch between its monomeric and dimeric forms [22].

General principles of PPI can be derived from systematic comparative studies. Zheng *et al.* [23•] successfully predicted the quaternary structures of 12 protein–protein complexes starting from unbound and/or unfolded monomers. Their model recovers the native interfaces of eight homodimers and four heterodimers reconstructing binding free energy surfaces that enable the analysis of possible protein–protein recognition mechanisms. Their results suggest that protein–protein association is well represented by funnelled energy landscapes, similar to those employed for folding of monomers. Cross-docking experiments between large sets of arbitrary binding partners require reduced interface representations, one possibility among many being the description of protein shape as spherical harmonics. Using shape complementarity and physical properties in a cross-docking study of 314 probe and 198 target proteins, Martin and Lavery have shown that partner molecules dock in a non-random fashion and biologically pertinent recognition interfaces can be identified by such an approach [24].

Towards reliable modelling of large-scale aggregates: association–dissociation dynamics

In order to move towards the modelling of large and complex assemblies such as those encountered in a cell, it is important to adequately reproduce association–dissociation dynamics, including transient weak bindings either driven by enthalpic or entropic terms. A difficult task in the modelling of PPI is to account for flexible loops. Using a simple CG representation, De Simone *et al.* showed that large flexible loops actually hinder protein aggregation, suggesting that their role may be to maintain protein solubility [25]. The sampling of protein–protein aggregation can be accelerated by coarsening the phase space exploration as shown by Ravikumar *et al.* using a push–pull-release approach at a residue-level CG description of five well-characterized protein complexes to generate a comprehensive energy landscape picture of complex association [26]. Emperador *et al.* combine the use of discrete molecular dynamics with a multi-scale representation of protein–protein binding interfaces to introduce flexibility in the docking process [27]. They show a significant improvement of the predicted complexes, notably over rigid body docking, with particular benefits in the case of large conformational changes. Using a residue-level CG model of the actin–myosin complex with Debye–Huckel electrostatics implicitly representing ionic concentration and repulsive steric terms, Okazaki *et al.* showed that concentration and temperature dependent effects can be reproduced [28]. This study provides insights into the driving forces for protein–protein recognition. Modelling of actin filaments poses further challenges: Saunders and Voth [29] compare three F-actin models from the literature using CG molecular dynamics with a focus on structural and dynamic features. Although significant differences exist among the models, a coherent picture emerges when they are linked to different phases in the actin polymerization cycle. Coarse-grained modelling of the association pathway of β–β–α protein, a small and simple homotetrameric complex, provides an interesting test case for studying the underlying assembly mechanisms [30•]. The study

Figure 1



Protein aggregates of increasing complexity in solution. **(a)** Barnase–Barstar complex from [19] in green and red color solvated by a coarse-grained water box shown in transparent blue; image courtesy of N Basdevant, T Ha Duong and D Borgis. **(b)** Model of a monoclonal antibody from [31] with the all-atom structure in cyan and the coarse-grained model in red; image courtesy of A Chaudhri and GA Voth. **(c)** A model for the study of virus indentation revealing mechanical properties mediated by protein–protein interactions [34]. Coat proteins adopt slightly different folds of type a, b and c, colored pink, green and blue, respectively.

suggests the possibility that structures of protein assemblies adopt different forms in crowded environments, diluted solution and in a crystal structure. Modelling of monoclonal antibodies at high concentrations represents another challenge for coarse-graining self-association phenomena [31], requiring a reliable description of the antibody-antibody interface (Figure 1b). A different approach to treat concentrated protein solutions is to coarse grain the solvent representation, yet maintain an all-atom interface description using many-protein BD simulations [32].

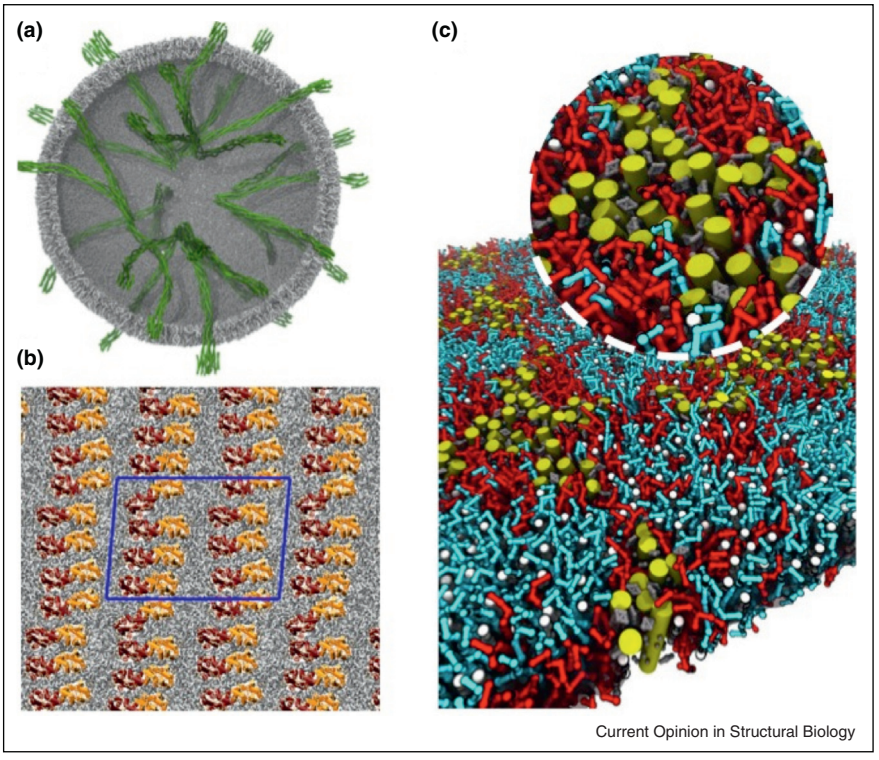
Viruses represent a long-standing benchmark for CG PPI studies as their shear size is very challenging, self-assembly is a typical feature and experimental reference data is abundant. They offer the possibility to exploit symmetry to guide the computational approach. Virus maturation is a process involving the concomitant rearrangement of many protein–protein interfaces. May *et al.* studied maturation pathways of bacteriophage HK97 with a biased CG potential suggesting that icosahedral motions govern the assembly process [33•]. With appropriate optimized CG models it was furthermore possible to predict inter-protein conformational flexibility and mechanical properties relevant to virus capsid deformation [34]. The study used an iteratively refined distance-based elastic network representation to maintain the inter-monomer arrangement within the assembly and derive accurate fluctuation dynamics of the ensemble (Figure 1c). In a previous study, atomistic and unbiased CG simulations were combined to characterize the geometrical interfaces at symmetry points within the capsid [35]. Cieplak and Robbins studied the mechanical response of 35 virus capsids to nanoindentation [36] and some studies indicate that even very coarse models can be used to represent mechanical stress in viral shells [37]. The HIV-1 capsid is another major target that has been modelled at a CG level [38] and more recently even at an all-atom level [39], providing a demanding testcase for protein–protein interactions.

Membrane proteins

Refinement of membrane protein complexes using experimental constraints

CG modelling is progressively used as an effective tool to optimize the structure of various membrane protein complexes. Similar to soluble protein complexes, optimization of membrane protein complexes is most effective in combination with experimental data. For instance, Steinbrecher *et al.* [40] used a number of biophysical assays together with CG modelling to study the structure and membrane interactions of the bacterial stress–response

Figure 2



Membrane protein aggregates. **(a)** Self-assembly of bacterial chemoreceptors in a small liposome from [49•]; image courtesy of BA Hall, JP Armitage, and MSP Sansom. A cross-section of the liposome is shown, with receptors depicted in green and lipids in grey. **(b)** Supramolecular organization of rhodopsins in disk membranes from [55•]; image courtesy of X. Periole. Top view of the membrane with pairs of rhodopsins shown in red/orange, lipids in grey. **(c)** Clustering of TM helices in multi component membranes under crowding conditions from [65]. Tilted view of the membrane with saturated lipids in cyan, unsaturated lipids in red, cholesterol in grey with white hydroxyl group, and WALP peptides in yellow with grey flanking residues.

peptide TisB, aiming to find out whether it forms pores or other proton-selective channels. On the basis of their findings, the authors postulate that antiparallel dimers could be assembled via a ladder of salt bridges that enable protons to pass along a wire of trapped water molecules across the membrane. Marius *et al.* [41] employed chemical cross-linking and fluorescence resonance energy transfer (FRET) measurements in conjunction with CG MD simulations to determine the oligomeric state of the membrane protein Fukutin-I. Their studies reveal that the N-terminal transmembrane domain of Fukutin-I exists as a dimer driven by interactions between a characteristic TXXSS motif. Oguchi *et al.* [42•] combine time-resolved amide hydrogen exchange and various biochemical assays together with CG modelling to predict the positioning and mechanistic role of M domains in ClpB hexamers. They show that M-domains nestle at the ClpB ring surface and contribute to maintaining a repressed ClpB activity state. Together, the results underline the vital nature of tight ClpB activity control and elucidate a regulated M-domain toggle control mechanism. Another example of predicting protein–protein binding modes is provided by the study of Kalli *et al.* [43] on the integrin–talin complex. Activation of the integrin αIIb/β3 dimer by the talin head domain was probed using multiscale MD simulations. The results, combined with various published experimental observations, suggest a model for the mechanism of inside-out activation of integrins by talin. Davies *et al.* [44] used electron cryotomography and CG simulations to investigate the structure and organization of ATP synthase dimers in mitochondrial membranes. The combined approach shows that isolated dimers induce a plastic deformation in the lipid bilayer, which is partially relieved by their side-by-side association. In many of these studies, in a so-called serial multi-scale approach, the favourable association interface determined from CG models is further optimized using atomistic simulations after a backmapping procedure. In addition to structural refinement, CG protein models also offer a direct connection to the time scales probed by state of the art kinetic experiments. Casuso *et al.* [45], for instance, computed the unbinding pathways of OmpF dimers to interpret high speed AFM experiments.

De novo prediction of membrane protein–protein interfaces

Ultimately one would like to be able to predict the structure of membrane protein complexes *de novo*. Like soluble proteins, the large number of possible interfaces still poses a major sampling problem. On the one hand, due to the two-dimensional nature of the membrane, prediction of membrane protein interfaces is easier. On the other hand, sampling of the desolvation of the protein–protein interface is problematic due to trapping of lipids. In case of single transmembrane (TM) helices, a growing number of studies demonstrate that the experimental interfaces can be reproduced using CG self-assembly (reviewed in [46]). Self-assembly simulations are typically performed using many independent replicas to probe the reproducibility of the interfaces formed, and the behaviour of WT and mutant proteins are compared providing additional insights into the packing motifs of TM helices. The ability to predict packing of TM helices has paved the way for CG modelling studies of self-assembly of larger protein complexes, in particular, protein complexes in which the interface is formed by single TM helices [47,48]. In a large scale study, Hall *et al.* [49•] simulated self-assembly of bacterial chemoreceptors embedded in a small liposome (Figure 2a). Notably, receptor dimers, representing the non-kinase activating form, were observed to self-assemble. The association process was driven by the large bending fluctuations of the aqueous domains of the receptor. Self-assembly studies of polytopic membrane proteins are still hampered by slow kinetics. Reversible sampling of protein–protein binding/unbinding events is extremely challenging for models that retain chemical specificity. Pioneering studies on respiratory chain complexes by Arnarez *et al.* [50] show complexation between cytochrome *bc1* and cytochrome *c* oxidase during self-assembly simulations, with a specific role attributed to cardiolipins in bridging the proteins together. However, on the multi-microsecond time scale of the simulations, equilibration of the protein–protein interfaces could not be achieved.

An alternative strategy aimed at *de novo* prediction of protein–protein complexes is through computation of the potential of mean force (PMF), that is the dimerization free energy profile as a function of protein–protein distance. The PMF can be used to predict specific binding modes and obtain insight into the thermodynamic driving forces for protein–protein aggregation. However, PMFs are notoriously difficult to calculate because of the sampling problems mentioned above. Even between simple TM helices convergence of the PMF requires sampling on the microsecond time scale, requiring the use of CG models [51,52,53•,54]. Calculation of PMFs between polytopic membrane proteins so far has therefore been limited to specific interfaces only. Recent studies on G-protein coupled receptors (GPCRs) [55•,56] reveal specific, favourable, association interfaces stabilized by energies of the order of 30–60 kJ mol^{-1}. Remarkably, Periole *et al.* [55•] found that the amount of protein burial (i.e. number of protein–protein contacts not exposed to lipids) does not correlate with the binding strength of the interface. This finding challenges the potential utility of buried accessible surface area as a predictor of the strength of membrane-embedded protein–protein interfaces, a strategy that works well for soluble proteins. To further increase the predictive power of these kinds of simulations, future studies should benefit from smart sampling techniques [57,58]. Another approach is the use of biomimetics (such as octanol or dodecane) that allow PMF calculation even at the all-atom level [59], or to resort to generic protein models for which many state conditions can be systematically probed [60]. Alternatively, implicit membrane models can be used which lead to an obvious speed-up [61,62].

Towards large scale membrane protein assembly

Many integral membrane proteins assemble in oligomeric structures in biological membranes. Again, reduced modelling approaches are beginning to address the key questions in the field, with an increasing number of studies based on the CG MARTINI force field [63] that combines parameters for lipids and proteins. Periole *et al.* [55•] carried out CG MD simulations over time scales reaching 100 μs for model membranes containing up to 64 molecules of the visual receptor rhodopsin, constituting one of the largest membrane patches simulated to date at near-atomistic resolution. Based on the observed interfaces, and supported by PMF calculations and chemical cross-linking data, a model was built for the supramolecular organization of the rows-of-dimers as observed in AFM images of the rod outer segment disk membrane (Figure 2b). In a joint experimental/modelling effort, van den Boogaart *et al.* [64] showed that syntaxin clustering is mediated by electrostatic interactions with the strongly anionic lipid phosphatidylinositol-4,5-bisphosphate (PIP2). Domanski *et al.* [65] found that, at lipid/protein ratios characteristic of real membranes, TM peptides can induce domain segregation resulting in large clusters of aggregated proteins (Figure 2c). Goose and Sansom [66] and Javanainen *et al.* [67] also simulated membranes under crowded conditions, with formation of extended clusters and networks of proteins that could lead to compartmentalisation of lipids in extreme cases. Other studies based on the MARTINI force field have addressed the driving forces for sorting and clustering of TM and membrane anchored proteins between ordered and disordered membrane domains [68,69,70•].

Although the above studies are able to show tendencies of membrane proteins to oligomerize under various conditions, equilibrium sampling of the supra-structural organization is currently not possible. To this end, coarser models are needed. Using less detailed CG models, an extensive systematic study on aggregation of model

membrane proteins, modulated by hydrophobic mismatch, membrane curvature, and protein class, was performed by Parton *et al.* [71]. Kovalenko *et al.* [72] developed a new BD-based method that takes into account explicit electrostatic interactions between proteins and an implicit membrane model. The approach was demonstrated for photosynthetic electron transport reaction kinetics taking place in the thylakoid membrane, involving hundreds of proteins. A different approach to upscaling was taken by Lai *et al.* [73]. Using a very coarse 16-site bead model, they studied self-association of hundreds of epsin N-terminal homology domains embedded on membrane tubes. The chemical specificity of the protein model was retained by matching the essential dynamics of a reference all-atom model with a heterogeneous elastic network. A similar large scale description of membrane protein organization can be obtained using a shape-based CG approach, in which all-atom simulations are used to calibrate effective interactions between entire secondary protein units. Yu and Schulten [74], for instance, used this method to study the membrane curvature induced by lattices of F-BAR domains.

Conclusions and outlook

The recent development in coarse-grained techniques has given a boost to computational modelling of protein–protein interfaces. The accessibility of long time scale (micro-millisecond range) has opened up the computation of binding free energies in self-assembly simulations, and spatially resolved PMFs are used to discriminate between different protein–protein interfaces. Nowadays multiscale modelling provides a powerful tool, especially in combination with experimental data, to refine and predict the structural and dynamic aspects of both soluble and membrane protein complexes, and to understand the physico-chemical nature of the driving forces.

We expect progress in combining dynamic CG models with protein–protein docking platforms. Inclusion of side chain flexibility, solvent entropy, and induced conformational changes are important factors to improve the accuracy of current docking approaches [75]. Using atomistic models is still prohibitively expensive, at least for high throughput assays, but CG models that retain chemical specificity are a powerful alternative. This specificity is particularly important for future applications of interest to the pharmaceutical industry, given that more and more PPI are found to be modulated by small molecules [76,77].

Another important area where advances should be expected is in hybrid multi-scale methods, that is models combining CG and atomistic levels of resolution concurrently in a single simulation. For instance, PPI could be efficiently simulated with only the surface residues in full atomic detail and protein interior and solvent environment treated at a CG level. An even more sophisticated scenario would have the proteins change resolution on the fly upon close contact. Although hybrid multi-scale methods are being developed by many groups [78,79], applications to real biomolecular processes involving PPI are still ahead. To deal with protein organization on scales reaching the size of the cell, the challenge is to develop supra CG models that still connect to the underlying structure. Current efforts in this direction include essential dynamics based CGing [73], ultra CG models with embedded states [80] and shape-based CG models [74]. More detailed CG models can be used in the parameterization, for example, protein–protein PMFs may define effective interactions between the supra-CG models with the prospect of fully resolved (i.e. distance and angle dependent) multi-dimensional protein–protein free energy landscapes. In that respect, it will be important to develop enhanced sampling methods to obtain converged protein–protein PMFs; promising novel methods in this field are window-exchange umbrella sampling [57] and the use of smart restraints [58].

Future developments of integrating experiment and CG simulation are particularly appealing for hypothesis generation, as the experimental data can be used to rapidly guide the CG models towards compatible solutions. A great variety of experimental data of varying precision is available, some of which is very straight forward to introduce in a simulation, for example distance constraints from FRET measurements or shape envelopes derived from CryoEM or SAXS experiments. The availability of osmotic second virial coefficients for a number of different proteins in a range of conditions of pH and ionic strength offers an excellent opportunity in the near future to undertake a systematic parameterization of CG models to match thermodynamic data on the interactions of proteins in aqueous solution [81]. The experiment-simulation feedback loop can be further improved by adding interactivity to the model construction process. A recent example using virtual reality and haptic manipulation was performed on the RecA nucleofilament complex [82].

Acknowledgments

MB and SJM thank the many colleagues who contributed original ideas and comments on the broad field of CG PPI. SJM acknowledges Xavier Periole, Helgi Ingólfsson, and Djurre de Jong for critical reading of the manuscript. MB thanks the French Agency for Research for funding (Grant ANR-10-BIOE-003). The work was further supported by the ''Initiative d'Excellence" program from the French State (Grant ''DYNAMO", ANR-11-LABX-0011).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Ellis RJ: **Macromolecular crowding: an important but neglected aspect of the intracellular environment**. *Curr Opin Struct Biol* 2001, **11**:114-119.

2. Engelman DM: **Membranes are more mosaic than fluid**. *Nature* 2005, **438**:578-580.

3. Phillip Y, Schreiber G: **Formation of protein complexes in crowded environments - From in vitro to in vivo**. *FEBS Letters* 2013, **587**:1046-1052.

4. Saunders MG, Voth GA: **Coarse-graining of multiprotein assemblies**. *Curr Opin Struct Biol* 2012, **22**:144-150.

5. Shinoda W, DeVane R, Klein ML: **Computer simulation studies of self-assembling macromolecules**. *Curr Opin Struct Biol* 2012, **22**:175-186.

6. Flores SC, Bernauer J, Shin S, Zhou RH, Huang XH: **Multiscale modeling of macromolecular biosystems**. *Brief Bioinformatics* 2012, **13**:395-405.

7. Ingólfsson H, Lopez CA, Uusitalo JJ, de Jong DH, Gopal S, Periole X, Marrink SJ: **The power of coarse-graining in biomolecular simulations**. *WIREs Comput Mol Sci* 2013 http://dx.doi.org/10.1002/wcms.1169. in press.

8. Kastritis PL, Bonvin AMJJ: **On the binding affinity of macromolecular interactions: daring to ask why proteins interact**. *J R Soc Interface* 2013, **10**:20120835 http://dx.doi.org/10.1098/rsif.2012.0835 1742-5662.

9. Coelho ED, Arrais JP, Oliveira JL: **From protein–protein interactions to rational drug design: are computational methods up to the challenge?** *Curr Top Med Chem* 2013, **13**:602-618.

10. King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, Andre I, Gonen T, Yeates TO, Baker D: **Computational design of self-assembling protein nanomaterials with atomic level accuracy**. *Science* 2012, **336**:1171-1174.

11. de Vries SJ, Zacharias M: **ATTRACT-EM: a new method for the computational assembly of large molecular machines using cryo-em maps**. *PloS One* 2012, **7**:e49733.

12. Liu S, Vakser IA: **DECK: distance and environment-dependent, coarse-grained, knowledge-based potentials for protein–protein docking**. *BMC Bioinformatics* 2011, **12**:280.

13. Ghoorah AW, Devignes M-D, Smail-Tabbone M, Ritchie DW: **Spatial clustering of protein binding sites for template based protein docking**. *Bioinformatics* 2011, **27**:2820-2827.

14. • Andreani J, Faure G, Guerois R: **InterEvScore: A novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution**. *Bioinformatics* 2013, **29**:1742-1749.

Multi-residue interactions are crucial for protein–protein binding and are included in an original scoring function. Furthermore evolutionary information accounts for the conservation of protein interfacial interactions. A benchmark documents the significant improvement of scoring results.

15. Karaca E, Bonvin AMJJ: **Advances in integrative modeling of biomolecular complexes**. *Methods* 2013, **59**:372-381.

16. Schneidman-Duhovny D, Rossi A, Avila-Sakar A, Kim SJ, Velazquez-Muriel J, Strop P, Liang H, Krukenberg KA, Liao M, Kim HM *et al.*: **A method for integrative structure determination of protein–protein complexes**. *Bioinformatics* 2012, **28**:3282-3289.

17. Dreyfus T, Doye V, Cazals F: **Assessing the reconstruction of macromolecular assemblies with toleranced models**. *Proteins: Struc Funct Bioinf* 2012, **80**:2125-2136.

18. Zacharias M: **Combining coarse-grained nonbonded and atomistic bonded interactions for protein modeling**. *Proteins: Struc Funct Bioinf* 2013, **81**:81-92.

19. Basdevant N, Borgis D, Ha-Duong T: **Modeling protein–protein recognition in solution using the coarse-grained force field SCORPION**. *J Chem Theory Comput* 2013, **9**:803-813.

20. Frembgen-Kesner T, Elcock AH: **Absolute protein–protein association rate constants from flexible, coarse-grained brownian dynamics simulations: the role of intermolecular hydrodynamic interactions in barnase-barstar association**. *Biophys J* 2010, **99**:L75-L77.

21. Solernou A, Fernandez-Recio J: **pyDockCG: New coarse-grained potential for protein protein docking**. *J Phys Chem B* 2011, **115**:6032-6039.

22. Ceres N, Pasi M, Lavery R: **A protein solvation model based on residue burial**. *J Chem Theory Comput* 2012, **8**:2141-2144.

23. • Zheng W, Schafer NP, Davtyan A, Papoian GA, Wolynes PG: **Predictive energy landscapes for protein–protein association**. *Proc Natl Acad Sci USA* 2012, **109**:19244-19249.

This study investigates fundamental principles of protein–protein association, showing that the binding process is described by a funneled model similar but not identical to the folding landscapes of monomeric proteins.

24. Martin J, Lavery R: **Arbitrary protein–protein docking targets biologically relevant interfaces**. *BMC Biophys* 2012, **5**:7.

25. De Simone A, Kitchen C, Kwan AH, Sunde M, Dobson CM, Frenkel D: **Intrinsic disorder modulates protein self-assembly and aggregation**. *Proc Natl Acad Sci USA* 2012, **109**:6951-6956.

26. Ravikumar KM, Huang W, Yang S: **Coarse-grained simulations of protein–protein association: an energy landscape perspective**. *Biophys J* 2012, **103**:837-845.

27. Emperador A, Solernou A, Sfriso P, Pons C, Lluis Gelpi J, Fernandez-Recio J, Orozco M: **Efficient relaxation of protein–protein interfaces by discrete molecular dynamics simulations**. *J Chem Theory Comput* 2013, **9**:1222-1229.

28. Okazaki K-i, Sato T, Takano M: **Temperature-enhanced association of proteins due to electrostatic interaction: a coarse-grained simulation of actin-myosin binding**. *J Am Chem Soc* 2012, **134**:8918-8925.

29. Saunders MG, Voth GA: **Comparison between actin filament models: coarse-graining reveals essential differences**. *Structure* 2012, **20**:641-653.

30. • Sieradzan AK, Liwo A, Hansmann UHE: **Folding and self-assembly of a small protein complex**. *J Chem Theory Comput* 2012, **8**:3416-3422.

The characterization of the self-assembly mechanism of a tetrameric model system reveals a surprising complexity, even for one of the smallest and simplest protein complexes.

31. Chaudhri A, Zarraga IE, Yadav S, Patapoff TW, Shire SJ, Voth GA: **The role of amino acid sequence in the self-association of therapeutic monoclonal antibodies: insights from coarse-grained modeling**. *J Phys Chem B* 2013, **117**:1269-1279.

32. Mereghetti P, Wade RC: **Atomic detail Brownian Dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions**. *J Phys Chem B* 2012, **116**:8523-8533.

33. • May ER, Feng J, Brooks CL III: **Exploring the symmetry and mechanism of virus capsid maturation via an ensemble of pathways**. *Biophys J* 2012, **102**:606-612.

This study investigates the maturation pathway of bacteriophage HK97 with a particular focus on symmetry. It provides mechanistic insight into virus capsid maturation based on a Cα-based Gō model where hydrogen bonding and side chain-side chain interactions are used to identify contacts.

34. Globisch C, Krishnamani V, Deserno M, Peter C: **Optimization of an elastic network augmented coarse grained model to study ccmv capsid deformation**. *PloS One* 2013, **8**:e60582.

35. Bereau T, Globisch C, Deserno M, Peter C: **Coarse-grained and atomistic simulations of the salt-stable cowpea chlorotic mottle virus (ss-ccmv) subunit 26-49: beta-barrel stability of the hexamer and pentamer geometries**. *J Chem Theory Comput* 2012, **8**:3750-3758.

36. Cieplak M, Robbins MO: **Nanoindentation of 35 virus capsids in a molecular model: relating mechanical properties to structure**. *PloS one* 2013, **8**:e63640.

37. Carrasco C, Luque A, Hernando-Perez M, Miranda R, Carrascosa JL, Serena PA, de Ridder M, Raman A, Gomez-Herrero J, Schaap IAT *et al.*: **Built-in mechanical stress in viral shells**. *Biophys J* 2011, **100**:1100-1108.

38. Grime JMA, Voth GA: **Early stages of the HIV-1 capsid protein lattice formation**. *Biophys J* 2012, **103**:1774-1783.

39. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang P: **Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics**. *Nature* 2013, **497**:643-646.

40. Steinbrecher T, Prock S, Reichert J, Wadhwani P, Zimpfer B, Burck J, Berditsch M, Elstner M, Ulrich AS: **Peptide–lipid interactions of the stress-response peptide TisB that induces bacterial persistence**. *Biophys J* 2012, **103**:1460-1469.

41. Marius P, Leung YM, Piggot TJ, Khalid S, Williamson PTF: **Probing the oligomeric state and interaction surfaces of Fukutin-I in dilauroylphosphatidylcholine bilayers**. *Eur Biophys J* 2012, **41**:199-207.

42. • Oguchi Y, Kummer E, Seyffer F, Berynskyy M, Anstett B, Zahn R, Wade RC, Mogk A, Bukau B: **A tightly regulated molecular toggle controls AAA+ disaggregase**. *Nat Struct Mol Biol* 2012, **19**:1338.

A nice example of an integrated modeling-experimental approach. Combination of time-resolved amide hydrogen exchange together with a number of biochemical assays suggests a flexible role of a molecular toggle in the AAA+ disaggregase protein complex, which is confirmed by CG simulations.

43. Kalli AC, Campbell, Sansom MSP: **Multiscale simulations suggest a mechanism for integrin inside-out activation**. *Proc Natl Acad Sci USA* 2011, **108**:11890-11895.

44. Davies KM, Anselmi C, Wittig I, Faraldo-Gomez JD, Kuhlbrandt W: **Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae**. *Proc Natl Acad Sci USA* 2012, **109**:13602-13607.

45. Casuso I, Khao J, Chami M, Paul-Gilloteaux P, Husain M, Duneau JP, Stahlberg H, Sturgis JN, Scheuring S: **Characterization of the motion of membrane proteins using high-speed atomic force microscopy**. *Nat Nanotech* 2012, **7**:525-529.

46. Psachoulia E, Marshall DP, Sansom MSP: **Molecular dynamics simulations of the dimerization of transmembrane alpha-helices**. *Acc Chem Res* 2010, **43**:388-396.

47. Prakash A, Janosi L, Doxastakis M: **GxxxG motifs, phenylalanine, and cholesterol guide the self-association of transmembrane domains of ErbB2 receptors**. *Biophys J* 2011, **101**:1949-1958.

48. Chng CP, Tan SM: **Leukocyte integrin αLβ2 transmembrane association dynamics revealed by coarse-grained molecular dynamics simulations**. *Proteins* 2011, **79**:2203-2213.

49. • Hall BA, Armitage JP, Sansom MSP: **Mechanism of bacterial signal transduction revealed by molecular dynamics of tsr dimers and trimers of dimers in lipid vesicles**. *PLoS Comp Biol* 2012, **8**:e1002685.

Large scale CG modeling study showing the propensity of bacterial chemoreceptors to aggregate either as dimers or trimers of dimers. The results indicate an activation mechanism whereby small motions of a single helix can be amplified to initiate large changes in the whole receptor structure.

50. Arnarez C, Mazat JP, Elezgaray J, Marrink SJ, Periole X: **Evidence for cardiolipin binding sites on the membrane-exposed surface of the cytochrome bc1**. *J Am Chem Soc* 2013, **135**:3112-3120.

51. Sengupta D, Marrink SJ: **Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes**. *Phys Chem Chem Phys* 2010, **12**:12987-12996.

52. Castillo N, Monticelli L, Barnoud J, Tieleman DP: **Free energy of WALP23 dimer association in DMPC, DPPC, and DOPC bilayers**. *Chem Phys Lip* 2013, **169**:95-105.

53. • Benjamini A, Smit B: **Lipid mediated packing of transmembrane helices - a dissipative particle dynamics study**. *Soft Matt* 2013, **9**:2673-2683.

This paper uses a CG model to challenge the notion that packing of TM helices is determined by specific interactions. Instead, the authors show that hydrophobic mismatch, through its effect on helix tilt, can explain many cross-angle distribution features.

54. Yoo J, Cui Q: **Membrane-mediated protein–protein interactions and connection to elastic models: a coarse-grained simulation analysis of gramicidin A association**. *Biophys J* 2013, **104**:128-138.

55. • Periole X, Knepp AM, Sakmar TP, Marrink SJ, Huber T: **Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers**. *J Am Chem Soc* 2012, **134**:10959-10965.

The first report on PMFs between fully solvated polytopic membrane proteins, GPCRs. The authors show that sampling times exceeding 1 ms are required to obtain converged profiles for specific binding interfaces. Comparing different interfaces, weak and strong binding interfaces are predicted and used to model the supramolecular rows-of-dimers organization of the rod outer segment disc membrane.

56. Johnston JM, Wang H, Provasi D, Filizola M: **Assessing the relative stability of dimer interfaces in G protein-coupled receptors**. *PLoS Comp Biol* 2012, **8**:e1002649.

57. Park S, Im W: **Two dimensional window exchange umbrella sampling for transmembrane helix assembly**. *J Chem Theory Comput* 2013, **9**:13-17.

58. Gumbart JC, Roux B, Chipot CJ: **Efficient determination of protein–protein standard binding free energies from first principles**. *J Chem Theory Comput* 2013, **9**:3789-3798.

59. Hsin J, LaPointe LM, Kazy A, Chipot C, Senes A, Schulten K: **Oligomerization state of photosynthetic core complexes is correlated with the dimerization affinity of a transmembrane helix**. *J Am Chem Soc* 2011, **133**:14071-14081.

60. de Meyer FJM, Rodgers JM, Willems TF, Smit B: **Molecular simulation of the effect of cholesterol on lipid-mediated protein–protein interactions**. *Biophys J* 2010, **99**:3629-3638.

61. Zhang LQ, Sodt AJ, Venable RM, Pastor RW, Buck M: **Prediction, refinement, and persistency of transmembrane helix dimers in lipid bilayers using implicit and explicit solvent/lipid representations: Microsecond molecular dynamics simulations of ErbB1/B2 and EphA1**. *Proteins: Struct Funct Bioinf* 2013, **81**:365-376.

62. Polyansky AA, Volynsky PE, Efremov RG: **Multistate organization of transmembrane helical protein dimers governed by the host membrane**. *J Am Chem Soc* 2012, **134**:14390-14400.

63. Marrink SJ, Tieleman DP: **Perspective on the Martini model**. *Chem Soc Review* 2013, **42**:6801-6822.

64. van den Bogaart G, Meyenberg K, Risselada HJ, Amin H, Willig KI, Hubrich BE, Dier M, Hell SW, Grubmuller H, Diederichsen U, Jahn R: **Membrane protein sequestering by ionic protein–lipid interactions**. *Nature* 2011, **479**:552-555.

65. Domanski J, Marrink SJ, Schäfer LV: **Transmembrane helices can induce domain formation in crowded model membranes**. *BBA Biomembr* 2012, **1818**:984-994.

66. Goose JE, Sansom MSP: **Reduced lateral mobility of lipids and proteins in crowded membranes**. *PLoS Comput Biol* 2013, **9**:e1003033.

67. Javanainen M, Hammaren H, Monticelli L, Jeon JH, Miettinen MS, Martinez-Seara H, Metzler R, Vattulainen I: **Anomalous and normal diffusion of proteins and lipids in crowded lipid membranes**. *Farad Discuss* 2013, **161**:397-417.

68. Parton DL, Tek A, Baaden M, Sansom MSP: **Formation of raft-like assemblies within clusters of influenza hemagglutinin observed by MD simulations**. *PLoS Comput Biol* 2013, **9**:e1003034.

69. de Jong DH, Lopez CA, Marrink SJ: **Molecular view on protein sorting into liquid-ordered membrane domains mediated by gangliosides and lipid anchors**. *Farad Discuss* 2013, **161**:347-363.

70. • Janosi L, Li Z, Hancock JF, Gorfe AA: **Organization, dynamics, and segregation of Ras nanoclusters in membrane domains**. *Proc Natl Acad Sci USA* 2012, **109**:8097-8102.

CG study on sorting and clustering of Ras membrane-bound proteins, showing that posttranslational lipid modification not only provides a hydrophobic membrane anchor but also plays a key role in the formation and spatiotemporal distribution of proteolipid nanodomains. The predicted cluster size distributions match well to those obtained from EM.

71. Parton DL, Klingelhoefer JW, Sansom MSP: **Aggregation of model membrane proteins, modulated by hydrophobic mismatch, membrane curvature, and protein class**. *Biophys J* 2011, **101**:691-699.

72. Kovalenko IB, Abaturova AM, Diakonova AN, Knyazeva OS, Ustinin DM, Khruschev SS, Riznichenko GY, Rubin AB: **Computer simulation of protein–protein association in photosynthesis**. *Math Model Nat Phenom* 2011, **6**:39-54.

73. Lai CL, Jao CC, Lyman E, Gallop JL, Peter BJ, McMahon HT, Langen R, Voth GA: **Membrane binding and self-association of the epsin N-terminal homology domain**. *J Mol Biol* 2012, **423**:800-817.

74. Yu H, Schulten K: **Membrane sculpting by F-BAR domains studied by molecular dynamics simulations**. *PLoS Comput Biol* 2013, **9**:e1002892.

75. Tuffery P, Derreumaux P: **Flexibility and binding affinity in protein–ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches**. *J R Soc Interface* 2012, **9**:20-33.

76. Fry DC: **Small-molecule inhibitors of protein–protein interactions: how to mimic a protein partner**. *Curr Pharm Des* 2012, **18**:4679-4684.

77. Villoutreix BO, Labbe CM, Lagorce D, Laconde G, Sperandio O: **A leap into the chemical space of protein–protein interaction inhibitors**. *Curr Pharm Des* 2012, **18**:4648-4667.

78. Wassenaar TA, Ingólfsson H, Prieß M, Marrink SJ, Schäfer LV: **Mixing Martini: electrostatic coupling in hybrid atomistic–coarse-grained biomolecular simulations**. *J Phys Chem B* 2013, **117**:3516-3530.

79. Poblete S, Praprotnik M, Kremer K, Delle Site L: **Coupling different levels of resolution in molecular simulations**. *J Chem Phys* 2010, **132**:114101.

80. Dama JF, Sinitskiy AV, McCullagh M, Weare J, Roux B, Dinner AR, Voth GA: **The theory of ultra-coarse-graining. 1. General principles**. *J Chem Theory Comput* 2013, **9**:2466-2480.

81. Stark AC, Andrews CT, Elcock AH: **Toward optimized potential functions for protein–protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the Martini coarse-grained force field**. *J Chem Theory Comput* 2013, **9**:4176-4185.

82. Saladin A, Amourda C, Poulain P, Ferey N, Baaden M, Zacharias M, Delalande O, Prevost C: **Modeling the early stage of DNA sequence recognition within RecA nucleoprotein filaments**. *Nucleic Acids Res* 2010, **38**:6313-6323.