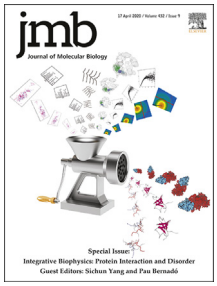
Review



Integrative Modelling of Biomolecular Complexes

P.I. Koukos and A.M.J.J. Bonvin

Computational Structural Biology Group, *Department of Chemistry, Faculty of Science, Utrecht University, 3584CH, Utrecht, the Netherlands*

Correspondence to A.M.J.J. Bonvin: *a.m.j.j.bonvin@uu.nl*
https://doi.org/10.1016/j.jmb.2019.11.009
Edited by Pau Bernado

Abstract

In recent years, the use of integrative, information-driven computational approaches for modeling the structure of biomolecules has been increasing in popularity. These are now recognized as a crucial complement to experimental structural biology techniques such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and cryo-electron microscopy (cryo-EM). This trend can be credited to a few reasons such as the increased prominence of structures solved by cryo-EM, the improvements in proteomics approaches such as cross-linking mass spectrometry (XL-MS), the drive to study systems of higher complexity in their native state, and the maturation of many computational techniques combined with the widespread availability of information-driven integrative modeling platforms.

In this review, we highlight recent works that exemplify how the use of integrative and/or information-driven approaches and platforms can produce highly accurate structural models. These examples include systems which present many challenges when studied with traditional structural biology techniques such as flexible and dynamic macromolecular assemblies and membrane-associated complexes.

We also identify some key areas of interest for information-driven, integrative modeling and discuss how they relate to ongoing challenges in the fields of computational structural biology. These include the use of coarse-grained force fields for biomolecular simulations—allowing for simulations across longer (time-) and bigger (size-dimension) scales—the use of bioinformatics predictions to drive sampling and/or scoring in docking such as those derived from coevolution analysis and finally the study of membrane and membrane-associated protein complexes.

Introduction

Biological macromolecules such as proteins and nucleic acids make up the majority of the machinery of life since they are responsible for performing most cellular functions. Although a lot of meaningful insights about these functions can be deduced by experimental work that falls under the umbrella of functional assays, these kinds of experiments (e.g., yeast two-hybrid assays) usually fail to reveal any direct information regarding the structure of the biomolecules involved in a given process. True understanding of the mechanism of action that underlies any cellular function can however only be gained by resolving at atomic detail the molecular structures of the components and assemblies involved, allowing us a glimpse at the molecular mechanisms at play [1].

Historically, the two main techniques used for experimental structure determination of biomolecules have been X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy [2]. More recently, cryo-electron microscopy (cryo-EM) has been added to the arsenal of structural biologists and has now overtaken NMR as the second most popular technique for obtaining molecular structures with 846 vs 395 models deposited in the PDB during 2018 for cryo-EM and NMR, respectively, although both are still lagging far behind X-ray crystallography (>10000 models deposited in 2018).

All three techniques have unique advantages and disadvantages that make them suitable for specific applications, with X-ray crystallography still being the method of choice for systems which do not contain flexible or disordered regions. On the opposite end, NMR can still capture valuable information about flexible systems as well as characterize dynamics under conditions that can be considered native-like. Solution-state NMR has, however, size limitations which only make it applicable for rather small systems when it comes to solving 3D structures. It does, however, allow to answer specific questions, in particular, related to the dynamics of large systems such as nucleosomes [3–7], proteasomes [8–12], mRNA signaling machinery [13–16], as well as systems with high clinical significance such as kinase and chaperone complexes [17–19], for which NMR has a long and well-documented history of serving as the primary data source driving the simulations [20]. Although solid-state NMR [21,22] does not suffer from size limitations, it still has difficulty in yielding atomic resolution–quality spatial information, especially 3D structures, despite recent methodological advancements in specific fields [23–25]. Cryo-EM is increasingly becoming one of the most popular ways of determining the structure of biomolecules and most importantly large complexes and macromolecular assemblies. However, it cannot yet routinely produce structural models of atomic resolution, the level of detail which is required to understand molecular mechanisms in depth, as can be seen from the recent statistics of the Electron Microscopy Data Resource (EMDataResource—https://www.emdataresource.org/statistics.html) and those of the European Bioinformatics Institute (EBI—https://www.ebi.ac.uk/pdbe/emdb/statistics_num_res.html/). The field is still undergoing rapid transformations reflecting its nascent state, with the absence of well-defined standard practices and ongoing instrumentation and software optimizations being highlighted as potential points of improvement that should lead to higher quality structures being made available through cryo-EM in the coming years [26]. The Electron Microscopy Data Bank (EMDB) [27] has recently sponsored two blind challenges whose stated goals were to emphasize the need for map and model validation standards and engage with the cryo-EM community toward the shared development of assessment benchmarks and best practices [28].

A careful reading of the relative strengths and weaknesses of the three techniques mentioned in the previous paragraph reveals an ideal use case for computational structural modeling which relies on the use of high-quality structural models solved with X-ray crystallography or NMR spectroscopy, for determining the finer structural details of interacting biomolecules, combined with the use of cryo-EM density maps for determining the overall topology and stoichiometry of the wider context of the complex. Indeed, we believe that the revolution cryo-EM ushered in the field of structural biology a few years ago is only going to lead to an increased demand for computational techniques that not only can make use of the data that are being made available through cryo-EM studies but also combine those with other types of data available through other techniques, to generate structural models that would normally be beyond the reach of any of those techniques taken on their own.

An additional reason necessitating the use of integrative approaches is the need to study biological systems in their proper context, not only as single-structures but also as macromolecular assemblies and high-order complexes as this is seen as a stepping-stone toward realizing a structural model of the cell in atomic or near atomic detail [29,30]. A complicating factor is that, to achieve this goal, experimental data measured under as close as possible physiological conditions need to be captured, once again requiring robust integrative modeling frameworks and protocols to unify the various sources of experimental information in cohesive structural models.

In addition to the three aforementioned structure determination techniques, a plethora of complimentary techniques is available that can provide some pieces of the puzzle for the biological systems under study. Prime examples are cross-linking mass spectrometry (XL-MS) and small-angle X-ray scattering (SAXS). XL-MS can be used to determine distances between specific residues of biomolecular complexes that can then be used in modeling because they allow for an upper distance bound between the residues they are targeting. Variations of the technique also enable the study of dynamics of complex populations in native-like conditions or even within living cells. SAXS, on the other hand, is the solution equivalent of X-ray crystallography and can provide low-resolution shape information about complexes in solution and, similarly to XL-MS, can also yield information regarding dynamic populations. Both of these techniques, along with the previously mentioned ones, will be discussed in detail in the next section.

Next to experimental methods, computational and bioinformatics approaches such as coevolution analyses can be used to identify residues at protein interfaces which evolve in tandem, allowing to use those residue pairs in integrative modeling directly in the simulation or during the scoring stage. Additional developments, such as the availability of coarse-grained force fields, allow for simulations across longer (time) and bigger (size) scales, enabling multiscale studies from the quantum to various levels of coarse-grained representations. These pave the way for continuous and mesoscale studies of biomolecular systems. Some docking codes now

also support modeling of membrane complexes with specially adapted implicit potentials. All these developments mean that systems of increasing complexity and high relevance can now be studied within reasonable computational costs.

Integrative and Information-driven Modeling

We have so far mentioned integrative and information-driven modeling without explicitly defining what constitutes such a modeling approach and distinguishes it from de novo or first-principles modeling. The main emphasis of this mini-review is on the use of biomolecular docking for modeling the 3D structure of biomolecular protein-protein complexes with a special focus at the end on membrane protein complexes. We concentrate on integrative methods that use some kind of information (typically a combination of various experimental and/or bioinformatics sources) to drive the modeling process, excluding the so-called template-based approaches that fall more under the umbrella of homology modeling.

Molecular docking refers to a set of techniques that allows us to predict the 3D structure of a biomolecular complex via simulation when starting from the 3D structures of its unbound (free) components [31]. Unlike de novo modeling, information-driven modeling centers on the concept of using experimentally determined (or predicted) data to guide the modeling process in the hope of sampling or selecting only the meaningful part of the conformational, interaction landscape of the complex. It thus bypasses the need to exhaustively sample the vast conformational space, which would cover a 6D space for a binary complex consisting of rigid molecules. Its complexity will, however, greatly increase when considering flexibility and/or modeling a larger number of subunits. Integrative modeling refers thus to the use of some docking protocol that combines multiple sources of information (e.g., cryo-EM density map and XL-MS–derived distance restraints) to generate a 3D model of the assembly under study [32,33].

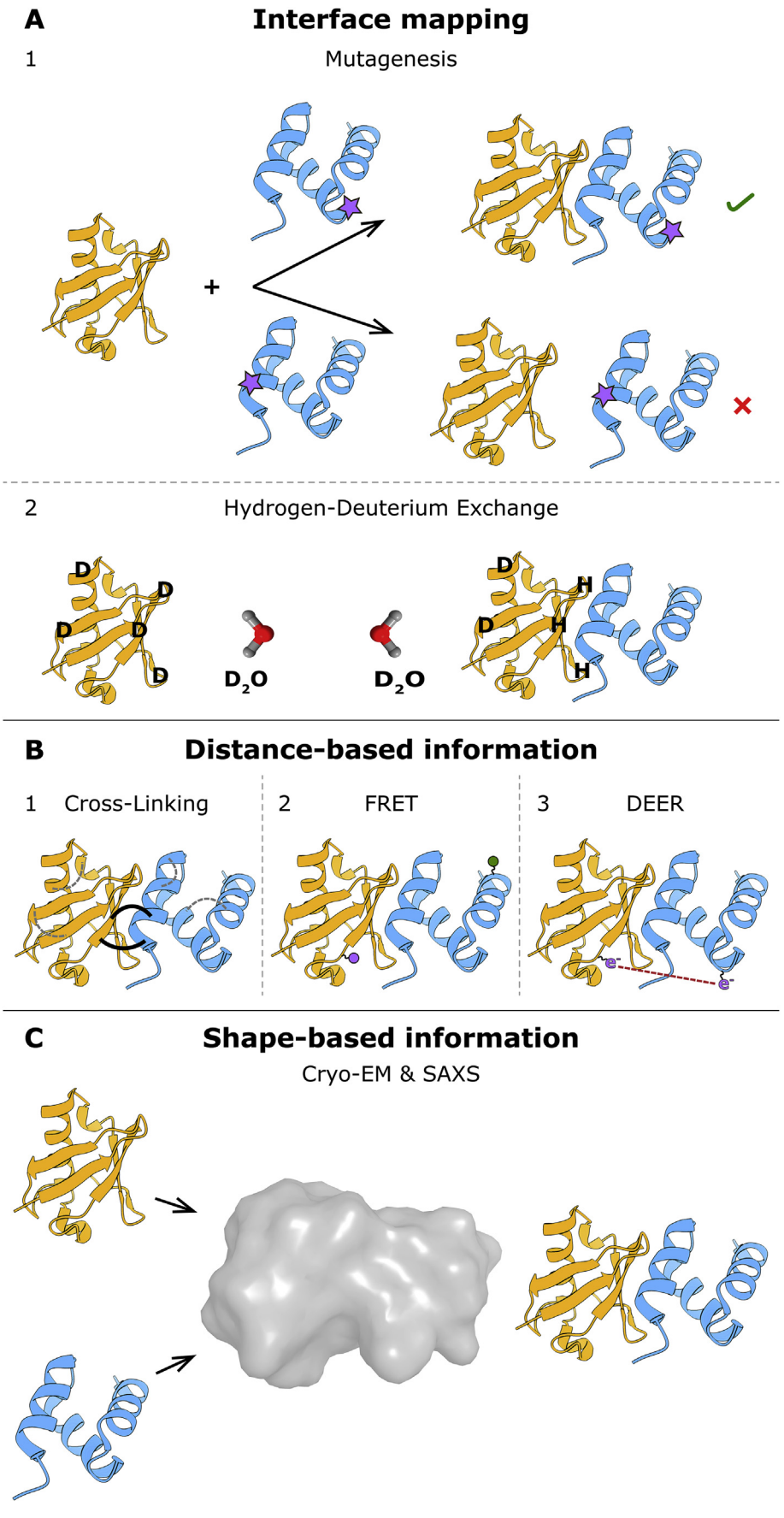
Docking has existed as a standalone field for close to 40 years [34,35] and is one of the main two computational methods which allows us to study the 3D structure of interacting biomolecules, the other being atomistic binding simulations based, for example, on molecular dynamics (MD) simulations [36,37]. Docking has seen a wide range of applications from structure-based drug design [38] to protein-protein interaction studies [39,40] and network biology [41,42]. Unlike atomistic simulations, the computational requirements of docking can be met quite easily [36], which allows us to generate models at a fraction of the time of what would be required by MD. Similar to MD and other biomolecular simulation approaches though, two factors govern its performance: Sampling and scoring. Simply put, sampling refers to the process that is used to generate the binding poses from the unbound conformations. Scoring, on the other hand, is the process which allows us to discriminate between good—or native-like—and bad—or non–native-like—models. In the context of integrative modeling, the information at hand can be used to guide the simulation toward specific conformations, thus affecting the way the sampling is performed, as a filter to select or discard models based on their agreement with the experimental data, thus affecting the way the scoring is performed, or both.

A detailed overview of the challenges of the various types of docking depending on the nature of the interacting biomolecules such as protein-protein [1,43], protein–nucleic acid [44–46], protein–small molecule [47,48] and protein-peptide [49] is beyond the scope of this mini-review as are the intricate details of the algorithms used by various docking programs to achieve good sampling and scoring performance. The latter is something that has been continuously evaluated over a period spanning almost 20 years in CAPRI (Critical Assessment of PRediction of Interactions) [50]—the blind docking experiment [51–55]. Recent CAPRI evaluations clearly demonstrate that the best strategy to model complexes is to follow a template-based approach when homologous complexes or interfaces can be identified from the PDB database [56,57]. Among the various docking software, several are supporting the use of data directly during sampling, such as HADDOCK (High Ambiguity Driven DOCKing) [58,59], the pioneer of information-driven docking, together with other widely used codes such as ATTRACT [60–63], Hex [64,65], IMP [66,67], LightDock [68,69], and ROSETTA [70,71]. In general, most docking codes under active development have added support for the use of information either to drive the simulation or, more commonly, as a way to filter the generated models [1].

The next section is going to focus on the various types of experimental information that can be used by integrative modeling frameworks.

Information Sources for Integrative Modeling

Of course, integrative modeling entirely depends on the availability of data to drive the simulation. In this section, we will expand on some of the most widely used types of information that can be used in an integrative capacity starting from simple experimental setups which do not require extensive


A
Interface mapping
1
Mutagenesis
+
2
Hydrogen-Deuterium Exchange
D
D
D
D
D
D
D2O
D2O
D
H
D
H
H
B
Distance-based information
1 Cross-Linking
2 FRET
3 DEER
C
Shape-based information
Cryo-EM & SAXS

expertise or instrumentation before proceeding to more complicated ones.

The most commonly used approaches are those that yield residue-level information. This kind of data can be obtained from mutagenesis, cross-linking—providing upper limits to the distance between the cross-linked residues, hydrogen-deuterium exchange (HDX), and NMR spectroscopy experiments. The next set of techniques can yield anything from low-resolution information to high-resolution structural models of macromolecular assemblies; Cryo-EM and SAXS belong to this category. Finally, computational techniques such as multiple sequence alignments, coevolution analysis, and metagenomics sequencing can yield high-quality information and interfacial and interacting residues.

Mutagenesis

Mutagenesis experiments [72–75] rest on the hypothesis that mutation of residues that are functionally important for complex formation will prevent the biomolecules (proteins specifically in this case) from interacting with each other and thus the complex from being formed (Fig. 1, panel A.1). It has been used to map the interfaces or binding sites of interacting proteins [76–78]. The benefits of mutagenesis experiments are the relative ease with which the experiment can be performed, with a large variety of detection methods possible, and the fact that it provides residue-level information which constitutes high-quality data that can significantly aid the modeling process compared with assays which can only provide qualitative information with regard to whether two biomolecules are interacting or not. The main downside is that, owing to the indirect nature of the experiment, it needs to be combined with functional and folding assays to ensure that lack of complex formation is a result of the mutation that was introduced and not of the incorrect folding of one of the partners. Another complicating factor is allosteric effects, which can be very challenging to detect. Recent improvements to existing high-throughput mutagenesis pipelines which minimize experimental errors should enable the rapid creation of mutant libraries, which, in turn, will allow quick screening of hundreds of mutations [79,80]. Despite these advancements and the benefits conveyed in targeted mutagenesis as demonstrated for example in the CRISPR/CAS9 system [81,82], we do not expect mutagenesis data to become a dominant part of integrative modeling protocols. It will however remain a valuable source of information, even more these days where next-generation sequencing has boosted the amount of genomic information available, including the identification of disease-related mutations.

Hydrogen-deuterium exchange

HDX is based on the principle of constant exchange of protons between biomolecules and water in which they are dissolved. In a typical HDX experiment (Fig. 1, panel A.2), the solvent (H_2O) is exchanged for D_2O, which means that the exchangeable protons of the protein (e.g., the backbone amide protons) will—at some point—exchange their proton for deuterium. The rate at which this exchange event takes place is determined by the stability of the hydrogen bond network the proton is part of, its solvent accessibility, and the chemical characteristics of the residue it belongs to as well as those in its immediate vicinity [83]. Time-resolved measurements of these exchange events allow us to calculate the so-called protection factors for every exchanging amide [84], which in turn can be used to map protein-protein interfaces. These protection factors need to be determined separately for the unbound monomers and the complex. Owing to the properties of

Fig. 1. Schematic representation of some of the experimental methods which can be used in integrative modeling. **Panel A** shows methods which can be used to map interfaces of interacting biomolecules. **Panel A.1** shows the experimental setup for a mutagenesis experiment combined with a binding assay. Mutations of residues which lie at the interface of the two proteins prevent complex formation. Mutated sites are shown as purple stars. **Panel A.2** shows the experimental setup for an HDX experiment during which the exchange rates for both the free forms of the proteins and their complex are compared to detect the regions which are occluded at the interface of the complex due to slower exchanging protons. **Panel B** shows methods which can be used to calculate residue-based distances. **Panel B.1** shows a complex which has undergone cross-linking with the intermolecular cross-links shown as continuous black lines and the intramolecular ones as dotted gray lines. **Panel B.2** shows a complex to which fluorophore dyes have been attached at specific residues allowing us to calculate the distance between the target residues with FRET. The donor and acceptor fluorophores are shown as purple and green circles, respectively. **Panel B.3** shows a complex to which spin labels have been attached enabling calculation of intermolecular distances with DEER spectroscopy. The labels are as shown as purple radicals. **Panel C** shows shape-based methods. The free structures of the complex are combined with shape-based information about the complex structure which can be derived from cryo-EM densities or SAXS shapes. The surface representation of the complex was generated with PyMOL [240]. All molecular graphics structures were created with ChimeraX [241]. The complex shown is the Ubiquitin-UBA domain from Cbl-b ubiquitin ligase (PDB entry 2oob). Ubiquitin is colored orange and UBA light blue.

deuterium, detection can be performed with either NMR or—much more commonly—mass spectrometry (MS). An important point regarding the mapped regions when MS is the detection method is that, owing to the nature of the technique, it is not individual residues that are identified but peptide fragments whose length usually ranges between 5 and 10 residues complicating somewhat the way they can be used during computational modeling. Benefiting from methodological improvements in liquid chromatography (LC) and MS, along with major progress in analysis software, HDX is increasing in popularity, also aided by the relatively simple experimental setup it requires [85]. The nature of HDX experiments means these can be applied broadly for the study of any biomolecule with exchangeable protons. Although HDX data are not used as often in integrative modeling of complexes, we expect that situation to change in the coming years. Of particular note is the fact that the HDX community, in anticipation of this increased interest in the field, has taken steps to codify practices ranging from sample preparation, measurement, and data analysis to publication and dissemination of data in standardized formats [85]. These efforts are of particular importance as they allow easier integration in modeling paradigms which combine multiple experimental sources of information [86].

Chemical cross-linking

Chemical cross-linking, most often combined with mass spectrometry for detection purposes—XL-MS—refers to the chemical linking of residues (most often lysine or cysteine) which are located on the surface of proteins using compounds which consist of two reactive heads and a (flexible) spacer of known maximal length [87,88]. After the cross-linking reagent has been added to the protein/cell sample and the sample has been washed, it is subjected to trypsination (or treatment with another protease) and the peptide fragments are detected via mass spectrometry [89] (Fig. 2, panel B.1). The benefit of XL-MS when compared to mutagenesis or HDX experiments is that the residue information is not ambiguous as it always comes in pairs (unless there are two lysines within the detected peptide fragment) and, in addition to the residues themselves, it provides information regarding their distance as the maximal spacer distance is known *a priori*. Recent improvements in cross-linking protocols and reagents along with widespread availability

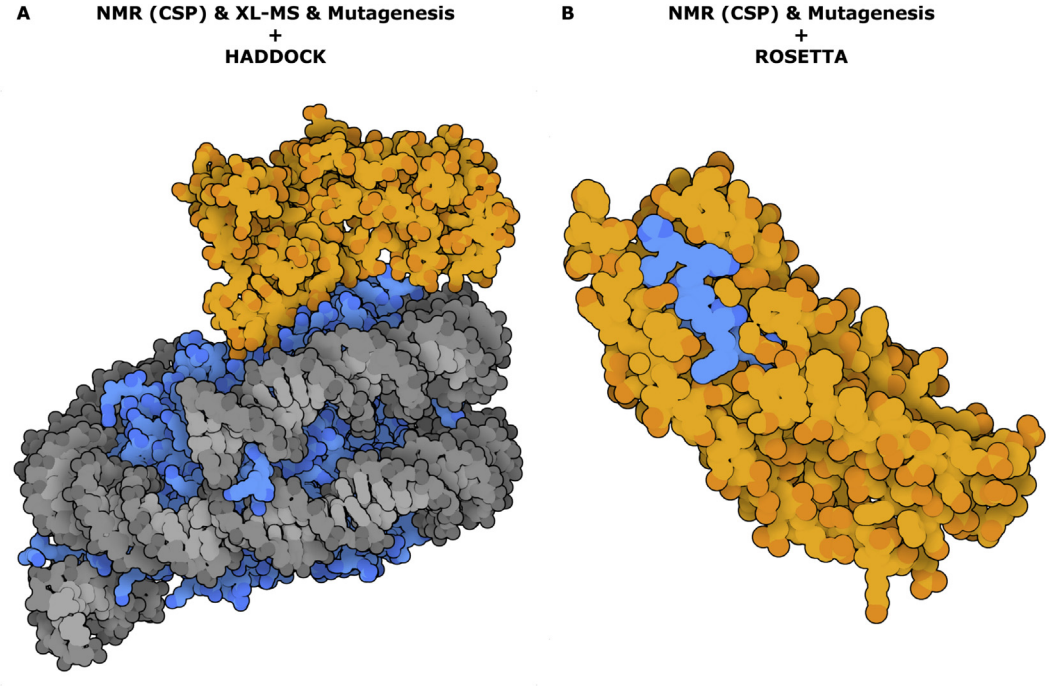


Fig. 2. Structural models determined with integrative approaches. Panel A shows a rendering of the nucleosome complex bound to UbcH5c and RNF168-RING domain. The model was determined with HADDOCK using NMR-derived spatial restraints (CSPs) in combination with mutagenesis and XL-MS data (PDB-dev entry 29). The DNA is shown in gray, the histones in blue, and the UbcH5c and RNF168-RING domain in orange. Panel B shows a rendering of the ghrelin peptide bound to its G-protein–coupled receptor (GHSR) (PDB-dev entry 24). The model was determined with ROSETTA using NMR-derived spatial restraints (CSPs) and mutagenesis data. GHSR is colored orange with the peptide in light blue with some clipping to enable visualization of the binding pocket. Both models are illustrated with the program ‘‘illustrate’’ by David Goodsell. Only the top model from each submission is shown.

of proteomics facilities as well as the high-throughput nature of modern MS should also be counted among the benefits of XL-MS. All these allow for rapid and semiautomated retrieval of the distance profiles after sample preparation is complete [90]. Additional advancements have been made in spectral analysis and database search algorithms [91–93] generating high-quality, quantitative XL-MS data which can be used to monitor the structure and dynamics of macromolecules and their assemblies in solution [94–96]. The wide variety of residues and chemical types which can be targeted for cross-linking ensures the wide applicability of the technique. Additionally, the combination of spacers of different lengths can provide distance information across varying scales, allowing us to capture data about both short and longer distances all of which can be used during the modeling process. Impressively, these experiments can be performed in intact cells or intact cell compartments, allowing us to extract information about the native state of the system under study [97–99]. Two major challenges of using cross-link data in docking are the fact that the cross-links captured might reflect multiple conformational states of the complex or assembly under study, the inherent difficulty of distinguishing between intra- and intermolecular cross-links—or cross-links between two residues of the same protein and residues of different proteins when dealing with symmetrical systems, and the high reactivity of the reagents which can capture nonnative, encounter complexes. Despite these, we expect XL-MS to further develop in the coming years, and XL-MS–derived distance restraints to become increasingly prevalent in integrative modeling as integrative modeling software also develops ways of dealing with these shortcomings in a consistent way. Similar to the HDX community, multiple leading MS groups have decided to establish community guidelines regarding best practises in sample preparation, measurement, data analysis, model validation, and result reporting [100–103]. Development of software which can automatically group cross-links into the conformational states to which they correspond and identify potential false positives would be a valuable addition. This would allow docking software which can make use of distance restraints to determine different structural models for the different states more closely reflecting the behavior of the system under study in solution. The DisVis standalone program and web server [104,105] can already perform the task of identifying potential false positives through enumeration of violations of distance restraints after exhaustive rotational and translational sampling, and developments to allow clustering of distinct distances into conformational states are already under way.

Nuclear magnetic resonance spectroscopy

Until fairly recently, NMR spectroscopy was one of the two ways (the other way of course being X-ray crystallography) of obtaining high-resolution structures of biomolecules. Its application to complexes was limited as solution-state NMR cannot routinely deal with complexes whose size is greater than 40 kDa [106]. The most easily obtainable measurements for macromolecular complexes are chemical shift perturbations (CSPs) which allow for the mapping of the interacting regions of biomolecules [107] at the residue/atomic level. In addition to CSPs, additional restraints can be recorded from NMR experiments such as, for example, intermolecular NOEs, residual dipolar couplings (RDCs), and relaxation anisotropy [108,109] which reveal details regarding the relative orientation of two interacting biomolecules. Paramagnetic probes [110] can be used to provide additional information to standard NMR experiments in the form of long-distance atomic information [111], to probe interacting surfaces of biomolecules [112,113] and study dynamics [114]. Unlike solution NMR, solid-state NMR (ssNMR) [115,116] has no theoretical limitations on the size of the systems which can be studied; however, obtaining atomic-resolution structures can be complicated because of the spectral complexity [117,118]. More recently, ssNMR has been applied successfully for the study of transmembrane systems ranging from the KcsA ion channel [25,119] to small peptide–based antibiotics which interfere with the lipid-II cycle [24]. Of course, the application of NMR—whether in solid or solution state—to small and flexible systems such as peptides is not new even when considering membrane-embedded peptides [120,121]. One of the recent developments in the field of NMR has been the ability to study molecules in cells allowing for qualitative comparisons between native and nonnative species or analysis of conformational heterogeneity across different cell types [122–125]. A limiting factor of NMR is the often costly procedure of preparing samples for NMR as well as the relative difficulty in analyzing and interpreting the experimental measurements. In light of these observations, we expect NMR to continue to factor significantly in integrative modeling over the coming years mainly owing to the undeniable benefit and unique ability of NMR of being able to study dynamics at atomic resolution in real time across different time scales. This is particularly attractive when compared with techniques such as XL-MS which can only estimate dynamics as a result of conformational heterogeneity observed in the distance profiles.

Cryo-electron microscopy

The techniques which fall under the umbrella of cryo-EM have been revolutionizing structural and integrative biology for a few years now. This is for the most part due to advancements in detector technology, automation, and software [126,127]. Cryo-EM–derived structures can now match or even surpass structures of the same system obtained with X-ray crystallography [128]. This is also reflected in the number of near atomic-resolution structures deposited in the EMDB [129] with a third of the structures deposited in 2018 having a resolution of 4 Å or better (https://www.emdataresource.org)—a trend which further improved in 2019. Single particle analysis (SPA) constitutes the overwhelming majority of cryo-EM experiments undertaken in recent years [128]. The typical cryo-EM SPA experiment constitutes of loading an aqueous solution containing the biological sample on a grid mesh, blotting to remove excess solution and to form a thin layer, and covering it with a thin carbon film after which the sample-loaded grid is plunge frozen. The particles are then imaged with an electron beam using sufficiently low doses to prevent radiation damage. Many 2D images are collected, aligned, and then used to computationally reconstruct the molecule in 3D [130,131]. When the resolution is not sufficient to determine the molecular structure at atomic or near-atomic resolution, various rigid or flexible fitting protocols can be used to fit existing structural models of the components of the assembly into the EM-derived map [132–135]. In the time period preceding the resolution revolution, these inherently integrative protocols were the most common way of generating structural models with cryo-EM (Fig. 1, panel C). More recently, popular codes such as ATTRACT, IMP (which supported cryo-EM data from day 1), HADDOCK, and ROSETTA have added support for cryo-EM–derived density maps during the modeling process [61,67,105,136–139].

However, the ever-increasing performance gains in terms of resolution for structures solved with cryo-EM pose some interesting questions for the field of integrative modeling, specifically is there a place for integrative approaches in an era where atomic resolution models for a wide variety of systems and molecular weights can routinely be obtained with cryo-EM data alone? We believe the answer is yes, for multiple reasons. First and most importantly, methodological limitations make it difficult to study small and/or flexible systems with cryo-EM. Some recently solved structures show, however, promising results in that direction [140,141] even potentially allowing us to study interactions between drugs and proteins [142]. Secondly, even in the cases where a high-resolution model has been obtained, the resolution distribution might not be uniform with some parts of the molecule having lower resolution than others. This nonuniformity can arise as a result of structural heterogeneity, nonisotropic distribution of sampled orientations, or even processing artefacts. Some groups have already suggested alternative ways of measuring resolution that take into account significant local variations from the reported mean value [143,144]. Sample structural heterogeneity is usually considered a limiting factor for cryo-EM as it makes the averaging of aligned images more difficult and results in lower-resolution models. Whereas most cryo-EM sample preparation protocols emphasize the importance of structural homogeneity to be able to generate high-resolution models, some recently described approaches embrace the importance of structural heterogeneity as an inherent property of dynamic systems such as biological samples, allowing for identification of distinct conformational states [145–147]. It is expected that identification of these distinct states will allow to simultaneously estimate the conformational landscape and thermodynamic behavior of the system. Such results would be very desirable when attempting to describe the intermediate states of a cellular process or when studying systems for which high structural variability is expected [148].

We conclude that despite the impressive advances made recently in the field of cryo-EM, we expect the importance of integrative approaches in the context of cryo-EM to increase. Integrative modeling might be used either as a way to validate the structural models, as a way to aid the modeling process for systems which are difficult to study with cryo-EM alone, or to model parts of the cryo-EM maps that might not reach sufficient resolution for de novo structure determination. The modeling of such systems from cryo-EM data can significantly benefit from the inclusion of additional data (e.g., from XL-MS or NMR [149,150]). The importance of integrative approaches can also be seen by recent studies which favorably compare integrative models with high-resolution structures of the same complexes made available years later by cryo-EM [33]. In light of these observations, we expect cryo-EM to play a prominent—if not dominant—role in many aspects of integrative modeling in the forthcoming years.

Small-angle X-ray scattering

Biological SAXS is the solution equivalent to X-ray crystallography. It is another field which has been undergoing a renaissance in recent years with more improvements expected in the next few years [151]. In a basic SAXS experiment, a macromolecular solution is bombarded with X-ray beams and the scattering pattern is recorded by a detector placed in close proximity to the sample. The most basic information that can be extracted from the measurement is the scattering curve which is extracted from the distance profile between all sample atoms which can in turn be used to construct a low-resolution

shape or envelope of the system under study [152]. The potential for SAXS data to be useful in integrative studies was realized early [153] with protocols resembling those that are used for fitting X-ray– or NMR-derived structural models into medium- to low-resolution EM density maps where the unbound structures of the components of the complex were docked against each other and the shape of the resulting complex was scored against the SAXS-calculated shape [154,155] or directly against the scattering curve [151,156–162]. More recently, protocols that can make use of the shape information to guide the docking toward conformations that agree with the SAXS shape have been described [163] (Fig. 1, panel C). The maturity of SAXS protocols, the standardization of guidelines for publishing SAXS data, the relative ease with which samples can be prepared, the automated manner of data acquisition and analysis, as well as the high-throughput nature of BIO SAXS are some of the factors which make SAXS a very attractive option for probing macromolecular interactions under solution conditions without a size limitation, but sample purity and homogeneity are important aspects to be able to derive reliable structural data [164]. In addition to calculating low-resolution shapes of macromolecular complexes, SAXS can be used to qualitatively and quantitatively compare samples, probe conformational differences, assembly states, folding status, and, in some cases, even refine flexible, low-resolution regions of structures determined with X-ray crystallography [152]. All these factors combine to paint a very favorable picture of SAXS in its current and future states. The ability to probe dynamics in solution without size limitations while at the same time deriving shape-based restraints which can either be used to restrain the sampling of docking simulations or filter out non–native-like solutions when scoring generated models are counted among its greatest strengths. We only expect the contribution of SAXS in the field of integrative modeling to further proliferate.

Other experimental sources of information

In addition to the experimental methods that have already been mentioned, structural biologists have access to a plethora of other methods giving various levels of experimental information about the interacting biomolecules. Covering all of these techniques as well as the ways in which the data that can be derived from them for use in integrative modeling is beyond the scope of this review. We will mention though two techniques standing out because of their high importance for the field of modeling, both of which can provide distance information between residues of the interacting biomolecules. The first is Förster Energy Resonance Transfer (FRET). It allows to detect the energy transfer between donor and acceptor fluorophores allowing for the calculation of long-range distances between those parts [165] (Fig. 1, panel B.2). It does, however, require covalent attachment of dyes to specific parts of the molecules. FRET data have been used successfully in integrative modeling efforts either alone or in combination with data from other sources to determine the structure of biomolecular complexes and study their dynamics [166–170]. Similarly to FRET, the double electron-electron resonance technique (DEER) is a spectroscopic approach which enables the calculation of long-distances between interacting spin labels that have been attached to specific residues (most commonly cysteines) [171,172] (Fig. 1, panel B.3). It has been applied widely to study systems of varying sizes and composition including small protein-protein complexes to large molecular machines and RNA-containing complexes [173–176].

Bioinformatics and computational approaches

Perhaps, some of the most interesting advancements in the field of integrative modeling in recent years originate from bioinformatics and computational techniques which, on their own, cannot be classified as integrative but whose output can be combined in integrative modeling frameworks just similar to experimental data. In this section, we are only going to provide a succinct overview of recent developments in the field, emphasizing three areas: The coming of age of coevolution, the appearance of membrane-specific modeling tools, and the use of coarse-graining approaches.

Coevolution

Coevolution rests on the observation that sometimes mutations at specific positions in a protein sequence correlate with mutations at other positions of the same or interacting proteins, the hypothesis being that if such residues "coevolve," they might be in spatial proximity. When a mutation is introduced in one of the interacting pairs, a compensatory mutation arises in the other because of evolutionary pressure relating to functional or conformational importance of that residue pair [177]. This information can be used in the structure determination of proteins [178] but most importantly for integrative modeling purposes. The concept can be quite easily extended to protein residues which belong to different proteins forming a complex or being part of a larger molecular assembly [179]. Methods such as EVcomplex, GREMLIN [180,181], and InterEvDock [182] have been applied successfully in docking simulations [183–185]. Of particular note is the recent development of InterEvDock2, a free and fully automated web server which allows the user to input sequences instead of

structures, submit multimeric next to monomeric components, and automatically derive coevolution-based restraints to use for scoring the models generated during the simulation [186]. The use of coevolution-based data does not stop with protein folding and determination of soluble protein-protein interfaces though. More recently, it has been used to determine transmembrane protein interaction sites [187,188], identify new protein-protein interaction networks [189], and novel protein contact maps making use of metagenomics data [190]. The robust state of the coevolution community in combination with the intuitive nature of the output data makes us confident that the use of coevolution-derived spatial restraints is only going to become more prevalent in the near future. One potential limiting factor for the use of coevolution-based restraints for docking is the need for extensive and diverse sequencing data to get deep enough alignments, although deep learning methods are becoming more robust with respect to the alignment depth [191]. This limitation can, for example, hamper their applicability for the study of mammalian systems, for which sequencing data are not as exhaustive compared with those of bacteria and yeast.

Membrane modeling

Another field which has attracted attention recently and has seen many developments is membrane protein modeling. It is traditionally considered as one of the most difficult kinds of systems to study with experimental structural biology methods because of the nature of the lipid bilayer which requires that, either it is dissolved with detergents and reconstituted or that native or native-like membrane mimetics are used. The former is easier and has been used with success for X-ray and NMR but raises questions about the effect the detergent has on the 3D structure. The use of native or native-like membrane mimetics is much closer to physiological conditions, which means that any structure determined this way should be closer to its counterpart in the cellular environment, but this introduces many challenges in sample preparation and measurements. A relatively recent advancement enabling studies of membrane proteins and their complexes in native-like and even native environments is the advent of styrene–maleic acid (SMA) copolymers which can be used in combination with synthetic liposomes or native membranes to solubilize patches of protein-containing lipid bilayers without the adverse effects of detergents [192]. These, so-called, SMALPs (SMA-lipid particles) have already been used together with MS to determine the stoichiometry [193], acquire atomic resolution structures of membrane protein complexes using X-ray crystallography and cryo-EM [194,195], as well as study their dynamics with solid-state NMR [196,197]. Computational methods therefore remain an attractive alternative for the study of membrane-bound or membrane-associated proteins and their complexes [198]. The simplest—yet most effective—way in which membrane protein modeling has been made easier in the recent years comes from a higher availability of representative 3D structures in the PDB, thanks, in no small part, to advances in cryo-EM [199]. This, in combination with the availability of membrane-specific homology modeling, tools such as MEDELLER [200] and Memoir [201], which implement protocols similar to and inspired by one of the most popular homology modeling tools—MODELLER [202]—enables the creation of structural models that strongly approximate native ones [203]. These can be used to confidently model structures which have not yet been determined experimentally. These models can act as the starting point for further investigation, which usually involves some degree of integrative modeling, for example, rigid/flexible fitting in low-medium resolution cryo-EM maps or embedding into membranes and studying the system by MD. This wider availability of transmembrane (TM) protein structures is also reflected in the enrichment of entries in databases that deal with membrane proteins exclusively [199], such as the manually curated *mpstruc* (membrane proteins of known 3D structure—https://blanco.biomol.uci.edu/mpstruc/), which annotates all non-redundant proteins in the PDB. The latter also serves as the starting point for the classification system the PDB uses for identifying entries as TM, for OPM (orientations of proteins in membranes—https://opm.phar.umich.edu/) [204], which computes the membrane insertion angle, tilt and width for transmembrane and membrane-associated proteins, and for MemProtMD (http://memprotmd.bioch.ox.ac.uk/) which inserts proteins in lipid bilayers via self-assembly with coarse-grained MD simulations and also makes available the preequilibrated membrane bilayer–protein structures [205,206]. The plethora of G-protein–coupled receptor (GPCR) structures which have been solved recently is of major importance not only to the structural biology community but to areas of pharmaceutical research as well owing to the importance of GPCRs in many diseases [207]. These structures, along with important details regarding the method that was used to determine them, the conditions under which the experiments were performed, and various aggregated statistics and analyses are collected in GPCRdb (GPCR database—https://gpcrdb.org/) [208]. Coarse-grained MD force fields such as MARTINI [209–212], which was originally developed for membrane, have also been extended to include proteins and nucleic acids. These allow to simulate larger systems and/or reach longer time scales. The MARTINI force field has recently been implemented into HADDOCK for the modeling of

proteins and nucleic acids complexes [213,214]. Other docking codes such as ATTRACT [61], CABS-dock [215,216], the Integrative Modeling Platform (IMP) [66,67], and PyRy3D (http://pyry3d.icm.edu.pl/) also support coarse-graining. Despite these significant advances, the only docking codes which currently offer the ability to dock TM proteins with specific implicit membrane potentials are (to the best of our knowledge) ROSETTA [217,218], DOCK/PIERR [219,220], and Memdock [221]. More recently a generic, ready-to-dock benchmark of membrane protein complexes accompanied by docking decoys for the purpose of training membrane-specific scoring functions was made available [222,223]. The lack of widely available explicit support for the docking of membrane proteins has resulted in some creative integrative modeling with, for example, researchers using HADDOCK to probe the interaction between the K-RAS4B oncogenic protein when complexed with the Cmpd2 inhibitor and lipid nanodisks making use of NMR-derived restraints to drive the simulation [224]. In summary, we believe it is high time the field of membrane complexes modeling is given the attention it deserves by the docking community as all the ingredients for successful integrative modeling are in place, with experimental methods providing good template structures for modeling as well as experimental restraints, computational tools similar to coevolution providing additional data to drive the docking, and plentiful implicit or explicit implementations of membrane bilayers allowing for studies at different representation levels.

Perspectives

We have highlighted some key areas of experimental and computational structural biology and identified the ones which, we believe, will factor significantly in the coming years for the field of integrative modeling in general and molecular docking specifically. Despite these advances, there are, however, also some areas for which we believe developments have been lacking. Chief among these is the fact that many docking codes can still not make use of information during the simulation, instead only in the scoring stage, and therefore cannot be considered integrative approaches, with some exceptions existing, e.g., ATTRACT, HADDOCK, IMP, PyRy3D, and RosettaDock to cite the most known ones. Another limiting factor is the fact that the number of distinct subunits which can be included in the simulation is still limited, with most codes, except a few, supporting only one receptor and one ligand [1,186,225], i.e., binary complexes.

Another aspect of integrative modeling is that being able to combine multiple sources of information into a single docking run does not necessarily mean that the resulting models benefit from the included information. The reason for this is that information needs to be combined in a probabilistically sound way, that is, in a way that reflects the uncertainty of the original measurements and properly propagates it [66,226]. Perhaps, the most well-known examples of software which properly accounts for this and weights the multiple data sources used in the modeling through a Bayesian framework are IMP [66,67] and the Inferential Structure Determination Software (ISD) [227]. IMP has most famously been used for the determination of an integrative model for the nuclear pore complex [228], which was validated last year when the cryo-EM structure for the entire complex was solved with a final resolution of 28 Å [229]. ISD [227], originally developed for NMR [230], has recently been extended and applied to challenging systems such as membrane proteins, bacterial pili and chromosomes [231–233], and also large macromolecular assemblies using shape-based (SAXS or cryo-EM) data [234].

Another alternative to one-stop integrative modeling software is the combination of multiple codes in easy-to-use and cohesive workflows which hide the technical details away from the end users and allow for seamless flow of information between different packages. Some encouraging work in this direction has already started with CROSS-ID [103], a package for the analysis and visualization of XL-MS data which is part of XlinkX [98,102] and makes uses of DisVis for the visualization and validation of cross-link data. Another interesting initiative is the BioExcel consortium as one of their stated goals is to promote integration among several flagship computational biology/chemistry packages such as HADDOCK and GROMACS [235,236].

Finally, next to development in integrative software, proper description and archival of integrative models is an important area which is benefiting from the advent of PDB-dev [32,237], a portal developed by wwPDB in collaboration and consultation with experts in the field of integrative modeling. Its aim is to act as a hub to collect structural models, and all their associated data, that have been determined by integrative approaches. Two examples of integrative models deposited into PDB-dev obtained with various software and data types are shown in Fig. 2.

Conclusions

In this review, we have discussed aspects of integrative modeling and in particular recent developments related to the various types of information that can be used to aid the modeling process. We conclude that the future for integrative modeling software is bright as the availability and quality of data are only going to increase as will the ability of algorithms and hardware to handle that data efficiently and meaningfully. There still remain,

however, long-standing challenges, such as accurate binding affinity prediction and accurately modeling large conformational changes. These have been challenges in the biomolecular simulation world since the very first days of the field [1,225]. Our ability to model the structure of biomolecules as well as biomolecular complexes has been continuously evaluated over a period spanning more than 20 years in the CASP (Critical Assessment of Structure Prediction) [238] and CAPRI (Critical Assessment of PRediction of Interactions) [239] experiments, with the first focusing on single protein structure prediction (with a multimer component) and the latter on protein complexes. In recent iterations of the challenge, CASP has featured a data-assisted category for which some information about the target system is disseminated to the participating groups, thus evaluating the ability of software to incorporate information in the prediction and its outcome. In CAPRI so far, only once was a SAXS scattering profile provided. The field would clearly benefit from truly integrative blind modeling challenges as such blind challenges have been and will remain important catalysts for further development and advances.

Funding

This work was supported by the European H2020 e-Infrastructure grants BioExcel (Grants No. 675728 and 823830) and from the Dutch Foundation for Scientific Research (NWO) (TOP-PUNT Grant 718.015.001).

Conflicts of Interest

The authors declare no competing financial interest.

Received 14 September 2019;
Received in revised form 12 November 2019;
Accepted 13 November 2019
Available online 26 November 2019

Keywords:
Molecular simulations;
Docking;
Interactions;
Structural biology;
Membrane proteins

References

[1] J.P.G.L.M. Rodrigues, A.M.J.J. Bonvin, Integrative computational modeling of protein interactions, FEBS J. 281 (2014) 1988–2003, https://doi.org/10.1111/febs.12771.

[2] M.P. Rout, A. Sali, Principles for integrative structural biology studies, Cell 177 (2019) 1384–1403, https://doi.org/10.1016/j.cell.2019.05.016.

[3] H. Kato, H. van Ingen, B.-R. Zhou, H. Feng, M. Bustin, L.E. Kay, Y. Bai, Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR, Proc. Natl. Acad. Sci. 108 (2011) 12283–12288, https://doi.org/10.1073/pnas.1105848108.

[4] I. Corbeski, K. Dolinar, H. Wienk, R. Boelens, H. van Ingen, DNA repair factor APLF acts as a H2A-H2B histone chaperone through binding its DNA interaction surface, Nucleic Acids Res. 46 (2018) 7138–7152, https://doi.org/10.1093/nar/gky507.

[5] V. Horn, H. van Ingen, Recognition of nucleosomes by chromatin factors: lessons from data-driven docking-based structures of nucleosome-protein complexes, in: Epigenetics [Working Title], IntechOpen, 2018, https://doi.org/10.5772/intechopen.81016.

[6] S. Xiang, U.B. le Paige, V. Horn, K. Houben, M. Baldus, H. van Ingen, Site-specific studies of nucleosome interactions by solid-state NMR spectroscopy, Angew. Chem. Int. Ed. 57 (2018) 4571–4575, https://doi.org/10.1002/anie.201713158.

[7] C.L. van Emmerik, H. van Ingen, Unspinning chromatin: revealing the dynamic nucleosome landscape by NMR, Prog. Nucl. Magn. Reson. Spectrosc. 110 (2019) 1–19, https://doi.org/10.1016/j.pnmrs.2019.01.002.

[8] R. Sprangers, L.E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR, Nature 445 (2007) 618–622, https://doi.org/10.1038/nature05512.

[9] T.L. Religa, R. Sprangers, L.E. Kay, Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR, Science 328 (2010) 98–102, https://doi.org/10.1126/science.1184991.

[10] A.M. Ruschak, L.E. Kay, Proteasome allostery as a population shift between interchanging conformers, Proc. Natl. Acad. Sci. 109 (2012) E3454–E3462, https://doi.org/10.1073/pnas.1213640109.

[11] R. Huang, F. Pérez, L.E. Kay, Probing the cooperativity of thermoplasma acidophilum proteasome core particle gating by NMR spectroscopy, Proc. Natl. Acad. Sci. 114 (2017) E9846–E9854, https://doi.org/10.1073/pnas.1712297114.

[12] J.L. Kitevski-LeBlanc, T. Yuwen, P.N. Dyer, J. Rudolph, K. Luger, L.E. Kay, Investigating the dynamics of destabilized nucleosomes using Methyl-TROSY NMR, J. Am. Chem. Soc. 140 (2018) 4774–4777, https://doi.org/10.1021/jacs.8b00931.

[13] S.A. Fromm, V. Truffault, J. Kamenz, J.E. Braun, N.A. Hoffmann, E. Izaurralde, R. Sprangers, The structural basis of Edc3- and Scd6-mediated activation of the Dcp1: Dcp2 mRNA decapping complex, EMBO J. 31 (2012) 279–290, https://doi.org/10.1038/emboj.2011.408.

[14] S.A. Fromm, J. Kamenz, E.R. Nöldeke, A. Neu, G. Zocher, R. Sprangers, In vitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery, Angew. Chem. Int. Ed. 53 (2014) 7354–7359, https://doi.org/10.1002/anie.201402885.

[15] S. Schütz, E.R. Nöldeke, R. Sprangers, A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping, Nucleic Acids Res. 45 (2017) 6911–6922, https://doi.org/10.1093/nar/gkx353.

[16] M.A. Cvetkovic, R. Sprangers, Methyl TROSY spectroscopy to study large biomolecular complexes, in: Mod.

Magn. Reson., Springer International Publishing, Cham, 2018, pp. 453–467, https://doi.org/10.1007/978-3-319-28388-3_45.
[17] T. Saleh, P. Rossi, C.G. Kalodimos, Atomic view of the energy landscape in the allosteric regulation of Abl kinase, Nat. Struct. Mol. Biol. 24 (2017) 893–901, https://doi.org/10.1038/nsmb.3470.
[18] T. Saleh, C.G. Kalodimos, Enzymes at work are enzymes in motion, Science 355 (2017) 247–248, https://doi.org/10.1126/science.aal4632.
[19] C. Huang, P. Rossi, T. Saio, C.G. Kalodimos, Structural basis for the antifolding activity of a molecular chaperone, Nature 537 (2016) 202–206, https://doi.org/10.1038/nature18965.
[20] H. van Ingen, A.M.J.J. Bonvin, Information-driven modeling of large macromolecular assemblies using NMR data, J. Magn. Reson. 241 (2014) 103–114, https://doi.org/10.1016/j.jmr.2013.10.021.
[21] M. Renault, A. Cukkemane, M. Baldus, Solid-state NMR spectroscopy on complex biomolecules, Angew. Chem. Int. Ed. 49 (2010) 8346–8357, https://doi.org/10.1002/anie.201002823.
[22] A.E. Aliev, R.V. Law, Chapter 7. Solid state NMR spectroscopy, in: Nucl. Magn. Reson., 2014, pp. 286–344, https://doi.org/10.1039/9781849738125-00286.
[23] A. Marchanka, T. Carlomagno, Solid-state NMR spectroscopy of RNA, in: Methods Enzymol., 2019, pp. 333–371, https://doi.org/10.1016/bs.mie.2018.08.029.
[24] J. Medeiros-Silva, S. Jekhmane, A.L. Paioni, K. Gawarecka, M. Baldus, E. Swiezewska, E. Breukink, M. Weingarth, High-resolution NMR studies of antibiotics in cellular membranes, Nat. Commun. 9 (2018) 3963, https://doi.org/10.1038/s41467-018-06314-x.
[25] S. Jekhmane, J. Medeiros-Silva, J. Li, F. Kümmerer, C. Müller-Hermes, M. Baldus, B. Roux, M. Weingarth, Shifts in the selectivity filter dynamics cause modal gating in K+ channels, Nat. Commun. 10 (2019) 123, https://doi.org/10.1038/s41467-018-07973-6.
[26] F.L. Kyrilis, A. Meister, P.L. Kastritis, Integrative biology of native cell extracts: a new era for structural characterization of life processes, Biol. Chem. 400 (2019) 831–846, https://doi.org/10.1515/hsz-2018-0445.
[27] C.L. Lawson, A. Patwardhan, M.L. Baker, C. Hryc, E.S. Garcia, B.P. Hudson, I. Lagerstedt, S.J. Ludtke, G. Pintilie, R. Sala, J.D. Westbrook, H.M. Berman, G.J. Kleywegt, W. Chiu, EMDataBank unified data resource for 3DEM, Nucleic Acids Res. 44 (2016) D396–D403, https://doi.org/10.1093/nar/gkv1126.
[28] C.L. Lawson, W. Chiu, Comparing cryo-EM structures, J. Struct. Biol. 204 (2018) 523–526, https://doi.org/10.1016/j.jsb.2018.10.004.
[29] I.A. Vakser, E.J. Deeds, Computational approaches to macromolecular interactions in the cell, Curr. Opin. Struct. Biol. 55 (2019) 59–65, https://doi.org/10.1016/j.sbi.2019.03.012.
[30] N. Soni, M.S. Madhusudhan, Computational modeling of protein assemblies, Curr. Opin. Struct. Biol. 44 (2017) 179–189, https://doi.org/10.1016/j.sbi.2017.04.006.
[31] T. Lengauer, M. Rarey, Computational methods for biomolecular docking, Curr. Opin. Struct. Biol. 6 (1996) 402–406, https://doi.org/10.1016/S0959-440X(96)80061-3.
[32] B. Vallat, B. Webb, J.D. Westbrook, A. Sali, H.M. Berman, Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules, Structure 26 (2018) 894–904.e2, https://doi.org/10.1016/j.str.2018.03.011.
[33] M. Braitbard, D. Schneidman-Duhovny, N. Kalisman, Integrative structure modeling: overview and assessment, Annu. Rev. Biochem. 88 (2019) 113–135, https://doi.org/10.1146/annurev-biochem-013118-111429.
[34] S.J. Wodak, J. Janin, Computer analysis of protein-protein interaction, J. Mol. Biol. 124 (1978) 323–342, https://doi.org/10.1016/0022-2836(78)90302-9.
[35] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, A geometric approach to macromolecule-ligand interactions, J. Mol. Biol. 161 (1982) 269–288, https://doi.org/10.1016/0022-2836(82)90153-X.
[36] A. Tovchigrechko, C.A. Wells, I.A. Vakser, Docking of protein models, Protein Sci. 11 (2002) 1888–1896, https://doi.org/10.1110/ps.4730102.
[37] N. Plattner, S. Doerr, G. De Fabritiis, F. Noé, Complete protein–protein association kinetics in atomic detail revealed by molecular dynamics simulations and Markov modelling, Nat. Chem. 9 (2017) 1005–1011, https://doi.org/10.1038/nchem.2785.
[38] P. Śledź, A. Caflisch, Protein structure-based drug design: from docking to molecular dynamics, Curr. Opin. Struct. Biol. 48 (2018) 93–102, https://doi.org/10.1016/j.sbi.2017.10.010.
[39] I.A. Vakser, Protein-protein docking: from interaction to interactome, Biophys. J. 107 (2014) 1785–1793, https://doi.org/10.1016/j.bpj.2014.08.033.
[40] I. Anishchenko, P.J. Kundrotas, A.V. Tuzikov, I.A. Vakser, Protein models: the Grand Challenge of protein docking, Proteins Struct. Funct. Bioinf. 82 (2014) 278–287, https://doi.org/10.1002/prot.24385.
[41] R. Mosca, A. Céol, P. Aloy, Interactome3D: adding structural details to protein networks, Nat. Methods 10 (2013) 47–53, https://doi.org/10.1038/nmeth.2289.
[42] K. Luck, D.-K. Kim, L. Lambourne, K. Spirohn, B.E. Begg, W. Bian, R. Brignall, T. Cafarelli, F.J. Campos-Laborie, B. Charloteaux, D. Choi, A.G. Cote, M. Daley, S. Deimling, A. Desbuleux, A. Dricot, M. Gebbia, M.F. Hardy, N. Kishore, J.J. Knapp, I.A. Kovács, I. Lemmens, M.W. Mee, J.C. Mellor, C. Pollis, C. Pons, A.D. Richardson, S. Schlabach, B. Teeking, A. Yadav, M. Babor, D. Balcha, O. Basha, C. Bowman-Colin, S.-F. Chin, S.G. Choi, C. Colabella, G. Coppin, C. D'Amata, D. De Ridder, S. De Rouck, M. Duran-Frigola, H. Ennajdaoui, F. Goebels, L. Goehring, A. Gopal, G. Haddad, E. Hatchi, M. Helmy, Y. Jacob, Y. Kassa, S. Landini, R. Li, N. van Lieshout, A. MacWilliams, D. Markey, J.N. Paulson, S. Rangarajan, J. Rasla, A. Rayhan, T. Rolland, A. San-Miguel, Y. Shen, D. Sheykhkarimli, G.M. Sheynkman, E. Simonovsky, M. Tasan, A. Tejeda, J.-C. Twizere, Y. Wang, R.J. Weatheritt, J. Weile, Y. Xia, X. Yang, E. Yeger-Lotem, Q. Zhong, P. Aloy, G.D. Bader, J.D. Las Rivas, S. Gaudet, T. Hao, J. Rak, J. Tavernier, V. Tropepe, D.E. Hill, M. Vidal, F.P. Roth, M.A. Calderwood, A reference map of the human protein interactome, BioRxiv (2019) 605451, https://doi.org/10.1101/605451.
[43] N. Moitessier, P. Englebienne, D. Lee, J. Lawandi, C.R. Corbeil, Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go, Br. J. Pharmacol. 153 (2009) S7–S26, https://doi.org/10.1038/sj.bjp.0707515.
[44] C. Nithin, P. Ghosh, J. Bujnicki, Bioinformatics tools and benchmarks for computational docking and 3D structure

prediction of RNA-protein complexes, Genes 9 (2018) 432, https://doi.org/10.3390/genes9090432.
[45] A. Krüger, F.M. Zimbres, T. Kronenberger, C. Wrenger, Molecular modeling applied to nucleic acid-based molecule development, Biomolecules 8 (2018) 83, https://doi.org/10.3390/biom8030083.
[46] J. Luo, W. Wei, J. Waldispühl, N. Moitessier, Challenges and current status of computational methods for docking small molecules to nucleic acids, Eur. J. Med. Chem. 168 (2019) 414–425, https://doi.org/10.1016/j.ejmech.2019.02.046.
[47] S.F. Sousa, P.A. Fernandes, M.J. Ramos, Protein-ligand docking: current status and future challenges, Proteins Struct. Funct. Bioinf. 65 (2006) 15–26, https://doi.org/10.1002/prot.21082.
[48] G.M. Morris, M. Lim-Wilby, Molecular docking, in: Methods Mol. Biol.,, 2008, pp. 365–382, https://doi.org/10.1007/978-1-59745-177-2_19.
[49] N. London, B. Raveh, O. Schueler-Furman, Peptide docking and structure-based characterization of peptide binding: from knowledge to know-how, Curr. Opin. Struct. Biol. 23 (2013) 894–902, https://doi.org/10.1016/j.sbi.2013.07.006.
[50] J. Janin, K. Henrick, J. Moult, L. Ten Eyck, M.J.E. Sternberg, S. Vajda, I. Vakser, S.J. Wodak, CAPRI: a critical assessment of PRedicted interactions, Proteins Struct. Funct. Genet. 52 (2003) 2–9, https://doi.org/10.1002/prot.10381.
[51] R. Méndez, R. Leplae, M.F. Lensink, S.J. Wodak, Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures, Proteins Struct. Funct. Bioinf. 60 (2005) 150–169, https://doi.org/10.1002/prot.20551.
[52] M.F. Lensink, R. Méndez, S.J. Wodak, Docking and scoring protein complexes: CAPRI 3rd Edition, Proteins Struct. Funct. Bioinf. 69 (2007) 704–718, https://doi.org/10.1002/prot.21804.
[53] M.F. Lensink, S.J. Wodak, Docking and scoring protein interactions: CAPRI 2009, Proteins Struct. Funct. Bioinf. 78 (2010) 3073–3084, https://doi.org/10.1002/prot.22818.
[54] M.F. Lensink, S.J. Wodak, Docking, scoring, and affinity prediction in CAPRI, Proteins Struct. Funct. Bioinf. 81 (2013) 2082–2095, https://doi.org/10.1002/prot.24428.
[55] M.F. Lensink, S. Velankar, S.J. Wodak, Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition, Proteins Struct. Funct. Bioinf. 85 (2017) 359–377, https://doi.org/10.1002/prot.25215.
[56] K.A. Porter, I. Desta, D. Kozakov, S. Vajda, What method to use for protein–protein docking? Curr. Opin. Struct. Biol. 55 (2019) 1–7, https://doi.org/10.1016/j.sbi.2018.12.010.
[57] M.F. Lensink, G. Brysbaert, N. Nadzirin, S. Velankar, R.A.G. Chaleil, T. Gerguri, P.A. Bates, E. Laine, A. Carbone, S. Grudinin, R. Kong, R.-R. Liu, X. Xu, H. Shi, S. Chang, M. Eisenstein, A. Karczynska, C. Czaplewski, E. Lubecka, A. Lipska, P. Krupa, M. Mozolewska, Ł. Golon, S. Samsonov, A. Liwo, S. Crivelli, G. Pagès, M. Karasikov, M. Kadukova, Y. Yan, S.-Y. Huang, M. Rosell, L.A. Rodríguez-Lumbreras, M. Romero-Durana, L. Díaz-Bueno, J. Fernandez-Recio, C. Christoffer, G. Terashi, W.-H. Shin, T. Aderinwale, S. Raghavendra Maddhuri Venkata Subram, D. Kihara, D. Kozakov, S. Vajda, K. Porter, D. Padhorny, I. Desta, D. Beglov, M. Ignatov, S. Kotelnikov, I.H. Moal, D.W. Ritchie, I. Chauvot de Beauchêne, B. Maigret, M.-D. Devignes, M.E.R. Echartea, D. Barradas-Bautista, Z. Cao, L. Cavallo, R. Oliva, Y. Cao, Y. Shen, M. Baek, T. Park, H. Woo, C. Seok, M. Braitbard, L. Bitton, D. Scheidman-Duhovny, J. DapkŪnas, K. Olechnovič, Č. Venclovas, P.J. Kundrotas, S. Belkin, D. Chakravarty, V.D. Badal, I.A. Vakser, T. Vreven, S. Vangaveti, T. Borrman, Z. Weng, J.D. Guest, R. Gowthaman, B.G. Pierce, X. Xu, R. Duan, L. Qiu, J. Hou, B. Ryan Merideth, Z. Ma, J. Cheng, X. Zou, P.I. Koukos, J. Roel-Touris, F. Ambrosetti, C. Geng, J. Schaarschmidt, M.E. Trellet, A.S.J. Melquiond, L. Xue, B. Jiménez-García, C.W. Noort, R. V Honorato, A.M.J.J. Bonvin, S.J. Wodak, Blind prediction of homo- and hetero- protein complexes: the CASP13-CAPRI experiment, Proteins Struct. Funct. Bioinf. 0 (2019) 25838, https://doi.org/10.1002/prot.25838.
[58] C. Dominguez, R. Boelens, A.M.J.J. Bonvin, HADDOCK: a Protein–Protein docking approach based on biochemical or biophysical information, J. Am. Chem. Soc. 125 (2003) 1731–1737, https://doi.org/10.1021/ja026939x.
[59] G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries, A.M.J.J. Bonvin, The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes, J. Mol. Biol. 428 (2016) 720–725, https://doi.org/10.1016/j.jmb.2015.09.014.
[60] M. Zacharias, Protein-protein docking with a reduced protein model accounting for side-chain flexibility, Protein Sci. 12 (2003) 1271–1282, https://doi.org/10.1110/ps.0239303.
[61] S.J. de Vries, M. Zacharias, ATTRACT-EM: a new method for the computational assembly of large molecular machines using cryo-EM maps, PLoS One 7 (2012), e49733, https://doi.org/10.1371/journal.pone.0049733.
[62] C.E.M. Schindler, S.J. de Vries, A. Sasse, M. Zacharias, SAXS data alone can generate high-quality models of protein-protein complexes, Structure 24 (2016) 1387–1397, https://doi.org/10.1016/j.str.2016.06.007.
[63] S.J. de Vries, I. Chauvot de Beauchêne, C.E.M. Schindler, M. Zacharias, Cryo-EM data are superior to contact and interface information in integrative modeling, Biophys. J. 110 (2016) 785–797, https://doi.org/10.1016/j.bpj.2015.12.038.
[64] D.W. Ritchie, V. Venkatraman, Ultra-fast FFT protein docking on graphics processors, Bioinformatics 26 (2010) 2398–2405, https://doi.org/10.1093/bioinformatics/btq444.
[65] G. Macindoe, L. Mavridis, V. Venkatraman, M.-D. Devignes, D.W. Ritchie, HexServer: an FFT-based protein docking server powered by graphics processors, Nucleic Acids Res. 38 (2010) W445–W449, https://doi.org/10.1093/nar/gkq311.
[66] F. Alber, F. Förster, D. Korkin, M. Topf, A. Sali, Integrating diverse data for structure determination of macromolecular assemblies, Annu. Rev. Biochem. 77 (2008) 443–477, https://doi.org/10.1146/annurev.biochem.77.060407.135530.
[67] D. Russel, K. Lasker, B. Webb, J. Velázquez-Muriel, E. Tjioe, D. Schneidman-Duhovny, B. Peterson, A. Sali, Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies, PLoS Biol. 10 (2012), e1001244, https://doi.org/10.1371/journal.pbio.1001244.
[68] B. Jiménez-García, J. Roel-Touris, M. Romero-Durana, M. Vidal, D. Jiménez-González, J. Fernández-Recio, LightDock: a new multi-scale approach to protein–protein docking, Bioinformatics 34 (2018) 49–55, https://doi.org/10.1093/bioinformatics/btx555.

[69] J. Roel-Touris, A.M.J.J. Bonvin, B. Jiménez-García, LightDock goes information-driven, Bioinformatics (2019), https://doi.org/10.1093/bioinformatics/btz642.
[70] C. Wang, P. Bradley, D. Baker, Protein–protein docking with backbone flexibility, J. Mol. Biol. 373 (2007) 503–519, https://doi.org/10.1016/j.jmb.2007.07.050.
[71] S. Chaudhury, M. Berrondo, B.D. Weitzner, P. Muthu, H. Bergman, J.J. Gray, Benchmarking and analysis of protein docking performance in rosetta v3.2, PLoS One 6 (2011), e22477, https://doi.org/10.1371/journal.pone.0022477.
[72] R.A. Flavell, D.L. Sabo, E.F. Bandle, C. Weissmann, Site-directed mutagenesis: effect of an extracistronic mutation on the in vitro propagation of bacteriophage Qbeta RNA, Proc. Natl. Acad. Sci. 72 (1975) 367–371, https://doi.org/10.1073/pnas.72.1.367.
[73] D. Shortle, D. Nathans, Local mutagenesis: a method for generating viral mutants with base substitutions in preselected regions of the viral genome, Proc. Natl. Acad. Sci. 75 (1978) 2170–2174, https://doi.org/10.1073/pnas.75.5.2170.
[74] D. Shortle, D. DiMaio, D. Nathans, Directed mutagenesis, Annu. Rev. Genet. 15 (1981) 265–294, https://doi.org/10.1146/annurev.ge.15.120181.001405.
[75] K.L. Heckman, L.R. Pease, Gene splicing and mutagenesis by PCR-driven overlap extension, Nat. Protoc. 2 (2007) 924–932, https://doi.org/10.1038/nprot.2007.132.
[76] F.F. Vajdos, C.W. Adams, T.N. Breece, L.G. Presta, A.M. de Vos, S.S. Sidhu, Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis, J. Mol. Biol. 320 (2002) 415–428, https://doi.org/10.1016/S0022-2836(02)00264-4.
[77] E.W. Yu, J.R. Aires, G. McDermott, H. Nikaido, A periplasmic drug-binding site of the AcrB multidrug efflux pump: a crystallographic and site-directed mutagenesis study, J. Bacteriol. 187 (2005) 6804–6815, https://doi.org/10.1128/JB.187.19.6804-6815.2005.
[78] A. Ashkenazi, L.G. Presta, S.A. Marsters, T.R. Camerato, K.A. Rosenthal, B.M. Fendly, D.J. Capon, Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis, Proc. Natl. Acad. Sci. 87 (1990) 7150–7154, https://doi.org/10.1073/pnas.87.18.7150.
[79] A. Bill, E.M. Rosethorne, T.C. Kent, L. Fawcett, L. Burchell, M.T. van Diepen, A. Marelli, S. Batalov, L. Miraglia, A.P. Orth, N.A. Renaud, S.J. Charlton, M. Gosling, L.A. Gaither, P.J. Groot-Kormelink, High throughput mutagenesis for identification of residues regulating human prostacyclin (hIP) receptor expression and function, PLoS One 9 (2014), e97973, https://doi.org/10.1371/journal.pone.0097973.
[80] F.M. Heydenreich, T. Miljuš, R. Jaussi, R. Benoit, D. Milić, D.B. Veprintsev, High-throughput mutagenesis using a two-fragment PCR approach, Sci. Rep. 7 (2017) 6787, https://doi.org/10.1038/s41598-017-07010-4.
[81] Y. Luo, D. Zhu, Z. Zhang, Y. Chen, X. Sun, Integrative analysis of CRISPR/Cas9 target sites in the human HBB gene, BioMed Res. Int. 2015 (2015) 1–9, https://doi.org/10.1155/2015/514709.
[82] B. Zhang, X. Yang, C. Yang, M. Li, Y. Guo, Exploiting the CRISPR/Cas9 system for targeted genome mutagenesis in petunia, Sci. Rep. 6 (2016) 20315, https://doi.org/10.1038/srep20315.
[83] S.W. Englander, N.R. Kallenbach, Hydrogen exchange and structural dynamics of proteins and nucleic acids, Q. Rev. Biophys. 16 (1983) 521–655, https://doi.org/10.1017/S0033583500005217.
[84] L. Konermann, J. Pan, Y.-H. Liu, Hydrogen exchange mass spectrometry for studying protein structure and dynamics, Chem. Soc. Rev. 40 (2011) 1224–1234, https://doi.org/10.1039/C0CS00113A.
[85] G.R. Masson, J.E. Burke, N.G. Ahn, G.S. Anand, C. Borchers, S. Brier, G.M. Bou-Assaf, J.R. Engen, S.W. Englander, J. Faber, R. Garlish, P.R. Griffin, M.L. Gross, M. Guttman, Y. Hamuro, A.J.R. Heck, D. Houde, R.E. Iacob, T.J.D. Jørgensen, I.A. Kaltashov, J.P. Klinman, L. Konermann, P. Man, L. Mayne, B.D. Pascal, D. Reichmann, M. Skehel, J. Snijder, T.S. Strutzenberg, E.S. Underbakke, C. Wagner, T.E. Wales, B.T. Walters, D.D. Weis, D.J. Wilson, P.L. Wintrode, Z. Zhang, J. Zheng, D.C. Schriemer, K.D. Rand, Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments, Nat. Methods 16 (2019) 595–602, https://doi.org/10.1038/s41592-019-0459-y.
[86] M. Rey, V. Sarpe, K.M. Burns, J. Buse, C.A.H. Baker, M. van Dijk, L. Wordeman, A.M.J.J. Bonvin, D.C. Schriemer, Mass spec studio for integrative structural biology, Structure 22 (2014) 1538–1548, https://doi.org/10.1016/j.str.2014.08.013.
[87] A. Sinz, C. Arlt, D. Chorev, M. Sharon, Chemical cross-linking and native mass spectrometry: a fruitful combination for structural biology, Protein Sci. 24 (2015) 1193–1209, https://doi.org/10.1002/pro.2696.
[88] A. Leitner, M. Faini, F. Stengel, R. Aebersold, Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines, Trends Biochem. Sci. 41 (2016) 20–32, https://doi.org/10.1016/j.tibs.2015.10.008.
[89] A.J.R. Heck, Native mass spectrometry: a bridge between interactomics and structural biology, Nat. Methods 5 (2008) 927–933, https://doi.org/10.1038/nmeth.1265.
[90] A.N. Holding, XL-MS: protein cross-linking coupled with mass spectrometry, Methods 89 (2015) 54–63, https://doi.org/10.1016/j.ymeth.2015.06.010.
[91] J. Rappsilber, S. Siniossoglou, E.C. Hurt, M. Mann, A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry, Anal. Chem. 72 (2000) 267–275, https://doi.org/10.1021/ac991081o.
[92] A. Maiolica, D. Cittaro, D. Borsotti, L. Sennels, C. Ciferri, C. Tarricone, A. Musacchio, J. Rappsilber, Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching, Mol. Cell. Proteom. 6 (2007) 2200–2211, https://doi.org/10.1074/mcp.M700274-MCP200.
[93] J. Rappsilber, The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes, J. Struct. Biol. 173 (2011) 530–540, https://doi.org/10.1016/j.jsb.2010.10.014.
[94] Z.A. Chen, J. Rappsilber, Protein dynamics in solution by quantitative crosslinking/mass spectrometry, Trends Biochem. Sci. 43 (2018) 908–920, https://doi.org/10.1016/j.tibs.2018.09.003.
[95] M. Schneider, A. Belsom, J. Rappsilber, Protein tertiary structure by crosslinking/mass spectrometry, Trends Bio-

chem. Sci. 43 (2018) 157–169, https://doi.org/10.1016/j.tibs.2017.12.006.
[96] Z.A. Chen, J. Rappsilber, Quantitative cross-linking/mass spectrometry to elucidate structural changes in proteins and their complexes, Nat. Protoc. 14 (2019) 171–201, https://doi.org/10.1038/s41596-018-0089-3.
[97] H. Zhang, X. Tang, G.R. Munske, N. Tolic, G.A. Anderson, J.E. Bruce, Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry, Mol. Cell. Proteom. 8 (2009) 409–420, https://doi.org/10.1074/mcp.M800232-MCP200.
[98] F. Liu, D.T.S. Rijkers, H. Post, A.J.R. Heck, Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry, Nat. Methods 12 (2015) 1179–1184, https://doi.org/10.1038/nmeth.3603.
[99] D. Fasci, H. van Ingen, R.A. Scheltema, A.J.R. Heck, Histone interaction landscapes visualized by crosslinking mass spectrometry in intact cell nuclei, Mol. Cell. Proteom. 17 (2018) 2018–2033, https://doi.org/10.1074/mcp.RA118.000924.
[100] C. Iacobucci, C. Piotrowski, R. Aebersold, B.C. Amaral, P. Andrews, K. Bernfur, C. Borchers, N.I. Brodie, J.E. Bruce, Y. Cao, S. Chaignepain, J.D. Chavez, S. Claverol, J. Cox, T. Davis, G. Degliesposti, M.-Q. Dong, N. Edinger, C. Emanuelsson, M. Gay, M. Götze, F. Gomes-Neto, F.C. Gozzo, C. Gutierrez, C. Haupt, A.J.R. Heck, F. Herzog, L. Huang, M.R. Hoopmann, N. Kalisman, O. Klykov, Z. Kukačka, F. Liu, M.J. MacCoss, K. Mechtler, R. Mesika, R.L. Moritz, N. Nagaraj, V. Nesati, A.G.C. Neves-Ferreira, R. Ninnis, P. Novák, F.J. O'Reilly, M. Pelzing, E. Petrotchenko, L. Piersimoni, M. Plasencia, T. Pukala, K.D. Rand, J. Rappsilber, D. Reichmann, C. Sailer, C.P. Sarnowski, R.A. Scheltema, C. Schmidt, D.C. Schriemer, Y. Shi, J.M. Skehel, M. Slavin, F. Sobott, V. Solis-Mezarino, H. Stephanowitz, F. Stengel, C.E. Stieger, E. Trabjerg, M. Trnka, M. Vilaseca, R. Viner, Y. Xiang, S. Yilmaz, A. Zelter, D. Ziemianowicz, A. Leitner, A. Sinz, First community-wide, comparative cross-linking mass spectrometry study, Anal. Chem. 91 (2019) 6953–6961, https://doi.org/10.1021/acs.analchem.9b00658.
[101] F. Liu, P. Lössl, R. Scheltema, R. Viner, A.J.R. Heck, Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification, Nat. Commun. 8 (2017) 15473, https://doi.org/10.1038/ncomms15473.
[102] O. Klykov, B. Steigenberger, S. Pektas, D. Fasci, A.J.R. Heck, R.A. Scheltema, Efficient and robust proteome-wide approaches for cross-linking mass spectrometry, Nat. Protoc. 13 (2018) 2964–2990, https://doi.org/10.1038/s41596-018-0074-x.
[103] S.C. de Graaf, O. Klykov, H. van den Toorn, R.A. Scheltema, I.D. Cross-, Analysis and visualization of complex XL–MS-driven protein interaction networks, J. Proteome Res. 18 (2019) 642–651, https://doi.org/10.1021/acs.jproteome.8b00725.
[104] G.C.P. van Zundert, A.M.J.J. Bonvin, DisVis: quantifying and visualizing accessible interaction space of distance-restrained biomolecular complexes: fig. 1, Bioinformatics 31 (2015) 3222–3224, https://doi.org/10.1093/bioinformatics/btv333.
[105] G.C.P. van Zundert, M. Trellet, J. Schaarschmidt, Z. Kurkcuoglu, M. David, M. Verlato, A. Rosato, A.M.J.J. Bonvin, The DisVis and PowerFit web servers: explorative and integrative modeling of biomolecular complexes, J. Mol. Biol. 429 (2017) 399–407, https://doi.org/10.1016/j.jmb.2016.11.032.
[106] L.E. Kay, NMR studies of protein structure and dynamics, J. Magn. Reson. 173 (2005) 193–207, https://doi.org/10.1016/j.jmr.2004.11.021.
[107] A. Cavalli, X. Salvatella, C.M. Dobson, M. Vendruscolo, Protein structure determination from NMR chemical shifts, Proc. Natl. Acad. Sci. 104 (2007) 9615–9620, https://doi.org/10.1073/pnas.0610313104.
[108] D. Fushman, D. Cowburn, Model-independent analysis of 15 N chemical shift anisotropy from NMR relaxation data. Ubiquitin as a test example, J. Am. Chem. Soc. 120 (1998) 7109–7110, https://doi.org/10.1021/ja980565j.
[109] D. Fushman, D. Cowburn, The effect of noncollinearity of 15N-1H dipolar and 15N CSA tensors and rotational anisotropy on 15N relaxation, CSA/dipolar cross correlation, and TROSY, J. Biomol. NMR 13 (1999) 139–147, https://doi.org/10.1023/A:1008349331773.
[110] A.M. Petros, L. Mueller, K.D. Kopple, NMR identification of protein surfaces using paramagnetic probes, Biochemistry 29 (1990) 10041–10048, https://doi.org/10.1021/bi00495a005.
[111] E. Matei, A.M. Gronenborn, 19 F paramagnetic relaxation enhancement: a valuable tool for distance measurements in proteins, Angew. Chem. Int. Ed. 55 (2016) 150–154, https://doi.org/10.1002/anie.201508464.
[112] C. Öster, S. Kosol, C. Hartlmüller, J.M. Lamley, D. Iuga, A. Oss, M.-L. Org, K. Vanatalu, A. Samoson, T. Madl, J.R. Lewandowski, Characterization of protein–protein interfaces in large complexes by solid-state NMR solvent paramagnetic relaxation enhancements, J. Am. Chem. Soc. 139 (2017) 12165–12174, https://doi.org/10.1021/jacs.7b03875.
[113] K. Kato, T. Yamaguchi, Paramagnetic NMR probes for characterization of the dynamic conformations and interactions of oligosaccharides, Glycoconj. J. 32 (2015) 505–513, https://doi.org/10.1007/s10719-015-9599-1.
[114] V. Venditti, N.L. Fawzi, Probing the atomic structure of transient protein contacts by paramagnetic relaxation enhancement solution NMR, Methods Mol. Biol. (2018) 243–255, https://doi.org/10.1007/978-1-4939-7386-6_12.
[115] R.G. Griffin, Dipolar recoupling in MAS spectra of biological solids, Nat. Struct. Biol. 5 (1998) 508–512, https://doi.org/10.1038/749.
[116] A. Lange, S. Becker, K. Seidel, K. Giller, O. Pongs, M. Baldus, A concept for rapid protein-structure determination by solid-state NMR spectroscopy, Angew. Chem. Int. Ed. 44 (2005) 2089–2092, https://doi.org/10.1002/anie.200462516.
[117] A. Samoson, E. Lippmaa, A. Pines, High resolution solid-state N.M.R. Mol. Phys. 65 (1988) 1013–1018, https://doi.org/10.1080/00268978800101571.
[118] R. Kaptein, G. Wagner, Integrative methods in structural biology, J. Biomol. NMR 73 (2019) 261–263, https://doi.org/10.1007/s10858-019-00267-z.
[119] K.M. Visscher, J. Medeiros-Silva, D. Mance, J.P.G.L.M. Rodrigues, M. Daniëls, A.M.J.J. Bonvin, M. Baldus, M. Weingarth, Supramolecular organization and functional implications of K + channel clusters in membranes, Angew. Chem. Int. Ed. 56 (2017) 13222–13227, https://doi.org/10.1002/anie.201705723.
[120] T.A. Cross, S.J. Opella, Solid-state NMR structural studies of peptides and proteins in membranes, Curr. Opin. Struct.

Biol. 4 (1994) 574–581, https://doi.org/10.1016/S0959-440X(94)90220-8.
[121] A. McDermott, Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR, Annu. Rev. Biophys. 38 (2009) 385–403, https://doi.org/10.1146/annurev.biophys.050708.133719.
[122] D.I. Freedberg, P. Selenko, Live cell NMR, Annu. Rev. Biophys. 43 (2014) 171–192, https://doi.org/10.1146/annurev-biophys-051013-023136.
[123] P. Selenko, G. Wagner, Looking into live cells with in-cell NMR spectroscopy, J. Struct. Biol. 158 (2007) 244–253, https://doi.org/10.1016/j.jsb.2007.04.001.
[124] F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H.M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, Structural disorder of monomeric α-synuclein persists in mammalian cells, Nature 530 (2016) 45–50, https://doi.org/10.1038/nature16531.
[125] S. Narasimhan, S. Scherpe, A. Lucini Paioni, J. van der Zwan, G.E. Folkers, H. Ovaa, M. Baldus, DNP-supported solid-state NMR spectroscopy of proteins inside mammalian cells, Angew. Chem. Int. Ed. 58 (2019) 12969–12973, https://doi.org/10.1002/anie.201903246.
[126] W. Kuhlbrandt, The resolution revolution, Science 343 (2014) 1443–1444, https://doi.org/10.1126/science.1251652.
[127] G. McMullan, A.R. Faruqi, R. Henderson, Direct electron detectors, in: Resolut. Revolut. Recent, Adv. CryoEM, 2016, pp. 1–17, https://doi.org/10.1016/bs.mie.2016.05.056.
[128] D.N. Kim, K.Y. Sanbonmatsu, Tools for the cryo-EM gold rush: going from the cryo-EM map to the atomistic model, Biosci. Rep. 37 (2017), BSR20170072, https://doi.org/10.1042/BSR20170072.
[129] C.L. Lawson, M.L. Baker, C. Best, C. Bi, M. Dougherty, P. Feng, G. van Ginkel, B. Devkota, I. Lagerstedt, S.J. Ludtke, R.H. Newman, T.J. Oldfield, I. Rees, G. Sahni, R. Sala, S. Velankar, J. Warren, J.D. Westbrook, K. Henrick, G.J. Kleywegt, H.M. Berman, W. Chiu, EMDataBank.org: unified data resource for CryoEM, Nucleic Acids Res. 39 (2011) D456–D464, https://doi.org/10.1093/nar/gkq880.
[130] Y. Cheng, N. Grigorieff, P.A. Penczek, T. Walz, A primer to single-particle cryo-electron microscopy, Cell 161 (2015) 438–449, https://doi.org/10.1016/j.cell.2015.03.050.
[131] Y. Cheng, Single-particle cryo-EM at crystallographic resolution, Cell 161 (2015) 450–457, https://doi.org/10.1016/j.cell.2015.03.049.
[132] M. Orzechowski, F. Tama, Flexible fitting of high-resolution X-ray structures into cryoelectron microscopy maps using biased molecular dynamics simulations, Biophys. J. 95 (2008) 5692–5705, https://doi.org/10.1529/biophysj.108.139451.
[133] L.G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics, Structure 16 (2008) 673–683, https://doi.org/10.1016/j.str.2008.03.005.
[134] E. Villa, K. Lasker, Finding the right fit: chiseling structures out of cryo-electron microscopy maps, Curr. Opin. Struct. Biol. 25 (2014) 118–125, https://doi.org/10.1016/j.sbi.2014.04.001.
[135] A. Brown, F. Long, R.A. Nicholls, J. Toots, P. Emsley, G. Murshudov, Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions, Acta Crystallogr. Sect. D Biol. Crystallogr. 71 (2015) 136–153, https://doi.org/10.1107/S1399004714021683.
[136] G.C.P. van Zundert, A.M.J.J. Bonvin, Fast and sensitive rigid-body fitting into cryo-EM density maps with PowerFit, AIMS Biophys. 2 (2015) 73–87, https://doi.org/10.3934/biophy.2015.2.73.
[137] G.C.P. van Zundert, A.S.J. Melquiond, A.M.J.J. Bonvin, Integrative modeling of biomolecular complexes: HADDOCKing with cryo-electron microscopy data, Structure 23 (2015) 949–960, https://doi.org/10.1016/j.str.2015.03.014.
[138] G.C.P. van Zundert, A.M.J.J. Bonvin, Defining the limits and reliability of rigid-body fitting in cryo-EM maps using multi-scale image pyramids, J. Struct. Biol. 195 (2016) 252–258, https://doi.org/10.1016/j.jsb.2016.06.011.
[139] R.Y.R. Wang, Y. Song, B.A. Barad, Y. Cheng, J.S. Fraser, F. DiMaio, Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta, Elife 5 (2016), https://doi.org/10.7554/eLife.17219.
[140] X. Fan, J. Wang, X. Zhang, Z. Yang, J.-C. Zhang, L. Zhao, H.-L. Peng, J. Lei, H.-W. Wang, Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution, Nat. Commun. 10 (2019) 2386, https://doi.org/10.1038/s41467-019-10368-w.
[141] Y. Liu, D.T. Huynh, T.O. Yeates, A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold, Nat. Commun. 10 (2019) 1864, https://doi.org/10.1038/s41467-019-09836-0.
[142] M.A. Herzik, M. Wu, G.C. Lander, High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM, Nat. Commun. 10 (2019) 1032, https://doi.org/10.1038/s41467-019-08991-8.
[143] G. Cardone, J.B. Heymann, A.C. Steven, One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions, J. Struct. Biol. 184 (2013) 226–236, https://doi.org/10.1016/j.jsb.2013.08.002.
[144] A. Kucukelbir, F.J. Sigworth, H.D. Tagare, Quantifying the local resolution of cryo-EM density maps, Nat. Methods 11 (2014) 63–65, https://doi.org/10.1038/nmeth.2727.
[145] M. Bonomi, M. Vendruscolo, Determination of protein structural ensembles using cryo-electron microscopy, Curr. Opin. Struct. Biol. 56 (2019) 37–45, https://doi.org/10.1016/j.sbi.2018.10.006.
[146] A. Moscovich, A. Halevi, J. Andén, A. Singer, Cryo-EM Reconstruction of Continuous Heterogeneity by Laplacian Spectral Volumes, 2019. http://arxiv.org/abs/1907.01898. (Accessed 23 July 2019).
[147] M. Bonomi, R. Pellarin, M. Vendruscolo, Simultaneous determination of protein structure and dynamics using cryo-electron microscopy, Biophys. J. 114 (2018) 1604–1613, https://doi.org/10.1016/j.bpj.2018.02.028.
[148] J. Frank, Single-particle imaging of macromolecules by cryo-electron microscopy, Annu. Rev. Biophys. Biomol. Struct. 31 (2002) 303–319, https://doi.org/10.1146/annurev.biophys.31.082901.134202.
[149] M. Kaplan, A. Cukkemane, G.C.P. van Zundert, S. Narasimhan, M. Daniëls, D. Mance, G. Waksman, A.M.J.J. Bonvin, R. Fronzes, G.E. Folkers, M. Baldus, Probing a cell-embedded megadalton protein complex by DNP-supported solid-state NMR, Nat. Methods 12 (2015) 649–652, https://doi.org/10.1038/nmeth.3406.
[150] B. Bardiaux, F. Cordier, S. Brier, A. López-Castilla, N. Izadi-Pruneyre, M. Nilges, Dynamics of a type 2 secretion system pseudopilus unraveled by complementary ap-

proaches, J. Biomol. NMR 73 (2019) 293–303, https://doi.org/10.1007/s10858-019-00246-4.
[151] H.D.T. Mertens, D.I. Svergun, Structural characterization of proteins and complexes using small-angle X-ray solution scattering, J. Struct. Biol. 172 (2010) 128–141, https://doi.org/10.1016/j.jsb.2010.06.012.
[152] C.A. Brosey, J.A. Tainer, Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology, Curr. Opin. Struct. Biol. 58 (2019) 197–213, https://doi.org/10.1016/j.sbi.2019.04.004.
[153] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, Q. Rev. Biophys. 40 (2007) 191–285, https://doi.org/10.1017/S0033583507004635.
[154] D. Schneidman-Duhovny, S. Kim, A. Sali, Integrative structural modeling with small angle X-ray scattering profiles, BMC Struct. Biol. 12 (2012) 17, https://doi.org/10.1186/1472-6807-12-17.
[155] D. Schneidman-Duhovny, M. Hammel, Modeling structure and dynamics of protein complexes with SAXS profiles, Methods Mol. Biol. (2018) 449–473, https://doi.org/10.1007/978-1-4939-7759-8_29.
[156] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Rep. Prog. Phys. 66 (2003) 1735–1782, https://doi.org/10.1088/0034-4885/66/10/R05.
[157] M.V. Petoukhov, D.I. Svergun, Global rigid body modeling of macromolecular complexes against small-angle scattering data, Biophys. J. 89 (2005) 1237–1250, https://doi.org/10.1529/biophysj.105.064154.
[158] G. Tria, H.D.T. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, IUCrJ 2 (2015) 207–217, https://doi.org/10.1107/S205225251500202X.
[159] D. Franke, C.M. Jeffries, D.I. Svergun, Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra, Nat. Methods 12 (2015) 419–422, https://doi.org/10.1038/nmeth.3358.
[160] E. Karaca, A.M.J.J. Bonvin, On the usefulness of ion-mobility mass spectrometry and SAXS data in scoring docking decoys, Acta Crystallogr. Sect. D Biol. Crystallogr. 69 (2013) 683–694, https://doi.org/10.1107/S0907444913007063.
[161] C. Pons, M. D'Abramo, D.I. Svergun, M. Orozco, P. Bernadó, J. Fernández-Recio, Structural characterization of protein–protein complexes by integrating computational docking with small-angle scattering data, J. Mol. Biol. 403 (2010) 217–230, https://doi.org/10.1016/j.jmb.2010.08.029.
[162] B. Jiménez-García, C. Pons, D.I. Svergun, P. Bernadó, J. Fernández-Recio, pyDockSAXS: protein–protein complex structure by SAXS and computational docking, Nucleic Acids Res. 43 (2015) W356–W361, https://doi.org/10.1093/nar/gkv368.
[163] E. Karaca, J.P.G.L.M. Rodrigues, A. Graziadei, A.M.J.J. Bonvin, T. Carlomagno, M3: an integrative framework for structure determination of molecular machines, Nat. Methods 14 (2017) 897–902, https://doi.org/10.1038/nmeth.4392.
[164] G.L. Hura, A.L. Menon, M. Hammel, R.P. Rambo, F.L. Poole II, S.E. Tsutakawa, F.E. Jenney Jr., S. Classen, K.A. Frankel, R.C. Hopkins, S. Yang, J.W. Scott, B.D. Dillard, M.W.W. Adams, J.A. Tainer, Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), Nat. Methods 6 (2009) 606–612, https://doi.org/10.1038/nmeth.1353.
[165] M. Dimura, T.O. Peulen, C.A. Hanke, A. Prakash, H. Gohlke, C.A. Seidel, Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems, Curr. Opin. Struct. Biol. 40 (2016) 163–185, https://doi.org/10.1016/j.sbi.2016.11.012.
[166] M. Bonomi, R. Pellarin, S.J. Kim, D. Russel, B.A. Sundin, M. Riffle, D. Jaschob, R. Ramsden, T.N. Davis, E.G.D. Muller, A. Sali, Determining protein complex structures based on a bayesian model of in vivo förster resonance energy transfer (FRET) data, Mol. Cell. Proteom. 13 (2014) 2812–2823, https://doi.org/10.1074/mcp.M114.040824.
[167] S. Kilic, S. Felekyan, O. Doroshenko, I. Boichenko, M. Dimura, H. Vardanyan, L.C. Bryan, G. Arya, C.A.M. Seidel, B. Fierz, Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1α, Nat. Commun. 9 (2018) 235, https://doi.org/10.1038/s41467-017-02619-5.
[168] K. Lehmann, S. Felekyan, R. Kühnemuth, M. Dimura, K. Tóth, C.A.M. Seidel, Multiple interaction modes of the nucleosomal histone H3 N-terminal tail revealed by high precision single-molecule FRET, Biophys. J. 116 (2019) 468a–469a, https://doi.org/10.1016/j.bpj.2018.11.2530.
[169] B. Hellenkamp, S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kühnemuth, S. Rezaei Adariani, B. Ambrose, M. Aznauryan, A. Barth, V. Birkedal, M.E. Bowen, H. Chen, T. Cordes, T. Eilert, C. Fijen, C. Gebhardt, M. Götz, G. Gouridis, E. Gratton, T. Ha, P. Hao, C.A. Hanke, A. Hartmann, J. Hendrix, L.L. Hildebrandt, V. Hirschfeld, J. Hohlbein, B. Hua, C.G. Hübner, E. Kallis, A.N. Kapanidis, J.-Y. Kim, G. Krainer, D.C. Lamb, N.K. Lee, E.A. Lemke, B. Levesque, M. Levitus, J.J. McCann, N. Naredi-Rainer, D. Nettels, T. Ngo, R. Qiu, N.C. Robb, C. Röcker, H. Sanabria, M. Schlierf, T. Schröder, B. Schuler, H. Seidel, L. Streit, J. Thurn, P. Tinnefeld, S. Tyagi, N. Vandenberk, A.M. Vera, K.R. Weninger, B. Wünsch, I.S. Yanez-Orozco, J. Michaelis, C.A.M. Seidel, T.D. Craggs, T. Hugel, Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study, Nat. Methods 15 (2018) 669–676, https://doi.org/10.1038/s41592-018-0085-0.
[170] C. Möckel, J. Kubiak, O. Schillinger, R. Kühnemuth, D. Della Corte, G.F. Schröder, D. Willbold, B. Strodel, C.A.M. Seidel, P. Neudecker, Integrated NMR, fluorescence, and molecular dynamics benchmark study of protein mechanics and hydrodynamics, J. Phys. Chem. B 123 (2019) 1453–1480, https://doi.org/10.1021/acs.jpcb.8b08903.
[171] M. Pannier, S. Veit, A. Godt, G. Jeschke, H. Spiess, Dead-time free measurement of dipole–dipole interactions between electron spins, J. Magn. Reson. 142 (2000) 331–340, https://doi.org/10.1006/jmre.1999.1944.
[172] G. Jeschke, DEER distance measurements on proteins, Annu. Rev. Phys. Chem. 63 (2012) 419–446, https://doi.org/10.1146/annurev-physchem-032511-143716.
[173] N. Fehr, C. Dietz, Y. Polyhach, T. von Hagens, G. Jeschke, H. Paulsen, Modeling of the N-terminal section and the lumenal loop of trimeric light harvesting complex II (LHCII)

by using EPR, J. Biol. Chem. 290 (2015) 26007–26020, https://doi.org/10.1074/jbc.M115.669804.
[174] B.M. Stadtmueller, M.D. Bridges, K.-M. Dam, M.T. Lerch, K.E. Huey-Tubman, W.L. Hubbell, P.J. Bjorkman, DEER spectroscopy measurements reveal multiple conformations of HIV-1 SOSIP envelopes that show similarities with envelopes on native virions, Immunity 49 (2018) 235–246, https://doi.org/10.1016/j.immuni.2018.06.017, e4.
[175] J. Glaenzer, M.F. Peter, G. Hagelueken, Studying structure and function of membrane proteins with PELDOR/DEER spectroscopy – the crystallographers' perspective, Methods 147 (2018) 163–175, https://doi.org/10.1016/j.ymeth.2018.03.002.
[176] G. Masliah, C. Maris, S.L. König, M. Yulikov, F. Aeschimann, A.L. Malinowska, J. Mabille, J. Weiler, A. Holla, J. Hunziker, N. Meisner-Kober, B. Schuler, G. Jeschke, F.H. Allain, Structural basis of si RNA recognition by TRBP double-stranded RNA binding domains, EMBO J. 37 (2018), e97089, https://doi.org/10.15252/embj.201797089.
[177] H. Szurmant, M. Weigt, Inter-residue, inter-protein and inter-family coevolution: bridging the scales, Curr. Opin. Struct. Biol. 50 (2018) 26–32, https://doi.org/10.1016/j.sbi.2017.10.014.
[178] L. Sutto, S. Marsili, A. Valencia, F.L. Gervasio, From residue coevolution to protein conformational ensembles and functional dynamics, Proc. Natl. Acad. Sci. 112 (2015) 13567–13572, https://doi.org/10.1073/pnas.1508584112.
[179] G. Uguzzoni, S. John Lovis, F. Oteri, A. Schug, H. Szurmant, M. Weigt, Large-scale identification of coevolution signals across homo-oligomeric protein interfaces by direct coupling analysis, Proc. Natl. Acad. Sci. 114 (2017) E2662–E2671, https://doi.org/10.1073/pnas.1615068114.
[180] S. Balakrishnan, H. Kamisetty, J.G. Carbonell, S.-I. Lee, C.J. Langmead, Learning generative models for protein fold families, Proteins Struct. Funct. Bioinf. 79 (2011) 1061–1078, https://doi.org/10.1002/prot.22934.
[181] H. Kamisetty, S. Ovchinnikov, D. Baker, Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era, Proc. Natl. Acad. Sci. 110 (2013) 15674–15679, https://doi.org/10.1073/pnas.1314045110.
[182] J. Yu, M. Vavrusa, J. Andreani, J. Rey, P. Tufféry, R. Guerois, InterEvDock: a docking server to predict the structure of protein–protein interactions using evolutionary information, Nucleic Acids Res. 44 (2016) W542–W549, https://doi.org/10.1093/nar/gkw340.
[183] T.A. Hopf, C.P.I. Schärfe, J.P.G.L.M. Rodrigues, A.G. Green, O. Kohlbacher, C. Sander, A.M.J.J. Bonvin, D.S. Marks, Sequence co-evolution gives 3D contacts and structures of protein complexes, Elife 3 (2014), https://doi.org/10.7554/eLife.03430.
[184] J. Yu, J. Andreani, F. Ochsenbein, R. Guerois, Lessons from (co-)evolution in the docking of proteins and peptides for CAPRI Rounds 28-35, Proteins Struct. Funct. Bioinf. 85 (2017) 378–390, https://doi.org/10.1002/prot.25180.
[185] S. Ovchinnikov, H. Kamisetty, D. Baker, Robust and accurate prediction of residue–residue interactions across protein interfaces using evolutionary information, Elife 3 (2014), https://doi.org/10.7554/eLife.02030.
[186] C. Quignot, J. Rey, J. Yu, P. Tufféry, R. Guerois, J. Andreani, InterEvDock2: an expanded server for protein docking using evolutionary and biological information from homology models and multimeric inputs, Nucleic Acids Res. 46 (2018) W408–W416, https://doi.org/10.1093/nar/gky377.
[187] B. Zeng, P. Hönigschmid, D. Frishman, Residue coevolution helps predict interaction sites in α-helical membrane proteins, J. Struct. Biol. 206 (2019) 156–169, https://doi.org/10.1016/j.jsb.2019.02.009.
[188] P.J. Stansfeld, Computational studies of membrane proteins: from sequence to structure to simulation, Curr. Opin. Struct. Biol. 45 (2017) 133–141, https://doi.org/10.1016/j.sbi.2017.04.004.
[189] Q. Cong, I. Anishchenko, S. Ovchinnikov, D. Baker, Protein interaction networks revealed by proteome coevolution, Science 365 (2019) 185–189, https://doi.org/10.1126/science.aaw6718.
[190] Q. Wu, Z. Peng, I. Anishchenko, Q. Cong, D. Baker, J. Yang, Protein contact prediction using metagenome sequence data and residual neural networks, Bioinformatics (2019), https://doi.org/10.1093/bioinformatics/btz477.
[191] R. Shrestha, E. Fajardo, N. Gil, K. Fidelis, A. Kryshtafovych, B. Monastyrskyy, A. Fiser, Assessing the accuracy of contact predictions in CASP13, Proteins Struct. Funct. Bioinf. 0 (2019), https://doi.org/10.1002/prot.25819.
[192] J.M. Dörr, S. Scheidelaar, M.C. Koorengevel, J.J. Dominguez, M. Schäfer, C.A. van Walree, J.A. Killian, The styrene–maleic acid copolymer: a versatile tool in membrane research, Eur. Biophys. J. 45 (2016) 3–21, https://doi.org/10.1007/s00249-015-1093-y.
[193] N. Hellwig, O. Peetz, Z. Ahdash, I. Tascón, P.J. Booth, V. Mikusevic, M. Diskowski, A. Politis, Y. Hellmich, I. Hänelt, E. Reading, N. Morgner, Native mass spectrometry goes more native: investigation of membrane protein complexes directly from SMALPs, Chem. Commun. 54 (2018) 13702–13705, https://doi.org/10.1039/C8CC06284F.
[194] K.S. Simon, N.L. Pollock, S.C. Lee, Membrane protein nanoparticles: the shape of things to come, Biochem. Soc. Trans. 46 (2018) 1495–1504, https://doi.org/10.1042/BST20180139.
[195] M. Parmar, S. Rawson, C.A. Scarff, A. Goldman, T.R. Dafforn, S.P. Muench, V.L.G. Postis, Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure, Biochim. Biophys. Acta Biomembr. 1860 (2018) 378–383, https://doi.org/10.1016/j.bbamem.2017.10.005.
[196] J.M. Dörr, M.C. Koorengevel, M. Schäfer, A.V. Prokofyev, S. Scheidelaar, E.A.W. van der Cruijsen, T.R. Dafforn, M. Baldus, J.A. Killian, Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K + channel: the power of native nanodiscs, Proc. Natl. Acad. Sci. 111 (2014) 18607–18612, https://doi.org/10.1073/pnas.1416205112.
[197] J. Radoicic, S.H. Park, S.J. Opella, Macrodiscs comprising SMALPs for oriented sample solid-state NMR spectroscopy of membrane proteins, Biophys. J. 115 (2018) 22–25, https://doi.org/10.1016/j.bpj.2018.05.024.
[198] J.G. Almeida, A.J. Preto, P.I. Koukos, A.M.J.J. Bonvin, I.S. Moreira, Membrane proteins structures: a review on computational modeling tools, Biochim. Biophys. Acta Biomembr. 1859 (2017) 2021–2039, https://doi.org/10.1016/j.bbamem.2017.07.008.

[199] K. Shimizu, W. Cao, G. Saad, M. Shoji, T. Terada, Comparative analysis of membrane protein structure databases, Biochim. Biophys. Acta Biomembr. 1860 (2018) 1077–1091, https://doi.org/10.1016/j.bbamem.2018.01.005.
[200] S. Kelm, J. Shi, C.M. Deane, MEDELLER: homology-based coordinate generation for membrane proteins, Bioinformatics 26 (2010) 2833–2840, https://doi.org/10.1093/bioinformatics/btq554.
[201] J.-P. Ebejer, J.R. Hill, S. Kelm, J. Shi, C.M. Deane, Memoir: template-based structure prediction for membrane proteins, Nucleic Acids Res. 41 (2013) W379–W383, https://doi.org/10.1093/nar/gkt331.
[202] A. Šali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815, https://doi.org/10.1006/jmbi.1993.1626.
[203] D.M. Nikolaev, A.A. Shtyrov, M.S. Panov, A. Jamal, O.B. Chakchir, V.A. Kochemirovsky, M. Olivucci, M.N. Ryazantsev, A comparative study of modern homology modeling algorithms for rhodopsin structure prediction, ACS Omega 3 (2018) 7555–7566, https://doi.org/10.1021/acsomega.8b00721.
[204] M.A. Lomize, A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, OPM: orientations of proteins in membranes database, Bioinformatics 22 (2006) 623–625, https://doi.org/10.1093/bioinformatics/btk023.
[205] P.J. Stansfeld, J.E. Goose, M. Caffrey, E.P. Carpenter, J.L. Parker, S. Newstead, M.S.P. Sansom, MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes, Structure 23 (2015) 1350–1361, https://doi.org/10.1016/j.str.2015.05.006.
[206] T.D. Newport, M.S.P. Sansom, P.J. Stansfeld, The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions, Nucleic Acids Res. 47 (2019) D390–D397, https://doi.org/10.1093/nar/gky1047.
[207] A.S. Hauser, M.M. Attwood, M. Rask-Andersen, H.B. Schiöth, D.E. Gloriam, Trends in GPCR drug discovery: new agents, targets and indications, Nat. Rev. Drug Discov. 16 (2017) 829–842, https://doi.org/10.1038/nrd.2017.178.
[208] C. Munk, E. Mutt, V. Isberg, L.F. Nikolajsen, J.M. Bibbe, T. Flock, M.A. Hanson, R.C. Stevens, X. Deupi, D.E. Gloriam, An online resource for GPCR structure determination and analysis, Nat. Methods 16 (2019) 151–162, https://doi.org/10.1038/s41592-018-0302-x.
[209] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI force field: coarse grained model for biomolecular simulations, J. Phys. Chem. B 111 (2007) 7812–7824, https://doi.org/10.1021/jp071097f.
[210] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, S.-J. Marrink, The MARTINI coarse-grained force field: extension to proteins, J. Chem. Theory Comput. 4 (2008) 819–834, https://doi.org/10.1021/ct700324x.
[211] D.H. de Jong, G. Singh, W.F.D. Bennett, C. Arnarez, T.A. Wassenaar, L.V. Schäfer, X. Periole, D.P. Tieleman, S.J. Marrink, Improved parameters for the martini coarse-grained protein force field, J. Chem. Theory Comput. 9 (2013) 687–697, https://doi.org/10.1021/ct300646g.
[212] C. Arnarez, J.J. Uusitalo, M.F. Masman, H.I. Ingólfsson, D.H. de Jong, M.N. Melo, X. Periole, A.H. de Vries, S.J. Marrink, Dry martini, a coarse-grained force field for lipid membrane simulations with implicit solvent, J. Chem. Theory Comput. 11 (2015) 260–275, https://doi.org/10.1021/ct500477k.
[213] J. Roel-Touris, C.G. Don, R.V. Honorato, J.P.G.L.M. Rodrigues, A.M.J.J. Bonvin, Less is more: coarse-grained integrative modeling of large biomolecular assemblies with HADDOCK, J. Chem. Theory Comput. 15 (2019) 6358–6367, https://doi.org/10.1021/acs.jctc.9b00310.
[214] R. V Honorato, J. Roel-Touris, A.M.J.J. Bonvin, MARTINI-based protein-DNA coarse-grained HADDOCKing, Front. Mol. Biosci. 6 (2019) 102, https://doi.org/10.3389/fmolb.2019.00102.
[215] M. Kurcinski, M. Jamroz, M. Blaszczyk, A. Kolinski, S. Kmiecik, CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site, Nucleic Acids Res. 43 (2015) W419–W424, https://doi.org/10.1093/nar/gkv456.
[216] M. Kurcinski, M. Pawel Ciemny, T. Oleniecki, A. Kuriata, A.E. Badaczewska-Dawid, A. Kolinski, S. Kmiecik, CABS-dock standalone: a toolbox for flexible protein–peptide docking, Bioinformatics 35 (2019) 4170–4172, https://doi.org/10.1093/bioinformatics/btz185.
[217] R.F. Alford, J. Koehler Leman, B.D. Weitzner, A.M. Duran, D.C. Tilley, A. Elazar, J.J. Gray, An integrated framework advancing membrane protein modeling and design, PLoS Comput. Biol. 11 (2015), e1004398, https://doi.org/10.1371/journal.pcbi.1004398.
[218] J. Koehler Leman, B.K. Mueller, J.J. Gray, Expanding the toolkit for membrane protein modeling in Rosetta, Bioinformatics 33 (2016) btw716, https://doi.org/10.1093/bioinformatics/btw716.
[219] S. Viswanath, D.V.S. Ravikant, R. Elber, DOCK/PIERR: web server for structure prediction of protein–protein complexes, Methods Mol. Biol. (2014) 199–207, https://doi.org/10.1007/978-1-4939-0366-5_14.
[220] S. Viswanath, L. Dominguez, L.S. Foster, J.E. Straub, R. Elber, Extension of a protein docking algorithm to membranes and applications to amyloid precursor protein dimerization, Proteins Struct. Funct. Bioinf. 83 (2015) 2170–2185, https://doi.org/10.1002/prot.24934.
[221] N. Hurwitz, Di Schneidman-Duhovny, H.J. Wolfson, Memdock: an α-helical membrane protein docking algorithm, Bioinformatics 32 (2016) 2444–2450, https://doi.org/10.1093/bioinformatics/btw184.
[222] P.I. Koukos, I. Faro, C.W. van Noort, A.M.J.J. Bonvin, A membrane protein complex docking benchmark, J. Mol. Biol. 430 (2018) 5246–5256, https://doi.org/10.1016/j.jmb.2018.11.005.
[223] P. Koukos, A. Bonvin, HADDOCK Membrane Protein-Protein Complex Models, 2018, https://doi.org/10.15785/SBGRID/618.
[224] Z. Fang, C.B. Marshall, T. Nishikawa, A.D. Gossert, J.M. Jansen, W. Jahnke, M. Ikura, Inhibition of K-RAS4B by a unique mechanism of action: stabilizing membrane-dependent occlusion of the effector-binding site, Cell Chem. Biol. 25 (2018) 1327–1336, https://doi.org/10.1016/j.chembiol.2018.07.009, e4.
[225] N.S. Pagadala, K. Syed, J. Tuszynski, Software for molecular docking: a review, Biophys. Rev. 9 (2017) 91–102, https://doi.org/10.1007/s12551-016-0247-1.
[226] M. Bonomi, S. Hanot, C.H. Greenberg, A. Sali, M. Nilges, M. Vendruscolo, R. Pellarin, Bayesian weighing of electron cryo-microscopy data for integrative structural modeling,

Structure 27 (2019) 175–188, https://doi.org/10.1016/j.str.2018.09.011, e6.
[227] W. Rieping, Inferential structure determination, Science 309 (2005) 303–306, https://doi.org/10.1126/science.1110428.
[228] K. Tamura, I. Hara-Nishimura, The molecular architecture of the plant nuclear pore complex, J. Exp. Bot. 64 (2013) 823–832, https://doi.org/10.1093/jxb/ers258.
[229] S.J. Kim, J. Fernandez-Martinez, I. Nudelman, Y. Shi, W. Zhang, B. Raveh, T. Herricks, B.D. Slaughter, J.A. Hogan, P. Upla, I.E. Chemmama, R. Pellarin, I. Echeverria, M. Shivaraju, A.S. Chaudhury, J. Wang, R. Williams, J.R. Unruh, C.H. Greenberg, E.Y. Jacobs, Z. Yu, M.J. de la Cruz, R. Mironska, D.L. Stokes, J.D. Aitchison, M.F. Jarrold, J.L. Gerton, S.J. Ludtke, C.W. Akey, B.T. Chait, A. Sali, M.P. Rout, Integrative structure and functional anatomy of a nuclear pore complex, Nature 555 (2018) 475–482, https://doi.org/10.1038/nature26003.
[230] W. Rieping, M. Nilges, M. Habeck, ISD: a software package for Bayesian NMR structure calculation, Bioinformatics 24 (2008) 1104–1105, https://doi.org/10.1093/bioinformatics/btn062.
[231] S.A. Shahid, B. Bardiaux, W.T. Franks, L. Krabben, M. Habeck, B.-J. van Rossum, D. Linke, Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals, Nat. Methods 9 (2012) 1212–1217, https://doi.org/10.1038/nmeth.2248.
[232] B. Habenstein, A. Loquet, S. Hwang, K. Giller, S.K. Vasa, S. Becker, M. Habeck, A. Lange, Hybrid structure of the type 1 pilus of uropathogenic Escherichia coli, Angew. Chem. Int. Ed. 54 (2015) 11691–11695, https://doi.org/10.1002/anie.201505065.
[233] S. Carstens, M. Nilges, M. Habeck, Inferential structure determination of chromosomes from single-cell hi-C data, PLoS Comput. Biol. 12 (2016), e1005292, https://doi.org/10.1371/journal.pcbi.1005292.
[234] Y.-L. Chen, M. Habeck, Data-driven coarse graining of large biomolecular structures, PLoS One 12 (2017), e0183057, https://doi.org/10.1371/journal.pone.0183057.
[235] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718, https://doi.org/10.1002/jcc.20291.
[236] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, Software 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[237] S.K. Burley, G. Kurisu, J.L. Markley, H. Nakamura, S. Velankar, H.M. Berman, A. Sali, T. Schwede, J. Trewhella, PDB-Dev, A prototype system for depositing integrative/hybrid structural models, Structure 25 (2017) 1317–1318, https://doi.org/10.1016/j.str.2017.08.001.
[238] J. Moult, K. Fidelis, A. Kryshtafovych, T. Schwede, A. Tramontano, Critical assessment of methods of protein structure prediction (CASP)-Round XII, Proteins Struct. Funct. Bioinf. 86 (2018) 7–15, https://doi.org/10.1002/prot.25415.
[239] J. Janin, K. Henrick, J. Moult, L. Ten Eyck, M.J.E. Sternberg, S. Vajda, I. Vakser, S.J. Wodak, CAPRI: a critical assessment of PRedicted interactions, Proteins Struct. Funct. Genet. 52 (2003) 2–9, https://doi.org/10.1002/prot.10381.
[240] The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC., ((n.d.)).
[241] T.D. Goddard, C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: meeting modern challenges in visualization and analysis, Protein Sci. 27 (2018) 14–25, https://doi.org/10.1002/pro.3235.