Integrating Cryo-EM and NMR data

James A. Geraets[a], Karunakar R. Pothula[a], Gunnar F. Schröder[a,b,*]

[a] *Institute of Biological Information Processing (IBI-7: Structural Biochemistry) and JuStruct, Jülich Center for Structural Biology, Forschungszentrum Jülich, 52425 Jülich, Germany*

[b] *Physics Department, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany*

Abstract

Single-particle cryo-electron microscopy is increasingly used as a technique to determine the atomic structure of challenging biological systems. Recent advances in microscope engineering, electron detection, and image processing have allowed the structural determination of bigger and more flexible targets than possible with the complementary techniques X-ray crystallography and NMR spectroscopy. However, there exist many biological targets for which atomic resolution cannot be achieved currently with cryo-electron microscopy, making unambiguous determination of the protein structure impossible. Although determining the structure of large biological systems using solely NMR is often difficult, highly complementary experimental atomic-level data for each molecule can be derived from the spectra, and used in combination with cryo-electron microscopy data. We review here strategies with which both techniques can be synergistically combined, in order to reach detail and understanding unattainable by each technique acting alone; and the types of biological systems for which such an approach would be desirable.

Key words: Structural determination, Cryo-EM, NMR, hybrid methods, macromolecular complexes

Conflict of Interest: none.

*Corresponding author

Email addresses: gu.schroeder@fz-juelich.de (juelich,duesseldorf)

1. Introduction

NMR spectroscopy has been used for a long time as a powerful tool to determine atomic structures of proteins and other macromolecules. The analysis of NMR spectra provides information on the local chemical environment of atoms from which local restraints can be derived that can be used to determine the three-dimensional atomic structure of a protein. The size limitation for routine structure determination is typically $\sim$ 30 kDa in liquid-state NMR (lsNMR), which relies on the fast tumbling of molecules, although in some cases much larger proteins have been studied by NMR [1].

For magic-angle spinning NMR (referred to here as solid-state NMR or ssNMR) this size limitation does not hold (proteins can be larger than 100 kDa), but other limitations pose challenges for the structural interpretation of ssNMR data [2]. NMR in general is extremely sensitive to even small changes in the local structure, which makes it a unique tool to study structural details such as protonation states, binding of ions, presence of different rotamers, etc., which are invisible to or at least very difficult to determine by most other experimental techniques. This is particularly helpful in drug discovery, where ligand binding modes need to be determined with very high accuracy. A particular strength of the NMR technique lies in the fact that proteins can be studied in solution at physiological temperatures and therefore (potentially functionally important) protein dynamics can be observed [3].

On the other hand, in single-particle cryo-electron microscopy (cryo-EM), individual protein molecules are embedded in a thin (20–100 nm) layer of vitrified ice. A large number of images of these single proteins, ideally in random orientations, are then reconstructed into a 3D density map which enables the building of an atomic model if the resolution of the reconstruction is high enough (better than $\sim$ 4.5 Å). For several years the resolution in cryo-EM was rather limited to just below 1 nm, which allowed docking of known structures into the EM density maps, for example to ascertain the placement of individually determined (by X-ray crystallography or NMR) subunits within a large protein

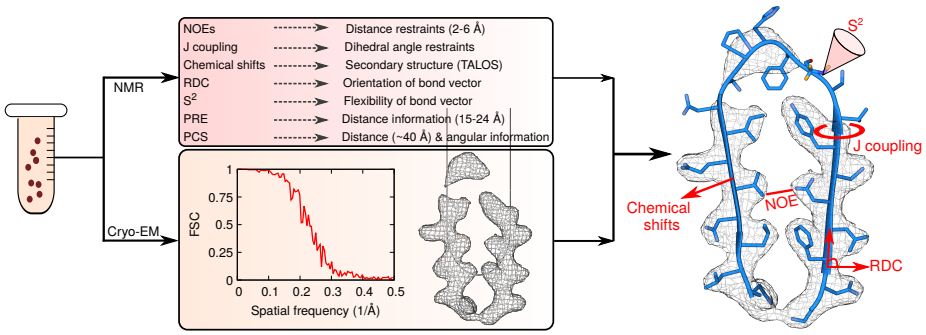

Figure 1: Schematic representation of integrating NMR with cryo-EM. Brief details of information obtained from NMR and cryo-EM are shown.

complex. However, the recent technological breakthroughs in cryo-EM now enable the determination of molecular structures quite routinely to resolutions of 3–4 Å for many protein complexes.

As NMR provides the most detailed information on local length scales and cryo-EM is most accurate at larger length scales, the techniques are complementary, and it seems promising to combine both techniques to study the structure of proteins and protein complexes.

2. Information from NMR and Cryo-EM Experiments

The most informative measures obtained from NMR experiments (see also Fig. 1) are NOE (nuclear Overhauser effect) distance restraints (up to 6 Å), dihedral angle restraints (from J-coupling), and chemical shifts which can for example be used to identify secondary structure (typically through TALOS [4]) but can also be used directly as restraints in MD simulations [5]. S^2 order parameters report on angular motional freedom of internuclear vectors. Only few measures yield information on longer length scales, in particular PCS (pseudo-contact shifts), PRE (paramagnetic relaxation enhancement) and RDC (residual dipolar coupling), which are less common and more difficult to obtain: PCS and PRE require the use of paramagnetic labels, whereas RDC can also use paramagnetic labels or can alternatively be performed in aligning media. PRE can

be used to obtain distance information in the range of 15–24 Å, PCS provides long-range ($\sim$ 40 Å) distance and angular information, and RDC yields information on the orientation of bond vectors. Restraints derived from these three complementary techniques are therefore particularly useful to identify potential differences between solution state and the structural state observed in cryo-EM or also X-ray crystallography [6].

Single-particle cryo-EM yields density maps of which the resolution is determined by the spectral signal-to-noise ratio as measured by the Fourier-shell correlation (FSC) (see Fig. 1). The information in the density map is most reliable at low spatial frequency and decreases towards higher spatial frequencies as the signal-to-noise ratio decreases. The local resolution can vary drastically within a single density map due to the flexibility of domains and due to stronger negative effects of image alignment errors at the periphery of the particles. A single density map is an ensemble average over a large number of particle images (usually 20 000–100 000). However, the single particle images could further provide additional information on conformational variability, either through classification into distinct states or by analysing variance and covariance, which is highly valuable for the determination of structural ensembles.

3. Combining NMR and EM

A widely-used joint application of cryo-EM and NMR, analogous to that between cryo-EM and X-ray crystallography, is the docking of protein domains solved by NMR into density maps solved by cryo-EM, followed by refinement to determine accurate atomic models. The use of this technique pre-dates the "resolution revolution" of cryo-EM, as it is possible for both medium- or low-resolution density maps resolved by EM. This straight-forward approach has been used for example in the case of virus particles [7], actin filaments [8], inflammasome structure [9] and a recent pilus filament [10]. A recent review on integrating cryo-EM and NMR highlighted several such applications [11]. However, to optimally make use of complementarity, the structure should ideally

be determined by using both restraints from NMR and cryo-EM at the same time.

In the case of low- or mid-resolution cryo-EM maps, hybrid methods employing NMR-derived restraints have become a robust approach for building accurate structural models. Starting with the first type-III secretion system (in *Shigella flexneri*) to the recent TET2 enzyme complex, it has been highlighted how details enabled by NMR—principally secondary structure information and long range distance restraints—have improved the atomic details of cryo-EM maps at resolution worse than 4 Å. Demers et al. showed that with 996 ssNMR distance constraints and cryo-EM density map of 7.7 Å, the structure of the type-III secretion system needle of *S. flexneri* could be determined to a precision of 0.4 Å RMSD [12••]. Recently, Gauto et al. presented a strategy to combine NMR with cryo-EM, which allowed them to solve a complex structure of the 468 kDa TET2 aminopeptidase dodecamer [13••]. The EM density map at a resolution of 4.1 Å could not be traced unambiguously and ssNMR and lsNMR experiments yielded 516 distance restraints and 544 backbone dihedral angle restraints, which alone did not suffice to solve the structure either. By first identifying α-helices in both the EM map and the assigned chemical shifts an initial model could be built, which was then refined iteratively. The hybrid cryo-EM/NMR approach resolved putatively functional loops (residues 120–138) that were unmodelable in the TET2 crystal structure (Fig. 2a–d). The positive charge and positioning of the loops in the centre of the catalytic chamber suggests that they presumably play a role in guiding peptides entering the chamber—which lead with their positively charged N-terminus [14]—away from the centre and toward the twelve negatively charged proteolytic sites within the compartment. Determination of the loop structure facilitates detailed analysis of the ensemble by MD, which is reliant on the correct positioning of the loops in the starting conformation.

There are certain classes of molecules for which a similar approach can be more routinely applied. One such instance is large RNA molecules, for which the inherent flexibility of the molecule can make crystallography or cryo-EM

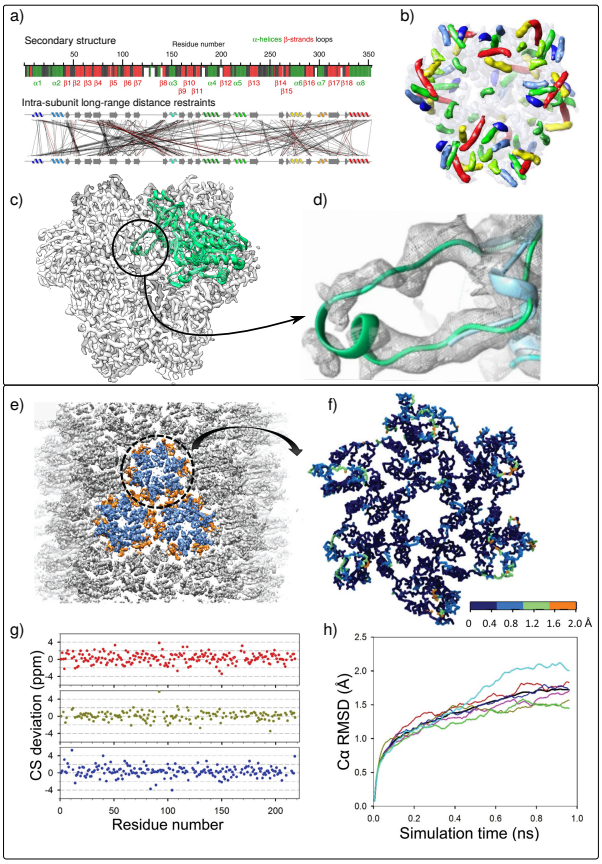

Figure 2: Panels (a–d) show the integration approach of NMR and cryo-EM for the TET2 dodecamer complex [13••]. Figures (a–d) taken with permission from [13••]. a) Experimentally detected secondary of TET2 from ssNMR resonance assignments and the TALOS-N software and long range intra-subunit interactions from lsNMR and ssNMR. b) α-helices detected in the 8 Å resolution EM map of TET2 dodecamer complex. c) Ensemble of 10 structures of one monomer overlaid in the EM map. The missing loop region, residues 120–138, in the crystal structure is highlighted with black circle. d) Zoom in of EM map around the loop missing in the crystal structure (blue) compared with one of the refined models using NMR and cryo-EM data. Panels (e–h) show the refinement of integrating NMR chemical shifts in the MD simulations [15••]. Figures (e–h) taken with permission from [15••]. e) Cryo-EM density map of the HIV-1 capsid tubular assembly at 5 Å resolution. f) Root-mean-square fluctuations (RMSFs) of the ensemble of the capsid protein hexamer obtained from integrating the chemical shift in the MD simulations. g) Differences of the chemical shifts between experimental and refined models for all the residues (Cα in red, Cβ in dark green, and C in blue). h) The Cα root-mean-square deviations (RMSD) for 6 capsid protein chains between the starting MDFF model and chemical shift biased models along the trajectory.

challenging [11, 16]. However, NMR is able to resolve atomistic local distance restraint ensembles in these situations, albeit with limited information on global tertiary structure of the RNA molecule and its flexibility, which can be provided by cryo-EM. An example of this approach is that taken by Zhang and colleagues to resolve the HIV-1 RNA duplex [16] by integrating NMR-derived distance restraints where the cryo-EM resolution was limited to 9 Å due to the internal flexibility of the duplex.

The combination of NMR and cryo-EM can be used to probe the structure and functional mechanisms of natively disordered ligands for large complexes that would otherwise be difficult to resolve by cryo-EM alone. For example, in a study of the $\sim$ 1.5 MDa human anaphase-promoting complex (APC/C), a combined EM/NMR approach was used to identify that an intrinsically-disordered inhibiting peptide, EMI1, is seen to have multiple disperse interactions with multiple binding sites, serving to dynamically modulate the function of the $\mathrm{APC/C^{CDH1}}$ receptor complex [17]. Initially, NMR studies indicated that other than a 45 aa zinc-binding region, the EMI1 inhibitory region (C-terminus, 143 aa) is substantially natively disordered. Difference mapping of EM reconstructions with complexes constructed using recombinant mutant EMI1 fragments, some with a WD40 β-propeller marker insertion (from *S. cerevisiae* Doa1), were used to identify the binding positions of parts of EMI1 in the complex. From this, a surface on the EMI1 linker region was determined to be in a structurally important position for inhibiting APC/C, and its specificity for inhibition was confirmed recombinantly. Another functionally important surface was identified by alanine-scanning mutagenesis within the stable zinc-binding region, for which an NMR structure was solved. These two surfaces, along with the essential EMI1 D-box region, in spite of individually weak interactions, synergistically inhibit $\mathrm{APC/C^{CDH1}}$ by both blocking the substrate-binding site and also mediating the ubiquitin chain elongation. The position of the inhibitor was later confirmed with a cryo-EM structure of the whole complex [18].

Similarly, Iadanza and colleagues recently presented a β_2-microglobulin (β_2m) amyloid structure [19]. When visualized by cryo-EM, the fibrils were heteroge-

neous and displayed a wide range of morphologies, whereas in contrast a single set of resonances was observed by ssNMR for residues within the core region of β_2m, indicating a conserved common subunit structure within the polymorphs (a rare instance of peak doubling was thought to correspond to local perturbations between the polymorphs). The ssNMR and cryo-EM data collectively allowed a unique structural model for the β_2m fibril to be built, unveiling the intra- and inter-subunit stabilizing interactions, namely the canonical amyloid cross-β-structure down the fibrillar axis supported by β-stacking, subunit stabilization by hydrophobic packing, and an intramolecular steric zipper and disulfide bond. Therefore even for the polymorphs unable to be reconstructed by cryo-EM, detailed structural information was determined for the common subunit, itself able to realize several polymorph geometries.

Perilla et al. [15••] proposed a new approach to incorporate the NMR chemical shifts as a linear potential in MD force fields for the cryo-EM structure refinement of a HIV-1 capsid tubular assembly with a 5 Å resolution map (Fig. 2e). This approach showed significant improvement in the refinement of flexible loops in comparison to MD flexible fitting (MDFF), especially in the CypA-binding loop and the loop connecting helix 8 and 9 (Fig. 2f). Figure 2g shows the validation of this approach as the differences of the chemical shifts between experimental and the chemical-shift-biased model are small, and the sampling reaches a steady phase with root-mean-square deviations of 1 Å after 0.2 ns simulation (Fig. 2h).

In all these examples, conditions for sample preparations of cryo-EM and NMR were different, which needs to considered when analyzing the data. Since lsNMR has an upper size limit that is lower than the typical lower size limit for cryo-EM, lsNMR data would have to be collected for smaller subunits of a larger complex studied by cryo-EM, which makes combination of data more difficult since the two experiments do not observe the protein(s) in exactly the same state. However, there are cases where the same sample has been studied by ssNMR and cryo-EM, in particular helical assemblies such as amyloid fibrils [20].

One such example of same specimen is the ssNMR/cryo-EM structure of

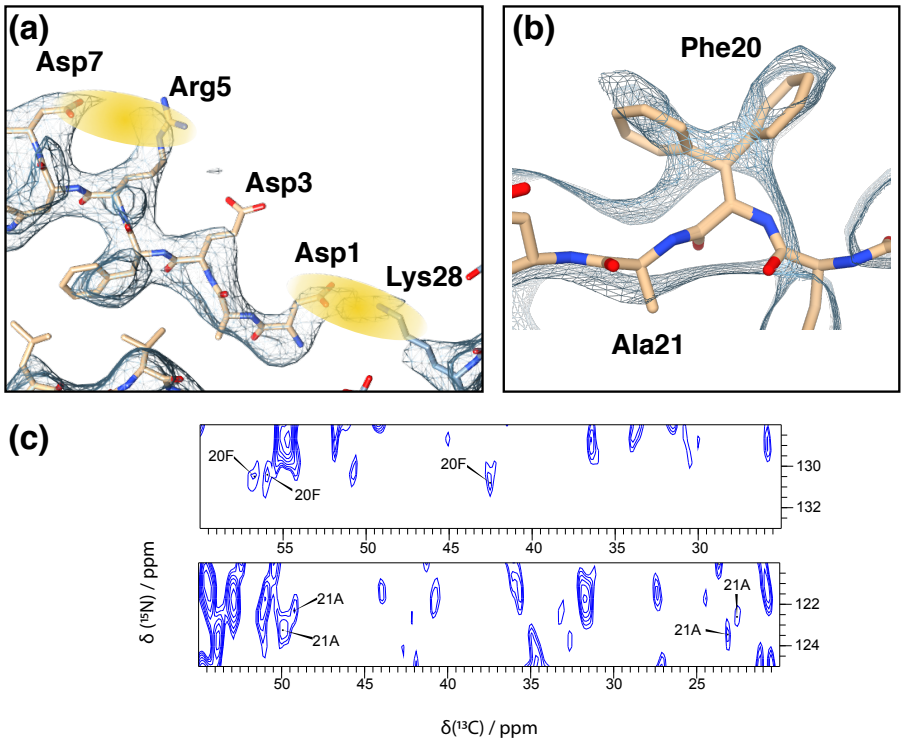

Figure 3: Showing structural details of an Aβ-fibril obtained by ssNMR not detectable from the cryo-EM data. a) Protonation states are determined by NMR which identify salt-bridges between Asp7 and Arg5 as well as between Asp1 and Lys28. b) An alternative rotamer is suggested by the density map and could be clearly confirmed by a doubling of the corresponding NMR resonance peaks of Phe20 and the neighboring Ala21 shown in c). Figures (b) and (c) taken with permission from [20].

the Aβ(1–42) amyloid fibrils [20], where an atomic model was determined in a map of 4.0 Å resolution, with long-range non-sequential contacts identified by ssNMR supporting the model. Protonation states (invisible to cryo-EM at the resolution of 4 Å) and therefore also salt bridges could be determined by ssNMR (see Fig. 3a). Full site-specific NMR resonance assignments could be obtained for all 42 residues and most residues only exhibit one set of resonances, which indicates that the fibril is highly ordered and structurally homogeneous. In addition an alternative rotamer of Phe20 (Fig. 3b,c) could be confirmed by

ssNMR which at the resolution of 4 Å could not have been assigned reliably by the EM density map alone.

In the same direction, Baker et al. proposed a sample preparation method for membrane proteins to use same specimens for cryo-EM and NMR techniques [22•]. With this new approach, they determined the structural and conformational differences of YidC in native membranes from purified and reconstituted YidC. The authors also highlighted that this approach could substantially reduce cost and time for the preparation of the samples.

4. Combining cryo-EM and NMR and Simulation

The traditional approach [23] to combine experimental data with prior knowledge of proteins is to minimize a hybrid energy function

$$E = E_{MM} + wE_{exp}\,,$$

where the molecular mechanics force-field E_{MM} describes prior knowledge of proteins (e.g. stereo-chemical restraints), and E_{exp} describes the fit of the model to the experimental data. When using different types of restraints, which are schematically visualized and listed in Figure 1, the main question is how to choose the different weights w_r in

$$E = E_{MM} + \sum_r w_r E_{exp,r}\,.$$

Usually the optimal weight, w, is determined by cross-validation, which has the purpose of preventing overfitting. The Bayesian formalism instead provides a more general approach to determine this weight directly from the experimental errors, provided they are known with sufficient accuracy. The Bayesian approach to protein structure determination has been introduced by Rieping et al. [24], was recently formulated for cryo-EM data [25••], and was in a similar way implemented into the Integrative Modelling Platform (IMP) [26] by Bonomi et al. [27•], which is a software that enables integrating data from diverse biochemical and biophysical experiments for atomic model building.

In general, we aim for the determination of an ensemble of possible structures, because the measured samples also do contain ensembles of structures. When an ensemble is constructed to describe the experimental data, there are however two different interpretations of what the structural ensemble represents: uncertainty or true dynamics [21]. The choice of the approach to determine the ensemble obviously depends on what the ensemble is supposed to represent. The methods presented in the following interpret the ensemble as true structural dynamics.

When combining experimental data with MD simulations for the determination of ensembles, the maximum entropy (MaxEnt) principle [28] has been found useful, which minimally biases the distribution obtained from an unrestrained MD simulation with respect to the entropy of the distribution [29, 30]. Restraints can be introduced into an MD simulation to create a MaxEnt solution [31, 32, 33, 34]. It has been shown that ensemble-restrained MD simulations yield a MaxEnt solution [35, 36]. An implementation of this approach is available through PLUMED-ISDB [5, 37]. NMR provides also information on the time-scales of motions. To include such time-resolved data into MD simulations, Capelli et al. [38••] propose an approach that is based on the maximum-caliber principle [39] and uses replica-averaged simulations with a time-dependent potential.

The MaxEnt principle can be combined with a Bayesian formulation to take into account a specific error model that describes the experimental errors, as was done in the metainference method [40•]. Another class of such approaches are reweighting schemes, where first an ensemble is created either by MD simulations or Monte-Carlo calculations, and then afterwards all structures in the ensemble are weighted such that the ensemble average best fits the experimental data. Reweighting approaches have been developed that also includes MaxEnt restraints and Bayes formalism [41, 42, 43, 44•, 45, 46]. The reweighting approach requires that all relevant structures are well sampled in an MD (or MC) simulation, which can be difficult to achieve for large protein complexes [47].

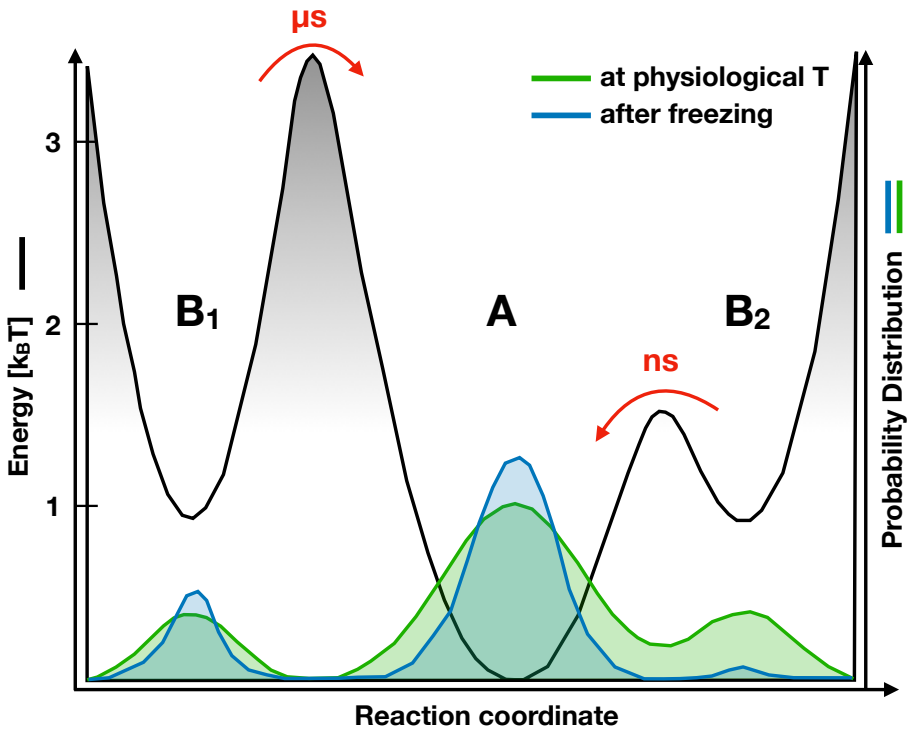

Figure 4: Schematic plot of an energy landscape (black, left axis) and the corresponding Boltzmann distribution (green, right axis) in solution state (room temperature). When freezing this distribution on a microsecond time-scale the resulting non-equilibrium distribution depends on the height of the energy barriers (blue, right axis).

5. NMR and cryo-EM report on different ensembles

In cryo-EM proteins are frozen in a thin film of vitrified ice. The freezing process happens on a microsecond time-scale (it needs to be sufficiently fast to prevent the formation of ice crystals). The obtained frozen conformational distribution is therefore narrower than the Boltzmann distribution at physiological temperatures (in solution state), which means low-energy states are more populated than at native condition. Figure 4 schematically illustrates how the conformational distribution changes from a room temperature Boltzmann distribution to a distribution of quickly frozen conformations. Since the freezing happens fast, the frozen distribution is trapped in a non-equilibrium state, which

means whether a state is visible in the cryo-EM data set depends not only on free energy differences between the states, but also on the free energy barriers, i.e. transition rates, between the states. For example state B2 (see Fig. 4) is separated from a low energy minimum (state A) by a low energy barrier (with a corresponding transition rate on the nanosecond time-scale) and can relax towards state A during the cooling time of microseconds. However, state B1 needs to overcome a higher energy barrier (with a corresponding transition rate on the microsecond time-scale) and therefore will be trapped and its population will not change much during the freezing process. This difference in the conformational distribution between NMR and cryo-EM needs to be taken into account in the model building process.

6. Outlook

The relative weights, w_r, for combining different restraints have to be chosen according to the reliability of the information that they represent. In a Bayesian description these weights are formally obtained from Likelihood distributions representing the error. The accurate modeling of the errors is therefore very important, but not yet completely implemented in existing MaxEnt and/or Bayesian approaches [48], which is particularly true for cryo-EM data. The error of a cryo-EM density map is usually quantified by the spectral signal-to-noise ratio (determined by the FSC). The reliability of the density map also varies strongly between different regions, measured by local resolution, which means the corresponding weight, w, should be position dependent.

Both NMR and cryo-EM also provide additional information on conformational dynamics. In principle NMR can yield for example S^2 order parameters, RDC [49] orientation restraints, and peak widths [50, 51] which report on the conformational ensemble. Information on dynamics in single-particle cryo-EM is usually represented by multiple conformations as well as by density variance and covariances for each of these conformational states. In the implementations of the MaxEnt principle mentioned above the ensemble is predicted (e.g. by MD

simulation) and experimental data is used only as ensemble-averaged restraints. However, since the experiments provide also information on the ensemble, the predicted ensemble needs to be combined with experimental data to determine the most likely structural ensemble. In the Bayesian framework this requires also the modelling of the errors of the simulation and the errors of the experimental ensemble data, such as for example errors of density covariances.

The error of the conformational distribution predicted by simulation comes from both incomplete conformational sampling as well as from inaccuracies in the force field (an approximate description of the energy landscape). The force field inaccuracies depend non-linearly on the parameters (such as partial charges, force constants, etc). Quantifying the errors of the simulation and of the experimental data on dynamics is difficult but needs to be done for a complete Bayesian formulation, which is still an open problem and corresponding computational approaches still need to be developed.

Acknowledgements

We thank Henrike Heise for discussions. This work is supported by the Helmholtz Association Initiative and Networking Fund under project number ZT-I-0003.

References

[1] D. P. Frueh, A. C. Goodrich, S. H. Mishra, S. R. Nichols, NMR methods for structural studies of large monomeric and multimeric proteins, Curr. Opin. Struct. Biol. 23 (2013) 734–739. doi:10.1016/j.sbi.2013.06.016.

[2] R. Tycko, Solid-state NMR studies of amyloid fibril structure, Annu. Rev. Phys. Chem. 62 (2011) 279–299. doi:10.1146/annurev-physchem-032210-103539.

[3] D. B. Good, S. Wang, M. E. Ward, J. Struppe, L. S. Brown, J. R. Lewandowski, V. Ladizhansky, Conformational dynamics of a seven

transmembrane helical protein Anabaena Sensory Rhodopsin probed by solid-state NMR, J. Am. Chem. Soc. 136 (2014) 2833–2842. doi:10.1021/ja411633w.

[4] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR 44 (2009) 213–223. doi:10.1007/s10858-009-9333-z.

[5] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird, Comput. Phys. Commun. 185 (2014) 604–613. doi:10.1016/j.cpc.2013.09.018.

[6] A. Carlon, E. Ravera, W. Andrałojć, G. Parigi, G. N. Murshudov, C. Luchinat, How to tackle protein structural data from solution and solid state: An integrated approach, Prog. Nucl. Magn. Reson. Spectrosc. 92–93 (2016) 54–70. doi:10.1016/j.pnmrs.2016.01.001.

[7] A. A. Rizzo, M. M. Suhanovsky, M. L. Baker, L. C. Fraser, L. M. Jones, D. L. Rempel, M. L. Gross, W. Chiu, A. T. Alexandrescu, C. M. Teschke, Multiple functional roles of the accessory I-domain of bacteriophage P22 coat protein revealed by NMR structure and CryoEM modeling, Structure 22 (2014) 830–841. doi:10.1016/j.str.2014.04.003.

[8] V. E. Galkin, A. Orlova, D. S. Kudryashov, A. Solodukhin, E. Reisler, G. F. Schröder, E. H. Egelman, Remodeling of actin filaments by ADF/cofilin proteins, Proc. Natl. Acad. Sci. 108 (2011) 20568–20572. doi:10.1073/pnas.1110109108.

[9] A. Lu, V. G. Magupalli, J. Ruan, Q. Yin, M. Atianand, M. R. Vos, G. Schröder, K. Fitzgerald, H. Wu, E. Egelman, Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes, Cell 156 (2014) 1193–1206. doi:10.1016/j.cell.2014.02.008.

[10] B. Bardiaux, G. C. de Amorim, A. L. Rico, W. Zheng, I. Guilvout, C. Jollivet, M. Nilges, E. H. Egelman, N. Izadi-Pruneyre, O. Francetic,

Structure and assembly of the enterohemorrhagic *Escherichia coli* Type 4 Pilus, Structure 7 (2019) 1082–1093.e5. doi:10.1016/j.str.2019.03.021.

[11] P. Cuniasse, P. Tavares, E. V. Orlova, S. Zinn-Justin, Structures of biomolecular complexes by combination of NMR and cryoEM methods, Curr. Opin. Struct. Biol. 43 (2017) 104–113. doi:10.1016/j.sbi.2016.12.008.

[12••] J.-P. Demers, B. Habenstein, A. Loquet, S. Kumar Vasa, K. Giller, S. Becker, D. Baker, A. Lange, N. G. Sgourakis, High-resolution structure of the *Shigella* type-III secretion needle by solid-state NMR and cryo-electron microscopy, Nat. Commun. 5 (2014) 4976. doi:10.1038/ncomms5976.
A large helical assembly is determined combining ssNMR and low-resolution cryoEM data with structure prediction using Rosetta.

[13••] D. F. Gauto, L. F. Estrozi, C. D. Schwieters, G. Effantin, P. Macek, R. Sounier, A. C. Sivertsen, E. Schmidt, R. Kerfah, G. Mas, J.-P. Colletier, P. Güntert, A. Favier, G. Schoehn, P. Schanda, J. Boisbouvier, Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex, Nat. Commun. 10 (2019) 2697. doi:10.1038/s41467-019-10490-9,
The authors present an approach to combine NMR and EM data to solve a structure that could not have been solved with either method alone. The method identifies secondary structure elements in the density map and also in the sequence based on NMR chemical shift assignments and then finds corresponding elements through NMR-derived distance restraints.

[14] L. Borissenko, M. Groll, Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes, J. Mol. Biol. 346 (2005) 1207–1219. doi:10.1016/j.jmb.2004.12.056.

[15••] J. R. Perilla, G. Zhao, M. Lu, J. Ning, G. Hou, I.-J. L. Byeon, A. M. Gronenborn, T. Polenova, P. Zhang, CryoEM structure refinement by integrating NMR chemical shifts with molecular dynamics simulations, J. Phys. Chem. B 121 (2017) 3853–3863. doi:10.1021/acs.jpcb.6b13105, HIV-1 capsid protein determined by a combination of MAS NMR from which chemical shifts were obtained and used as restraints in the combined MD driven refinement against the cryo-EM map.

[16] K. Zhang, S. C. Keane, Z. Su, R. N. Irobalieva, M. Chen, V. Van, C. A. Sciandra, J. Marchant, X. Heng, M. F. Schmid, D. A. Case, S. J. Ludtke, M. F. Summers, W. Chiu, Structure of the 30 kDa HIV-1 RNA dimerization signal by a hybrid cryo-EM, NMR, and molecular dynamics approach, Structure 26 (2018) 490–498. doi:10.1016/j.str.2018.01.001.

[17] J. J. Frye, N. G. Brown, G. Petzold, E. R. Watson, C. R. R. Grace, A. Nourse, M. A. Jarvis, R. W. Kriwacki, J.-M. Peters, H. Stark, B. A. Schulman, Electron microscopy structure of human APC/C$^{\mathrm{CDH1}}$–EMI1 reveals multimodal mechanism of E3 ligase shutdown, Nat. Struct. Mol. Biol. 20 (2013) 827–835. doi:10.1038/nsmb.2593.

[18] L. Chang, Z. Zhang, J. Yang, S. H. McLaughlin, D. Barford, Atomic structure of the APC/C and its mechanism of protein ubiquitination, Nature 522 (2015) 450–454. doi:10.1038/nature14471.

[19] M. G. Iadanza, R. Silvers, J. Boardman, H. I. Smith, T. K. Karamanos, G. T. Debelouchina, Y. Su, R. G. Griffin, N. A. Ranson, S. E. Radford, The structure of a β2-microglobulin fibril suggests a molecular basis for its amyloid polymorphism, Nat. Commun. 9 (2018) 4517. doi:10.1038/s41467-018-06761-6.

[20] L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. G. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold, G. F. Schröder, Fibril structure of amyloid-β(1–42) by cryo-electron microscopy, Science 358 (2017) 116–119. doi:10.1126/science.aao2825.

[21] G. F. Schröder, Hybrid methods for macromolecular structure determination: experiment with expectations, Curr. Opin. Struct. Biol. 31 (2015) 20–27. doi:10.1016/j.sbi.2015.02.016.

[22•] L. A. Baker, T. Sinnige, P. Schellenberger, J. de Keyzer, C. A. Siebert, A. J. M. Driessen, M. Baldus, K. Grünewald, Combined 1H-detected solid-state NMR spectroscopy and electron cryotomography to study membrane proteins across resolutions in native environments, Structure 26 (2018) 161–170. doi:10.1016/j.str.2017.11.011,
This paper is about a sample preparation method for the structural and functional study of membrane proteins in their native environment, where the same specimens could be used for both techniques. With new sample preparation approach they showed that the conformation and likely dynamics of YidC in native membranes differs from purified and reconstituted YidC, and observe corresponding RNC binding differences.

[23] A. Jack, M. Levitt, Refinement of large structures by simultaneous minimization of energy and R factor, Acta Crystallogr. Sect. A 34 (1978) 931–935. doi:10.1107/s0567739478001904.

[24] W. Rieping, M. Habeck, M. Nilges, Inferential structure determination, Science 309 (2005) 303–306. doi:10.1126/science.1110428.

[25••] M. Habeck, Bayesian modeling of biomolecular assemblies with cryo-EM maps, Front. Mol. Biosci. 4 (2017) 15. doi:10.3389/fmolb.2017.00015,
Describes the application of the inferential structure determination approach to cryo-EM data and Markov chain Monte Carlo algorithms for sampling model structures from the posterior distribution.

[26] 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. doi:10.1371/journal.pbio.1001244.

[27•] 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. doi:10.1016/j.str.2018.09.011,
The authors describe a Bayesian scoring function for cryo-EM data similar to Habeck (2017) implemented in IMP.

[28] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957) 620–630. doi:10.1103/physrev.106.620.

[29] M. Bonomi, G. T. Heller, C. Camilloni, M. Vendruscolo, Principles of protein structural ensemble determination, Curr. Opin. Struct. Biol. 42 (2017) 106–116. doi:10.1016/j.sbi.2016.12.004.

[30] A. Cesari, S. Reißer, G. Bussi, Using the maximum entropy principle to combine simulations and solution experiments, Computation 6 (2018) 15. doi:10.3390/computation6010015.

[31] J. W. Pitera, J. D. Chodera, On the use of experimental observations to bias simulated ensembles, J. Chem. Theory Comput. 8 (2012) 3445–3451. doi:10.1021/ct300112v.

[32] W. Boomsma, J. Ferkinghoff-Borg, K. Lindorff-Larsen, Combining experiments and simulations using the maximum entropy principle, PLoS Comput. Biol. 10 (2014) e1003406. doi:10.1371/journal.pcbi.1003406.

[33] A. D. White, G. A. Voth, Efficient and minimal method to bias molecular simulations with experimental data, J. Chem. Theory Comput. 10 (2014) 3023–3030. doi:10.1021/ct500320c.

[34] D. B. Amirkulova, A. D. White, Recent advances in maximum entropy biasing techniques for molecular dynamics, Mol. Simul. 45 (2019) 1285–1294. doi:10.1080/08927022.2019.1608988,

[35] A. Cavalli, C. Camilloni, M. Vendruscolo, Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle, J. Chem. Phys. 138 (2013) 094112. doi:10.1063/1.4793625.

[36] B. Roux, J. Weare, On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method, J. Chem. Phys. 138 (2013) 084107. doi:10.1063/1.4792208.

[37] M. Bonomi, C. Camilloni, Integrative structural and dynamical biology with PLUMED-ISDB, Bioinformatics 33 (2017) 3999–4000. doi:10.1093/bioinformatics/btx529.

[38••] R. Capelli, G. Tiana, C. Camilloni, An implementation of the maximum-caliber principle by replica-averaged time-resolved restrained simulations, J. Chem. Phys. 148 (2018) 184114. doi:10.1063/1.5030339,
The authors show that replica-averaged simulations restrained using a time-dependent potential are equivalent to the principle of maximum caliber, the dynamic version of the principle of maximum entropy. This approach may allow integration of time-resolved data with MD simulations.

[39] E. T. Jaynes, The minimum entropy production principle, Annu. Rev. Phys. Chem. 31 (1980) 579–601. doi:10.1146/annurev.pc.31.100180.003051.

[40•] M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo, Metainference: A Bayesian inference method for heterogeneous systems, Sci. Adv. 2 (2016) e1501177. doi:10.1126/sciadv.1501177,
The metainference method is presented, which combines the maximum entropy principle with a Bayesian formulation.

[41] H. T. A. Leung, O. Bignucolo, R. Aregger, S. A. Dames, A. Mazur, S. Bernèche, S. Grzesiek, A rigorous and efficient method to reweight

very large conformational ensembles using average experimental data and to determine their relative information content, J. Chem. Theory Comput. 12 (2015) 383–394. doi:10.1021/acs.jctc.5b00759.

[42] G. Hummer, J. Köfinger, Bayesian ensemble refinement by replica simulations and reweighting, J. Chem. Phys. 143 (2015) 243150. doi:10.1063/1.4937786.

[43] J. Köfinger, B. Różycki, G. Hummer, Inferring structural ensembles of flexible and dynamic macromolecules using Bayesian, maximum entropy, and minimal-ensemble refinement methods, in: M. Bonomi, C. Camilloni (Eds.), Biomolecular Simulations, Methods Mol. Biol., volume 2022, Springer, 2019, pp. 341–352. doi:10.1007/978-1-4939-9608-7_14.

[44•] J. Köfinger, L. S. Stelzl, K. Reuter, C. Allande, K. Reichel, G. Hummer, Efficient Ensemble Refinement by Reweighting, J. Chem. Theory Comput., 15(5), (2019) 3390–3401. doi:10.1021/acs.jctc.8b01231,
Two efficient methods for reweighting ensembles with experimental data are presented. NMR J-coupling data are used to reweight an ensemble of an Ala-5 peptide.

[45] S. Bottaro, T. Bengtsen, K. Lindorff-Larsen, Integrating molecular simulation and experimental data: A Bayesian/maximum entropy reweighting approach, bioRxiv (2018) 457952. doi:10.1101/457952, preprint.

[46] H. Lou, R. I. Cukier, Reweighting ensemble probabilities with experimental histogram data constraints using a maximum entropy principle, J. Chem. Phys. 149 (2018) 234106. doi:10.1063/1.5050926.

[47] R. Rangan, M. Bonomi, G. T. Heller, A. Cesari, G. Bussi, M. Vendruscolo, Determination of structural ensembles of proteins: Restraining vs reweighting, J. Chem. Theory Comput. 14 (2018) 6632–6641. doi:10.1021/acs.jctc.8b00738.

[48] D. Schneidman-Duhovny, R. Pellarin, A. Sali, Uncertainty in integrative structural modeling, Curr. Opin. Struct. Biol. 28 (2014) 96–104. doi:10.1016/j.sbi.2014.08.001.

[49] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, B. L. de Groot, Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution, Science 320 (2008) 1471–1475. doi:10.1126/science.1157092.

[50] R. H. Havlin, R. Tycko, Probing site-specific conformational distributions in protein folding with solid-state NMR, Proc. Natl. Acad. Sci. 102 (2005) 3284–3289. doi:10.1073/pnas.0406130102.

[51] A. König, D. Schölzel, B. Uluca, T. Viennet, Ü. Akbey, H. Heise, Hyperpolarized MAS NMR of unfolded and misfolded proteins, Solid State Nucl. Magn. Reson. 98 (2019) 1–11. doi:10.1016/j.ssnmr.2018.12.003.