Breaking free from the crystal lattice: Structural biology in solution to study protein degraders

Kevin Haubrich¹,a, Valentina A. Spiteri¹,a, William Farnaby¹, Frank Sobott² and Alessio Ciulli¹

Abstract
Structural biology offers a versatile arsenal of techniques and methods to investigate the structure and conformational dynamics of proteins and their assemblies. The growing field of targeted protein degradation centres on the premise of developing small molecules, termed degraders, to induce proximity between an E3 ligase and a protein of interest to be signalled for degradation. This new drug modality brings with it new opportunities and challenges to structural biologists. Here we discuss how several structural biology techniques, including nuclear magnetic resonance, cryo-electron microscopy, structural mass spectrometry and small angle scattering, have been explored to complement X-ray crystallography in studying degraders and their ternary complexes. Together the studies covered in this review make a case for the invaluable perspectives that integrative structural biology techniques in solution can bring to understanding ternary complexes and designing degraders.

Addresses
¹ Centre for Targeted Protein Degradation & Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dundee, UK
² School of Molecular and Cellular Biology & Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK

Corresponding author: Ciulli, Alessio (a.ciulli@dundee.ac.uk)
(Haubrich K.), (Spiteri V.A.), (Farnaby W.), (Sobott F.), (Ciulli A.)

¹ These authors contributed equally.

Introduction
Small-molecule degraders such as Proteolysis targeting chimeras (PROTACs) and molecular glues work by forming a ternary complex with an E3 ligase and a neo-substrate. Within such complexes, degraders either introduce de novo, or stabilize existing transient, protein-protein interactions (PPIs). This results in targeted ubiquitination and subsequent degradation of the neo-substrate (Figure 1) [1–3]. Designing drugs with the guidance of structural data has been the mainstay of pharmaceutical research for decades, with crystal structures offering a “view” into how to rationally design and optimise compounds. Targeted protein degradation (TPD) presents new challenges and opportunities for structure-based drug design, as it necessitates a departure from solely optimising binary protein/ligand interactions and requires additional insights into the ternary complex structure, dynamics and PPIs [4].

While protein interactions and protein complexes have been studied by the established structural techniques of X-ray crystallography and cryo-Electron Microscopy (EM), these studies can be limited to more stable systems and made more challenging due to the often dynamic nature of PPIs, both in terms of the lifetime of the complex as well as conformational flexibility [5,6]. This is especially problematic for ternary complexes induced by degraders, as they may involve weak, transient interactions that get enhanced by the forced proximity of the proteins. Nevertheless, key studies have successfully demonstrated structure-based PROTAC design driven by high resolution ternary structures solved by X-ray crystallography [4,7–9]. A potential limitation of using static X-ray structures to describe dynamic systems is that crystals can act as a filter that capture a single state, from an ensemble of states that exist in solution. This homogeneity is important for solving high resolution crystal structures, where conformationally heterogeneous parts of a structure do not yield interpretable electron density. However, it can be difficult to ascertain whether the crystallisable conformer fully reflects the dominant or
This caveat is supported by recent molecular dynamics simulations, which showed that the conformational landscape of degrader ternary complexes may be much more diverse than the poses observed in crystal structures. This can lead to cases where the crystal structures of ternary complexes with different degraders bound may appear very similar, but their global attributes, including both binding affinities and functional consequences to missense surface mutations can differ significantly.

To shine a light on what is unseen in crystal structures, orthogonal techniques should be explored to study the structure and conformational heterogeneity of macromolecular complexes in solution. Several alternative techniques have been employed to investigate transient and dynamic interactions across different fields, most...
notably in the study of protein/RNA complexes [14–18]. These techniques include nuclear magnetic resonance spectroscopy (NMR), small angle scattering (SAXS) and structural mass spectrometry (MS). Individually, these techniques have their own limitations that may either render them unable to provide data of comparable atomic resolution to crystallography or cryo-EM, or restrict the size of system under study. Nonetheless, the techniques complement X-ray and cryo-EM as well as each other, and offer an integrative approach that can facilitate insights into the structure and dynamics of protein complexes. The purpose of this review is therefore to highlight the use of these techniques in the field of targeted protein degradation and make a case for the wider use of solution structural methods and integrative structural biology, as complementary approaches to crystallography for the study of ternary complexes.

**Nuclear magnetic resonance spectroscopy**

NMR is the only solution technique capable of delivering atomic resolution structures [19]. Despite this, very few NMR studies of ternary complexes have been reported and NMR studies on PROTACs and molecular glues to date are largely limited to ligand-observed studies. Structural studies by NMR have recently given unique insights into the ‘chameleonity’ of PROTACs, rationalising how PROTACs remain cell permeable despite their high molecular weight and bifunctional nature made of two binding ligands joined by a linker. Nuclear Overhauser Effect Spectroscopy (NOESY) based structural ensembles of PROTACs, in chloroform or water/DMSO mixtures to mimic the hydrophobicity of different environments in the cytosol and cell membranes, demonstrated that some PROTACs can fold by stacking their ligase and neo-substrate binding moieties on top of each other, thereby shielding hydrophilic regions and allowing the PROTAC to cross cell membranes [20]. NMR is well-established as an extremely sensitive technique to detect even transient protein/ligand interactions. $^1$H saturation transfer difference (STD) NMR was used to confirm the weak interaction of the molecular glue indisulam with the E3 ligase substrate receptor DCAF15 in the absence of a neo-substrate [21]. Competitive NMR with a $^{19}$F-labelled probe was used to estimate affinities and cooperativity of ternary complex formation for VHL/Bromodomain recruiting and VHL-homodimerising PROTACs [22].

To our knowledge, the only study to date employing protein NMR on PROTAC-induced ternary complexes used $^1$H/$^15$N-HSQC spectra of cIAP1$^{Bir3}$ in the ternary complex overlayed well with that of the binary cIAP1$^{Bir3}$/degrader complex, consistent with no stable PPIs forming in solution. However, diffusion observed NMR (DOSY) still confirmed the formation of a ternary complex through an increase in the diffusion coefficient.

More studies of ternary complexes have likely been prevented by the size limitations of protein-observed NMR, historically making NMR studies of complexes exceeding 35 kDa difficult due to excessive line broadening and spectral crowding. However, over the past 20 years technical innovations such as Transverse relaxation optimized spectroscopy (TROSY), selective labelling and structure calculations relying on long distance restraints such as paramagnetic relaxation enhancement (PREs), pseudocontact shifts (PCS) or residual dipolar coupling (RDCs) rather than NOESYs have allowed NMR to largely overcome these limitations [16,24–26]. To date the largest protein structures solved de novo by NMR exceed 80 kDa (e.g., malate synthase G or Translocase Motor SecA) [27,28] and it has been shown to be capable of studying protein/protein interfaces and dynamics of complexes several hundred kilodaltons in size (e.g., 390 kDa RNP complex box C/D or even the 20S proteasome) [15,29]. This brings even the largest ternary complexes in range of NMR studies.

NMR could be useful in the TPD field to map or validate weak PPIs in solution using e.g., cross-saturation experiments or chemical shift perturbations in SEATROSYs even for large complexes [30,31]. Most NMR techniques require stable isotope labelling (most commonly $^{15}$N, $^{13}$C, $^2$H), which can be costly and difficult to achieve in some expression systems. However, the reliance on specific labelling patterns also allows for detailed studies of specific, selectively labelled subcomponents in the context of the full complex, giving NMR unique capabilities for structural studies. For example, it would be possible to study the conformational ensemble of a PROTAC in the presence of its interacting proteins, if these are masked by deuteration, and thereby gaining insights into conformational selection upon complex formation.

**Structural mass spectrometry**

Structural mass spectrometry (MS) refers to a collection of MS techniques that provide information on the higher-order structure of proteins and their complexes [32]. Native MS approaches including ion-mobility (IM)-MS are used to characterize stoichiometry, size, shape, and distribution of conformations and interactions (while retaining the native structure of the protein/protein complexes, by gentle ionisation methods such as nano-electrospray. On the other hand, labelling techniques such as hydrogen-deuterium exchange (HDX)-MS take advantage of the rapid
exchange of deuterium with solvent exposed amide hydrogens in deuterated buffer. The exchange is then quenched, the protein rapidly digested and the peptides are analysed by MS. Mass analysis is used to determine increased mass due to deuterium uptake. For structural interpretation the peptides are mapped on a three-dimensional model.

Several studies have investigated degrader ternary complexes by MS. The first of these studies by Zorba et al. (2018) focused on understanding the role of cooperativity in cereblon-recruiting degraders targeting BTK [33]. The study utilised HDX-MS and despite comprehensive experiments, Zorba et al. were not able to identify any interacting residues. This may have been due to the use of a truncated cereblon construct [33]. Later, a study by Eron et al. (2021) utilised a longer cereblon construct (only missing N-terminal residues 1—39) and demonstrated that the BRD4 (BD1) degrader CFT-1297 was able to induce HDX-MS protection of several residues at the BD1-cereblon interface, which corresponds to its binding site [34]. However, in a similar experiment with a different degrader, dBET6, HDX-MS protection at the same interface could not be detected, suggesting a lack of stable, highly populated ternary complex interfaces found in solution in this case. Both studies focussed on cereblon and suggest a more complex conformational landscape for cereblon-recruiting degraders. Indeed, Watson et al. (2022) recently used HDX-MS and cryo-EM to investigate the allosteric effect of molecular glue binding to cereblon that results in a “closed” versus “open” conformation of cereblon. They found that the addition of the ligand resulted in reduced solvent exposure, relative to the apo state, at the interface of the Lon and the thalidomide binding domain (TBD), consistent with a transition from an open to a closed conformation. A dramatic change in protection was observed at the “sensor loop” in the TBD. This data, together with extensive cryo-EM structural investigations (vide infra), shone light into an allosteric mechanism by which ligand binding induces the sensor loop to detach from the helical bundle in cereblon and adopt an upright beta-hairpin arrangement. The detachment allows the TBD to become untethered and interact with the Lon domain of cereblon, resulting in the closed conformation that favours neo-substrate recruitment.

In another study, Dixon et al. (2022) [11] used HDX-MS with VHL-recruiting PROTAC degraders of SMARCA2/4 to show clear and significant protection of residues at the protein-protein interface, consistent with previously determined ternary complex cocrystal structures [7]. This HDX-MS data was fed into in silico docking studies and improved the ability to predict ternary complexes, demonstrating the utility of this method for enriching computational techniques [11].

In a departure from HDX-MS, Beveridge et al. (2020) showed that native mass spectrometry can effectively predict degrader efficacy [35]. A clear correlation was observed between a higher fraction of ternary complex observed in the MS spectra and favourable thermodynamic and kinetic parameters such as increased cooperativity and longer half-lives of the ternary complex, as determined in a previous study using surface-plasmon resonance (SPR) and fluorescence polarization (FP) binding assays [36]. In their study, Beveridge et al. focussed predominantly on the well-characterised BRD:MZ1/AT1:VCB systems [4,36] (Figure 1a). It would be interesting to uncover if similar correlations could be observed using native-MS for other systems. Finally, a study by Song et al. (2021) used ion-mobility MS to show that BRD4:MZ1:VCB was able to populate a distribution of several distinct conformations [37], some of which had not been observed crystallographically [4]. Ion mobility resolves conformational states as ions travel through the drift tube, with more compact structures travelling faster than those with a larger cross-sectional area. Moreover, the study employed gas-phase fragmentation methods such as collision induced dissociation (CID) and electron capture dissociation (ECD) to allow for the mapping of ligand interaction sites. Using CID, it was observed that the more compact ternary complexes have protein-protein interactions that allow a binary VCB:BRD4 complex to persist when MZ1 is eliminated. This implies the significance of those PPIs in stabilising the ternary complexes. It should be noted that the VCB:BRD4 binary complexes have not been observed in solution in the absence of MZ1 [4].

Together these articles demonstrate how structural MS techniques can augment our understanding of ternary complexes and provide information on the whole conformational landscape (via IM-MS). Through HDX-MS we can gather information on the protein-protein interface which can help us understand systems that suffer from an absence of X-ray/EM data.

**Small angle X-ray scattering**

Small Angle X-ray scattering (SAXS) is a high-throughput, in solution technique that requires minimal sample preparation after the protein has been purified, with most physiologically relevant buffers being suitable. SAXS is one of the few techniques that can be used to examine highly flexible proteins and for systems within a wide range of sizes 1—1000 nm [38]. Moreover, where sample availability may be limiting, a good scattering curve can be obtained from dilute solutions. The main caveat for SAXS is that the data obtained is low resolution and often needs to be supplemented by data obtained via higher resolution methods, e.g. existing structures of sub-complexes. It can also be challenging to interpret heterogeneous distributions unless there is prior knowledge of the system. So far, this technique has not been widely exploited by the TPD community, however its use can augment in silico searches as
<table>
<thead>
<tr>
<th>Technique</th>
<th>Resolution</th>
<th>Feasible system size limits</th>
<th>Is Flexibility an issue?</th>
<th>Protein requirements</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray Crystallography</td>
<td>Atomistic/high</td>
<td>Limited by flexibility</td>
<td>Yes</td>
<td>Moderate-high</td>
<td>Atomistic resolution allows protein-ligand and protein-protein interactions to be scrutinised. High throughput due to high degree of automatization Accessible to non-experts Still the gold standard for binary protein/ligand interactions and rigid systems</td>
<td>Crystallisation of the system is dependent on favourable crystal contacts that filter for a single conformation. Ternary complexes with inherent flexibility and/or lower cooperativity may be difficult to crystallise.</td>
<td>[4,7,10,62]</td>
</tr>
<tr>
<td>NMR</td>
<td>Atomistic/high</td>
<td>Up to 40 kDa for atomic resolution experiments. Up to 1 MDa with TROSY experiments and selective labelling.</td>
<td>No</td>
<td>Moderate-high</td>
<td>Atomistic resolution in solution that allows for the observation of dynamics. Gives insight to PROTAC conformations. Can map/validate weak PPI. Labelling allows for the selective analysis of particular components in the system.</td>
<td>Isotope labelling can be challenging and expensive. Atomsic resolution can be challenging to achieve for big complexes. Low throughput for structure elucidation</td>
<td>[20–23]</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Atomistic/high-moderate</td>
<td>38 kDa (theoretical) -5 MDa</td>
<td>Potentially makes data analysis complicated</td>
<td>Low</td>
<td>Can achieve good resolution on large protein complexes.</td>
<td>Sample and grid optimisation can be challenging. High resolution for larger complexes may be difficult to achieve without crossing linking. Low contrast for small proteins. Heterogeneous conformer populations can be hard to interpret due to data being averaged.</td>
<td>[21,44]</td>
</tr>
<tr>
<td>SAXS</td>
<td>Low</td>
<td>1–1000 nm</td>
<td>No</td>
<td>Low</td>
<td>Resolves conformational and constitutional heterogeneity High throughput and can use most physiologically relevant buffers Good scattering curve can also be obtained for dilute samples.</td>
<td>Resolution is low and may need to be supplemented by higher resolution data or other data on the system to help interpretation.</td>
<td>[11]</td>
</tr>
</tbody>
</table>
demonstrated by Dixon et al. (2022) [11]. In this study, SAXS data was used to validate molecular dynamics derived conformational ensembles of SMARCA2/4:ACBI1:VHL ternary complexes [7] by producing theoretical SAXS curves for each conformer in the ensemble. Additionally, it was recently shown that SAXS can be used as a structural screening tool, as SAXS is sensitive to changes in size and shape that allows for estimation of the relative proportion of ternary complex in solution and classification based on structural similarity [39]. We believe that there is significant scope for SAXS to be utilised to screen for ternary complexes in TPD due to its high-throughput and structural sensitivity.

Cryo-electron microscopy
Although not a true solution technique as it relies on capturing the sample in a layer of vitrified ice, cryo-EM shares many of the desired properties of these techniques and crucially, unlike crystallography, retains information about conformational heterogeneity in solution. In addition, due to recent technological advancements it is now capable of delivering near atomic-resolution structures of comparable quality to crystallography for systems of sufficient size [40–42]. In the TPD field, these capabilities have been used to obtain a 3.5 Å structure of the ternary complex of DDB1-DCAF15 and RBM39, induced by the molecular glue indisulam [21]. Another study published around the same time reported a 4.4 Å cryo-EM map of the DDB1-DCAF15/RBM39 complex induced by the glue E7820, however the resolution of the cryo-EM maps were insufficient to reliably interpret E7820 binding, and required a separate co-crystal structure of DDB1-DCAF15:E7820:RBM39 to build an unambiguous model of the ternary complex [43]. Recently, a series of cryo-EM structures using iterative heterogeneous classification gave new insights into the conformational dynamics of cereblon apo-form and in the presence of molecular glues [44]. This highlights the ability of cryo-EM to resolve heterogeneous conformations in solution while maintaining high resolution, allowing for insights into heterogenous systems unobtainable by crystallography. For now, these additional capabilities come with increased costs in terms of the necessary equipment, expertise, and experiment time, as cryo-EM has not yet reached the level of automation, throughput and convenience found in crystallography. A significant factor here is that cryo-EM still suffers from limitations due to charge and surface effects, preferred particle orientations and the need for classification and averaging, all of which require careful and time-consuming optimisation of the sample preparation and typically limit its applicability to complexes of considerable size and limited flexibility [45].

Bringing it all together in silico
Structural biology has long benefitted and will continue to advance from being able to model protein systems in
The potential impact of combining *in silico* modelling with experimental methods is more significant when it comes to solution structural biology techniques. These techniques often suffer from limited information content and may require combining multiple approaches [46]. This has led to the development of software packages for integrative modelling such as the Integrative Modeling Platform (IMP) or HADDOCK, that allow for refinement of structural models against sparse restraints obtained from diverse methods, including shape and distance restraints as well as interaction surfaces [47,48]. Although the TPD community has already shown an enthusiastic interest in utilising computational methods to develop a better understanding of the degrader’s mode of action, and to create theoretical/predictive models to design more effectively potent and faster degraders, it has only started to recognize the power of integrative modelling [11,12,49—52].

**Conclusion**

The ability to reliably generate high-resolution data with X-ray crystallography to rationally design drugs has revolutionized our methods to discover and optimize drug molecules. This has been possible due to the efforts of the X-ray crystallography community in developing tools ranging from automated pipelines for crystallization, data collection and processing to model building and data curation and visualisation. These not only allow for high-throughput structure determination, but also make crystallography accessible to the non-expert [53—56]. Currently no other structural biology technique offers a comparable user experience, except for SAXS that benefits from highly automated beamlines, software pipelines and public repositories [57—59]. Although SAXS and alternative techniques can generate a wealth of information, analysis and interpretation of the data typically requires expert knowledge. From being able to solve high-resolution structures of multi-subunit complexes with cryo-EM, studying transient interactions by NMR or HDX-MS and characterizing conformational heterogeneity of proteins with native ion mobility mass spectrometry, these solution-based structural biology techniques offer unique insights into dynamic systems and transient interactions that may be invisible to X-ray crystallography (Table 1). However, for these techniques to be impactful continued efforts are necessary to make them more accessible to a non-expert user, increase throughput and allow for easy visualisation and interpretation of results. This is all the more true as these techniques are best used as part of an integrative approach that balances the strengths and weaknesses of different techniques and leverages previous knowledge through *in silico* modelling. While this review focused on techniques that have so far been successfully used in the TPD field, additional techniques are available that can be utilised as part of an integrative approach or to answer specific questions. For example small-angle neutron scattering (SANS) using contrast matching could be useful to study the position and conformation of individual subunits in the context of a larger complex [60]. Electron paramagnetic resonance (EPR)-based techniques such as DEER or RIDME measure the distance distribution between two paramagnetic labels and can thereby give direct insight into the conformational heterogeneity of a complex [61]. As the field grows it will be interesting to see methods for studying the structure and dynamics of ternary complexes evolve.

**Declaration of competing interest**

The Ciulli laboratory receives or has received sponsored research support from Almirall, Amgen, AmphiTherapeutics, Boehringer Ingelheim, Eisai, Merck KGaA, Nurix Therapeutics, Ono Pharmaceutical and Tocris-Biotechne. A.C. is a scientific founder, shareholder, and advisor of AmphiTherapeutics, a company that is developing targeted protein degradation therapeutic platforms.

**Data availability**

No data was used for the research described in the article.

**Acknowledgements**

Research in the Ciulli laboratory on targeted protein degraders and PROTACs receives or has received funding from the European Research Council (ERC) under the European Union’s Seventh Framework Programme (FP7/2007-2013) as a Starring Grant (grant agreement No. ERC-2012-StG-31460 DrugE2CRLs to A.C.), and the Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement no. 875510 (EUIopen project). The IMI2 Join Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme, EFPIA companies and Associated Partners: KTH, OICR, Diamond and McGill.

**References**

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

* of special interest
** of outstanding interest


The cell permeability of PROTACs, that usually far exceed the degradation rate of the target protein, can be rationalized by studying the structural and functional plasticity of the ternary complexes using hydrogen–deuterium exchange mass spectrometry and computational modeling: Implications for structure-based design. ACS Chem Biol 2021, 16:2226–2243.

By using HDX-MS and protein docking the authors were able to build a 3D model of ternary complexes for two cereblon recruiting degraders, dBET6 and CFT-1297. They demonstrate that the highest scoring models differ from published crystal structures of dBET6-mediated ternary complexes. Moreover, they were able to demonstrate the significant conformational flexibility of the ternary complexes.

The guide for structure-based design of PROTACs is to use mass spectrometry methods that allow monitoring of the interactions for a system with only limited PPI contacts observed in the crystal structure.


This paper reports the first structure of apo-Cereblon as well as several structures of glue-bound cereblon obtained by iterative heterogeneous classification of single particle cryo-EM data. It observes the presence of an “open” and “closed” conformation of cereblon and that molecular glues drive the equilibrium towards the closed conformation. It thereby nicely shows the potential of cryo-EM to resolve conformational heterogeneity and inform degrader drug design.


