DOI:
10.1016/j.tibs.2016.04.008

Document Version
Peer reviewed version

AUTHOR QUERY FORM

ELSEVIER	**Journal:** TIBS **Article Number:** 1252	**Please e-mail your responses and any corrections to:** **E-mail:** corrections.esnl@elsevier.thomsondigital.com

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Please note that the figure resolution in these proofs is lower than the original figures to avoid problems with emailing large file sizes. The figure resolution in the final publication will be equivalent to the original figure submission.
Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

Location in article	Query / Remark: **click on the Q link to go** **Please insert your reply or correction at the corresponding line in the proof**
Q1	The author names have been tagged as given names and surnames (surnames are highlighted in teal color). Please confirm if they have been identified correctly. Please check this box or indicate your approval if you have no corrections to make to the PDF file ☐

Thank you for your assistance.

Trends in Biochemical Sciences

CellPress

Forum

Hybrid Mass Spectrometry: Towards Characterization of Protein Conformational States

Zainab Ahdash,[1] Euan Pyle,[1,2] and Argyris Politis[1,*]

A current challenge in structural biology is to unravel the conformational states of protein complexes. Hybrid mass spectrometry (MS) has emerged as a key tool to study the structural dynamics of large protein complexes unattainable by traditional methods. Here, we discuss recent advances in hybrid MS allowing characterization of challenging biological systems.

Hybrid approaches in structural biology integrate diverse methods to study the structure and dynamics of proteins and their complexes. In this context, MS-based hybrid methods have emerged as a pivotal tool for investigating the molecular architectures and dynamic interactions of challenging protein complexes [1]. As such, hybrid MS methods contribute complementary structural and dynamical information that can subsequently be integrated with sophisticated computational tools for interrogating structural models of protein complexes. A core method in such methods is native MS, which has seen significant advances over the past decade and is now used as a primary tool to uncover protein complex topologies and protein–protein interactions [2]. Combining native MS with ion mobility (IM) unravels low-resolution structural information, allowing the examination of structural models generated by computational approaches [3]. These methods can complement powerful chemical modification methods, such as chemical cross-linking [4,5], covalent labelling [6], and hydrogen deuterium exchange (HDX) [7], which, when followed by MS, allow mapping of protein–protein interactions and solvent accessibility. The main advantage of MS-based methods over traditional structural approaches is their ability to study heterogeneous and low-abundant complexes. Their computational data can be integrated to bring together complementary mass spectrometric experiments performed on the same sample, enabling a topological map of protein complexes and their associations with other biomolecules.

Integrative structure determination using advanced modelling was introduced by the Sali group [8] and, since then, it has been used for modelling large macromolecular complexes, such as the proteasome [9], ribosomes [10], and others. A milestone in integrative modelling approaches was the merging of heterogeneous data from different MS-based methods [1] enabling atomic-level characterization of dynamic (sub)complexes and their transient associations with other biomolecules. Recently, IM has been coupled with MS (IM-MS), which, in conjunction with molecular dynamics (MD) simulations, was used to assign flexible regions within a V-type ATP synthase from *Thermus thermophilus* [11]. This pioneering study illustrated the potential of integrating mass spectrometric with computational strategies to study protein complex dynamics. The significant advancement of hybrid methods involving chemical modification followed by MS enabled their application to highly dynamic transmembrane proteins, such as the G-protein-coupled receptors (GPCRs) and their signalling partners [12]. It is anticipated that emerging MS-based hybrid strategies will be used in the future for studying the conformational dynamics of increasingly complex systems, such as membrane protein complexes, and other challenging targets that continually frustrate structural biologists. Such dynamic complexes and their intermediate assemblies are normally present in very low amounts and, thus, are unsuitable for traditional structural approaches.

The hybrid MS strategy couples any purification protocol with different MS experiments and computational and data analysis (Figure 1). Initially, the purified protein complex of interest is aliquoted and subjected to different MS experiments yielding complementary information about protein complex structure and dynamics. In particular, IM-MS techniques are allowed information regarding the structural dynamics and conformational heterogeneity of protein assemblies. IM-MS measures the shape of different ions based on the time taken to traverse in the drift cell of a mass spectrometer. By converting the arrival time distributions (ATDs) of different ions into the rotationally averaged collision cross section (CCS), we can derive topological information about protein complexes, which allows the interrogation of structural models generated by computational methods. Ions with multiple conformations result in broad, Gaussian-shaped ATDs, whereas single conformations correspond to narrow ATD peaks [11]. Additional Information about protein dynamics is attained by monitoring the exchange of labile protons for bulk deuterium using HDX followed by MS [7]. HDX-MS experiments report time-dependent changes in weighted average peptide masses, which yields rates of HDX within different regions of the protein. Such information about the uptake of deuterium over time is used to map the binding interfaces and solvent accessibility on the surface of the investigated system. Covalent labelling of accessible residues followed by MS allows mapping of accessible residues on the surface of protein complexes, enabling information complementary to that from HDX-MS [6]. Comparative cross-linking MS allows the mapping of the changes that occur in

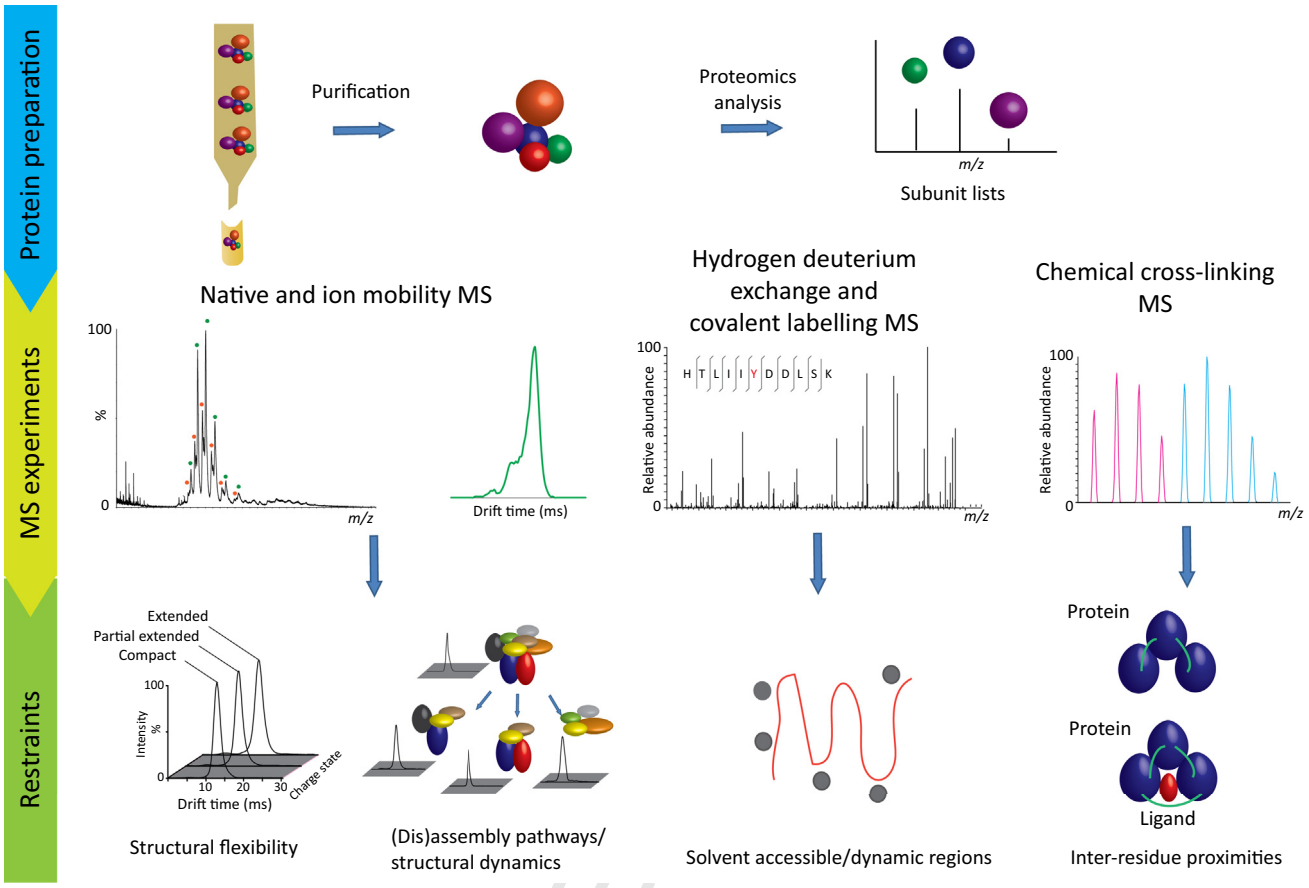

Figure 1. Mass Spectrometry (MS)-Based Hybrid Strategy for Generating Modelling Restraints. The protein complex of interest is purified using appropriate purification protocols, followed by proteomics experiments that provide a detailed list of the components within the protein complex. Native MS coupled with ion mobility (IM)-MS yield topological information of macromolecular complexes, while inspecting the profile of arrival time distributions (ATDs) enables additional information about their conformational heterogeneity. Labelling strategies followed by MS, such as hydrogen deuterium exchange (HDX)-MS and covalent labelling complement such information, allowing mapping of surface-accessible residues and the identification of dynamic regions on the surface of proteins. Comparative cross-linking analysis provides information about inter-residue proximities in response to stimuli and can be used for capturing protein conformational states.

cross-linked peptides as a response to conformational dynamics triggered by a transition state. The information generated is converted into modeling restraints, which are subsequently used to feed a computational workflow for building 3D models of macromolecular complexes. Integrating the complementary MS experiments with computational algorithms holds great promise for the study of protein complexes that are not amenable to study by traditional structural techniques. In particular, unravelling the conformational states of dynamic protein complexes intractable by crystallographic and other approaches is currently the major goal of integrative structural biology. To this end, the combination of various MS approaches brought together by integrative modelling tools will become a primary method for studying increasingly large and dynamic assemblies; their complementarity to other structural biology methods will make inroads to novel methodological developments and integrative strategies.

To bring together diverse data sets that otherwise will be reported independently, a computational workflow that utilises the restraints generated from the different experiments is used [1,8]. This workflow is usually divided into distinct steps, starting from determining the individual components of the assemblies and ending by building a protein assembly model describing its conformational states (Figure 2). Initially, the protein complex components representing the individual complex subunits and/or the building blocks are selected using available structural information. The starting complex components can be represented by atomic-level structures (e.g., X-ray crystallography, NMR, or homology modelling), molecular maps or images, coarse-grained structural models, or hybrid representations typically made up by a combination of the above. With the starting

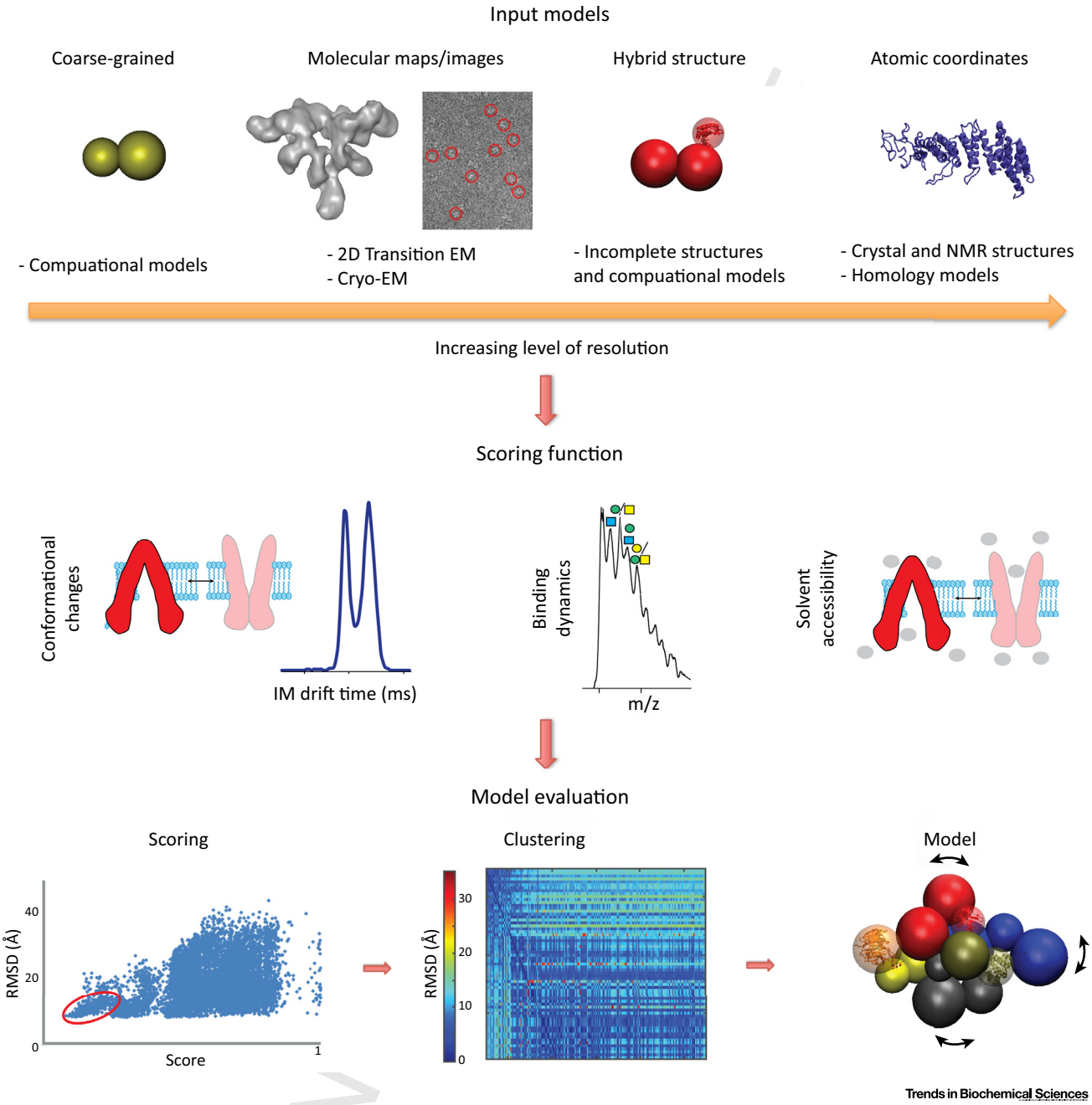

Figure 2. Computational Strategy for Building 3D Models of Native Protein Complexes. The experimental information generated is encoded into modelling restraints, which are brought together using a scoring function. The scoring function guides the sampling algorithms, allowing the conformational space of protein assemblies on timescales comparable to experiments to be sampled. This enables the generation of models consistent with the input experimental data.

structures in hand, a scoring function is used to guide the sampling of protein complex configurational space. The scoring function essentially sums up the modeling restraints generated to find the best match between the input data and the large number of models generated using sampling algorithms available. Such algorithms are able to search for protein complex conformational states by computationally handling the dynamic rearrangements of the individual subunits as they assemble into functional complexes. Finally, data analysis strategies can be applied to refine the models built and to judge the uniqueness of the solutions reached. Overall, such state-of-the art workflow can enable the building of

accurate 3D model structures, representing the best fit to the acquired experimental measurements.

Concluding, MS-based hybrid approaches combined with sophisticated modelling tools have emerged as a key strategy for investigating the structure and conformational states of native macromolecular complexes. It is expected that the application of such strategies will continue to increase, particularly because the size, complexity, and dynamics of target systems is increasing. A particularly exciting direction is the ability of this method to map the dynamic interactions between proteins and other biomolecules. This will require the continuing development of experimental and computational methods to capture the conformational heterogeneity of protein assemblies. For instance, extending the individual technologies and computational tools for targeting membrane-embedded proteins and their dynamic associations with lipids and substrates will make inroads in understanding complex functional states, underlying efforts for drug discovery and human health.

Acknowledgement

A.P. acknowledges funding from the Wellcome Trust (Grant number: 109854/Z/15/Z).

[1]Department of Chemistry, King's College London, London, UK
[2]Department of Chemistry, Imperial College London, London, UK

*Correspondence: argyris.politis@kcl.ac.uk (A. Politis).

http://dx.doi.org/10.1016/j.tibs.2016.04.008

References

1. Politis, A. *et al.* (2014) A mass spectrometry-based hybrid method for structural modelling of protein complexes. *Nat. Methods* 11, 403–406
2. Hernandez, H. and Robinson, C.V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protocols* 2, 715–726
3. Hall, Z. *et al.* (2012) Structural modeling of heteromeric protein complexes from disassembly pathways and ion mobility-mass spectrometry. *Structure* 20, 1596–1609
4. Rappsilber, J. (2011) The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* 173, 530–540
5. Leitner, A. *et al.* (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* 9, 1634–1649
6. Mendoza, V.L. and Vachet, R.W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* 28, 785–815
7. Konermann, L. *et al.* (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* 40, 1224–1234
8. Alber, F. *et al.* (2007) Determining the architectures of macromolecular assemblies. *Nature* 450, 683–694
9. Lasker, K. *et al.* (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1380–1387
10. Greber, B.J. *et al.* (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. *Science* 348, 303–308
11. Zhou, M. *et al.* (2014) ATP binding reduces conformational flexibility in a rotary ATPase – Evidence from ion mobility mass spectrometry. *Nat. Chem.* 6, 208–215
12. Shukla, A.K. *et al.* (2014) Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* 512, 218–222