International Journal of Mass Spectrometry 426 (2018) 29–37

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Full Length Article

Deconvolution of ion mobility mass spectrometry arrival time distributions using a genetic algorithm approach: Application to α_1-antitrypsin peptide binding

Ganesh N. Sivalingam[a], Adam Cryar[a], Mark A. Williams[b], Bibek Gooptu[c], Konstantinos Thalassinos[b,a,*]

[a] Institute of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London, WC1E 6BT, United Kingdom
[b] Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, University of London, London, WC1E 7HX, United Kingdom
[c] Leicester Institute of Structural and Cellular Biology and NIHR BRC-Respiratory, Henry Wellcome Building, Lancaster Road, Leicester, LE1 7RH, United Kingdom

ARTICLE INFO

Article history:
Received 2 August 2017
Received in revised form
19 December 2017
Accepted 5 January 2018
Available online 31 January 2018

Keywords:
Ion mobility mass spectrometry
Travelling wave ion mobility
Collision induced dissociation
Software' Data processing
Genetic algorithm

ABSTRACT

Ion mobility mass spectrometry (IM-MS) is a fast and sample-efficient method for analysing the gas phase conformation of proteins and protein complexes. Subjecting proteins to increased collision energies prior to ion mobility separation can directly probe their unfolding behaviour. Recent work in the field has utilised this approach to evaluate the effect of small ligand binding upon protein stability, and to screen compounds for drug discovery. Its general applicability for high-throughput screening will, however, depend upon new analytical methods to make the approach scalable. Here we describe a fully automated program, called Benthesikyme, for summarising the ion mobility results from such experiments. The program automatically creates collision induced unfolding (CIU) fingerprints and summary plots that capture the increase in collision cross section and the increase in conformational flexibility of proteins during unfolding. We also describe a program, based on a genetic algorithm, for the deconvolution of arrival time distributions from the CIU data. This multicomponent analysis method was developed to require as little user input as possible. Aside from the IM-MS data, the only input required is an estimate of the number of conformational families to be fitted to the data. In cases where the appropriate number of conformational families is unclear, the automated procedure means it is straightforward to repeat the analysis for several values and optimize the quality of the fit. We have employed our new methodology to study the effects of peptide binding to α_1-antitrypsin, an abundant human plasma protein whose misfolding exemplifies a group of conformational diseases termed the serpinopathies. Our analysis shows that interaction with the peptide stabilises the protein and reduces its conformational flexibility. The previously unresolved patterns of unfolding detected by the deconvolution algorithm will allow us to set up a fully automated screen for new ligand molecules with similar properties.

1. Introduction

1.1. Ion mobility and native conformation

Ion mobility mass spectrometry (IM-MS) is increasingly used in the field of structural biology to address a number of challenging questions relating to the structure and dynamics of proteins and protein complexes [1,2]. IM-MS offers several advantages compared to more established tools used in the field, namely the ability to study proteins of high flexibility [3–5] and polydispersity [6], the ability to monitor and separate distinct conformational families co-existing in solution [7,8] and all this whilst having modest sample requirements. IM-MS separates ions, typically generated from nano-electrospray ionization (nESI), based on their mass, charge and collision cross section (CCS). A large number of studies have shown that under carefully controlled experimental conditions and at low collision energies, the structures from IM-MS are closely related to those in solution for globular proteins and protein complexes [9–11].

* Corresponding author at: Institute of Structural and Molecular Biology, Division of Biosciences, University College London, Gower Street, London, WC1E 6BT, United Kingdom.
E-mail address: k.thalassinos@ucl.ac.uk (K. Thalassinos).

https://doi.org/10.1016/j.ijms.2018.01.008
1387-3806/

1.2. Gas phase unfolding of proteins

By intentionally increasing the internal energy of an ion prior to IM-MS analysis, however, a number of interesting observations can be carried out. In some of the early applications of IM-MS to study protein structure, the groups of Jarrold, Clemmer and Bowers used injection energy studies to probe the unfolding of a number of model and disease-related proteins [12–14]. These experiments revealed the relative stability of different proteins. Proteins held together with disulphide bonds were more resistant to unfolding compared to those without them. Injection energy studies have also been used to probe the oligomeric state for a number of proteins involved in protein aggregation diseases. Oligomers dissociate to lower aggregation states and eventually to monomer at increasing injection energies [15].

1.3. Unfolding of protein complexes

Collision induced dissociation (CID) followed by IM-MS analysis has also been used to study large protein complexes [10]. Unlike CID of peptides, which results in the fragmentation of the peptide bond, for non-covalent complexes the increase in energy results in unfolding, typically of the smallest surface accessible subunit and so the term collision induced unfolding (CIU) is used in this case. This unfolding is followed by an asymmetric dissociation of the complex with many charges from the precursor being transferred to the smallest subunit, which is now able to accommodate more charges as it has a larger surface area than prior to its dissociation from the complex [16]. More recent work, however, has shown that the above process depends on the tertiary structure of the protein complex [17] and on the charge state selected for CIU [18]. Since its initial description, CIU has been used to study a number of different processes, including gas phase protein stability due to anion and cation binding [19,20], peptide binding [21], and lipid binding selectivity [22].

1.4. Protein ligand binding and drug screening

Another major application area of the CIU-IM-MS approach is examining the effect of ligand binding on the protein or protein complex structure. Comparison of the arrival time distributions obtained at increased collision energies between apo- and ligand-bound forms, for a number of proteins [23] and protein complexes, enabled characterisation of the changes in structural stability induced by ligand binding. The influence of different ligand-bound states has also been examined; illustrating the power of this approach in studying heterogeneous systems having multiple ligand binding stoichiometries [24,25].

Ligand-specific CIU signatures may report other consequences of ligand binding. This has been demonstrated for small molecule inhibitors of the protein kinase domain of the Abl protein. These are defined by their binding specificity as type I (non-specific for the active conformer of the Abl kinase domain) or type II (active state specific) inhibitors. Although binding of either type of ligand to the protein did not substantially alter the conformation at low collision energies, they displayed distinct unfolding signatures. These 'reference' type I and type II signatures were then compared to the CIU signatures of novel ligands to predict their inhibitory mechanism [26].

1.5. The need for computational tools

Consequently, the unfolding of proteins and protein complexes in the gas phase followed by IM-MS analysis provides valuable information that cannot be easily obtained by other structural methods. The development of commercial IM-MS instrumentation [27] has also led to an expansion in the number of laboratories that have access to this technology. As a consequence of an increased IM-MS user base, and the development of new types of data acquisition methods such as CIU, new computational methods to process these data are required. Programs to process the IM-MS data [28,29], and CIU data [30,31] have been recently developed. It is particularly desirable to report data from IM-MS studies in a standardised, quantitative manner to allow them to be compared. This requires the definition of parameters that reduce the dimensionality of IM-MS data whilst retaining important information about the ensemble behaviours that IM-MS is particularly well suited to characterise.

Here we describe a new computational approach for the analysis of CIU experiments. We present the benefit of summarising the data using the intensity weighted mean (IWM_{ATD}) and standard deviation ($IWSD_{ATD}$) of each arrival time distribution (ATD). The results are a quantitative method for representing the degree of unfolding in terms of change in CCS. We also describe a new genetic-based algorithm for the deconvolution and subsequent multicomponent analysis of CIU data, capable of dealing with the conformational heterogeneity of unfolding proteins. Deconvolution allows accurate determination of the centre of conformational populations. It also facilitates more accurate calculation of the abundance of each resolvable conformational family in comparison to existing methodology, which uses peak heights. To demonstrate the efficacy of these methods, three proteins that have been previously well-characterised by native MS methods have been analysed; myoglobin, lysozyme and β-lactoglobulin, as well as the protein α_1-antitrypsin in both its apo-form and bound to a peptide previously shown to inhibit its aggregation.

Benthesikyme, our software for summarizing the data is available from our website (http://www.homepages.ucl.ac.uk/~ucbtkth/resources.html) while the genetic algorithm program is available on request.

2. Materials and methods

2.1. Mass spectrometry sample preparation

Proteins were analysed in 200 mM ammonium acetate (Sigma Aldrich, St. Louis, MO). Buffer exchange was carried out using BioRad (Hercules, CA) BioSpin 6 columns, with additional concentration and dilution steps using Amicon Ultra 0.5 ml centrifugal filter devices (Millipore UK Ltd, Watford, UK). The concentration of these samples was monitored using a Qubit 2.0 fluorometer (Life Technologies, Carlsbaαd, CA). Samples were analysed after diluting to a final concentration of 10 μM.

2.2. Peptide binding to α_1-antitrypsin

α_1-antitrypsin was buffer exchanged into 250 mM ammonium acetate (pH 7) by 3 rounds of dilution-concentration using 10 kDa Millipore Amicron Ultra centrifuge filters. The Ac-TTAI-NH2 peptide was also made up in ammonium acetate (pH 7). A molar ratio of 2:1 peptide:protein was used and the sample incubated for 72 h at 37.5 °C before analysis. The final concentration of α_1-antitrypsin was 15 μM.

2.3. IM-MS analyses

Experiments were performed on a Synapt HDMS mass spectrometer (Waters Corp., Manchester, UK) [27]. The Synapt geometry has two collision cells bracketing the ion mobility device: the trap, which is situated before the IM device, and the transfer which is located after the IM device. To carry out CIU experiments, it is the voltage in the trap cell that is increased. The instrument