International Journal of Mass Spectrometry 345–347 (2013) 54–62

Contents lists available at SciVerse ScienceDirect

International Journal of Mass Spectrometry

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

Amphitrite: A program for processing travelling wave ion mobility mass spectrometry data

Ganesh N. Sivalingam[a], Jun Yan[a], Harpal Sahota[b], Konstantinos Thalassinos[a,*]

[a] Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, UK
[b] Institute of Structural and Molecular Biology, Crystallography, Birkbeck College, London, UK

ARTICLE INFO

Article history:
Received 3 June 2012
Received in revised form 4 September 2012
Accepted 13 September 2012
Available online 6 October 2012

Keywords:
Ion mobility mass spectrometry
Travelling wave ion mobility
Software
Data processing

ABSTRACT

Since the introduction of travelling wave (T-Wave)-based ion mobility in 2007 a large number of research laboratories have embraced the technique, particularly those working in the field of structural biology. The development of software to process the data generated from this technique, however, has been limited. We present a novel software package that enables the processing of T-Wave ion mobility data. The program can deconvolute components in a mass spectrum and uses this information to extract corresponding arrival time distributions (ATDs) with minimal user intervention. It can also be used to automatically create a collision cross section (CCS) calibration and apply this to subsequent files of interest. A number of applications of the software, and how it enhances the information content extracted from the raw data, are illustrated using model proteins.

1. Introduction

Ion mobility is a gas-phase technique that separates ions as they travel through a counter flowing neutral target gas under the influence of an applied electric field. The time it takes an ion to traverse the cell is related to its mass, charge, and the rotationally averaged collision cross section (CCS) of an ion [1–3]. Ion mobility coupled to mass spectrometry (IM-MS) is a powerful analytical technique that was initially only available in a few laboratories capable of building such specialised instruments. The primary means of performing IM-MS separations was based on drift cell technology [4].

Shortly after the description of a commercial instrument that was modified for IM-MS measurements [5], the introduction of travelling wave (T-Wave) ion mobility separation [6], incorporated in a commercial quadrupole time-of-flight instrument (Synapt HDMS, Waters Corp.) [7], further popularised the technique. In addition to the high mass accuracy obtainable, the Synapt can be used to carry out ion mobility-tandem mass spectrometry experiments by performing collision induced dissociation (CID) before and/or after the mobility cell. A second generation instrument, the Synapt G2, was introduced in 2011 with an up to four-fold increase in the T-Wave ion mobility resolution, as expressed in terms of $\Omega/\Delta\Omega$ [8], where Ω is the rotationally averaged CCS. Another attractive feature of the Synapt instruments is that they can be modified for high mass operation by the incorporation of a 32 K quadrupole, allowing the selection and transmission of high m/z species.

T-Wave ion mobility mass spectrometry (TWIM-MS) has so far been used to study a number of synthetic and biological molecules such as polymers [9–11], carbohydrates [12], peptides [13,14] and lipids [15,16]. The majority of applications, however, have been within the structural biology field as TWIMS-MS has clear advantages over other established techniques within this area. Proteins that exhibit too much conformational flexibility or that are too large to study by established techniques such as X-ray and NMR respectively, can still be amenable to analysis by means of TWIM-MS. In addition, TWIM-MS can be used to separate and study co-existing populations present in solution [17] in contrast to the majority of other biophysical techniques that can only provide information regarding the population average. TWIM-MS has been used to probe the conformation of soluble proteins and proteins bound to various ligands [18–22], protein complexes [23–26], proteins involved in misfolding and aggregation [18,27,28] intact viruses, [29,30], and membrane proteins [31,32]. In conjunction with CID, TWIM-MS has also been used to probe the structural stability of such molecules [20,33,34].

For a large number of the applications mentioned, there is no requirement to convert arrival time distributions (ATDs) to CCS in order to answer the biological question studied. Obtaining a CCS, however, is essential in cases where the CCSs are used as a way of filtering computer generated models [35–37]. Classical IM-MS instrumentation uses a drift cell mobility separation device. While the physical principles behind drift cell IM-MS are well understood and can be used to obtain a CCS for each ion studied [38], the same

* Corresponding author. Tel.: +44 20 7679 2197; fax: +44 20 7679 7193.
E-mail address: k.thalassinos@ucl.ac.uk (K. Thalassinos).

1387-3806/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijms.2012.09.005

is not true for the T-Wave-based device. Despite initial attempts to characterise the T-Wave device, [39] the ion motion in the device is still not fully understood and cannot be used to derive CCSs directly from the arrival time (t_d) data, however, a number of protocols have been developed which allow the calibration of the T-Wave against standards of known cross section [13,22,25,40]. A number of such standards are available in the form of peptides [41], proteins [41], protein complexes [42,43] and drug-like molecules [44].

Despite the advances in both TWIM-MS instrumentation development and the growing applications, advances in the software to process such data has been limited. The only software currently available is Driftscope (Waters Corp.) which involves extensive manual user interaction. A user has to identify the peaks in the mass spectrum, which can be challenging especially when dealing with spectra containing more than one components such as heterogeneous protein complexes, then use these to reconstruct the corresponding ATD. From this distribution the drift time(s) of maximum intensity are extracted for further analysis. This manual intervention can be labour intensive and can also introduce errors in the analysis. While programs to process intact protein and protein complex MS data [45–48] are appearing in the literature, there is still no program for the automatic processing of TWIM-MS data.

In this work we present a novel software for the processing of TWIMS-MS data. The software automates the deconvolution of the MS data and automatically extracts ATDs from the raw data files. It also allows for the facile creation of a calibration that can then be applied to entire data sets automatically. The software can be used to create CCS vs. *m*/*z* heat maps that can be overlaid between different experimental conditions, something that allows for a more in-depth probing of the structural changes taking place between different conditions. Having a program do these analyses allows for the standardisation of the data processing, making the entire process more objective and reproducible between different practitioners. A number of different uses of the program, with a particular focus, on commonly encountered structural biology applications are illustrated using model proteins.

2. Materials and methods

2.1. Sample sources and preparation procedures

cytochrome *c* from equine heart, myoglobin from equine heart, alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae*, bovine serum albumin (BSA), and concanavalin A from *Canavalia ensiformis* were purchased from Sigma Aldrich (St. Louis, MO). Serum amyloid P component (SAP) from human serum was purchased from CalBioChem, Merck Biosciences Ltd. (Darmstadt, Germany). For native experiments, protein samples were buffer exchanged into 250 mM ammonium acetate, and concentrated to 20 μM using Amicon Ultra 0.5 ml centrifugal filters (Millipore UK Ltd, Watford, UK). For denaturing experiments, protein samples were buffer exchanged into a 49:49:2 (v:v:v) ratio of H_2O: methanol: acetic acid, and concentrated to 20 μM using Amicon Ultra 0.5 ml centrifugal filters.

2.2. TWIMS-MS

Mass spectrometry experiments were carried out on a first generation Synapt HDMS (Waters Corp., Manchester, UK) mass spectrometer [7]. The instrument was mass calibrated using a 33 μM solution of Cesium Iodide in 250 mM ammonium acetate. 2.5 μl aliquots of samples were delivered to the mass spectrometer by means of nanoESI using gold-coated capillaries, prepared in house [49]. Typical instrumental parameters were as follows unless otherwise specified: source pressure 5.5 mbar, capillary voltage 1.10 kV, cone voltage 40 V, trap energy 8 V, transfer energy 6 V, bias voltage 15 V. IMS pressure 5.18×10^{-1} mbar, IMS wave velocity 250 m/s, IMS wave height 6 V, and trap pressure 4.07×10^{-2} mbar.

2.3. Experimental procedures

cytochrome *c* was analysed with a bias voltage of 30 V, with denatured myoglobin being used as a calibrant for obtaining CCS. For the heating experiment ADH was incubated at 60 °C for 30 min in a heat block. The sample was removed from the heat block and immediately introduced to the mass spectrometer. Instrumental parameters were optimised as follows: source pressure 4.50 mbar, cone voltage 60 V, trap energy 15 V, transfer energy 12 V, and IMS wave height 7 V. BSA and concanavalin A were used as CCS calibrants. For the collision unfolding experiment the native fold of cytochrome *c* was disrupted by increasing the bias voltage by 10 V at a time, from 10 V, until reaching 80 V. Instrumental parameters were optimised as follows: source pressure 3.55 mbar, cone voltage 30 V, and IMS wave height 7 V. Denatured myoglobin and ADH were used as CCS calibrants. For the mixing experiment, ADH, BSA, and concanavalin A were mixed in an equimolar ratio, and the instrumental parameters were optimised as follows: trap energy 60 V, transfer energy 30 V, bias voltage 22 V. IMS wave height 7 V. SAP, BSA, and concanavalin A were used as CCS calibrants.

2.4. Software development

During a TWIM-MS experiment ion arrival time distributions (ATDs) are recorded by synchronizing the oa-TOF acquisition with the gated release of a packet of ions from the trap T-Wave. For each packet of ions 200 mass spectra are acquired at a rate dependent on the pusher frequency.

Amphitrite handles the data in the form of a $n \times 200$ matrix (where n is the number of *m*/*z* bin increments), with individual vectors to describe the associated axes. This matrix can be used to generate the full mass spectrum of all arrival times by summing down to a n-length vector. Additional manipulations can be carried out by selecting sections of the matrix by index, for example the arrival times of a particular ion could be extracted by supplying the lower and upper *m*/*z* limits, and then summing along the *m*/*z* axis. The manipulations of this matrix forms the basis of the functionality of the program.

The software was developed using the Python programming language [50]. Several Python modules were utilised for data analysis NumPy, SciPy [51] and Matplotlib [52], and the graphical user interface was developed using wxPython [53]. The initial conversion of a raw TWIM-MS file to an Amphitrite project file can only be run under Microsoft Windows, however, all other aspects of Amphitrite are cross platform compatible and installer binaries for Linux and Mac OS X systems are available on the website http://www.homepages.ucl.ac.uk/~ucbtkth/amphitrite.html. The software was developed on a 3.4 GHz quad-core processor machine with 16 GB memory running Ubuntu 12.04. Processing times quoted are for a 2011 MacBook Air with a 1.7 GHz dual-core processor and 4 GB memory.

3. Results and discussion

Until now, Driftscope (Waters Corp.) has been the sole program used to display and manipulate raw TWIM-MS data. The introduction of Amphitrite facilitates increased customisability of plots as well as the automation of previously labour-intensive, subjective and hence non-reproducible tasks. Using standard proteins we describe various examples of how the program can be used.