

Ion mobility–mass spectrometry analysis of large protein complexes

Brandon T Ruotolo, Justin L P Benesch, Alan M Sandercock, Suk-Joon Hyung & Carol V Robinson

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK. Correspondence should be addressed to B.T.R. (btr23@cam.ac.uk) or C.V.R. (cvr24@cam.ac.uk).

Published online 19 June 2008; doi:10.1038/nprot.2008.78

Here we describe a detailed protocol for both data collection and interpretation with respect to ion mobility–mass spectrometry analysis of large protein assemblies. Ion mobility is a technique that can separate gaseous ions based on their size and shape. Specifically, within this protocol, we cover general approaches to data interpretation, methods of predicting whether specific model structures for a given protein assembly can be separated by ion mobility, and generalized strategies for data normalization and modeling. The protocol also covers basic instrument settings and best practices for both observation and detection of large noncovalent protein complexes by ion mobility–mass spectrometry.

INTRODUCTION

Large-scale interaction maps suggest a complex interplay of proteins within a myriad of functional assemblies^{1,2}. A critical step in assigning functions to these assemblies is to determine their structure³. This goal is challenging, as many of these assemblies exist in low-copy numbers within cells, are frequently heterogeneous and may interact only transiently. Consequently, structural information for many protein complexes is not readily accessible by using the classical tools of structural biology (e.g., X-ray crystallography, nuclear magnetic resonance spectroscopy). New approaches are being developed that involve integrating data from a number of lower-resolution experimental methods and by combining distance and interaction restraints from these methods with homology modeling, architectural or even atomic models are being generated⁴. These restraints can be derived from a variety of experimental measurements including MS of intact complexes, chemical cross-linking, fluorescence resonance energy transfer, small angle X-ray scattering, and analytical ultracentrifugation^{5,6}. One very recent addition to this series of biophysical tools is ion mobility separation coupled to mass spectrometry (IM-MS). IM is an established technique for studying shape and conformation in small molecules and individual proteins in the gas phase^{7–10} but has only recently been applied to intact protein complexes^{11,12}. When IM is coupled with MS, mass and consequently subunit composition can be determined simultaneously with the overall topology of protein complexes^{10,12,13}.

IM-MS analysis is performed by first ionizing the protein complex of interest. In our experiments, nano-electrospray ionization is used, typically requiring careful preparation procedures for most protein complexes. These procedures, as well as general practical aspects of sample preparation, are detailed in a protocol by Hernández and Robinson¹⁴. Although they are not discussed in detail here, knowledge of the materials and protocol steps described in that work are critical to the success of the protocol described below.

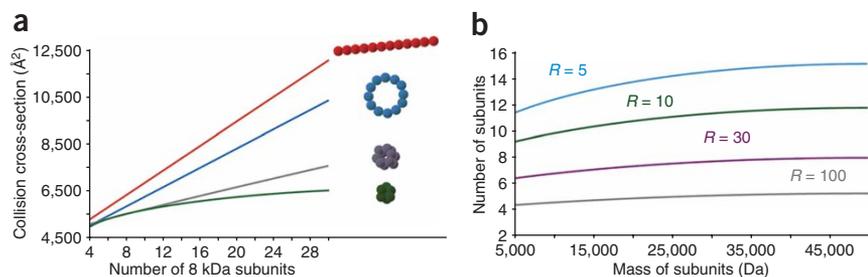
After ionization, ions are injected into a region containing neutral gas at a controlled pressure (e.g., 0.5 mBar of nitrogen gas). Under the influence of a relatively weak electric field, injected ions undergo IM separation^{7–9,15–17}. Large ions experience more collisions with neutrals and thus take more time (drift time or t_D)

to traverse the chamber than smaller ions. Under these conditions, ions will migrate through the neutrals and separate according to their ion-neutral collision cross-section (Ω). Additionally, ions having higher charge will experience greater separation field strengths and traverse the chamber more quickly. Consequently, ion mobility is often described as proportional to collision cross-section-to-charge ratio (Ω/z). After separation, ions are sampled by a mass spectrometer and analyzed according to their mass-to-charge (m/z) ratio. We have based the specific aspects of our protocol on our experience with the Synapt (Quadrupole-Ion Mobility-Time-of-Flight) HDMS instrument (Waters) as it is the only commercial IM-MS instrument currently available. However, we also describe several practical steps that are necessary for recording IM data of large protein complexes as well as methods for modeling and extracting quaternary structure information that apply generally to any IM-MS instrument or data set. We begin by posing a series of questions commonly encountered in this field of research.

Will separation of the most likely potential structures of a protein complex be possible?

Many protein complexes fall into structural classes that can be described with purely geometric modifiers (e.g., ring or spherical). For these cases, it is often helpful to evaluate the ability of IM to separate different protein structural archetypes before attempting more detailed data analysis (discussed below). For example, **Figure 1a** shows modeling data for four different protein complex topologies as a function of subunit mass and number. To generate these data, we used open source software for calculating the collision cross-section of trial structures (*MOBCAL*)^{18,19}, and altered the code so that spherical representations of whole subunits could be used in place of all-atom models (see protocol). It is worth noting that the collision cross-sectional dependence of similar structural trends has been discussed in the atomic cluster literature previously^{20–22}. The plot shows trend lines describing the cross-sectional increase predicted for the addition of subunits added to four different structural types (linear, ring, stacked double-ring and close-packed) composed of small (8 kDa) subunits. Although closely similar for small subunit numbers, the four trends diverge substantially at large subunit numbers, indicating that the

Figure 1 | Is separation of likely potential structures possible? (a) Simulated trends in collision cross-section for four different topologies as a function of the number of subunits. As the number of subunits increases in an assembly, structural families diverge, making ion mobility separation more facile at a fixed value for resolution. (b) Plot of subunit mass versus number of subunits comprising a protein assembly. The trends shown represent those subunit mass/number of subunit combinations where ring and collapsed structures are resolved (at half height, $R = t/\Delta t$). The numbers indicated correspond to the IM resolution necessary to resolve ring and collapsed structures along the corresponding trend. Above these lines the structures are also resolved at the indicated value of R .



structural families attributed to complexes composed of a larger number of subunits are more easily distinguished by IM separation.

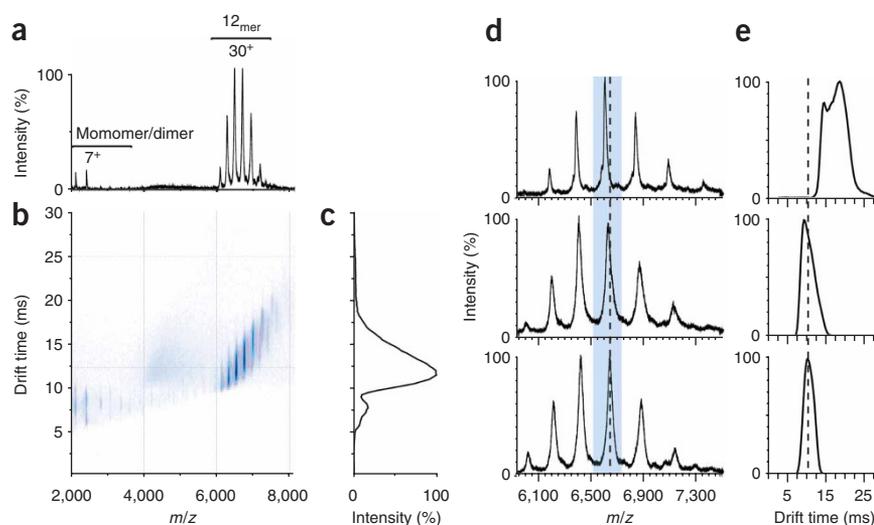
The separation power of an IM device is often characterized by resolution (R). Here, we define resolution as the centroid of the drift time distribution divided by the width of the distribution at half height ($t/\Delta t$). The higher the value of R achieved, the smaller the Ω differences that can be distinguished by IM separation. In addition to the topology and number of subunits, the mass of individual subunits will also influence the resolution necessary to separate different topologies. More massive subunits will, on average, increase the overall dimensions of the final quaternary structure adopted by the complex. **Figure 1b** plots subunit mass against subunit number and indicates several trend-lines that refer to R values that are sufficient to separate (at half height) a ring structure from a compact collapsed assembly at a given subunit number/mass combination. For example, at an R value of 10, only those structures containing a large number of small subunits are separable, and the chief effect of increasing IM resolution is the ability to separate ring/collapsed structures comprised of fewer subunits of similar size. Note that the influence of subunit mass on

the resolution required to separate protein complex topologies is more significant at lower values of R .

What are the general features of IM-MS data for protein complexes?

IM-MS data are inherently three-dimensional, consisting of mass, drift time and intensity (relative abundance) data for all the ions observed^{23–26}. A typical data set is illustrated in **Figure 2**, which shows a mass spectrum (a), an IM-MS contour plot (b) and a total integrated IM spectrum (c) for a series of ions corresponding to the small heat shock protein from wheat, TaHSP16.9. Signals corresponding to monomers, dimers and dodecamers are observed; however, the dodecamer is the dominant species in both the spectrum and in solution^{27–29}. Often, the interplay between IM and MS data is vital for the complete analysis and utility of both dimensions of information. For example, MS data are used to define the charge and composition of the ions observed, without which it is often difficult to both normalize and interpret IM data correctly (see protocol).

Figure 2 | Multidimensional IM-MS data representation and consequences of using too much acceleration voltage before IM separation. (a) Mass spectrum compiled from all ions observed. (b) Plot of drift time versus m/z for the small heat shock protein complex (sHSP) TaHSP16.9 (aqueous solution, 200 mM ammonium acetate buffer, 2 μ M protein complex), illustrating the multidimensional nature of the data produced by ion mobility–mass spectrometry. Ions within the spectrum can be assigned to charge states of the TaHSP16.9 dodecamer (202.8 kDa) ranging from 28⁺ to 33⁺ (data acquired on a Waters Synapt HDMS quadrupole–ion mobility–time-of-flight instrument; see **Table 2** for typical instrument settings). (c) Ion mobility arrival time distribution for all ions observed. An exponential intensity gradient was used to generate the contour plot in b. (d) A series of three mass spectra acquired at increasing values of activating voltage (in the cone and ion trap regions before the ion mobility separation device) for the TaHSP16.9 dodecamer (202.8 kDa). The blue shaded region highlights the peak corresponding to the 31⁺ charge state at low activation voltage (lower spectrum). As activation voltage is increased, the peaks comprising the mass spectrum decrease both in width and mass and, in the process, more closely relate to the sequence mass of the protein complex. The mass increase recorded for the ions (mass in excess of the expected sequence) was highest for low activation conditions (0.74%, lower spectrum) and lowest for high activation conditions (0.1%, upper spectrum). (e) Ion mobility arrival time distributions for the 30⁺ charge state of TaHSP16.9 from each of the activation voltages are shown on the right. Low activation voltages are required to observe compact states of the 12mer, and the ion increases in size dramatically at higher voltages, indicative of unfolding. As such, optimal conditions for both mass and mobility measurements for protein assemblies are often incompatible.



Generally, plots of drift time versus m/z , as shown in **Figure 2b**, reveal trends in the data that can be indicative of either the charge state or molecular class of the ions observed^{23–26}. For most protein complex ions, we have observed that the signals for a charge state series corresponding to a monodisperse protein assembly display a good correlation ($R^2 > 0.99$) to a linear relationship between drift time and m/z allowing for polydisperse samples to be identified readily³⁰. In addition, plotting data in a format similar to **Figure 2b** often highlights the presence of post-mobility cell fragmentation^{31,32}. If protein complexes are separated by ion mobility as intact assemblies and are activated before mass measurement, any fragmentation products generated will appear at the same drift time as the original parent ion. If analyzing samples containing more than one protein complex, or stoichiometry of a single protein, post-mobility cell fragmentation can lead to erroneous assignment of fragment ions as alternative protein topologies¹³. Therefore, three-dimensional representations of IM-MS data are of critical importance for the initial stages of interpretation.

How are instrument conditions balanced for optimal IM separation and mass measurement?

Previous reports have detailed the conditions necessary to observe intact protein assemblies by MS^{14,33–37}. These conditions often include increased pressures in the source region of the instrument and carefully reducing the amount of ‘collisional heating’ (also termed, ‘activation’) experienced by protein complex ions to avoid dissociation³⁸. However, as discussed previously, some activation is usually employed to desolvate the protein complex ions and achieve optimum mass accuracy^{39,40}. **Figure 2d** shows a series of mass spectra for TaHSP16.9 acquired using increasing accelerating voltage into the storage ion guide located before the IM separator. The mass accuracy achieved for the protein complex, as indicated in the figure, increases under more activating conditions as expected (see protocol). The corresponding IM data recorded for the 30⁺ charge state of TaHSP16.9 are shown in **Figure 2e**. As activation energy is increased beyond a threshold value, the drift time observed for the intact protein complex increases. We attribute this increase in drift time to the generation of multiple unfolded states of the protein complex as the internal energy of the protein complex is altered by collisional activation¹³. These unfolded states are unlikely to correspond to any accessible solution-state structures, as gas-phase activation occurs in an environment where solvation

and Coulombic forces have vastly different magnitudes compared to solution⁴¹. Therefore, to obtain drift time data consistent with solution-phase structures, careful control of the voltages used to accelerate ions before IM separation is required. In addition, to achieve both high mass accuracy and IM data consistent with solution-phase structure, data must be acquired over a range of acceleration voltages rather than a single optimized set of parameters (see protocol).

How is drift time data converted to collision cross-section?

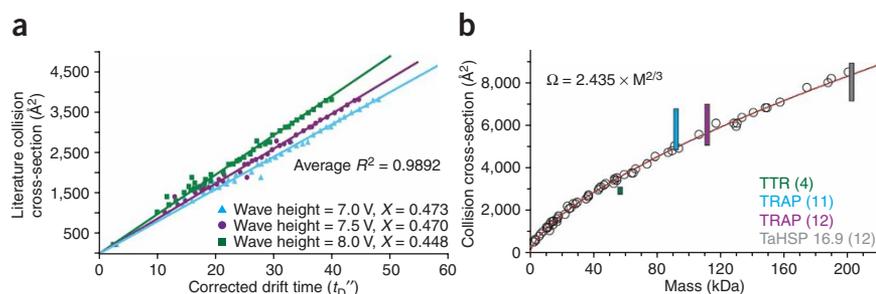
With careful measurement of pressure (P , in torr) and temperature (T , in Kelvin), the following equation can be used to convert drift times (t_D , in seconds and corrected for time spent outside the drift cell) to collision cross-section values (Ω output in m²) using a standard drift tube separator (constant electric field):

$$\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_b T)^{1/2}} \left[\frac{1}{m_1} + \frac{1}{m_N} \right]^{1/2} t_D E \frac{760}{L} \frac{T}{273.2} \frac{1}{N} \quad (1)$$

where k_b is the Boltzmann constant, z is the ion charge, e (C) is the elementary charge, m_1 is the mass of the ion, m_N is the mass of the neutral gas (both in kg), E is the electric field strength (V/m), L is the length of the drift region (m) and N is the neutral gas number density (m^{-3}). Several outstanding reviews have described detailed procedures for making accurate cross-section measurements using standard drift tubes^{8,9,17}. In cases where gas purity, pressure and temperature cannot be measured accurately, calibrating the drift time measurements using ions of known collision cross-sections is preferable⁴². This method of calibration, rather than absolute measurement, is also preferred for travelling wave-type IM separators that use time-varying electric fields within the drift region to propel the ions toward the detector^{43,44}. During the course of a measurement, the instrument operator can vary both the magnitude and velocity of voltage ‘waves’ to optimize IM separation. Over the course of analyzing many ions with known collision cross-sections, it was found that Ω was proportional to t_D^X , where X is an empirically determined parameter that depends upon many variables, including the height and velocity of the voltage ‘waves’ used to propel ions through the IM separation region⁴⁵.

In our laboratory, we use the procedure described below (see protocol) to calibrate travelling wave drift times to collision cross-sections^{45,46}. The output of this procedure is illustrated in **Figure 3a**, where calibration curves are shown at three different

Figure 3 | Plots used for calibrating IM drift times to collision cross-sections. **(a)** Calibration curves, combining data from bradykinin (human), ubiquitin (bovine), cytochrome *c* (equine) and myoglobin (equine), displayed as linear plots of collision cross-section and corrected drift time (units of milliseconds raised to the power of X). Plots are shown for three magnitudes of the wave that propels the ions through the travelling wave drift cell used in the device (7 V: blue; 7.5 V: violet; 8 V: green). The average correlation coefficient for all three fits is displayed as well as the exponential factor (X) determined as the best fit for each wave height. The units of the dependent axis are determined in Step 15 of the protocol. **(b)** Plot of literature values for collision cross-section from references 45 and 46 against molecular mass. The plot can be used as a means of validating a drift time calibration to larger values of collision cross-section by predicting the value for a roughly spherical ion of the appropriate molecular weight. The best-fit relationship is displayed on the graph and indicates an average density for proteins of $\sim 0.48 \text{ Da } \text{\AA}^{-3}$ in the gas phase. Blue, violet, gray and green bars indicate the range of cross-sections recorded for all the charge states observed in the mass spectrum of TRAP 11mer, TRAP 12mer, TaHSP16.9 and TTR, respectively.



wave-heights (magnitude of the voltage ‘wave’). The slope of the resulting calibration curve, and the exponential factor X , depends upon wave ‘height’, as shown in **Figure 3a** (see protocol step 15 for detailed description of corrected drift time). Typically, we run experiments over a range of wave heights to rule out the influence of electric field (i.e., dipole alignment) on the separation. Currently, there are few potential calibrant ions that have reported collision cross-sections in excess of 3,500 Å². Therefore, we usually calibrate the data with a mixture of protein ions (including equine myoglobin, equine cytochrome *c* and human ubiquitin) that define the relationship between collision cross-section and drift time for smaller ions and extrapolate this relationship to larger values. The dependence of X on the wave height or wave velocity is an active area of research. The protocol given here provides the best empirical fit to the data as well as collision cross-section measurements that are in close agreement with literature values. In general, to perform higher-confidence extrapolations to large collision cross-section values, a greater number of points are preferred to define the calibration relationship.

An extensive series of charge-reduced and low-charge state collision cross-sections extracted from the literature for a large data set of proteins, including the three discussed above, are shown in **Figure 3b**^{47,48}. The fit shown is based on the relationship between the projected area (collision cross-section) and density-dependant mass of spherical-type protein complexes having a nearly constant packing density ($\rho = 0.48 \text{ \AA}^3/\text{Da}$). We typically use the relationship shown in **Figure 3b** as a first approximation of the near-spherical collision cross-section of collapsed protein complexes and also to validate calibration curves for large molecular weight species (see protocol for details of validation). Although generally useful, it is important to note that the relationship shown in **Figure 3b** is used only as a rough guide to the collision cross-section of collapsed protein structures and supplements the molecular modeling approaches described below. Protein complexes measured by our group are plotted together with the literature data shown in **Figure 3b** to indicate the agreement of the calibrated data with previous results and trends. Our data are plotted as boxes to indicate the overall spread of collision cross-sections recorded from all charge states generated by nano-electrospray ionization¹².

A wide range of densities have been used in the literature to describe the packing efficiencies of gas-phase proteins and protein complexes^{38,49,50} and the density used here is toward the lower extreme of this range. As such, it is not surprising that some complexes exhibit collision cross-sections that appear substantially more compact than predicted by the general relationship shown in **Figure 3b**. For example, the collision cross-sections recorded for the higher charge states of TaHSP16.9 extend as low as 7,200 Å², 85% of the collision cross-section by the relationship shown in **Figure 3b**. At the other extreme, larger than expected (>130% of the predicted value) collision cross-sections for the tryptophan–RNA-binding attenuation protein (TRAP) 11mer have been interpreted as evidence for a ring-type topology for the gas-phase protein assembly¹².

How is molecular modeling employed to analyze data?

A critical component of IM-MS data analysis relies upon attempting to fit computational models to the collision cross-section and mass information gathered from experiment. For small ions, a large number of computational tools can be employed to determine the

structure or ensemble of structures that exist at a given internal temperature. For example, the gas-phase structures of small peptides are typically determined by using a simulated-annealing approach to exhaustively search the configuration space of the molecule, combined with energy calculations to determine the lowest-energy configuration. Energy determinations range in precision from force field–based relative energetics, for large systems, to density functional theory or *ab initio* approaches for smaller systems. Many varieties on this basic theme have appeared in the literature^{7–9,17,42,51–59}.

Owing to their size, the analysis of multiprotein complexes requires a paradigm shift from the methods described above. The computational approaches currently used in our laboratory to assign structure to gas-phase multiprotein complex ions are summarized in **Figure 4**. Although only three tracks are shown, a continuum of structural models and approaches should be considered for modeling a protein assembly, beginning with sphere-type coarse grain models and ending with atomistic models. For our measurements, performed at an IM resolution of ~ 10 , atomistic models tend to be less useful than those that have some level of coarse-graining, as these models are more commensurate with the information content of the measurement and allow more thorough topology searching. Atomic-level structures for subunits, subcomplexes or intact assemblies are, however, useful as templates for generating coarse grain structures (see protocol) and provide an opportunity for more focused folding and docking calculations in the context of the intact complex. This approach has proved useful for assessing the extent of unfolding experienced by activated protein complexes in the gas phase (similar to **Fig. 2d,e**)¹³.

What are the limitations of IM-MS for protein complex analysis?

The calibration procedure described below imposes several limitations on the IM-MS technique for protein complexes. Although we have been able to achieve accurate results for complexes approaching 1 MDa through careful calibration and replicate measurements, the ability of the calibration protocol to produce high-precision collision cross-section measurements for very large protein complexes (> 500 kDa) is limited by the current pool of calibrant ions available (see **Fig. 3** and Step 12) and their associated precision (see Step 20). Measured collision cross-sections larger than 3,500 Å² would allow more precise calibrated measurements and pave the way for future IM-MS studies of large protein complexes.

Some of the most important limitations of IM-MS technology are derived from the ionization event used to generate protein complex ions. Currently, using nano-electrospray, it is extremely difficult to generate ions that correspond to hydrophobic membrane-bound protein assemblies. There have been some recent reports, however, that provide some indication that careful control of solubilizing molecules in solution and IM-MS instrument parameters may allow membrane protein complexes to be studied more widely in the future⁶⁰. Clear and assignable MS information is critical for the interpretation of IM data and these and other limitations are already discussed in a previous protocol¹⁴. Given interpretable MS data, the structural information provided by IM-MS is limited by the IM resolution (R) achieved for the complex of interest. Currently, it is difficult to find examples of IM resolution >10–15 for large protein complexes^{13,50}; however, much higher values have been reported for smaller molecules^{61,62}.

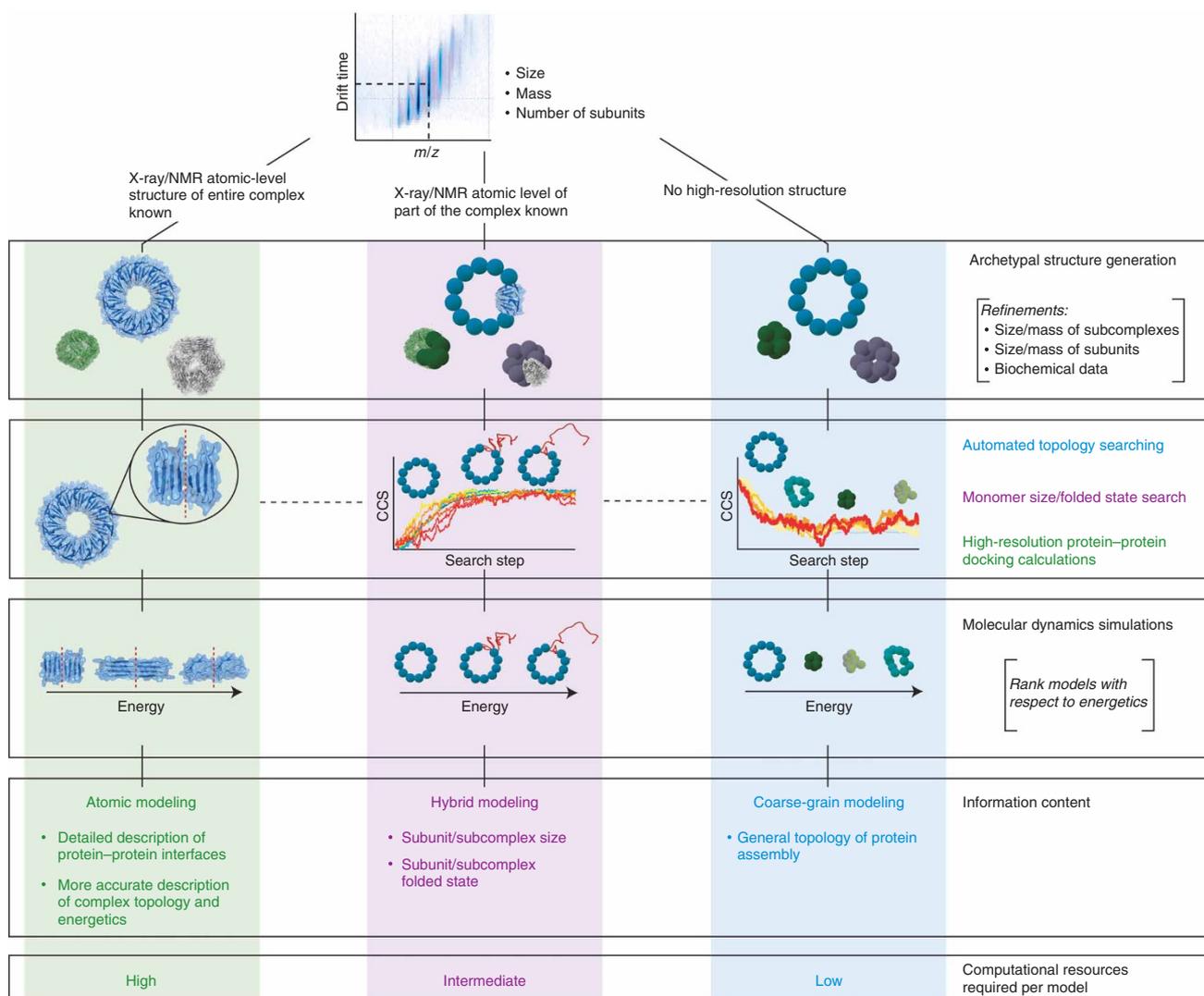


Figure 4 | Decision tree describing how model structures for protein complexes can be generated and compared with collision cross-section measurements. After measuring the minimal information required to begin the modeling process (shown top), at least three tracks are available for modeling information. The choice between these tracks depends on the amount of high-resolution structural information already available for the protein complex, computational resources available and the information required from the simulation as output (i.e., focusing on the folded state of a monomer within the complex rather than the complex as a whole). A continuum of tracks actually exists between the three shown here. For example, higher resolution coarse-graining can be achieved than the spheroid representations shown in the track highlighted in blue.

Increasing the maximum IM resolution achievable for proteins and protein complexes is an active area of research. We envision that the information provided by IM-MS could be integrated as restraints,

along with other low-resolution structural information, to provide higher-fidelity structure representations of protein assemblies, including those comprised of more than one type of subunit^{4,5}.

MATERIALS

REAGENTS

- Equine myoglobin (from horse heart) (Sigma, cat. no. M1882; 1 g, pdb: 1wla)
- Equine cytochrome *c* (from horse heart) (Sigma, cat. no. C-2506; 1 g, pdb: 1hrc)
- Bovine ubiquitin (from red blood cells) (Sigma, cat. no. U6253; 25 mg, pdb: 1ubq (Human))
- Tetrameric transthyretin (from human plasma) (Sigma, cat. no. P1742; 1 mg, pdb: 1dvq)
- Ar (pureshield)
- N₂ (> 99.9%, vol/vol)
- Alternative gases: Xe (> 99.9%), SF₆ (> 99.9%) (Boc Gases)

EQUIPMENT

- IM-mass spectrometer (e.g., Synapt HDMS, Q-IM-o-ToF with modified 32,000 *m/z* quadrupole option, Waters)
- MOBCAL software, Jarrod Group, <http://www.indiana.edu/~nano/Software.html> (free)
- Modified MOBCAL software, Robinson Group (free, contact by e-mail)
- MAESTRO software, <http://www.schrodinger.com/> (free to academic users)
- PYMOL software, <http://PYMOL.sourceforge.net/> (open-source access, but need to do own compilation, maintenance and support)

- HEX software, <http://www.csd.abdn.ac.uk/hex/> (free to academic or government-funded laboratories)
- Clustering software: hierarchical clustering using stored dissimilarities, Robinson Group (free, contact by e-mail)
- Topology search algorithm: for generating plausible structures that are intermediate between two archetypal protein topologies, Robinson Group (free, contact by e-mail)
- The following alternative software packages can also be used in various sections of the protocol:
 - Sigma software (alternative code for cross-section calculations), Bowers Group UC Santa Barbara, http://bowers.chem.ucsb.edu/theory_analysis/cross-sections/sigma.shtml (free, contact by e-mail)
 - Vmd, <http://www.ks.uiuc.edu/Research/vmd/> (free with registration)

PROCEDURE

Optimization of IM separation for large protein complexes

1| As a test complex, we recommend using transthyretin (TTR, tetramer, 56 kDa). A solution of this protein complex should be prepared at a concentration of 5–10 μM in aqueous ammonium acetate solution for subsequent steps of this protocol. Refer to the associated *Nature Protocols* article¹⁴, to generate appropriate protein solution and spray conditions for observing intact protein complexes by MS. Note that activation to increase mass resolution and accuracy can be detrimental to IM data, as noted above.

2| Choose a neutral gas for use in the IM separation and activation (Trap/Transfer) areas of the instrument. More massive gases will provide longer drift times (if used in the IM region of the instrument) and larger activation energies (if used in the Trap/Transfer region of the instrument). In our experiments, we use N_2 for IM separation and Ar for the Trap/Transfer region—although benefits of using a heavier gas (e.g., Xe or SF_6) in the Trap/Transfer region have been reported⁶³, but these gases are substantially more expensive.

3| Start from the standard settings given in **Table 1** for the focusing/IM separation voltages. These settings were derived on our instrument by optimization using the TTR sample discussed in Step 1 and should provide you with adequate IM resolution to begin optimization.

4| Ion mobility–mass spectrometry for TTR has been published previously¹³, and the settings described in **Table 1** should provide a spectrum similar to that data. The IM resolution of the features observed should be roughly as shown in **Figure 2** ($\sim 10, t/\Delta t$). If multiple features for a given peak are observed with anomalously narrow peak widths ($> 30, t/\Delta t$), refer to TROUBLESHOOTING table. See **Figure 5** and related discussion in the ‘Anticipated results’ section for a pictorial representation of common problems.

? TROUBLESHOOTING

5| From the initial values used to generate a stable signal for the protein complex, lower the following voltages: cone voltage, extractor cone, bias voltage and trap collision voltage, and monitor the centroid of the drift time distribution for the ions of interest. Activation of the complex should be minimized to avoid unfolding of the protein complex. The above voltages should be decreased (in ~ 10 V steps) until the drift time observed for the ions of interest does not change by more than 3% between steps.

▲ CRITICAL STEP This step is crucial to avoid measurement of ions having undergone gas-phase unfolding.

6| Remove the wave height ramping function described in **Table 1** and attempt to separate the complex at a fixed wave height and wave velocity value.

▲ CRITICAL STEP This step is important for calibration (see ‘Analyzing IM-MS Data with Molecular Modeling’, below), as the current calibration procedure for the instrument requires input data acquired at fixed wave heights.

7| From the minimum value of the voltages determined in Step 5, increase the Trap collision voltage in a stepwise manner, recording data at ~ 10 V intervals. Continue until the signal-to-noise ratio for the intact protein complex is lower than 3:1, due to the gas-phase dissociation of the intact complex³⁸. This step, along with Step 5, ensures that both the optimum IM conditions (at low Trap collision voltage) and MS resolution/mass accuracy (at high trap collision voltage) is achieved (see **Fig. 2**). Note, at this point, higher mass accuracy and desolvation can often be achieved without losing the IM resolution achieved in Steps 4–6 by activating ions in the Transfer region of the instrument. As, by this method, activation takes place after IM and before MS analysis, IM resolution will remain unaffected.

Calibration of travelling-wave IM drift times

8| Prepare calibrant solutions by diluting equine cytochrome *c*, equine myoglobin and bovine ubiquitin in 49:49:2 methanol/water/acetic acid at a concentration of 10 μM .

9| Record IM-MS data for an unknown protein complex over a range of different wave heights to separate the ions. This step is designed to indicate the presence of any field-strength-dependent effects on the drift of the ions under investigation.

TABLE 1 | Standard instrument conditions for the synapt HDMS (Q-IM-o-ToF) for analyzing large protein complexes

Pressure		
Region	Value (mbar)	Comment
Backing	5.5	Pressure listed is for optimal transmission of protein complexes between 50 and 200 kDa in molecular weight and will be adjusted higher or lower for larger or smaller mass species, respectively
Trap/Transfer	$2-8 \times 10^{-2}$	Varies depending upon degree of fragmentation desired and will influence transmission/detection of some very high m/z fragment ions. Adjustments will be necessary. Using Ar
IM	0.5	Using N ₂
Voltage		
Region	Value (V)	Comment
Trap collision energy	8–50	Use low-end value to avoid gas-phase unfolding protein complexes, high-end for enhanced mass accuracy
Transfer collision energy	8–50	Adjusted to either avoid (low voltage) or induce (high voltage) protein complex dissociation
Bias voltage	20–60	Low-end to avoid protein complex unfolding, high-end for enhanced mass accuracy
Cone voltage	30–150	As above, use low-end value for gentle conditions, high-end for enhanced mass accuracy and desolvation
Extractor cone	0.3–10	Use higher voltages for increased desolvation and higher mass accuracy
Capillary voltage	1,000–1,900	
IM DC entrance	10	
Transfer DC exit	9	
IM parameters		
Wave height	Ramp 3–30 V	
Wave velocity	250 m s ⁻¹	
Transfer T-wave parameters		
Wave height	8 V	
Wave velocity	200 m s ⁻¹	
Other		
MS Profile	1,000 m/z (dwell 10%), ramp 20%, 12,000 m/z (dwell 30%) ramp 40%, 24,000 m/z	Should be varied to view desired mass range with maximum transmission efficiency
All RF amplitudes Maximized		
Pusher interval	Optimized for mass range	Can be used to optimize both overall drift time range and IM bin-width, as the IM dimension is composed of a large number of time-of-flight experiments in serial

10 | Use precisely the same instrument conditions (including pressures) for all elements downstream of the trapping ion guide to acquire data for the three calibrant proteins, the solutions for which were generated in Step 8.

▲ CRITICAL STEP Altering voltages that affect the recorded drift time (i.e., any element of the IM separation stage or post-IM ion transfer stage) between calibration runs and measurements of unknowns can cause significant errors in calibrated measurements.

11 | Correct calibrant drift times (acquired using a single wave-height value) for mass-dependent flight time, calculated by the equation $t'_D = t_D - \left[\frac{C\sqrt{m/z}}{1000} \right]$, where t'_D is the corrected drift time in ms, t_D is the experimental drift time in ms,

PROTOCOL

m/z is the mass-to-charge ratio of the observed ion and C is a constant. The constant C can be found within the control software of the Synapt Q-IM-o-ToF instrument and is designated as the 'EDC (Enhanced Duty Cycle) delay coefficient'. C varies slightly from instrument to instrument, typically between 1.4 and 1.6. Note that the C values from the EDC setup are only valid if the exit lens of the transfer T-wave guide and transfer ion optics are the same as used when setting up the EDC calibration. To ensure this, once the system has been tuned for best mobility performance, select the Trapping tab from the tune page and select 'Use EDC'. Next insert the Transfer DC Exit voltage value from the TriWave DC tab view into the 'Extract Height (V)' setting. Put the m/z value of a calibrant ion species into the EDC mass box. Now open the 'Acquisition settings' window from the 'System' menu on the tune page. While monitoring the intensity of the selected m/z ion on the tune page, adjust the 'EDC Delay coefficient' (C value) to maximize the signal intensity. Once this has been achieved, the coefficient can be used in the correction term to get a reasonable approximation for the m/z -dependent flight time from the transfer T-wave guide to the pusher. The time required for ions to transit the Transfer region is also present in all drift times recorded. This time can be calculated by taking the length of the transfer T-wave (10 cm) divided by the wave velocity (indicated in the software) and subtracted from the total drift time recorded.

12| Take calibrant collision cross-sections (Ω , see **Table 2**)

and correct them for both ion charge state and reduced mass (μ) to generate Ω' ($\Omega' = \Omega / [\text{charge} \times (1/\mu)^{1/2}]$).

13| Create a plot of $\ln t'_D$ against $\ln \Omega'$.

14| Fit the plot to a linear relationship of the form: $\ln \Omega' = X \times \ln t'_D + \ln A$, where A is a fit-determined constant and X is referred to as the 'exponential factor' in this protocol. The correlation coefficient of the fit achieved in this step should be high ($R^2 > 0.98$).

15| Re-plot Ω versus a new corrected drift time (t''_D), where t_D is given by $[t'_D \times \text{charge} \times (1/\mu)^{1/2}]$. As in Step 14, the correlation coefficient of the calibration plot should be high ($R^2 > 0.98$). An anomalously low correlation coefficient could indicate an error in a previous step in the calibration protocol (Steps 8–15; see **Fig. 3a** for examples).

▲ CRITICAL STEP This approach provides a linear calibration plot, which provides a relationship between literature cross-section and T-wave drift time that is straightforward to extrapolate for measurements of large protein complex ions.

16| Use the plot generated in Step 15 to calibrate drift time data for unknowns.

17| Repeat Steps 11–16 for every wave-height value used to separate protein complex ions of interest.

18| Validate this calibration curve against the trend shown in **Figure 3b**. If cross-sectional values obtained are smaller (65%) or significantly larger (180%) than those predicted by the plot shown in **Figure 3b**, then refer to TROUBLESHOOTING table.

▲ CRITICAL STEP These upper and lower bounds are based both on our experience with a wide range of protein complexes and on their average packing densities in the gas-phase and is a crucial check of whether the calibration procedure has been correctly applied. It is important to note that the upper and lower bounds indicated may not hold for very high molecular weight species (~ 1 MDa).

? TROUBLESHOOTING

19| Attempt a collision cross-section measurement of TTR. Using Steps 8–17, we achieve a measurement of $2,900 \pm 213 \text{ \AA}^2$ for the 14–16+ charge states (ref. 13).

20| Estimate the error of the collision cross-section measurement. The total error of the calibrated collision cross-sections generated using Steps 8–17 above should be estimated and reported along with the average measurement. We assume that the total error of the calibrated measurement is a sum of the standard deviation of three or more replicate measurements

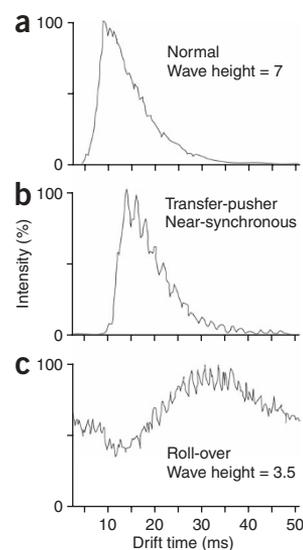


Figure 5 | Optimizing ion mobility separations. (a) Normal spectrum acquired for a concentrated CsI solution (100 mg ml^{-1} in water) with clusters covering a mass range from 500 to 10,000 m/z acquired using a static wave height of 7 V. (b) The same sample and conditions as in panel a; however, in this case, the transfer wave velocity and ToF pusher frequency are nearly synchronous, resulting in artificial 'beats' or 'waves' in the resulting ion mobility separation. (c) The same instrumental conditions as in panel a; however, wave height used for the IM separation is halved. In this case, the ions observed during the first portion of the separation shown are the same as those observed in the tailing edge of the ion mobility data shown in panel a.

(reproducibility, E_R), the average error of calibration curve generated in Step 15 (E_{Cal}) and the error carried by the protein standards used to calibrate the drift times of the unknowns (E_S , assumed to be $\sim 1\%$)⁶⁴. The expression E_{Cal} estimates the error associated with converting between measurements in N_2 and calibrant IM measurements in He. Provided that the relative cross-sectional differences between ions are the same between the two gases, the error in this step should be minimal ($< 2.5\%$)⁶⁵. The total error (E_T) is typically between ± 5 and 8% (in our experience) for protein complexes up to 500 kDa.

▲ CRITICAL STEP It is important to include error when comparing the measurement against calculated collision cross-section values for model structures.

Analyzing IM-MS data with molecular modeling

21| Generate a modified version of MOBICAL^{18,19}. Currently, we use two different versions of the MOBICAL code. One is a version of the code that is used for determining the collision cross-section of all-atom representations of a complex. For this version of MOBICAL, few alterations to the code are required. Increasing the 'len' variable is used to assure that the program can accept coordinate files that contain large numbers of atoms (len should be greater than or equal to the number of atoms in the coordinate file). We also alter the number of iterations used by the program to calculate cross-sections based on a single geometry file, as discussed in the accompanying instructions (mobcal.txt). The second version of the code is designed to determine the collision cross-section of coarse-grained (or hybridized) models of a complex. To generate this modified version of the code, use the following procedure: (i) open mobcal.f (source code) in a standard text editing program. (ii) On line 583 of the program, one of two sections that define atom masses and sizes begins. Replace atom radii and masses with the radius and mass of subunits comprising the complex of interest. To generate these sizes, there are several methods available. In the past, we have had success in estimating the radius either by (a) inputting a coordinate file corresponding to a high-resolution structure of a monomeric unit of the complex or (b) estimating the size of a compact monomer based on the relationship shown in **Figure 3b**. The same 'atom' radii and masses need to be modified in the section that begins on line 2,601 of the program code. (iii) Remove the Trajectory method calculation from the code by placing a 'c' in front of line 338.

Typically, we report the projection approximation value as the estimated collision cross-section of the model. We assume that scattering phenomena are accounted for adequately by appropriate selection of subunit radii. Subunit radii can be estimated in a variety of ways, including the crystal structures of the appropriate subunit (it is recommended to perform a trajectory method calculation with MOBICAL to generate an average radius for the subunit) or by average density estimates if no subunit structure is available (as in **Fig. 3b**).

▲ CRITICAL STEP Without this program, it will be difficult to quantitatively compare measured values to model structures.

22| Generate a coordinate file that corresponds to potential structure(s) of the protein complex under investigation. This coordinate file can be of any file format initially, but for compatibility with MOBICAL, it is recommended that the information other than Cartesian coordinates (e.g., secondary structure notation or subunit designations commonly found in .pdb files) be kept to a minimum. In some cases, we generate simplified coordinate systems by arranging carbon atoms at defined distances

TABLE 2 | Commonly used calibrant ions and their collision cross-sections⁴⁷.

Protein/peptide name (molecular mass, in Da)	m/z (charge)	Collision cross-section (in Å ²)
Bradykinin (1,059)	1,060 (1)	245
Cytochrome c (12,359)	1,746 (7)	1,247
	1,528 (8)	1,845 ^a
	1,528 (8)	2,061 ^a
	1,358 (9)	1,964 ^a
	1,358 (9)	2,215 ^a
	1,222 (10)	2,226
	1,111 (11)	2,303
	1,019 (12)	2,335
	940 (13)	2,391
	873 (14)	2,473
	815 (15)	2,579
	764 (16)	2,679
	719 (17)	2,723
680 (18)	2,766	
Equine myoglobin (16,952)	2,120 (8)	1,673
	1,885 (9)	1,758
	1,696 (10)	1,879 ^a
	1,696 (10)	2,796 ^a
	1,542 (11)	2,942
	1,414 (12)	3,044
	1,305 (13)	3,136
	1,212 (14)	3,143
	1,131 (15)	3,230
	1,061 (16)	3,313
	998 (17)	3,384
	943 (18)	3,489
	893 (19)	3,570
849 (20)	3,682	
808 (21)	3,792	
772 (22)	3,815	
Bovine ubiquitin (8,565)	1,429 (6)	1,041 ^a
	1,429 (6)	1,220 ^a
	1,429 (6)	1,525 ^a
	1,225 (7)	1,317
	1,225 (7)	1,580
	1,072 (8)	1,442
	1,072 (8)	1,622
	953 (9)	1,649
	858 (10)	1,732
	778 (11)	1,802

^aDenotes database values that often give the largest errors in our calibrations. If possible, we exclude all peaks that exhibit multiple distributions, as these peaks are metastable^{67–69}.

using packages such as MAESTRO or PYMOL. This approach often streamlines the process, as tools that define the symmetries of the coordinate systems generated are already built into the software (preferred to generating a coordinate system manually or with a simplified algorithm).

23| Modify the coordinate files generated in Step 22 for compatibility with MOBCAL. There is an example input file packaged with MOBCAL for comparison. Your input file must match this format exactly (e.g., spaces between number columns).

▲ CRITICAL STEP Many MOBCAL errors can be traced to incorrect formatting of the input file.

24| Run MOBCAL to calculate the collision cross-section of the input model. Refer to TROUBLESHOOTING table for common errors encountered.

? TROUBLESHOOTING

25| Evaluate the agreement of the model with experimental data.

26| Generate additional models for comparison against experimental data. This can be accomplished in a variety of ways. Different possible architectures of a protein assembly can be assembled based on biochemical data or literature sources. In some cases, this step involves docking several proteins of known solution structure together in different orientations. Manual docking can be accomplished using the Hex program, VMD, or other packages. Ideally, full docking calculations would be used to assess the stability of the final structure; however, this may not be possible in all instances (e.g., when high resolution data are not available for all subunits). IM-MS data can be used to test consistency against one of several hypotheses for a given protein complex structure. Molecular dynamics can also be used to generate potential candidate structures. We have developed a software approach that generates a large number of potential quaternary structures that are intermediate to two likely candidate structures (see ref. 12 for a complete description of the program, available on request). Typically, the output of this program is used in conjunction with clustering algorithms to generate a family of candidate structures that best fit the data.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 3**.

TABLE 3 | Troubleshooting table.

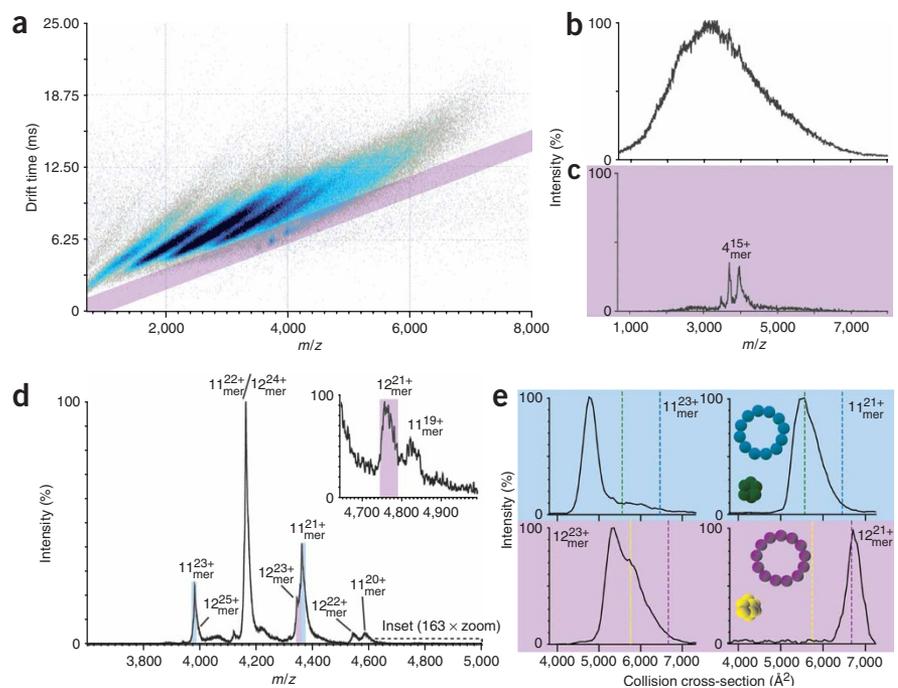
Problem	Possible reason(s)	Solution(s)
IM profile for test compounds shows too many peaks of anomalously high IM resolution ($R > 30$) (from Step 4, see Fig. 5)	Transfer T-wave or velocity is set incorrectly	If the transfer T-wave device delivers ions at a rate that is similar to the pusher rate of the ToF, then the resulting spectra can contain 'beats' or 'waves'. These can often be removed by adjusting the transfer T-wave velocity. The maximum/optimum wave for the transfer region is when the frequency of the wave (shown in a box next to the wave velocity on the tune page) is one-sixth of the pusher period of the ToF
Cannot achieve adequate IM resolution on test compounds	Trap ion gate set inappropriately	For large protein complexes, we typically set the trap voltage at its maximum value (30 V) to ensure that ions do not escape from the trapping ion guide before the injection pulse to the IM separator. Similarly, the height of the wave used to extract ions from the trap should also be maximized. Alternatively, higher gas pressures can be used in this region to enable lower voltages
	IM separation wave height/wave velocity set incorrectly	For larger ions, optimal separation is often achieved at higher wave heights and lower wave velocities
	Transfer T-wave set incorrectly	Refer to the solution section of problem 1. Additionally, the wave height/velocity must be set correctly so that ions are carried along with each wave. Typically, this translates to ≥ 5 V wave heights
	Protein ions are activated before IM separation	It is important to minimize the voltages used on the cone, extractor, trap collision voltage and bias voltage (IM injection voltage). Excess voltages in any of these regions may affect the structure/conformation of the protein complex and may lead to fragmentation



TABLE 3 | Troubleshooting table (continued).

Problem	Possible reason(s)	Solution(s)
The complete IM spectrum never appears in a single real-time update; instead, it takes several real-time updates to compile a complete IM spectrum with signal for all ions of interest	As with the problem above, this can occur to near-synchronized ToF extraction and occurs due to Transfer T-wave delivery rates	Alter the Transfer T-wave velocity until the effect is gone
Collision cross-sections of unknowns deviate considerably from projected values (from Step 18)	Inappropriate/too few calibrant ions used for generating calibration curve	Refer to Table 1 for those calibrant ions that have proven to be outliers in our calibration curves
	Calibration procedure followed incorrectly	Refer Steps 8–18 in the protocol section. If collision cross-sections of low charge states are to be used for calibration, it is important to ensure that these ions undergo minimal activation prior to IM separation. See Step 5 of the protocol
	Cross-sectional difference between unknown and calibrant ions is too large for extrapolation	Small variations in the slope of the calibration curve can lead to large deviations in calibrated cross-sections if those values are far removed (in cross-section) from the calibrant ions. Check the extrapolated values against Figure 3b (as described in protocol) to first assess the validity of the calibration curve. In our experience using tetrameric transthyretin ions (at a value of 2900 Å ² for 14 through 16+ charge states) as additional calibrants often improves the accuracy of calibrated cross-sections in excess of 10,000 Å ² . Note that including tetrameric transthyretin ions decreases the absolute precision of the measurement substantially, as the larger error associated with the collision cross-section of these ions (±7.3%) must be propagated in the reported composite error of the measurement
Ions covering a wide <i>m/z</i> range all appear to have the roughly the same <i>t_D</i>	Ions are being carried along on a single wave and are not mobility separated	Reduce the wave amplitude until the ions have distinct drift times
Some <i>m/z</i> species have signal both at the end and beginning of the drift time axis (see Fig. 5)	Ions from the previous mobility run are not exiting the IM guide before the next acquisition begins and are recorded in the subsequent run	Increase the IM T-wave amplitude or reduce the T-wave velocity until the signals do not appear at the beginning of the drift time axis, that is, all ions are exiting the IM guide during a single acquisition
Multiple models agree with experimental data	Measurements acquired with insufficient precision or IM resolution	If replicate measurements were acquired under different field strength conditions, check for trends as a function of separation field strength. Also, in general, increasing the number of calibration points used to generate calibration curves enhances measurement precision. To increase IM resolution, try running at higher wave-heights and pressures
	Models have been filtered insufficiently against other data sources	The number of models considered can be narrowed by either cross-section measurements on other related ions or by considering other biochemical data (e.g., high-resolution structures of subunits) to determine the likelihood of certain models
MOBCAL crashes with a 'segmentation fault'-type error message (from Step 24)	The coordinate file input to the program is too large (contains too many atoms or subunits)	Increase the value of the 'len' variable in the MOBCAL code and/or reduce the number of atoms present in coordinate set
MOBCAL crashes with an 'I/O'-type error message (from Step 24)	Input file formatted incorrectly or contains incorrect values	Verify all values in the input file (especially the number of atoms listed) and check the format of the file against the example provided

Figure 6 | Expected separation and structural results from IM-MS of protein complexes. **(a)** Plot of drift time versus m/z for human tetrameric transthyretin (MW = 56 kDa), prepared in an ammonium acetate buffer contaminated with sodium chloride unpurified after recovery from a cellular matrix (recombinant expression). Trends in the data are evident, including several bands at high drift time and 3–4 well-defined signals at low drift time (violet shaded area). **(b)** Total integrated mass spectrum for all ions. No well-defined peaks are observed for tetrameric transthyretin. **(c)** A mass spectrum compiled from just those ions that fall within the shaded area from panel **a**. The data shown here are sufficient for a high-confidence mass measurement and assignment of the signals as originating from tetrameric transthyretin (mass measured = 55.85 ± 0.1 kDa). **(d)** Mass spectrum of the *apo* tryptophan RNA-binding attenuation protein (TRAP) complex in both the 11- and 12-membered state. Charge state series for the 11mer includes 19–23⁺ ions, whereas the 12mer distribution extends from 21⁺ to 25⁺. Although some overlap between the two species is evident in the spectrum (most notably for the higher charge state species), mass spectrometry data and ion mobility data can be recorded for several separated charge states of the two protein assemblies. **(e)** Ion mobility drift time distributions converted to a collision cross-section axis for selected peaks from panel **a**. Data for the 11mer are shown on top (blue) and for the 12mer on the bottom (violet). Dashed lines represent model structures shown on the figure: collapsed 11mer (green) ring 11mer (blue) collapsed 12mer (yellow) ring 12mer (purple). As the 11mer, the 12mer also forms a stable ring structure, which is observed in the gas phase, although higher charge states of the complex form collapsed structures.



ANTICIPATED RESULTS

While performing IM-MS experiments it is possible to generate IM separation artifacts. Two such artifacts, unique to the Synapt T-wave IM separation device, are illustrated in **Figure 5**. **Figure 5a** shows an optimized trace for all drift times recorded for a solution of CsI using a wave height of 7 V to propel the ions through the IM separation stage. In **Figure 5b**, the Transfer wave velocity has been altered such that it is nearly synchronous with the frequency of the ToF extraction region. This results in a spectrum containing multiple artificial ‘beats’ or ‘waves’, as are visible in **Figure 5b**. As discussed above, this spectrum can be corrected by adjusting the frequency of the Transfer T-wave such that it is approximately one-sixth of the frequency of the ToF extraction region. Another IM artifact that can be observed if careful control of the instrument voltage settings is not exercised is shown in **Figure 5c**. If the wave height used to generate the IM data shown in **Figure 5a** is lowered by a factor of 2, the drift time plot changes its appearance dramatically (**Fig. 5c**). In this drift time plot, some ions appear at relatively short drift times that were not apparent in **Figure 5a**. This is because the wave height is too low to effect efficient separation over the IM separation time allowed for this experiment (50 ms) and some ions, those on the trailing edge of the distribution in **Figure 5a**, require more than one allotment of IM separation time to travel to the orthogonal extraction region of the ToF in **Figure 5c**. The apparent ‘beats’ or ‘waves’ in the data shown in **Figure 5c** are the result of poor signal to noise rather than any additional artifact of unoptimized IM separation.

We have observed that IM-MS is particularly powerful in the analysis of protein complexes for both (i) removing chemical noise from structural information and (ii) performing structural studies on protein complexes. As indicated in point number 1, proteins are often prepared in buffers containing salts and other stabilizing agents. These molecules generate chemical noise in MS experiments that is difficult to remove without extensive liquid or solid phase separations. **Figure 6a–c** shows an example where IM-MS is used to remove unwanted background ions from the mass spectrum of a sample of human TTR. The mass spectrum shown in **Figure 6b** illustrates what mass spectrometry alone would reveal about the sample. Although the mass spectrum contains a large amount of ion signal, there are no well-defined peaks. **Figure 6a** shows the same data, but as a contour plot that reveals both the IM and MS dimensions (as described in **Fig. 2**). By analyzing this three dimensional data set, it is apparent that the unstructured signal observed in the mass spectrum exhibits broad trends in the plot of t_D versus m/z shown in **Figure 6b**. The chemical noise, observed at higher drift times, can be identified as low-charge state salt clusters while the signals for the more highly charged TTR ions are observed at faster drift times than salt clusters of similar m/z . The TTR ion signal can be selected from the three-dimensional data set and plotted separately, as shown in **Figure 6c**. The data are now suitable for mass measurement and assignment. **Figures 6a–c** demonstrate the power of IM-MS for assessing protein complexes in a wider range of buffer components than are currently accessible by MS alone.



Figure 6d,e highlights the power of IM-MS data for performing structural studies on protein complexes (point 2 above). These data, on TRAP, show evidence of both 11- and 12-membered protein assemblies, as observed previously⁶⁶. IM-MS allows structural assignment of both the 11-mer and 12-mer even though both complexes coexist in solution. The precise set of solution conditions that favor the formation of dodecameric TRAP are currently unknown; however, we have observed dodecamer ions principally from samples analyzed after extended periods at room temperature (20 °C). Our data suggest that the 12-member TRAP complex is also ring-like (similar to the undecamer) at low charge states (21⁺). In addition, higher charge states appear to have progressively collapsed structures. Interestingly, this collapse occurs at similar values of charge-per-subunit values for the two complexes (between 1.72 and 1.81 charges-per-subunit values for undecamer and between 1.75 and 1.83 for the dodecamer)¹². Overall, the data presented in **Figure 6d,e** illustrate one of the primary advantages of an IM-MS approach to protein structure determination relative to classical spectroscopic methods. Classical tools of structural biology, for example, nuclear magnetic resonance and X-ray crystallography, require homogenous protein complexes containing mono-disperse protein populations in high concentrations (millimolar in some cases). The IM-MS technique described in this protocol has neither of these limitations and is applicable to heterogeneous, dynamic protein complexes.

ACKNOWLEDGMENTS Iain Campuzano, Kevin Giles, Jason Wildgoose, Robert Bateman (Waters) and Daniel Barsky (Lawrence Livermore National Laboratory) are thanked for their generous contributions to this manuscript. Eman Basha and Elizabeth Vierling (University of Arizona), Paul Gollnick (SUNY, Buffalo) and Margaret McCammon (University of Cambridge) are thanked for providing TaHSP16.9, TRAP and TTR, respectively. Ion Mobility–Mass Spectrometry Research at the University of Cambridge is funded by grants from the BBSRC, EPSRC and the Walters Kundert Trust. C.V.R. is a Professor of the Royal Society.

Published online at <http://www.natureprotocols.com>
Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

1. Gavin, A.C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
2. Gavin, A.C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006).
3. Sali, A., Glaeser, R., Earnest, T. & Baumeister, W. From words to literature in structural proteomics. *Nature* **422**, 216–225 (2003).
4. Taverner, T. *et al.* Subunit architecture of intact protein complexes from mass spectrometry and homology modeling. *Acc. Chem. Res.* (in the press) DOI: 10.1021/ar 700218q.
5. Robinson, C.V., Sali, A. & Baumeister, W. The molecular sociology of the cell. *Nature* **450**, 973–982 (2007).
6. Alber, F. *et al.* Determining the architectures of macromolecular assemblies. *Nature* **450**, 683–694 (2007).
7. von Helden, G., Wyttenbach, T. & Bowers, M.T. Conformation of macromolecules in the gas phase: use of matrix-assisted laser desorption methods in ion chromatography. *Science* **267**, 1483–1485 (1995).
8. Jarrold, M.F. Peptides and proteins in the vapor phase. *Annu. Rev. Phys. Chem.* **51**, 179–207 (2000).
9. Hoaglund-Hyzer, C.S., Counterman, A.E. & Clemmer, D.E. Anhydrous protein ions. *Chem. Rev.* **99**, 3037–3079 (1999).
10. Ruotolo, B.T. & Robinson, C.V. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.* **10**, 402–408 (2006).
11. Loo, J.A. *et al.* Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *J. Am. Soc. Mass Spectrom.* **16**, 998–1008 (2005).
12. Ruotolo, B.T. *et al.* Evidence for macromolecular protein rings in the absence of Bulk Water. *Science* **310**, 1658–1660 (2005).
13. Ruotolo, B.T. *et al.* Ion mobility–mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Ang. Chem. Int. Ed.* **46**, 8001–8004 (2007).
14. Hernández, H. & Robinson, C.V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2**, 1479–1489 (2007).
15. Mason, E.R. & McDaniel, E.W. *Transport Properties of Ions in Gases* (J. Wiley and Sons, New York, 1988).
16. McLean, J.A., Ruotolo, B.T., Gillig, K.J. & Russell, D.H. Ion mobility–mass spectrometry: a new paradigm for proteomics. *Int. J. Mass Spectrom.* **240**, 301–315 (2005).
17. Wyttenbach, T. & Bowers, M.T. Intermolecular interactions in biomolecular systems examined by mass spectrometry. *Ann. Rev. Phys. Chem.* **58**, 511–533 (2007).
18. Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C. & Jarrold, M.F. Structural information from ion mobility measurements: effects of the long-range potential. *J. Phys. Chem.* **100**, 16082–16086 (1996).

19. Shvartsburg, A.A. & Jarrold, M.F. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* **261**, 86–91 (1996).
20. Bowers, M.T. Cluster ions—carbon, met-cars, and sigma-bond activation. *Acc. Chem. Res.* **27**, 324–332 (1994).
21. Vonhelden, G., Gotts, N.G., Palke, W.E. & Bowers, M.T. Structures and energies of small carbon clusters—what experiment and theory have to say about C-8(+), C-9(+), and C-10(+). *Int. J. Mass Spectrom. Ion Proc.* **138**, 33–47 (1994).
22. Vonhelden, G., Gotts, N.G., Maitre, P. & Bowers, M.T. The structures of small iron carbon cluster anions—linear to planar to 3-dimensional. *Chem. Phys. Lett.* **227**, 601–608 (1994).
23. Hoaglund, C.S., Valentine, S.J., Sporleder, C.R., Reilly, J.P. & Clemmer, D.E. Three-dimensional ion mobility TOFMS analysis of electrosprayed biomolecules. *Anal. Chem.* **70**, 2236–2242 (1998).
24. Valentine, S.J., Counterman, A.E., Hoaglund, C.S., Reilly, J.P. & Clemmer, D.E. Gas-phase separations of protease digests. *J. Am. Soc. Mass Spectrom.* **9**, 1213–1216 (1998).
25. Ruotolo, B.T., Gillig, K.J., Stone, E.G. & Russell, D.H. Peak capacity of ion mobility mass spectrometry: separation of peptides in helium buffer gas. *J. Chrom. B.* **782**, 385–392 (2002).
26. Ruotolo, B.T. *et al.* Analysis of protein mixtures by matrix-assisted laser desorption ionization–ion mobility–orthogonal-time-of-flight mass spectrometry. *Int. J. Mass Spectrom.* **219**, 253–267 (2002).
27. van Montfort, R.L.M., Basha, E., Friedrich, K.L., Slingsby, C. & Vierling, E. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Struct. Biol.* **8**, 1025–1030 (2001).
28. Sobott, F., Benesch, J.L.P., Vierling, E. & Robinson, C.V. Subunit exchange of multimeric protein complexes—real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry. *J. Biol. Chem.* **277**, 38921–38929 (2002).
29. Benesch, J.L.P., Sobott, F. & Robinson, C.V. Thermal dissociation of multimeric protein complexes by using nano-electrospray mass spectrometry. *Anal. Chem.* **75**, 2208–2214 (2003).
30. Gangloff, M. *et al.* An indirect allosteric mechanism for activation of the Toll receptor by the dimeric ligand Spätzle. *J. Biol. Chem.* (2008) (in the press) DOI: 10.1074/jbc.m800112200.
31. Hoadlund-Hyzer, C.S., Li, J.W. & Clemmer, D.E. Mobility labeling for parallel CID of ion mixtures. *Anal. Chem.* **72**, 2737–2740 (2000).
32. Stone, E. *et al.* Surface-induced dissociation on a MALDI–ion mobility–orthogonal time-of-flight mass spectrometer: sequencing peptides from an ‘in-solution’ protein digest. *Anal. Chem.* **73**, 2233–2238 (2001).
33. Tahallah, N., Pinkse, M., Maier, C.S. & Heck, A.J.R. The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid Comm. Mass Spectrom.* **15**, 596–601 (2001).
34. Sobott, F., Hernández, H., McCammon, M.G., Tito, M.A. & Robinson, C.V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* **74**, 1402–1407 (2002).
35. van den Heuvel, R.H. & Heck, A.J.R. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **8**, 519–526 (2004).
36. Heck, A.J.R. & van den Heuvel, R.H.H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368–389 (2004).
37. Benesch, J.L.P., Ruotolo, B.T., Simmons, D.A. & Robinson, C.V. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem. Rev.* **107**, 3544–3567 (2007).



38. Benesch, J.L. & Robinson, C.V. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Curr. Opin. Struct. Biol.* **16**, 245–251 (2006).
39. Sobott, F. & Robinson, C.V. Characterising electrosprayed biomolecules using tandem-MS—the noncovalent GroEL chaperonin assembly. *Int. J. Mass Spectrom.* **236**, 25–32 (2004).
40. McKay, A.R., Ruotolo, B.T., Ilag, L.L. & Robinson, C.V. Mass measurements of increased accuracy resolve heterogeneous populations of intact ribosomes. *J. Am. Chem. Soc.* **128**, 11433–11442 (2006).
41. Patriksson, A., Marklund, E. & vanderSpoel, D. Protein structures under electrospray conditions. *Biochemistry* **46**, 933–945 (2007).
42. Ruotolo, B.T. & Russell, D.H. Gas-phase conformations of proteolytically derived protein fragments: Influence of solvent on peptide conformation. *J. Phys. Chem. B* **108**, 15321–15331 (2004).
43. Giles, K. *et al.* Applications of a travelling wave-based radio-frequency only stacked ring ion guide. *Rapid Comm. Mass Spectrom.* **18**, 2401–2414 (2004).
44. Pringle, S.D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* **261**, 1–12 (2007).
45. Wildgoose, J.L. *et al.* A comparison of travelling wave and drift tube ion mobility separations. *Abstr. Am. Soc. Mass Spectrom. Meet.* (2006).
46. Williams, J.P. & Scrivens, J.H. Coupling desorption electrospray ionisation and neutral desorption/extractive electrospray ionisation with a travelling-wave based ion mobility mass spectrometer for the analysis of drugs. *Rapid Comm. Mass Spectrom.* **22**, 187–196 (2008).
47. Clemmer Cross Section Database, <http://www.indiana.edu/~clemmer>.
48. Bacher, G. *et al.* Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses. *J. Mass Spectrom.* **36**, 1038–1052 (2001).
49. Kaufman, S.L., Kuchumov, A.R., Kazakevich, M. & Vinogradov, S.N. Analysis of a 3.6-MDa hexagonal bilayer hemoglobin from *Lumbricus terrestris* using a gas-phase electrophoretic mobility molecular analyzer. *Anal. Biochem.* **259**, 195–202 (1998).
50. Kaddis, C.S. *et al.* Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. *J. Am. Soc. Mass Spectrom.* **18**, 1206–1216 (2007).
51. Damsbo, M. *et al.* Application of evolutionary algorithm methods to polypeptide folding: comparison with experimental results for unsolvated Ac-(Ala-Gly-Gly)(5)-LysH(+). *Proc. Natl. Acad. Sci. USA* **101**, 7215–7222 (2004).
52. Gidden, J. & Bowers, M.T. Gas-phase conformations of deprotonated and protonated mononucleotides determined by ion mobility and theoretical modeling. *J. Phys. Chem. B* **107**, 12829–12837 (2003).
53. Jarrold, M.F. Helices and sheets *in vacuo*. *Phys. Chem. Chem. Phys.* **9**, 1659–1671 (2007).
54. Kinnear, B.S., Hartings, M.R. & Jarrold, M.F. The energy landscape of unsolvated peptides: helix formation and cold denaturation in Ac-A(4)G(7)A4+H+. *J. Am. Chem. Soc.* **124**, 4422–4431 (2002).
55. Kohtani, M. & Jarrold, M.F. The initial steps in the hydration of unsolvated peptides: water molecule adsorption on alanine-based helices and globules. *J. Am. Chem. Soc.* **124**, 11148–11158 (2002).
56. Liu, D.F., Wyttenbach, T. & Bowers, M.T. Hydration of mononucleotides. *J. Am. Chem. Soc.* **128**, 15155–15163 (2006).
57. Mao, Y., Woenckhaus, J., Kolafa, J., Ratner, M.A. & Jarrold, M.F. Thermal unfolding of unsolvated cytochrome *c*: experiment and molecular dynamics simulations. *J. Am. Chem. Soc.* **121**, 2712–2721 (1999).
58. Ruotolo, B.T., Tate, C.C. & Russell, D.H. Ion mobility-mass spectrometry applied to cyclic peptide analysis: conformational preferences of gramicidin S and linear analogs in the gas phase. *J. Am. Soc. Mass Spectrom.* **15**, 870–878 (2004).
59. Wyttenbach, T., Witt, M. & Bowers, M.T. On the question of salt bridges of cationized amino acids in the gas phase: glycine and arginine. *Int. J. Mass Spectrom.* **183**, 243–252 (1999).
60. Ilag, L.L., Ubarretxena-Belandía, I., Tate, C.G. & Robinson, C.V. Drug binding revealed by tandem mass spectrometry of a protein-micelle complex. *J. Am. Chem. Soc.* **126**, 14362–14363 (2004).
61. Srebalus, C.A., Li, J., Marshall, W.S. & Clemmer, D.E. Gas-phase separations of electrosprayed peptide libraries. *Anal. Chem.* **71**, 3918–3927 (1999).
62. Dugourd, P., Hudgins, R.R., Clemmer, D.E. & Jarrold, M.F. High-resolution ion mobility measurements. *Rev. Sci. Instrum.* **68**, 1122–1129 (1997).
63. Lorenzen, K., Versluis, C., van Duijn, E., van den Heuvel, R.H.H. & Heck, A.J.R. Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collision gases. *Int. J. Mass Spectrom.* **268**, 198–206 (2007).
64. Valentine, S.J. & Clemmer, D.E. H/D exchange levels of shape-resolved cytochrome *c* conformers in the gas phase. *J. Am. Chem. Soc.* **119**, 3558–3566 (1997).
65. Ruotolo, B.T., McLean, J.A., Gillig, K.J. & Russell, D.H. Peak capacity of ion mobility mass spectrometry: the utility of varying drift gas polarizability for the separation of tryptic peptides. *J. Mass Spectrom.* **39**, 361–367 (2004).
66. McCammon, M.G., Hernández, H., Sobott, F. & Robinson, C.V. Tandem mass spectrometry defines the stoichiometry and quaternary structural arrangement of tryptophan molecules in the multiprotein complex TRAP. *J. Am. Chem. Soc.* **126**, 5950–5951 (2004).
67. Badman, E.R., Hoaglund-Hyzer, C.S. & Clemmer, D.E. Monitoring structural changes of proteins in an ion trap over similar to 10–200 ms: unfolding transitions in cytochrome *c* ions. *Anal. Chem.* **73**, 6000–6007 (2001).
68. Myung, S., Badman, E.R., Lee, Y.J. & Clemmer, D.E. Structural transitions of electrosprayed ubiquitin ions stored in an ion trap over similar to 10 ms to 30 s. *J. Phys. Chem. A* **106**, 9976–9982 (2002).
69. Badman, E.R., Myung, S. & Clemmer, D.E. Evidence for unfolding and refolding of gas-phase cytochrome *c* ions in a Paul trap. *J. Am. Soc. Mass Spectrom.* **16**, 1493–1497 (2005).

