analytical chemistry

Article

pubs.acs.org/ac

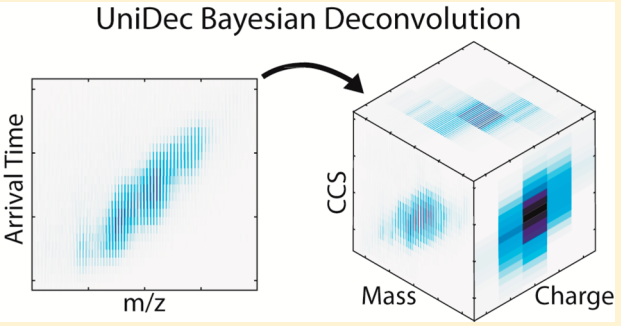
Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles

Michael T. Marty, Andrew J. Baldwin, Erik G. Marklund, Georg K. A. Hochberg, Justin L. P. Benesch, and Carol V. Robinson*

Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, Oxford OX1 3QZ, United Kingdom

Supporting Information

ABSTRACT: Interpretation of mass spectra is challenging because they report a ratio of two physical quantities, mass and charge, which may each have multiple components that overlap in m/z. Previous approaches to disentangling the two have focused on peak assignment or fitting. However, the former struggle with complex spectra, and the latter are generally computationally intensive and may require substantial manual intervention. We propose a new data analysis approach that employs a Bayesian framework to separate the mass and charge dimensions. On the basis of this approach, we developed UniDec (Universal Deconvolution), software that provides a rapid, robust, and flexible deconvolution of mass spectra and ion mobility-mass spectra with minimal user intervention. Incorporation of the charge-state distribution in the Bayesian prior probabilities provides separation of the m/z spectrum into its physical mass and charge components. We have evaluated our approach using systems of increasing complexity, enabling us to deduce lipid binding to membrane proteins, to probe the dynamics of subunit exchange reactions, and to characterize polydispersity in both protein assemblies and lipoprotein Nanodiscs. The general utility of our approach will greatly facilitate analysis of ion mobility and mass spectra.



Mass spectrometry (MS) is a powerful approach for structural and physical characterization of biomacromolecules.[1−7] However, for MS to revolutionize these fields as it has proteomics, data analysis techniques are needed to rapidly and reliably extract quantitative information from spectra. Interpretation of mass spectra is challenging because MS detects a ratio of mass to charge, two correlated physical quantities.

In electrospray ionization (ESI) MS, separation of mass and charge is possible because ESI produces multiply charged ions for a single molecular species. If peaks can be assigned to a single charge series, their masses and charges can be determined by solving a system of equations.[8] However, assignment becomes difficult when several charge state distributions overlap in m/z. As experimental approaches become more sophisticated and ever more heterogeneous systems are analyzed, rapid deconvolution of complex data has become a priority for MS-based structural biology.[9−12]

Deconvolution in the context of MS often refers to separating the interwoven charge and mass dimensions.[13] MS deconvolution approaches generally fall into three broad categories: peak assignment, isotopic, and simulation-based algorithms.[14,15]

Peak assignment algorithms extract a list of peaks from the spectrum and assign a charge to each.[8,14−18] Because of the data reduction achieved by peak selection, these algorithms are fast and produce simple outputs. However, these algorithms often struggle with complex spectra and do not generally produce quantitative results.

Isotopic algorithms require high-resolution data with resolved isotopologues.[13,19] Because the mass difference between peaks is known from the atomic masses of the isotopes, the charge can be inferred directly from the difference in m/z. However, these approaches are not generally applicable, as many spectra do not display isotopic resolution, and they are not generally quantitative.

Simulation algorithms address the deficiencies of peak assignment algorithms for complex systems.[10,11,20−22] Multiple hypothetical mass and charge distributions are generated from which a spectrum is simulated. The simulated spectrum that fits the data best based on a metric of minimum chi-squared[11] or maximum entropy[16,23,24] is then taken to be the most correct. Although this approach yields quantitative fits, current implementations tend to be computationally intensive and require substantial user guidance.

Here, we present a novel Bayesian deconvolution algorithm, which represents a means for rapid and quantitative interpretation of both mass and ion mobility (IM) spectra. We reframe the problem into a mathematical deconvolution, which is the process of retrieving the latent signal from a recorded signal that has been convolved with a given point-spread function. The approach is a special case of the Richardson-Lucy mathematical

Received: January 12, 2015
Accepted: March 23, 2015
Published: March 23, 2015

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

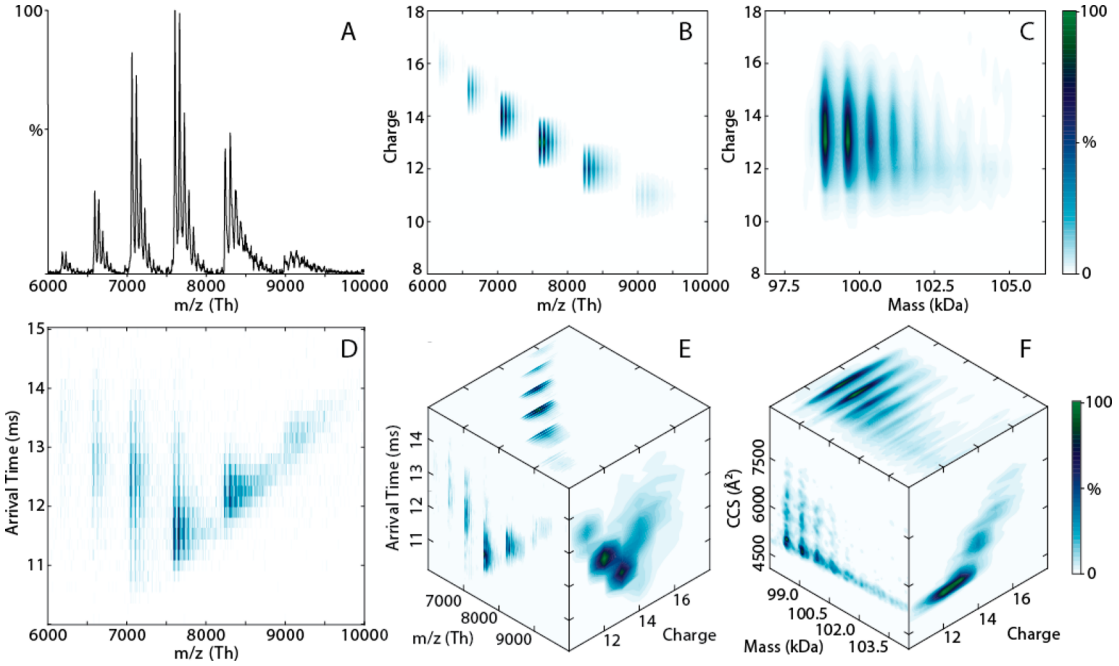


Figure 1. Description of UniDec outputs. The experimental IM-MS spectrum for AqpZ with bound POPC at 100 V collision voltage is shown in (D). Summing along the arrival time axis produces the mass spectrum shown in (A). UniDec deconvolution of (A) separates the charge dimensions into a m/z vs charge matrix and a mass vs charge matrix as shown in (B) and (C), respectively. UniDec deconvolution of the IM-MS spectrum produces a 3D matrix of m/z vs arrival time vs charge, which is transformed into a matrix of mass vs CCS vs charge. The projections of each 3D matrix along each axis are shown on the face of cubes in (E) and (F), respectively.

deconvolution algorithm.[25,26] A similar concept was previously applied to MS as a method of deconvolving extracted intensities from a peak list.[9] We have generalized and modified the algorithm to develop a new approach to deconvolving both mass spectra and ion mobility-mass spectra (IM-MS) with no reduction of the data.

The algorithm, UniDec (Universal Deconvolution), performs quantitative deconvolution and charge separation of complex spectra. It is fast, robust, and easily generalized to any number of dimensions or peak shapes. Remarkably, it does not require substantial user intervention or prior knowledge. Here, we describe the UniDec approach for studying intact protein complexes with both native MS and two-dimensional IM-MS. We explore its utility for studying the biophysics of heterogeneous systems such as membrane proteins, small heat-shock protein assemblies, and Nanodisc lipoprotein complexes. Overall, we show that UniDec provides a solution to analyzing challenging biochemical systems.

METHODS

Overview of Algorithm. Our approach assumes the spectrum can be described as a convolution between a peak shape and a set of weighted delta functions (referred to here as a "delta matrix"), which describe the contribution of specific molecular ions to the spectrum. A schematic description of the UniDec approach is shown in Figure S-1, Supporting Information, and a detailed description of the algorithm is provided in the Supplemental Methods, Supporting Information. For MS data, we implement a two-dimensional delta matrix containing m/z in one dimension and charge in the other (Figure 1B). Initially, we assume an equal probability for all charge states. The algorithm then proceeds by sequential iteration of three steps to determine the effective contribution of each m/z and charge to the overall spectrum.

First, a filter is applied to smooth the charge distributions in the delta matrix. This step modifies the Bayesian priors to include information from neighboring species. Physically, this approach assumes that a charge state series of an ion will be smooth, but no *a priori* assumptions are placed on the precise form of the charge state distribution. Thus, for a given m/z value, the probability of that intensity being assigned to a given charge state is related to the probability of the same m value appearing at $m/(z + 1)$ and $m/(z - 1)$. A similar filter can be applied for systems where known mass differences may be taken into account, an approach very suitable for oligomeric assemblies.

Second, the filtered delta matrix is summed along the charge axis to produce a one-dimensional m/z spectrum of weighted delta functions. The summation of delta functions is then convolved with the peak shape defined by the user to produce an m/z spectrum, which has the same dimensions as the data.

Finally, each element in the filtered delta matrix is multiplied by the ratio of the data to the projected spectrum at its specific m/z value. If the projected spectrum has a higher intensity than the data at a given m/z, which may be caused by overlapping peaks, the ratio will be less than unity, and the probability of each charge state at that m/z will be decreased. If the data is more intense than the projected spectrum, the ratio will be greater than unity, and the probability will be increased to compensate. There is no change if the data and projected spectrum are the same. It follows that the algorithm converges when the projected spectrum matches the data.

The result of the successive iteration of these three steps is a deconvolved matrix of delta functions. To better represent the data, each charge column of the matrix is convolved with the peak shape function. Because the m/z and charge are defined for each point in the delta matrix, the convolved delta matrix is then transformed into a matrix of mass vs charge (Figure 1C). An

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

analogous approach is used for IM-MS (Figure 1D−F) as detailed in the Supplemental Methods, Supporting Information.

The only parameters required by the algorithm are the peak width and the range of charge states to consider. Additional restraints on allowed masses may be provided to accelerate the deconvolution and to mitigate any overfitting errors. In general, the algorithm speed scales with the size of the m/z vs charge matrix and to a lesser degree with the size of the transformed mass vs charge matrix. The speed is not significantly affected by heterogeneity, allowing the algorithm to scale effectively to very complex systems. Uncertainties in the data, resulting from poor resolution, high noise, or untenable complexity, will provide a practical limit on the applicability of the algorithm. A detailed discussion of potential artifacts is provided in the Supplemental Methods, Supporting Information.

Deconvolution Parameters. We implemented UniDec in C and developed a user interface in Python to control the software and generate graphical outputs. A copy of the program is available on request or for download at unidec.chem.ox.ac.uk. Details of the implementation are provided in the Supplemental Methods, Supporting Information, together with a detailed description of the experimental methods and deconvolution parameters. Briefly, the key parameters used for each system are as follows. All spectra were deconvolved using a charge smooth filter with a width of one elementary charge except for those from Nanodiscs. An appropriate peak shape function was determined by fitting an isolated peak to Gaussian, Lorentzian, or split Gaussian/Lorentzian peak shapes. The split Gaussian/Lorentzian peak shape is defined as a Lorentzian on the high m/z side and a Gaussian on the other such that the fwhm is symmetrical about the maximum. The split Gaussian/Lorentzian distribution is a good model of native MS peak shapes when a long tail is observed on the high m/z side due to bound adduct species.[20]

Specific Applications. We describe here application of the UniDec approach to problems of increasing complexity: membrane protein AqpZ; small heat shock proteins HSP17.7, HSP16.5, and αB-crystallin; and lipoprotein Nanodiscs. MS and IM-MS data of aquaporin Z (AqpZ) with bound 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) obtained at 100 V accelerating potential into a dedicated collision cell were analyzed using UniDec by limiting the mass range to between 95 and 105 kDa.[27] An example of how the algorithm performs without mass limitations is shown in Figure S-2, Supporting Information. Data was smoothed in MassLynx 4.1 software (Waters Corp.) before analysis with Transform and MaxEnt, which used the same mass limitation.

Deconvolution of subunit exchange data from HSP17.7 was performed by limiting the allowed mass range to between 211 and 222 kDa. Tandem MS spectra of the isolated +47 charge state of HSP16.5 24-mers were summed across multiple collision voltages to compile an aggregate spectrum.[28] Deconvolution was performed by limiting the charge state between 10 and 49 and manually defining the +47 charge state, which was necessary because only one charge state was isolated in the MS/MS experiment. Collision induced dissociation (CID) spectra of αB-crystallin were obtained similarly. Masses were limited to within 3000 Da of a wide range of potential oligomer complexes ranging from 1 to 74 subunits of a 20 085 Da monomer. Charge was limited to between 5 and 84. In addition to the charge-smooth filter, a mass-smooth filter was applied to smooth the distribution of dimer units.

Nanodiscs with 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and POPC were analyzed with a linear drift cell Waters Synapt G1 ion mobility-mass spectrometer.[29] Data was deconvolved without a charge filter but by using a mass filter to smooth the distribution of lipids. Masses were limited to between 100 and 175 kDa. Conversion from arrival time to collision cross section (CCS) was performed using the Mason-Schamp equation as described previously,[27,29] using t_0 values calibrated from alcohol dehydrogenase analyzed under the same instrumental conditions.

RESULTS AND DISCUSSION

Deconvolution of Mass Spectra. To assess the accuracy of UniDec, we first describe the application to membrane protein AqpZ with bound POPC lipid. The spectrum contains multiple lipid adducts, yet the individual charge states are easily discerned (Figure 1A−C).[27] We compared the UniDec result with the MaxEnt and Transform algorithms included in MassLynx software (Waters Corp.) as shown in Figure S-2, Supporting Information. Similar to UniDec, MaxEnt provides a deconvolved zero-charge mass spectrum with few user-defined starting parameters.[16,23,24] As seen in Figure S-2C, Supporting Information, the raw output from UniDec agrees closely with the MaxEnt result for the same input data and restraints. However, the MaxEnt algorithm generally takes around 2 orders of magnitude longer to arrive at the result (0.15 s for UniDec vs ∼30 s for MaxEnt). The Transform algorithm relies on a user-defined mass window to simply sum up isolated charge states with no deconvolution. The UniDec reconvolved output agrees closely with the Transform result as seen in Figure S-2E, Supporting Information. These results demonstrate that UniDec provides accurate, rapid, and quantitative deconvolution and transformation of spectra with well-defined charge states with minimal user input.

Deconvolution of Subunit Exchange Experiments. Turning to more complex spectra, we explored the utility of UniDec for quantifying the quaternary dynamics of proteins in a case where charge and mass distributions are overlapping. We examined HSP17.7, a dodecameric small heat-shock protein that is known to exchange subunits at equilibrium (Figure 2). When

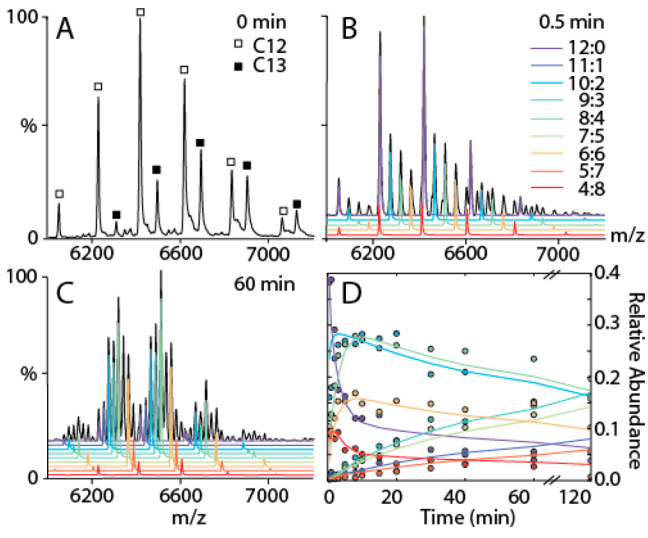


Figure 2. UniDec analysis of subunit exchange data for HSP17.7. The annotated mass spectrum is shown in (A) for the zero time point, (B) at 0.5 min, and (C) at 60 min. Extracted intensities for species annotated in (B) are shown in (D) as a function of equilibration time along with kinetic fits.

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

HSP17.7 labeled metabolically with ^{13}C (heavy) is mixed with its unlabeled (light) equivalent, two distinct charge series are observed, corresponding to homododecamers (Figure 2A).

Over time, continuous dissociation and reassociation of monomers and dimers results in the gradual appearance of peaks corresponding to hetero-oligomers comprising different numbers of light and heavy subunits.[30] At early time points, dodecamers composed predominately of an even number of both heavy and light subunits are observed (Figure 2B). However, after 60 min of incubation, 12-mers comprising odd numbers of both types of subunit are observed (Figure 2C). While subunit exchange is clearly occurring, overlapping peaks make it challenging to quantify directly the time-dependent evolution of the intensities of these various species, even in these well-resolved spectra.

Using UniDec for deconvolution, we quantified the relative abundances of all the species in the evolving mixture at each time point. We fit the extracted intensities using a previously described kinetic exchange model[31,32] that describes the sequential dissociation and reassociation based on separate rates for dissociation of dimers and monomers from the dodecamers (Figure 2D). These fits allow the extraction of kinetic constants, returning k_{off} = 0.0026 ± 0.0002 s^{-1} and 0.44 ± 0.07 s^{-1} for monomer and dimer, respectively. These results provide quantitative insight into how the dynamic oligomers interconvert, confirming that the monomer exchange is much slower than the dimer exchange. Such an analysis would not be possible without a reliable method to extract relative contributions of the various oligomers involved, which is provided here by UniDec.

Although analysis of overlapping spectra like this may be possible using simulation algorithms, UniDec streamlines the pipeline from raw data to kinetic fits. Because the peak shape is relatively constant, a single set of parameters is used for all of the fits, and the deconvolution does not need to be seeded with a starting distribution. Moreover, each deconvolution takes around 0.1 s, so an entire time series can be converted to relative concentrations in a matter of seconds. As experimental approaches continue to expand in both dimension and scale, UniDec enables analysis of large data sets, for example, taking triplicate measurements of kinetics as a function of temperature and pH, without additional computational overhead or extensive manual manipulation.

Deconvolution of CID Spectra. In addition to monitoring dynamic processes that occur in solution such as subunit exchange, there is an increasing demand to quantify dynamic processes that happen in the mass spectrometer as a result of collision induced unfolding and dissociation.[27,33,34] Collision-induced dissociation (CID) is used to study polydisperse ensembles because the loss of a highly charged monomer leaves behind a charge-stripped oligomer. The lower average charge increases the spacing between oligomers and reduces overlap between different species in the spectrum.[28,34] Mass spectra from CID experiments pose a particular challenge for quantitative deconvolution algorithms because reconstructing the underlying distribution of oligomers requires not only deconvolution of the mass components but also knowledge of the number of subunits lost to produce each stripped complex. Although this is often trivial for CID spectra of monodisperse protein complexes, it is challenging for polydisperse proteins where oligomeric distributions overlap between native and stripped complexes.

We employed the charge and mass separation afforded by UniDec to analyze CID experiments. Figure 3A shows a combined CID spectrum of the isolated 47+ charge state of

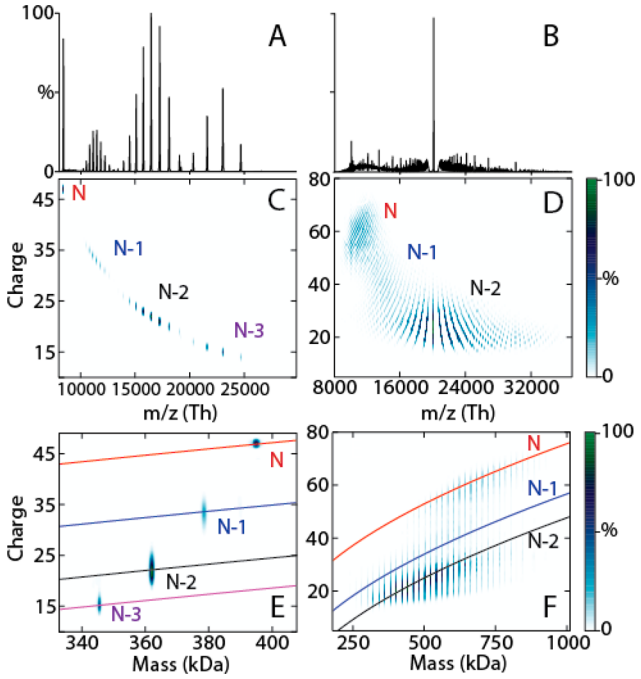


Figure 3. UniDec analysis of MjHSP16.5 (A, C, and E) and αB-crystallin (B, D, and F) CID spectra. Spectra combined from a number of collision voltages are shown in (A) and (B). Results in m/z vs charge dimensions are presented in (C) and (D) and in mass vs charge dimensions in (E) and (F). Lines plotted correspond to an offset from the predicted average native charge as a function of mass.

the monodisperse native HSP16.5 24-mer (N). Successive dissociation steps, which proceed by removal of monomers, lead to the appearance of 23 (N-1), 22 (N-2), and 21-mers (N-3) at higher m/z as annotated in Figure 3C.

UniDec deconvolution of the combined spectrum shows a lower average charge for each dissociation product as shown in Figure 3E. To assign the CID state of a particular stoichiometry, we fit the deconvolution results to an empirical relationship between mass and Z_A, the average charge for a complex in units of elementary charge: $Z_A = 0.0467m^{0.533} + F$ where m is the mass in Da and F is an offset from the native state.[11,12] For the HSP16.5 CID spectrum (Figure 3E), the +47 charge state has an offset of +2, indicating that the isolated state has a native charge. The dissociation products are well-separated with charge offsets of −10, −21, and −27, an effect caused by monomer units taking charges from the complex as they dissociate. Comparing the intensity of each CID state extracted as a function of collision voltage (Figure S-3A, Supporting Information) with manually extracted results[28] demonstrates that UniDec accurately extracts both charge distribution and intensities on monodisperse CID spectra.

Separation of the charge dimension now provides a novel avenue for exploring CID spectra from highly polydisperse proteins that cannot be routinely analyzed manually. Figure 3B shows a combined CID spectrum of polydisperse αB-crystallin. The m/z regions of each dissociation step were previously identified[34] but overlap in m/z (Figures 3D and 4A). However, deconvolution reveals distinct regions for native oligomers and the two successive dissociation steps as shown in Figure 3F. Using the relationship described above, we plotted the mass vs average charge relationship in Figure 3F with offset terms of +2, −17, and −26 for the native (N), singly stripped (N-1), and doubly stripped (N-2) oligomers, respectively. Summing

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

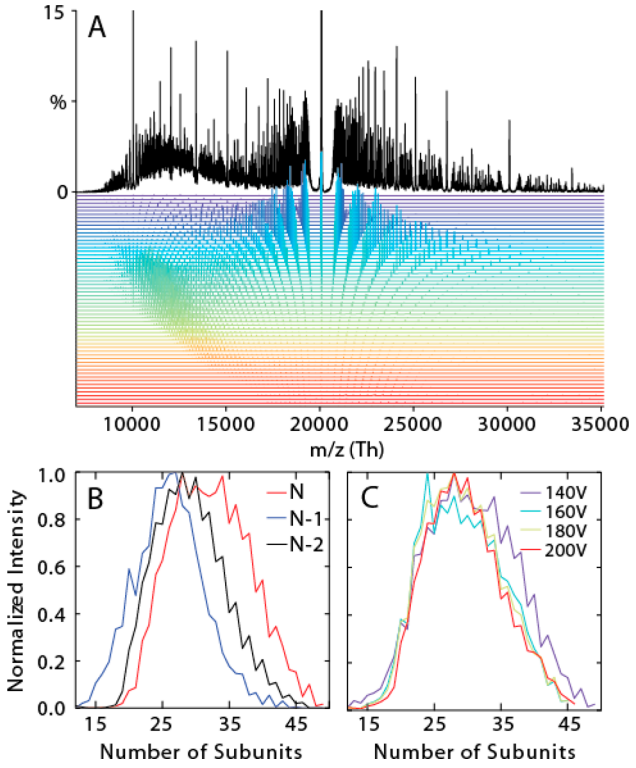


Figure 4. Deconvolution of αB-crystallin CID spectra. Composite mass spectrum (black, Figure 3B) over a range of CID voltages is presented in (A) with the separated oligomeric and charge states colored beneath. Oligomeric states range linearly from 12 at the top (purple) to 66 at the bottom (red). Extracted distributions, corrected for subunit loss, are presented in (B) for the native state (N) at 140 V (red), the first dissociation (N-1) at 160 V (blue), and the second dissociation (N-2) at 200 V (black), showing that the dissociated distributions are shifted toward smaller complexes. The overall oligomeric distribution is reconstructed at different collision energies by correcting for the loss of one monomer unit per dissociation step (C).

intensities along these charge curves for each dissociation state (plus or minus one charge state) yields the relative intensity of each isolated CID state (Figure S-3B, Supporting Information). Despite the large charge and m/z range, each spectrum deconvolved in 2−3 s.

The ability to define the CID state in mass vs charge space allows us to interrogate how the extracted distribution depends on the collision voltage. This is a key question because differing CID efficiencies between large and small oligomers could bias the extracted distribution in the CID products toward smaller oligomers. Using UniDec, we extracted the oligomeric distribution of αB-crystallin separately from the N, N-1, and N-2 oligomers at 140, 160, and 200 V, respectively (Figure 4B). This reveals that native oligomers have a slightly larger average number of subunits compared to the stripped oligomers. Because the overall distribution and thus the transmission efficiency is constant, the selective depletion of the smaller oligomeric complexes confirms that the larger complexes are more stable to CID, likely due to more internal degrees of freedom over which to distribute collisional energy.[35]

By correcting the stoichiometries to account for the monomer units lost by dissociation, we can fully reconstruct the initial distribution of oligomer sizes. Comparing the combined distribution across all CID states (Figure 4C), we observe that the extracted distribution is relatively constant from 140 to 200 V, despite transitioning from nearly fully N to almost completely populating the N-2 state (Figures S-3B, Supporting Information). By separating the mass and charge dimensions, UniDec enables the monitoring of CID in highly complex and overlapping spectra that is not possible through conventional data analysis methods.

Ion Mobility Mass Spectra. IM-MS provides an additional dimension to conventional MS by recording not only the m/z ratio but also the time taken to traverse a region of inert gas under the influence of a weak electric field. As with MS, the arrival time measured in IM is related to a ratio of two physical quantities, in this case the collision cross section (CCS) of the ion divided by its charge.[29,36] However, the experiment does not report on the CCS/z ratio directly. Arrival times must be converted into CCS using both experimental parameters and the charge on each ion. Charges in IM-MS spectra are usually assigned using the m/z dimension, but this is challenging in complex and overlapping spectra, where more sophisticated experimental approaches are required.[37]

UniDec provides a unique approach to interpreting IM-MS data by performing a simultaneous deconvolution of the m/z and arrival time dimensions. Examining the AqpZ spectrum with bound POPC presented in Figure 1D, we used UniDec to separate the IM-MS data into its constituent components. Here, the two-dimensional m/z vs arrival time data (Figure 1D) deconvolves into a three-dimensional m/z vs arrival time vs charge output (Figure 1E and Figure S-4A, Supporting Information). Because charge is known for each point in the matrix, the experimental observables of m/z and arrival time can be transformed into a three-dimensional matrix of physical quantities, mass vs CCS vs charge (Figure 1F).[29] Figure S-3B, Supporting Information, shows the projection along the mass axis into a charge vs CCS matrix. Here, the lower charge states remain in their native conformation while the higher charge states are unfolded as previously observed.[27,38] UniDec provides a rapid (around 1 s) deconvolution and transformation of the data with little user intervention.

Deconvolution of Nanodisc Ion Mobility Spectra. To test UniDec against complex IM-MS data with overlapping mass, charge, and CCS distributions, we explored the deconvolution of IM-MS spectra of Nanodiscs, which are nanoscale discoidal lipoprotein complexes similar to high-density lipoprotein particles.[39] Each Nanodisc complex contains two copies of the membrane scaffold protein belt and a number of lipids. Although the complexes are relatively monodisperse, there is an intrinsic distribution in the number of lipids per complex greater than ±5 lipids.[9] Overlap between this mass distribution and the charge state distribution makes spectral assignment of Nanodiscs challenging, especially because there is resonant overlap at defined m/z values.[9] An example IM-MS spectrum is shown in Figure 5A, and representative spectra as a function of collision voltage are provided in Supplementary Movies S-1 and S-2, ac5b00140_si_002.avi and ac5b00140_si_003.avi, respectively, Supporting Information. These spectra demonstrate the same constructive overlap pattern observed in conventional MS.[9] Applying UniDec to simultaneously deconvolve the m/z and arrival time dimensions, we discovered that UniDec is able to separate the underlying mass, charge, and CCS distributions (Figure 5D).

Despite the complexity and dimensionality, deconvolution only takes around 1−3 s per spectrum. A series of spectra were analyzed with increasing collision voltage to determine mass and CCS distributions as a function of collision voltage. The total

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

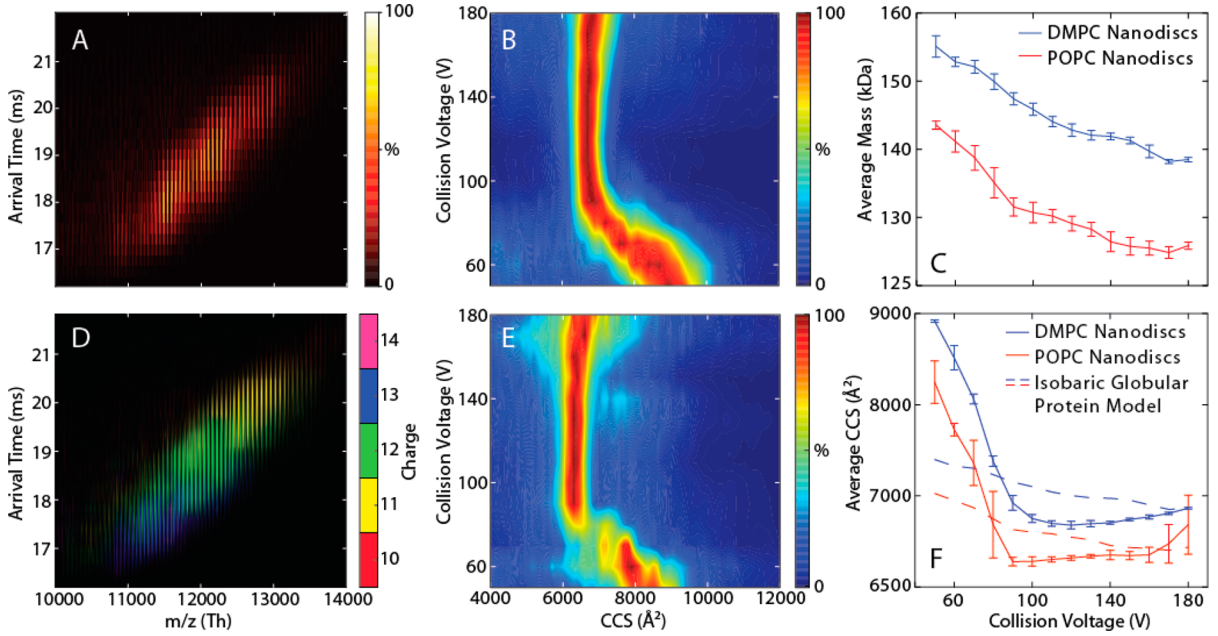


Figure 5. IM-MS deconvolution of Nanodiscs. The IM-MS spectrum of MSP1D1(−) DMPC Nanodiscs at 120 V CID is shown in (A), where the color bar indicates intensity. (D) The separation of the spectrum into its individual charge states. Here, color indicates the charge state and transparency indicates the intensity. Total CCS summed across masses and charge states is presented as a function of collision voltage for the average of three replicates for DMPC and POPC Nanodiscs in (B) and (E), respectively. Dissociation of lipids is shown for DMPC (blue) and POPC (red) Nanodiscs in (C), and the corresponding CCS collapse is presented in (F). Error bars are shown as the standard deviation of the mean of three replicate measurements. Dashed lines in (F) indicate the predicted CCS for a globular protein with masses from (C).

CCS distribution summed across all charge and mass states is shown for DMPC and POPC Nanodiscs (Figure 5B,E, respectively) with the average mass and CCS from three replicate measurements plotted in Figure 5C,F. Error bars report the standard deviation of the mean, demonstrating that the distributions are reproducible.

Knowledge of the mass and CCS distributions allowed us to examine the gas-phase behavior of Nanodiscs. The initial mass distributions agree with solution studies, showing 164 DMPC molecules and 131 POPC molecules in addition to the two membrane scaffold protein belts.[40,41] Using a scaled projection approximation method,[42] we calculated predicted CCS values based on a molecular dynamics model[43] and small-angle X-ray scattering data of Nanodisc complexes in solution (see Supplemental Methods, Supporting Information, for more detail).[44,45] We discovered that the predicted CCS for DMPC Nanodiscs of 8560 Å^2 agrees with the experimental value of 8920 Å^2 with a peak width (one standard deviation) of 800 Å^2 at a collision voltage of 50 V. POPC Nanodiscs have a predicted CCS of 8300 Å^2, which agrees closely with experimental values of 8250 ± 770 Å^2 at 50 V. These data suggest that Nanodiscs at low collision energy have not only masses that agree with the solution values but also CCS values that agree with a discoidal shape.

As the collision voltage increases, both POPC and DMPC Nanodiscs undergo CID and show both a progressive loss of lipids[9] and a corresponding decrease in CCS. To examine whether the decrease in CCS results from a structural rearrangement or is simply a consequence of the reduced mass, we calculated predicted CCS values for globular proteins of equivalent masses,[29,42] shown as dashed lines in Figure 5F. Nanodiscs begin with a structure larger than an isobaric globular protein. This is expected because lipids show a larger average CCS at a given mass than peptides.[46] Moreover, an idealized discoidal ellipsoid will have a larger CCS than an idealized sphere of the same volume.[47]

Although the loss of lipids contributes partially to the decrease in CCS, as evidenced by the downward slope in the dashed lines in Figure 5F, the mass loss alone is not sufficient to account for the CCS change. This suggests that a structural rearrangement produces collapse of the Nanodisc into a compact structure.

In contrast, most proteins show an increase in CCS with increased collision voltage due to protein unfolding, although certain topologies may show a slight collapse prior to unfolding.[38] The solution structure of Nanodiscs, however, contains no empty spaces to collapse into, suggesting this collapse is more likely the result of a decreasing shape factor. Given the changes observed in the shape factors of geometric solids,[47] we expect that this compact structure is more spherical than the initial discoidal ellipsoid.

Because UniDec enables deconvolution of both the m/z and arrival time dimensions, we are able to extract the CCS as a function of collision voltage and observe the structural change from a native-like to collapsed shape that is unseen in the mass spectra and very difficult to infer from the arrival times alone. This unusual behavior can only be revealed in systems as heterogeneous and complex as Nanodiscs if the charge is known for all relevant peaks, demonstrating the utility of UniDec to deconvolute complex IM-MS spectra.

■ CONCLUSION

UniDec is a fast, robust, and flexible approach to (IM-)MS data analysis. We have demonstrated its utility for systems of varying complexity, including membrane proteins, subunit exchange intermediates, and collision induced dissociation of polydisperse protein and lipoprotein complexes. UniDec robustly handles overlapping charge state distributions as well as multidimensional IM-MS data. The universal nature of the algorithm will

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376

show immediate utility in a wide range of MS experiments, including ligand binding, subunit exchange, and "top-down" experiments. As the complexity of systems expands, we anticipate that UniDec will become indispensable for analysis of multi-dimensional mass and ion mobility spectra.

■ ASSOCIATED CONTENT

Supporting Information

Theory describing algorithm, details on computational implementation, experimental methods, appendix describing Bayesian derivation of algorithm, supplemental figures, and movies of Nanodisc ion mobility-mass spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: carol.robinson@chem.ox.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Eamonn Reading and Andrew Aquilina for help in acquiring data for AqpZ and αB-crystallin, respectively, and Eman Basha and Elizabeth Vierling for providing HSP17.7. The authors thank Tim Allison for helpful discussions and code contributions. A.J.B. is a BBSRC David Phillip's fellow. E.G.M. is supported by Carl Trygger's Foundation for Scientific Research. G.K.A.H. is supported by an EPSRC studentship. J.L.P.B. holds a University Research Fellowship of the Royal Society. M.T.M. is funded by program grant G1000819 from the Medical Research Council. C.V.R. is funded by European Research Council Investigator Award (IMPRESSS, grant no. 268851).

■ REFERENCES

(1) Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem.* **2007**, *76*, 167−193.
(2) Marcoux, J.; Robinson, C. V. *Structure* **2013**, *21*, 1541−1550.
(3) Heck, A. J. R. *Nat. Methods* **2008**, *5*, 927−933.
(4) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. *Nat. Chem.* **2014**, *6*, 281−294.
(5) van den Heuvel, R. H. H.; Heck, A. J. R. *Curr. Opin. Chem. Biol.* **2004**, *8*, 519−526.
(6) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175−186.
(7) van Duijn, E. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 971−978.
(8) Mann, M.; Meng, C. K.; Fenn, J. B. *Anal. Chem.* **1989**, *61*, 1702−1708.
(9) Marty, M. T.; Zhang, H.; Cui, W. D.; Gross, M. L.; Sligar, S. G. *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 269−277.
(10) Morgner, N.; Robinson, C. V. *Anal. Chem.* **2012**, *84*, 2939−2948.
(11) Stengel, F.; Baldwin, A. J.; Bush, M. F.; Hilton, G. R.; Lioe, H.; Basha, E.; Jaya, N.; Vierling, E.; Benesch, J. L. P. *Chem. Biol.* **2012**, *19*, 599−607.
(12) Stengel, F.; Baldwin, A. J.; Painter, A. J.; Jaya, N.; Basha, E.; Kay, L. E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. P. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2007−2012.
(13) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *J. Am. Soc. Mass. Spectrom.* **2000**, *11*, 320−332.
(14) Tseng, Y.-H.; Uetrecht, C.; Yang, S.-C.; Barendregt, A.; Heck, A. J. R.; Peng, W.-P. *Anal. Chem.* **2013**, *85*, 11275−11283.
(15) Zheng, H.; Ojha, P. C.; McClean, S.; Black, N. D.; Hughes, J. G.; Shaw, C. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 429−436.
(16) Reinhold, B. B.; Reinhold, V. N. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 207−215.
(17) Hagen, J. J.; Monnig, C. A. *Anal. Chem.* **1994**, *66*, 1877−1883.
(18) Zhang, Z.; Marshall, A. G. *J. Am. Soc. Mass. Spectrom.* **1998**, *9*, 225−233.
(19) Fernandez-de-Cossio Diaz, J.; Fernandez-de-Cossio, J. *Anal. Chem.* **2012**, *84*, 7052−7056.
(20) Sivalingam, G. N.; Yan, J.; Sahota, H.; Thalassinos, K. *Int. J. Mass Spectrom.* **2013**, *345*−*347*, 54−62.
(21) van Breukelen, B.; Barendregt, A.; Heck, A. J. R.; van den Heuvel, R. H. H. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2490−2496.
(22) Hilton, G. R.; Hochberg, G. K.; Laganowsky, A.; McGinnigle, S. I.; Baldwin, A. J.; Benesch, J. L. *Philos. Trans. R. Soc. London, Ser. B: Biol. Sci.* **2013**, *368*, 20110405.
(23) Ferrige, A. G.; Seddon, M. J.; Green, B. N.; Jarvis, S. A.; Skilling, J.; Staunton, J. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 707−711.
(24) Ferrige, A. G.; Seddon, M. J.; Jarvis, S.; Skilling, J.; Aplin, R. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 374−377.
(25) Lucy, L. B. *Astron. J.* **1974**, *79*, 745.
(26) Richardson, W. H. *J. Opt. Soc. Am.* **1972**, *62*, 55−59.
(27) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. *Nature* **2014**, *510*, 172−175.
(28) Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13*, 597−605.
(29) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557−9565.
(30) Painter, A. J.; Jaya, N.; Basha, E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. P. *Chem. Biol.* **2008**, *15*, 246−253.
(31) Benesch, J. L. P.; Aquilina, J. A.; Baldwin, A. J.; Rekas, A.; Stengel, F.; Lindner, R. A.; Basha, E.; Devlin, G. L.; Horwitz, J.; Vierling, E.; Carver, J. A.; Robinson, C. V. *Chem. Biol.* **2010**, *17*, 1008−1017.
(32) Baldwin, A. J.; Lioe, H.; Robinson, C. V.; Kay, L. E.; Benesch, J. L. P. *J. Mol. Biol.* **2011**, *413*, 297−309.
(33) Zhong, Y.; Han, L.; Ruotolo, B. T. *Angew. Chem., Int. Ed.* **2014**, *53*, 9209−9212.
(34) Aquilina, J. A.; Benesch, J. L. P.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. *Proc. Natl. Acad. Sci.* **2003**, *100*, 10611−10616.
(35) Benesch, J. P. *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 341−348.
(36) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protoc.* **2008**, *3*, 1139−1152.
(37) Shepherd, D. A.; Marty, M. T.; Giles, K.; Baldwin, A. J.; Benesch, J. L. P. *Int. J. Mass Spectrom.* **2015**, *377*, 663−671.
(38) Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V. *J. Am. Chem. Soc.* **2012**, *134*, 3429−3438.
(39) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853−856.
(40) Marty, M. T.; Zhang, H.; Cui, W. D.; Blankenship, R. E.; Gross, M. L.; Sligar, S. G. *Anal. Chem.* **2012**, *84*, 8957−8960.
(41) Bayburt, T. H.; Sligar, S. G. *FEBS Lett.* **2010**, *584*, 1721−1727.
(42) Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. P. *Structure* **2015**, DOI: 10.1016/j.str.2015.02.010.
(43) Shih, A. Y.; Denisov, I. G.; Phillips, J. C.; Sligar, S. G.; Schulten, K. *Biophys. J.* **2005**, *88*, 548−556.
(44) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477−3487.
(45) Denisov, I. G.; McLean, M. A.; Shaw, A. W.; Grinkova, Y. V.; Sligar, S. G. *J. Phys. Chem. B* **2005**, *109*, 15580−15588.
(46) McLean, J. A. *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 1775−1781.
(47) Hewitt, D.; Marklund, E.; Scott, D. J.; Robinson, C. V.; Borysik, A. J. *J. Phys. Chem. B* **2014**, *118*, 8489−8495.

DOI: 10.1021/acs.analchem.5b00140
Anal. Chem. 2015, 87, 4370−4376