An Algorithm for Building Multi-State Classifiers Based on Collision Induced Unfolding Data

Daniel A. Polasky, Sugyan M. Dixit, Daniel D. Vallejo, Kathryn D. Kulju, and Brandon T. Ruotolo.*

University of Michigan Department of Chemistry, Ann Arbor, Michigan 48109, United States.

ABSTRACT: Collision-induced unfolding (CIU) has emerged as a valuable method for distinguishing iso-crossectional protein ions through their distinct gas-phase unfolding trajectories. CIU shows promise as a high-throughput, structure-sensitive screening technique with potential applications in drug discovery and biotherapeutic characterization. We recently developed a CIU classification workflow to support screening applications that utilized CIU data acquired from a single protein charge state to distinguish immunoglobulin (IgG) subtypes and membrane protein lipid binding. However, distinguishing highly similar protein structures, such as those associated with biotherapeutics, can be challenging. Here, we present an expansion of this classification method that includes CIU data from multiple charge states, or indeed any perturbation to protein structure that differentially affects CIU, into a combined classifier. Using this improved method, we are able to improve the accuracy of existing, single state classifiers for IgG subtypes and develop an activation-state sensitive classifier for selected Src kinase inhibitors when data from a single charge state was insufficient to do so. Finally, we employ the combination of multiple charge states and stress conditions to distinguish a highly similar innovator/biosimilar biotherapeutic pair, demonstrating the potential of CIU as a rapid screening tool for drug discovery and biotherapeutic analysis.

Native mass spectrometry (MS) and ion mobility-mass spectrometry (IM-MS) have been increasingly adopted techniques for the determination of protein-protein and protein-ligand contacts, stoichiometry, and shape.[1–3] Native IM-MS has seen rapid growth in the characterization of proteins[4,5] protein-ligand complexes,[6,7] and multi-protein complexes.[8] A significant challenge in these analyses remains the relatively low resolution of IM in the context of protein structure, limiting the ability of IM-MS to distinguish subtle, but biologically relevant, conformational variations that occur below the resolution limits of modern instrumentation. The activation of protein ions in the gas phase prior to IM separation in an effort to follow their subsequent structural transitions represents a useful method to distinguish such structural differences. This approach, termed collision-induced unfolding (CIU) when the ion activation is accomplished using collisions with an inert gas, has a rich history in the IM-MS analysis of protein structure[9–11] and has seen rapid growth for drug discovery[12–14] and biotherapeutic characterization.[15–20] The relative speed of CIU, combined with detailed comparative structure information, make it a promising technique for the development of structure-sensitive screening methods at medium to high throughput.

A number of reports have demonstrated proof-of-principle methods using CIU to distinguish ligand binding sites for kinases inhibitors,[13,14,21] quantifying cooperative binding of ligands within a protein complex,[22] and detecting protein allostery upon ligand attachment.[23] Screening approaches sensitive to these structural parameters are in great demand for a wide range of applications associated with protein biophysics. The relative comparison of CIU fingerprints under different conditions, for example following ligand binding to a target protein or after applying heat stress to a biotherapeutic, enables the determination of useful information about the structure of a protein and its response to perturbations.

Converting the complex datasets generated in CIU experiments into this structural information requires robust statistical methods. Several recent reports have developed quantitative methods to compare CIU fingerprints in support of these analyses.[24–29] For screening workflows in particular, supervised learning approaches show great promise. In these methods, "training" CIU data is acquired using known standards and used to generate a classifier that can then distinguish unknown CIU data. We recently developed CIUSuite 2, a software package that includes an automated workflow to construct classifiers for CIU data.[26] This approach was used to differentiate ligand and lipid binding modes in a membrane protein system[30] and shows promise for high-throughput screening and characterization of biotherapeutics.

Despite these successes, the current method is limited to the comparison of a single charge state of CIU data and relatively small quantities of training data. Native IM-MS experiments using electrospray ionization (ESI) typically generate multiple charge states, each with a unique CIU fingerprint. Recent work has demonstrated the benefits of including CIU information from multiple charge states in distinguishing the structures of monoclonal antibodies.[31] The incorporation of all information available from multiple charge states provides, in principle, great potential for improving CIU classification and screening methods without increasing data acquisition time. In this report, we describe the creation of a supervised classification algorithm that can accommodate CIU data from multiple protein 'states,' improve processing speed to enable processing of large datasets, and expand the scope of the classification workflow to include comparative analyses that move beyond the concept of using a single group of charge states

alone. We demonstrate the utility of these approaches to characterize ligand binding modes in a protein-inhibitor context and in distinguishing a highly similar innovator/biosimilar pair of biotherapeutic monoclonal antibodies.

Methods

Sample Preparation. SiLuLite SigmaMab Universal antibody standard, IgG1λ, and IgG4λ from human myeloma were purchased from Sigma-Aldrich and supplied as lyophilized powder (St. Louis, MO). Samples were reconstituted using Milli-Q water (Millipore) to a concentration of 2 mg/mL unless specified otherwise. Avastin® (Genentech, 25 mg/mL) and Avegra® (Biocad, 25 mg/mL) were purchased and supplied in solution formulation (158.6 mM Trehalose dehydrate, 40.9mM Sodium Phosphate, 0.16% Polysorbate 80, pH 6.2). Biotherapeutic samples were diluted to 1mg/mL using 0.9% bacteriostatic sodium chloride injection, USP. (Pfizer Inc. New York City, NY). Stressed samples were incubated at 40 ºC with 250 RPM orbital shaking for 4 weeks. All antibody samples were buffer exchanged into 200 mM ammonium acetate buffer using Micro Bio-spin 30 columns (Bio-Rad, Hercules, CA). Buffer exchanged samples were then diluted to a working concentration of 1 mg/mL (~6.7 μM).

Src kinase domain DNA was synthesized by GeneArt (Life Technologies, Grand Island, NY) using E. coli modified codons and subcloned into pET28a with a modified TEV-protease cleavable N-terminal 6x-His tag. The plasmid was transformed by electroporation into BL21 DE3 electrochemically competent cells with a YopH in pCDFDuet-1. Cell growth, protein expression, and purification were adapted from protocols previously developed for the c-Src kinase domain[32] without cleavage of the His-tag. Dasatinib, staurosporine, foretinib, and ponatinib were purchased from LC Laboratories (Woburn, MA). Protein was reconstituted and buffer exchanged into 200 mM ammonium acetate (Sigma-Aldrich, St. Louis, MO) at pH 7.0 using Micro Bio-Spin 6 columns (BioRad, Hercules, CA) to a final concentration of 10 μM. Samples were incubated at a ratio of 3:1 inhibitor:protein, on ice for 15 minutes prior to analysis by IM-MS.

CIU Acquisition. All CIU data were acquired using a Synapt G2 quadrupole-ion mobility-time-of-flight mass spectrometer (Q-IM-ToF MS) instrument (Waters, Milford, MA). Sample was transferred to a gold-coated borosilicate capillary needle (prepared in-house), and ions were generated by direct infusion using a nano-electrospray ionization (nESI) in positive mode. The electrospray capillary was operated at voltages of 1.5-1.7 kV with the sampling cone at 40 V. The backing pressure was set to 7.9-8.1 mbar for antibody samples or 5.0 mbar for kinase samples. The trap collision cell was pressurized to 4-5 $\times 10^{-2}$ mbar of argon gas, helium cell flow to $1.4\text{x}10^{3}$ mbar, traveling-wave IM separator to 3.4 mbar, and ToF MS to 1.5×10^{-6} mbar. IM wave height and wave velocity were 20 V and 150 m/s, respectively, for Src kinase domain or 40 V and 600 m/s for antibodies. CIU experiments were performed by ramping the collision voltage in the trap cell from 5 to 200 V (antibodies) or 10 to 125 V (Src kinase) in 5 V increments with a dwell time of 6 s at each collision voltage.

Data Processing and Classification. IM arrival time data was extracted from raw data for each charge state using TWIMExtract[33] and smoothed with CIUSuite 2[26] (Savitzky-Golay 2D smoothing, window 5, 2 iterations). An updated version of the CIUSuite 2 classification interface that recognizes user-specified labels across multiple states (*e.g.* charge states) was used to assemble the training data for each classifier. Note that all states used must be present in all training data used; for example, if charge states shift as a result of variations in ESI over the training data acquisition period, only charge states present in all inputs can be used as states for classification. Classifiers were generated in 'all data' mode with cross validation test sizes of 6, 1, and 3 for data presented in Figure 1, 2, and 3, respectively (adjusted to be approx. 1/3 of input dataset size in each case). Input data for Figure 1 and Figure 2 was normalized but not standardized; input data for Figure 3 was both normalized and standardized. The classification algorithm presented here is based on the original CIUSuite 2 algorithm, utilizing the 'scikit-learn' Python library,[34] with the following key differences: support for division of the input data into multiple states throughout the classification, addition of data standardization to improve classifier performance, and implementation of random sampling cross validation to allow large input training datasets to be used without prohibitive memory and computation costs. Input training data is standardized within each state and collision voltage by scaling to zero mean and unit variance. For input Gaussian data, each attribute of each Gaussian peak (centroid, width, and amplitude) is standardized separately because the initial values for centroids are typically much larger than those for width or amplitude. Thus, centroids are only standardized with centroids, widths with widths, and amplitudes with amplitudes to prevent one attribute from overwhelming the others as a result of larger input values. Standardized and labeled training data for each state is assessed separately by the univariate feature selection (UFS) method in CIUSuite 2, which uses ANOVA F-values to assess the variation within and between classes at each collision voltage. The highest scoring collision voltages are then chosen from amongst all states for cross validation and final classifier construction, meaning that a classifier can contain data from multiple states.

Cross validation is performed by holding back a portion of the training data (of configurable size), constructing a classifier with the remaining training data, then testing the withheld data (the 'test' data) to see if it is classified correctly. This accuracy can be used to assess the performance of the classifier, with high accuracies indicating a classifier capable of robustly distinguishing between the provided classes. Typically, high accuracy classifiers are those with several collision voltages that achieve high scores with low variation, as this indicates regions of the fingerprint that can reliably distinguish between classes. As in CIUSuite 2, cross validation involves adding 'features' in decreasing order of UFS score to determine the number of features that results in the most accurate classifier. In the workflow describe here, the features represent a single collision voltage from one of the states, so a particular voltage can be included multiple times if it scores highly in multiple states. The original CIUSuite 2 cross validation method tested all possible permutations of training and test data from an input dataset, which resulted in exponential time and memory cost with increasing dataset size and proved prohibitive for the larger datasets evaluated in this work. Random sampling from the possible input permutations was implemented to reduce this to a linear increase in performance cost by sampling only a user-specified number of the possible permutations, chosen at random. Following determination of the optimal number of features to include, final classifiers are generated as in CIUSuite 2.

Results and Discussion

Each charge state observed in a native IM-MS experiment undergoes a substantially different unfolding trajectory during CIU, providing potentially complementary information for a multi-state CIU-based classifier. To evaluate the utility of combining data from multiple charge states for CIU classification, we compared data acquired for the monoclonal antibody sub-classes IgG1 and IgG4, which differ only slightly in disulfide bonding pattern (Figure 1A). The native mass spectrum of IgG1 shows charge states from 22-26$^+$, with 24$^+$ being the most abundant (Figure 1B). The CIU fingerprints of IgG1 and IgG4 at the 24$^+$ charge state are quite similar, aside from minor differences in the second CIU feature in the range of 60-80 V (Figure 1A, bottom). Performing a single charge state comparison using the 24$^+$ charge state only, as would be done in the original CIUSuite 2 workflow, results in a feature selection plot showing minor differences in the 60-80 V region as expected, with minimal differences outside that region (Figure 1C). The classifier that can be trained from this data is of relatively low quality, achieving a maximum cross validation accuracy of 82% when using two features (70 and 75 V, Figure 1E). Assessing all charge states with the classification workflow, however, reveals that the 24$^+$ charge state, despite being the highest signal in the mass spectrum, is not the optimal CIU data to differentiate these two antibodies.

To examine all charge states, we perform feature selection sequentially for each, meaning that the 22$^+$ charge state of IgG1 is compared to the 22$^+$ of IgG4, and such binary comparisons were replicated across all mAb charge states. This analysis results in five feature selection plots, which can be overlaid to evaluate the potential information content each charge state (Figure 1D). The 22$^+$ charge state has the two highest scoring individual voltages (black, 85 and 90 V), followed by 80 V in the 23$^+$ charge state (blue), then 75 and 70 V in the 24$^+$ charge state (green) (Figure 1F). As in a standard, single charge state classification mode, cross validation is performed by incorporating the data into classifiers in decreasing order of feature selection score; but in classifiers derived from multiple charge states, the input data can originate in any of the charge states included in the analysis. The cross validation indicates that the optimal classifier in this case uses four collision voltages, two from the 22$^+$ charge state and one each from the 23$^+$ and 24$^+$ charge states, to achieve an accuracy of 95%, significantly improved over the 82% accuracy achieved by the classifier using just the 24$^+$ charge state (Figure 1F).

To complete the comparison, we generated single charge state classifiers for all five charge states and compared the cross validation accuracy at the optimal number of collision voltages for each classifier (Figure 1G). Given the pair of very high scores from the 22$^+$ charge state in the feature selection, it is not surprising that it results in the best single charge state classifier, and indeed achieves slightly higher accuracy than the combined classifier that considered all charge states (95% vs 94%). While it would be surprising for an individual charge state to outperform the combined classifier, the difference in accuracy between the 22+ and all-state classifiers is well within error, indicating that the classifiers have similar overall performance. The 23-26$^+$ charge states each individually achieve accuracies in the 80-90% range, each lower than the 22$^+$ or combined classifiers. Combining these four charge states into a combined classifier results in accuracy of 95% (data not shown), matching the performance of the full combined classifier and the individual 22+ state classifier. As such, the all charge state classification in this case functions primarily as a means to rapidly identify the optimal charge state and ensure it is incorporated into the final classifier. Indeed, the 22$^+$ ions are the lowest intensity signals included in the analysis, and would not be an obvious choice if using only IM-MS precursor data. In cases where several charge states achieve similar feature selection scores, however, combining data from multiple charge states can generate a superior classifier to any individual charge state.

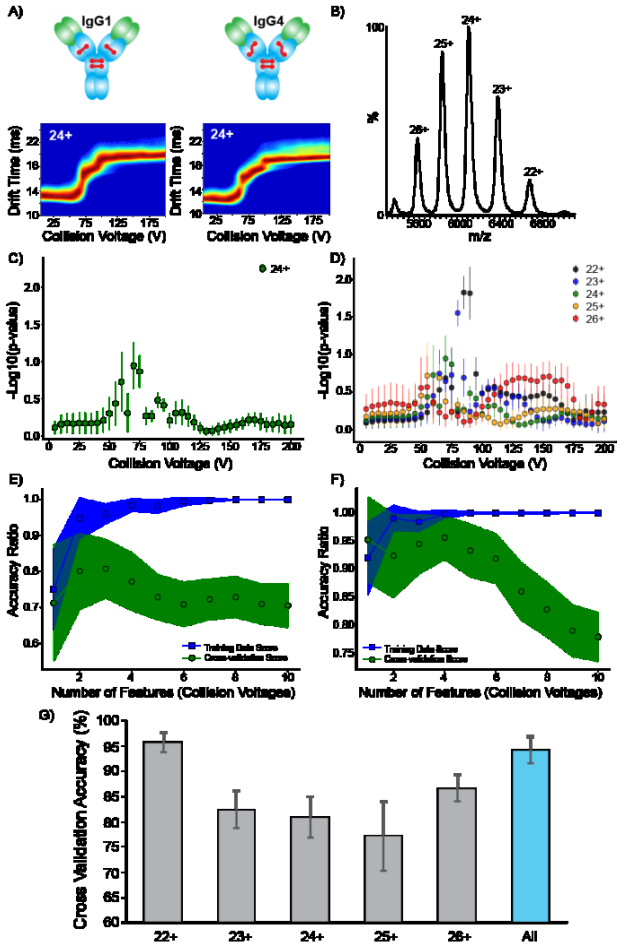

Figure 1. Multiple charge state classification of IgGs. A) IgG1 and IgG4 subtypes differ primarily in disulfide bond linkage, resulting in slightly different CIU fingerprints. B) Native mass spectrum of IgG1 with 22-26$^+$ charge states. C) UFS score plot distinguishing IgG1 and IgG4 at the 24$^+$ charge state only. D) UFS plot for all charge states of IgG1 and IgG4. E) Cross validation accuracies for 1-10 features from the 24$^+$ charge state alone and F) for all charge states incorporated into one classifier. G) Optimal cross validation accuracy from each individual charge state and all charge states combined.

We applied our multi-state classification workflow to a number of challenging proteins and complexes that had previously confounded CIU classification efforts using data from a single charge state. Src, a non-receptor protein tyrosine kinase, plays a key role in several cell signaling processes[35,36] and has been observed to be overexpressed in certain carcinomas and glioblastomas.[37] Several classes of inhibitors to kinases like Src are known to target different conformations of the kinase. Type I inhibitors like Dasatinib and Staurosporine bind to the active state, in which the DFG loop is in the "in" conformation, wrapping around the helices (green loop, Figure 2A).

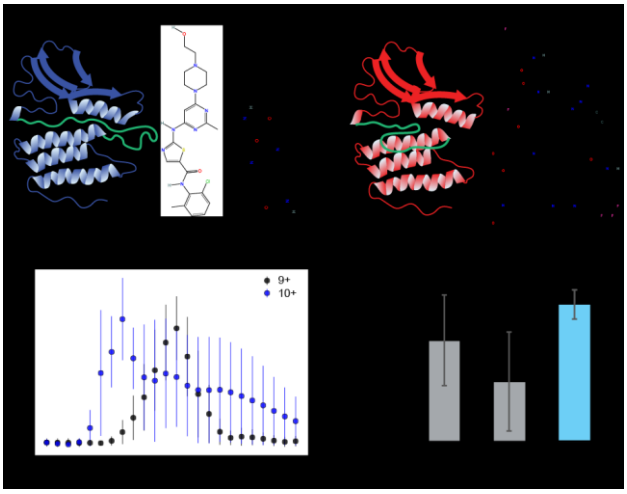

Figure 2. Multiple charge state classification of Src kinase. Type I and II kinase inhibitors target the active (A) or inactive (B) conformations of the kinase. C) UFS plot comparing Src CIU fingerprints with bound Type I (Dasatinib and Staurosporine) against Type II (Foretinib and Ponatinib) inhibitors at all charge states. D) Cross validation accuracy for individual charge state and combined classifiers using 2 features.

Type II inhibitors like Foretinib and Ponatinib bind the protein in the inactive conformation, in which the DFG loop is in the "out" conformation (green loop, Figure 2B). While single charge state classifiers and analogous methods have been successful in differentiating such tertiary structures within Abl,[13,25] a related kinase, differentiating these binding modes within Src using our previous single charge state classification method has proven challenging. Using the multi-state workflow developed here, we observe similar feature scores that distinguish Type I from Type II kinase inhibitors for both the 9^+ and 10^+ charge states (Figure 2C). As a result, the optimal classifier uses a single collision voltage each from 9^+ and 10^+, resulting in a cross validation accuracy of 98% (Figure 2D). Individual classifiers created using data from only 9^+ and 10^+ charge states separately exhibited lower accuracies in the 80-90% range (Figure 2D). The large error bars for the individual charge state classifier accuracies also indicate substantial uncertainty in their performance, with lower accuracy possible for external validation. Thus, the combined classifier using multiple charge states is superior in this case to any of the individual charge state classifiers, and enabled robust classification of ligand binding modes in a system that had proven challenging to classify with a single charge state alone.

Finally, we examined a biotherapeutic innovator/biosimilar pair, Avastin and Avegra, incorporating both multiple charge states and stress conditions into a multi-state classifier (Figure 3A). Assessing a biosimilar, or generic form of an innovator protein therapeutic, presents significant analytical challenges due to the typical size and complexity of monoclonal antibodies. Comparing higher order structure (HOS) information is particularly challenging without resorting to low-throughput, high-resolution structural biology techniques. As biosimilars, Avastin and Avegra are highly similar, and classification using CIU data across all charge states of the antibodies resulted in a low degree of differentiation, with the optimal classifier achieving accuracy of only 87% (Figure 3C). Charge states are not the only states that can be examined using our multi-state CIU data analysis algorithm, however. Our approach considers data acquired across any state that results in a different CIU

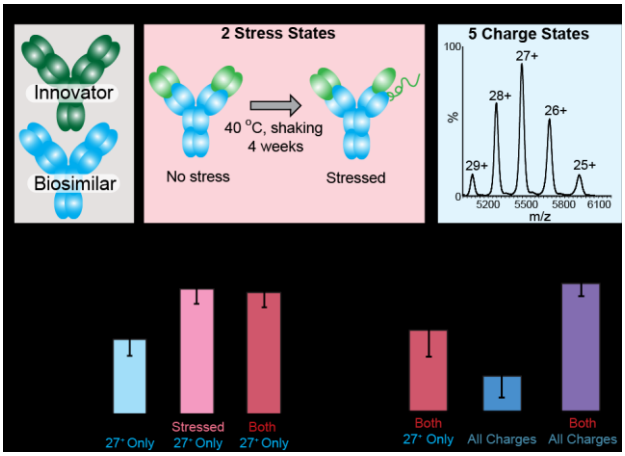

Figure 3. Stress-state classifiers distinguish Avastin and Avegra. A) Stress applied to innovator and biosimilar (center) and typical mass spectrum showing charge states. B) Best cross validation accuracy for classifiers using various stress states with a single charge state. C) Best cross validation accuracy for classifiers using all charges states of unstressed alone and both conditions (stressed and unstressed) Avastin and Avegra. The two-state classifier from panel B is included for comparison.

pathway, so long as it can be applied equally across the classes being compared. A key attribute monitored in biotherapeutics is the propensity to aggregate during transport and storage, which can be challenging to assess in the laboratory. Early warning methods for aggregation that detect structural changes following various types of stress (for example, heat or oxidation) are thus highly useful. Avastin and Avegra were stressed by heating to 40 ºC and applying orbital shaking for 4 weeks (Figure 3A). A two-state classifier was generated using the stressed and unstressed antibody data at a single charge state (27+, the most intense signal in the mass spectrum). The two-state classifier yielded improved performance (95% accuracy) compared to a single state classifier generated from the unstressed data alone (85% accuracy) (Figure 3B). However, the single state classifier for the stressed data resulted in similar performance to the two-state classifier from both stress states (Figure 3B). Finally, we combined both types of states into a multi-state classifier with 10 total states: all five charge states from each of the two stress states. This 10-state classifier achieved cross validation accuracy above 99% (Figure 3C), indicating very robust differentiation between Avastin and Avegra, outperforming the multi-state classifiers generated from all charge states of the unstressed data only, and from both stress states but a single charge state. A 5-state classifier using all charge states of only the stressed data performs similarly well, indicating that the stressed data is driving the performance of the combined, 10-state classifier. Our analysis indicates that Avastin and Avegra have different structural responses to the stress employed in this study, which can be utilized to develop a classifier capable of robustly distinguishing between them using our multi-state classification method and incorporating both charge states and stress states into classification.

Conclusions

CIU experiments generate rich datasets that have proven capable of distinguishing subtle differences in protein structures. Applying our multi-state classification workflow presented here to analyze all charge states observed within in a CIU experiment maximizes the detection of these subtle differences

by incorporating more of the experimental data into the statistical framework for classification. Improvements to the core algorithm have increased the accuracy of the classifiers developed through data standardization and have dramatically reduced the computational requirements for large datasets, enabling the extension of these algorithms to much larger training sets than analyzed previously. Finally, we demonstrate incorporating states other than protein charge states by generating a robust classifier to distinguish Avastin from its biosimilar Avegra by incorporating CIU data from heat-stressed samples. This work indicates the potential of the multi-state classification workflow to be used with a wide range of conditions or perturbations, as any change that causes differences in CIU for an analyte of interest can be incorporated into a classifier using this method. Incorporating differential responses to stimulus into CIU classification has the potential to make CIU sensitive to even more subtle structural differences and provide a rapid and informative workflow for evaluating protein structures.

ASSOCIATED CONTENT

AUTHOR INFORMATION

Corresponding Author

* E-mail: bruotolo@umich.edu

ORCID

Daniel A. Polasky: 0000-0002-0515-1735
Sugyan M. Dixit: 0000-0002-6313-7974
Brandon T. Ruotolo: 0000-0002-6084-2328

Author Contributions

D.P. and S.D. developed the algorithm. D.V. and K.K. acquired the experimental data. D.P. and B.R. wrote the manuscript. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank Jessica Rabuck-Gibbons for preparation of Src kinase domain, as well as Anna Schwendeman and Jukyung Kang for preparation of the Avastin and Avegra samples. The development of CIU data interpretation strategies in the Ruotolo lab is supported by the National Science Foundation, Chemical Measurement and Imaging Program in the Division of Chemistry, and with partial co-funding from the Division of Molecular and Cellular Biosciences (Award #1808541).

REFERENCES

(1) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What Is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 5–13.

(2) Lössl, P.; van de Waterbeemd, M.; Heck, A. J. The Diverse and Expanding Role of Mass Spectrometry in Structural and Molecular Biology. *EMBO J.* **2016**, *35* (24), 2634–2657.

(3) Zhong, Y.; Hyung, S.-J.; Ruotolo, B. T. Ion Mobility–Mass Spectrometry for Structural Proteomics. *Expert Rev. Proteomics* **2012**, *9* (1), 47–58.

(4) Tian, Y.; Ruotolo, B. T. The Growing Role of Structural Mass Spectrometry in the Discovery and Development of Therapeutic Antibodies. *Analyst* **2018**, *143* (11), 2459–2468.

(5) Terral, G.; Beck, A.; Cianférani, S. Insights from Native Mass Spectrometry and Ion Mobility-Mass Spectrometry for Antibody and Antibody-Based Product Characterization. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2016**, *1032*, 79–90.

(6) Bleiholder, C.; Bowers, M. T. The Solution Assembly of Biological Molecules Using Ion Mobility Methods: From Amino Acids to Amyloid β-Protein. *Annu. Rev. Anal. Chem* **2017**, *10* (1), 365–386.

(7) Liko, I.; Allison, T. M.; Yen, H.-Y.; Hopper, J. S. T.; Robinson, C. V. Using Native Mass Spectrometry to Inform Drug Discovery. *Drug Target Rev.* **2017**, *4* (2), 44–47.

(8) Mehmood, S.; Allison, T. M.; Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* **2015**, *66* (1), 453–474.

(9) Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. Naked Protein Conformations: Cytochrome c in the Gas Phase. *J. Am. Chem. Soc.* **1995**, *117* (40), 10141–10142.

(10) Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. *Curr. Opin. Chem. Biol.* **2018**, *42*, 93–100.

(11) Hopper, J. T. S.; Oldham, N. J. Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.

(12) Ruotolo, B. Searching for Conformationally-Selective Small Molecule Therapeutics Using Ion Mobility-Mass Spectrometry (227.1). *FASEB J.* **2014**, *28* (1 Supplement).

(13) Rabuck, J. N.; Hyung, S. J.; Ko, K. S.; Fox, C. C.; Soellner, M. B.; Ruotolo, B. T. Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry. *Anal. Chem.* **2013**, *85* (15), 6995–7002.

(14) Rabuck-Gibbons, J. N.; Keating, J. E.; Ruotolo, B. T. Collision Induced Unfolding and Dissociation Differentiates ATP-Competitive from Allosteric Protein Tyrosine Kinase Inhibitors. *Int. J. Mass Spectrom.* **2018**, *427*, 151–156.

(15) Tian, Y.; Han, L.; Buckner, A. C.; Ruotolo, B. T. Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. *Anal. Chem.* **2015**, *87* (22), 11509–11515.

(16) Tian, Y.; Ruotolo, B. T. Collision Induced Unfolding Detects Subtle Differences in Intact Antibody Glycoforms and Associated Fragments. *Int. J. Mass Spectrom.* **2018**, *425*, 1–9.

(17) Pisupati, K.; Tian, Y.; Okbazghi, S.; Benet, A.; Ackermann, R.; Ford, M.; Saveliev, S.; Hosfield, C. M.; Urh, M.; Carlson, E.; et al. A Multidimensional Analytical Comparison of Remicade and the Biosimilar Remsima. *Anal. Chem.* **2017**, *89* (9), 4838–4846.

(18) Hernandez-Alba, O.; Wagner-Rousset, E.; Beck, A.; Cianférani, S. Native Mass Spectrometry, Ion Mobility, and Collision-Induced Unfolding for Conformational Characterization of IgG4 Monoclonal Antibodies. *Anal. Chem.* **2018**, *90* (15), 8865–8872.

(19) Watanabe, Y.; Vasiljevic, S.; Allen, J. D.; Seabright, G. E.; Duyvesteyn, H. M. E.; Doores, K. J.; Crispin, M.; Struwe, W. B. Signature of Antibody Domain Exchange by Native Mass Spectrometry and Collision-Induced Unfolding. *Anal. Chem.* **2018**, *90* (12), 7325–7331.

(20) Tian, Y.; Lippens, J. L.; Netirojjanakul, C.; Campuzano, I. D. G.; Ruotolo, B. T. Quantitative Collision-Induced Unfolding Differentiates Model Antibody–Drug Conjugates. *Protein Sci.* **2019**, *28* (3), 598–608.

(21) Rabuck-Gibbons, J. N.; Lodge, J. M.; Mapp, A. K.; Ruotolo, B. T. Collision-Induced Unfolding Reveals Unique Fingerprints for Remote Protein Interaction Sites in the KIX Regulation Domain. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (1), 94–102.

(22) Niu, S.; Ruotolo, B. T. Collisional Unfolding of Multiprotein Complexes Reveals Cooperative Stabilization upon Ligand Binding. *Protein Sci.* **2015**, *24* (8), 1272–1281.

(23) Beveridge, R.; Migas, L. G.; Payne, K. A. P.; Scrutton, N. S.; Leys, D.; Barran, P. E. Mass Spectrometry Locates Local and Allosteric Conformational Changes That Occur on Cofactor Binding. *Nat. Commun.* **2016**, *7*, 12163.

(24) Migas, L. G.; France, A. P.; Bellina, B.; Barran, P. E. ORIGAMI: A Software Suite for Activated Ion Mobility Mass Spectrometry (AIM-MS) Applied to Multimeric Protein Assemblies. *Int. J. Mass Spectrom.* **2018**, *427*, 20–28.

(25) Eschweiler, J. D.; Rabuck-Gibbons, J. N.; Tian, Y.; Ruotolo, B. T. CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. *Anal. Chem.* **2015**, *87* (22), 11516–11522.

(26) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. CIUSuite 2: Next-Generation Software for the Analysis of Gas-

Phase Protein Unfolding Data. *Anal. Chem.* **2019**, *91* (4), 3147–3155.

(27) Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V. Quantifying the Stabilizing Effects of Protein-Ligand Interactions in the Gas Phase. *Nat. Commun.* **2015**, *6*, 8551.

(28) Sivalingam, G. N.; Yan, J.; Sahota, H.; Thalassinos, K. Amphitrite: A Program for Processing Travelling Wave Ion Mobility Mass Spectrometry Data. *Int. J. Mass Spectrom.* **2013**, *345–347*, 54–62.

(29) Sivalingam, G. N.; Cryar, A.; Williams, M. A.; Gooptu, B.; Thalassinos, K. Deconvolution of Ion Mobility Mass Spectrometry Arrival Time Distributions Using a Genetic Algorithm Approach: Application to A1-Antitrypsin Peptide Binding. *Int. J. Mass Spectrom.* **2018**, *426*, 29–37.

(30) Fantin, S. M.; Parson, K. F.; Niu, S.; Liu, J.; Ferguson-Miller, S. M.; Ruotolo, B. T. CIU Classifies Ligand Binding Behavior of Integral Membrane Translocator Protein TSPO. *Under Rev.*

(31) Vallejo, D. D.; Polasky, D. A.; Kurulugama, R. T.; Eschweiler, J. D.; Fjeldsted, J. C.; Ruotolo, B. T. A Modified Drift Tube Ion Mobility-Mass Spectrometer for Charge Multiplexed Collision Induced Unfolding. *Anal. Chem.* **2019**.

(32) Seeliger, M. A.; Young, M.; Henderson, M. N.; Pellicena, P.; King, D. S.; Falick, A. M.; Kuriyan, J. High Yield Bacterial Expression of Active C-Abl and c-Src Tyrosine Kinases. *Protein Sci.* **2005**, *14* (12), 3135–3139.

(33) Haynes, S. E.; Polasky, D. A.; Dixit, S. M.; Majmudar, J. D.; Neeson, K.; Ruotolo, B. T.; Martin, B. R. Variable-Velocity Traveling-Wave Ion Mobility Separation Enhancing Peak Capacity for Data-Independent Acquisition Proteomics. *Anal. Chem.* **2017**, *89* (11), 5669–5672.

(34) Pedregosa, F.; Varoquax, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O. Scikit-Learn: Machine Learning in Python. *J Mach Learn Res* **2011**, *12* (Oct), 2825–2830.

(35) Brown, M. T.; Cooper, J. A. Regulation, Substrates and Functions of Src. *Biochim. Biophys. Acta - Rev. Cancer* **1996**, *1287* (2–3), 121–149.

(36) Parsons, S. J.; Parsons, J. T. Src Family Kinases, Key Regulators of Signal Transduction. *Oncogene* **2004**, *23* (48 REV. ISS. 7), 7906–7909.

(37) Creedon, H.; Brunton, V. G. . Src Kinase Inhibitors: Promising Cancer Therapeutics? *Crit. Rev. Oncog.* **2012**, *17* (2), 145–159.

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Insert Table of Contents artwork here