Gábor Transform-Based Signal Isolation, Rapid Deconvolution, and Quantitation of Intact Protein Ions with Mass Spectrometry

Kayd L. Meldrum[1], Andrew K. Swansiger[1], Meghan M. Daniels[1], Wendi A. Hale[2], Crystal Kirmiz Cody[2], Xi Qiu[2], Michael Knierman[2], John Sausen[2], James S. Prell[1,3,*]

[1]Department of Chemistry and Biochemistry, 1253 University of Oregon, Eugene, OR, USA 97403-1253

[2]Agilent Technologies, Inc., 5301 Stevens Creek Blvd., Santa Clara, CA, USA, 95051

[3]Materials Science Institute, 1252 University of Oregon, Eugene, OR, USA 97403-1252

Submitted to *Analytical Chemistry*

21 February 2024

Revised 6 May 2024

[*]Corresponding author email: jprell@uoregon.edu

Abstract:

High-resolution mass spectrometry (HRMS) is a powerful technique for characterization and quantitation of complex biological mixtures, with several applications including clinical monitoring and tissue imaging. However, these medical and pharmaceutical applications are pushing the analytical limits of modern HRMS techniques requiring either further development in instrumentation or data processing methods. Here, we demonstrate new developments to the iFAMS (interactive Fourier-Transform Analysis for Mass Spectrometry) software including the first application of Gábor Transform (GT) to protein quantitation. Newly added automation tools detect signal from minimal user input and apply thresholds for signal selection, deconvolution, and baseline correction to improve objectivity and reproducibility of deconvolution. Additional tools were added to improve deconvolution of highly complex or congested mass spectra and are demonstrated here for the first time. The "Gábor Slicer" enables the user to explore trends in the Gábor spectrogram with instantaneous ion mass estimates accurate to 10 Da. The charge adjuster allows for easy visual confirmation of accurate charge state assignments and quick adjustment if necessary. Deconvolution refinement utilizes a second GT of isotopically resolved data to remove common deconvolution artifacts. To assess the quality of deconvolution from iFAMS, several comparisons are made to deconvolutions using other algorithms such as UniDec and an implementation of MaxEnt in Agilent MassHunter BioConfirm. Lastly, the newly added batch processing and quantitation capabilities of iFAMS are demonstrated and compared to a common extracted ion chromatogram (EIC) approach.

Introduction

Accurate quantitation of proteins and pharmaceutical drugs is important in many applications, from clinical monitoring of patient health and precision dosing to spatial mapping of biomolecules in biological tissues.[1–9] While ligand binding assays (LBAs) have been the cornerstone of quantitative clinical assays, their limited ability to distinguish between closely related proteins, proteoforms, and degradation and aggregation products has made it especially important to develop technologies that can efficiently separate and accurately characterize these molecules.[3,10,11] For this reason, liquid chromatography/high-resolution mass spectrometry (LC/HRMS) has rapidly grown in popularity for its ability to separate species by both retention time and mass and low sample requirements—often on the order of only a few picomoles. While LC/MS has been used for many years in combination with prior enzymatic digestion of protein species (e.g., in "bottom-up" proteomics),[11–13] mass accuracy within a few ppm combined with high resolution available on many modern instruments has made it possible to directly characterize and quantitate mixtures of multiple intact proteins by LC/HRMS. The exquisite separation, high speed, and high chemical specificity achievable with LC/HRMS have led to a rapid increase in its use for clinical protein and biotherapeutic quantitation,[14–20] where the presence of endogenous molecules or other drugs may interfere with accurate measurements, especially when dealing with limited sample volumes or low-abundance analytes.

While LC/MS methods can be extremely helpful in characterizing proteins from many biological samples, some complex samples still present major challenges, as some analytes may still co-elute and exhibit suppression effects in electrospray ionization (ESI), and instrument response factors for different analytes can be extremely difficult to determine without pure reference analytes.[1,12] For example, blood serum samples containing monoclonal antibody (mAb) drugs present a background of endogenous antibodies that may not separate well by LC and are strongly overlapped with the antibody drug signal in intact mass spectra.[1,11,21] In another example, current technology for nano-electrospray desorption ionization (nano-DESI) MS imaging of biological tissues do not couple easily to LC, thus signals for many intact proteins are often detected simultaneously and may overlap severely in the resulting mass spectra.[8,22,23] For improved separation of clinical samples, immunocapture methods, such as Melon gel filtration, are often used as a preliminary separation method before analysis by LC/MS.[11,24] For biological analytes as large as (or larger than) intact mAbs, the mass spectrum can often be congested when

the analyte constitutes a mixture of proteoforms or drug conjugation states that are not separated prior to or during the LC step. In addition to this spectral congestion, signal from each protein of interest is often spread over many different charge states as a result of the stochastic charging process in ESI. The shape of the charge state distribution itself may also drift in time and/or depend on experimental conditions. These effects are typically larger for denatured proteins introduced using common LC/MS protocols than for native-like proteins introduced by "static" nanoelectrospray ionization, and they can make correct charge state assignment difficult.[25]

To maximize the quantitative capabilities of MS for samples with highly complex mass spectra, post-acquisition data analysis methods have been developed, including several different types of deconvolution algorithms.[26] Deconvolution has the advantage that signals from many or all detected charges states of an analyte can be combined, in principle capturing more signal than alternative quantitation strategies in which abundances of only a few charge states are used. (Ref. 26 provides a detailed historical review and comparison of computational strategies for deconvolving complex mass spectra.) A recent study illustrates that quantitation using deconvolution can be consistent with direct quantitation from a mass spectrum.[27] However, no open-source, vendor-neutral deconvolution software packages with automated protein quantitation capabilities are available, which results from many open-source deconvolution software packages being highly sensitive to user-input parameters. Furthermore, robust deconvolution protocols that can be highly standardized and automated with minimal intervention from the user are needed to achieve the level of reproducibility and traceability required for routine clinical use.[28,29] Initially developed for qualitative analyses of highly complex mass spectra, the iFAMS (interactive Fourier-Transform Analysis for Mass Spectrometry) program uses Gábor Transform (GT) to separate overlapped analyte signals in mass spectra by their "frequencies" and deconvolve them while filtering out as much white noise and chemical interferent signal as possible. Previously released versions of iFAMS often require substantial user intervention, and the use of iFAMS in quantitation applications has not yet been explored.

In this paper, we outline new tools incorporated in the iFAMS software designed to increase automation and objectivity to the deconvolution process and allow for occasional user intervention, and we demonstrate the newly added quantitative capabilities of the iFAMS v. 6.3 ("iFAMS Quant") software. With the new automated signal-selection and baseline-correction

tools, the deconvolved mass spectra ("zero-charge" spectra) are easier and faster to generate and are much less sensitive to input parameters and baseline effects. We also analyze the quality of deconvolution through mass accuracy assessments, verification of charge state assignment, and comparison to other widely used deconvolution methods (an implementation of MaxEnt in the Agilent MassHunter BioConfirm v. 10.0 and the open-source, publicly available program UniDec from the Marty Group).[20,23,30–32] The theory behind these deconvolution methods is described and compared elsewhere.[26] New tools are introduced to aid in identification of and "instant" mass determination for related peaks in a charge distribution, validation of charge state assignments and identification and removal of deconvolution artifacts. Lastly, to demonstrate our program's capabilities, batch analysis and quantitation with iFAMS Quant is compared to the current industry standard for quantitation which utilizes extracted ion chromatograms to quantitate directly from the mass spectrum.[1,16,18,19]

Methods

Materials. Ultrapure (18 MΩ·cm) water was prepared using a Barnstead E-Pure Ultrapure Water Purification System (Thermo Scientific, Inc., Waltham, MA). Intact IgG1κ monoclonal antibody (NIST mAb, 10 mg/mL, SKU 8671) in 12.5 mmol/L L-histidine, 12.5 mmol/L L-histidine HCl (pH 6.0) was purchased from the National Institute of Standards and Technology (NIST, Gaithersburg, MD). Bovine serum albumin (BSA), ubiquitin from bovine erythrocytes, cytochrome c from equine heart, avidin from hen egg whites, myoglobin from equine skeletal muscle, and β-lactoglobulin A from bovine milk, all lyophilized in Tris buffer, were purchased from Millipore Sigma (St. Louis, MO). Ammonium acetate (100 mM, pH 7.4, Millipore Sigma, St. Louis, MO) and ammonium bicarbonate (50mM, pH 7.4, Honeywell Fluka) buffers were prepared from powder, pH corrected, and filtered by aspiration. Bio-Spin® size exclusion spin columns were purchased from Bio-Rad (Hercules, CA) for buffer exchange of protein samples.

Intact mouse trastuzumab data was collected and provided by Xi Qiu with Agilent Technologies, Inc. (Santa Clara, CA). The protein was obtained from Genentech (South San Francisco, CA), spiked into plasma, and purified with Streptavidin cartridges (Agilent Technologies, Santa Clara, CA). The antibodies were then measured using the same

instrumentation described below. More details on the experimental method are described previously.[16]

Protein sample preparation. For mass accuracy experiments, NIST mAb was diluted with ultrapure water to concentrations of 10, 100, and 500 µg/mL. For quantitation experiments, BSA was reconstituted in ultrapure water and buffer exchanged into 50 mM ammonium bicarbonate (pH 7.4) to a concentration of 20 µg/mL. NIST mAb was also buffer exchanged into ammonium bicarbonate buffer, spiked with BSA and diluted to 11 different calibrant concentrations ranging from 0.005-10 µg/mL (with final BSA concentrations of 10 ug/mL). For the multi-protein standard sample, all five proteins (ubiquitin, cytochrome c, avidin, myoglobin, and β-lactoglobulin) were reconstituted in ultrapure water, combined to final concentrations of about 100 nM, and buffer exchanged into ammonium acetate.

High-resolution intact protein liquid chromatography-mass spectrometry. For quantitative liquid chromatograph/mass spectrometry (LC/MS) experiments, an Agilent 1290 Infinity II LC was used with a PLRP-S reverse-phase column online with an Agilent 6545XT AdvanceBio quadrupole-time-of-flight (Q-TOF) mass spectrometer (generously loaned from Agilent Technologies, Santa Clara, CA). All mass spectrometry data were acquired using the Agilent Jet Stream electrospray ionization source online with the LC (for NIST mAb experiments) or a syringe pump (multi-protein mixture experiments; see above). For the NIST mAb sample used to assess mass accuracy, data were acquired up to 7000 *m*/*z* with a 2 GHz sampling rate ("High Mass Range Mode" in the Agilent MassHunter software) to ensure sampling of the entire charge state distribution. Quantitation data were instead acquired up to 3200 *m*/*z* with a 4 GHz sampling rate ("High Resolution Mode") to optimize resolution. For each NIST mAb sample, 1-10 microliters were injected onto the LC using a gradient starting with 18 MΩ·cm water with 0.1% formic acid (ACS reagent, puriss. p.a. grade, Sigma-Aldrich, St. Louis, MO) and ending with LC/MS-grade acetonitrile (from Agilent Technologies, Inc., Santa Clara, CA and Millipore Sigma, St. Louis, MO) with 0.1% formic acid. A table with further details of the LC gradients can be found in the Supporting Information (see Tables S1 and S2). Mass spectral data for LC retention times of 2.70-2.85 min, over which time interval both NIST mAb and BSA elute, were summed for quantitative analysis in Mass Hunter Qual software (Agilent Technologies, Santa Clara, CA) before further analysis with our in-house mass spectrometry data analysis program, iFAMS (see below). Replicate data sets were acquired on different days to

assess day-to-day reproducibility. A syringe pump flow rate of 400 μL/min was used for multi-protein mixture experiments, and details on the MS acquisition parameters can be found in the Supporting Information (see Table S3).

New computational tools in iFAMS v. 6.3. The version of iFAMS software used here (v. 6.3, called iFAMS Quant) builds off previous versions of iFAMS described elsewhere but incorporates new tools for automated protein signal isolation, baseline correction, and quantitation.[33–36] The protein signal isolation in the "guided search" tool works by identifying the charge states, total ion mass, and mass frequency ($1/\Delta m$) from an input of GT signal from two adjacent charge states. This algorithm then predicts the positions of other possible charge states of the same protein ion population in the GT and selects all GT signal peaks that pass built-in thresholds based on GT signal magnitude for deconvolution and further processing. Further details of deconvolution tools in iFAMS Quant can be found in the Supporting Information.

Quantitation using calibrant data is achieved by first identifying and integrating peaks in the deconvolved mass spectra. iFAMS Quant has adjustable parameters to aid in identifying peaks including minimum peak height, minimum peak spacing, and noise tolerance. Each of these parameters has a default value to streamline automated processing, though they can all be changed by the user as needed. For batch processing, identical deconvolution and integration parameters (or explicit user-defined bounds) are applied to each deconvolved mass spectra. For each spectrum processed, an integration peak-list file is generated, which can be loaded into the calibration menu in iFAMS. Calibration curves can be generated from a single calibrant mass or several masses for a combined peak area. The calibrant signal can also be normalized to an internal standard. Least-squares fitting of the calibrant data is achieved using built-in Python functions for linear, quadratic, and logistic curves, with logistic fits defined as

$$y = \frac{A}{1+e^{-Bx}} - C \quad (1)$$

and optimized using the Levenberg-Marquardt algorithm.[37] Several weighting options (here, defined as the weighting of squared difference terms in the least-square fitting procedure) are provided including "1/x" and "$1/x^2$."

Results and Discussion

Automated Gábor Transform filtration and deconvolution. In clinical and pharmaceutical applications, mass spectrometry-based quantitation of low protein signals in a noisy background

(e.g., blood serum) represents a major challenge, and deconvolution methods have the potential advantage of summing many signals that may be close to or below the noise level to produce deconvolved signals that can be quantitated. We investigated the extent to which GT-based filtering of mass spectra and deconvolution to yield a deconvolved mass distribution can aid in peak assignment, accurate mass determination, and analyte quantitation, including for the "NIST mAb" humanized recombinant IgG1κ monoclonal antibody standard from the National Institute of Standards and Technology.

In iFAMS Quant, users can specify negative or positive ion mode, then manually select charge states from a charge distribution to include in mass deconvolution or use an automatic charge state selector. To use the latter, the user draws a box around the fundamentals of two adjacent charge states in the GT spectrum, and the program uses the magnitude maxima within these boxes to extrapolate expected positions of related charge states. In the Gábor spectrogram, peaks belonging to the charge state series of a single protein are arranged in a predictable downward-chirped hyperbolic pattern as long as the protein mass does not vary significantly with charge (e.g., due to varying ligand adduction). Explicitly, the signal for each charge state appears at a frequency equal to the charge divided by some small, constant increment in mass (Δm, such as the mass difference between isotopes, glycoforms, or charge carriers). Since the m/z position is inversely dependent on charge, the series lie along reciprocal curves, defined as frequency = $n(m/\Delta m) \times (m/z)^{-1}$, where n is the harmonic number. Equivalently, these are curves of constant $n(M+zx)/\Delta m \approx nM/\Delta m$, where M is the mass of any ion with a signal lying along the curve, and x is the mass of the charge carrier (which, for large ions like proteins, can usually be ignored). This predictable pattern allows for accurate automated selection of other charge states belonging to the same protein based on the initial user-selected signals. Default threshold parameters for signal detection and peak centroiding were optimized using several types of protein mass spectra; these default parameters are listed in the *iFAMS Quant default signal selection and integration parameters* section in the Supporting Information. Since white noise in the mass spectrum is approximately uniformly distributed across the Gábor spectrogram, much or all high-frequency white noise is removed by this process in typical mass spectra. Low-frequency noise between the data selected for each harmonic is also removed. Thus, as explained elsewhere, GT selection with the above options implemented acts as a "notch" or "matched"

filter, in contrast to simple low-pass or smoothing filters, such as Savitzky-Golay or moving-average filters.

In Figure 1, a mass spectrum of intact, denatured NIST mAb at a modest concentration (10 µg/mL, which corresponds to ~67 fmol for a 1 µL sample) was analyzed using iFAMS with the default parameters, and peaks were assigned to the deconvolved mass spectrum using literature accurate measured masses for common glycoforms.[38] Integrated peak areas and computed mass centroids are displayed alongside the deconvolved mass spectrum. At this stage, the user can easily change the set of included charge states and/or harmonics to recalculate the deconvolved mass spectrum, mass centroids, and integrated peak values as desired.

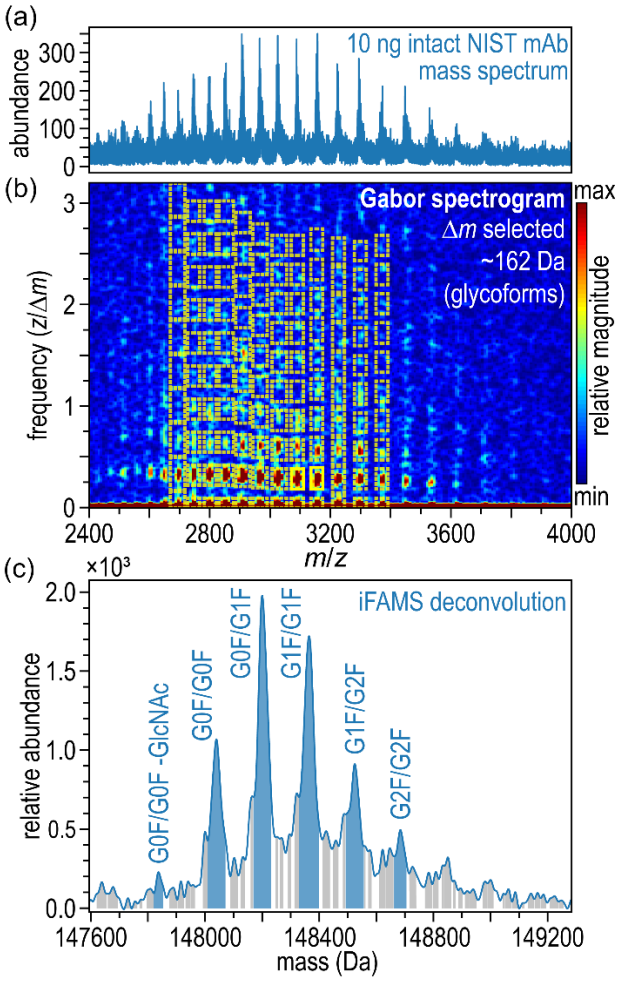


Figure 1. Representative iFAMS Quant deconvolution using automated peak selection. (a) Mass spectrum of 10 µg/mL intact IgG1κ monoclonal antibody (NIST mAb) from 1 µL injection. (b) Corresponding Gábor spectrogram with manually selected boxes in solid yellow and automatically selected boxes in dashed yellow. Default parameters selected 22 charge states (40-61+) with 9 harmonics. Illustrated here, selections were manually restricted to the 12 most abundant charge states (44-55+) with 9 harmonics. (c) Resulting iFAMS deconvolved mass

spectrum with peak integration shown with shaded regions. Blue peaks correspond to major glycoforms that were manually identified based on accurate mass and labeled with glycan composition.

Automatic baseline correction options. A non-zero, nearly constant or even curved baseline is a common feature of many experimental mass spectra, arising from a combination of detector noise, peak tails (e.g., due to collisions during mass measurement), electronic and chemical interferences, and other sources. Peculiar to GT-based mass spectral analysis, windowing artifacts arising from selecting sharply defined boxes of data for signal reconstruction can result in conspicuously curved baselines in the deconvolved mass spectrum. This curved baseline arises from a convolution of the Gaussian window shape used to generate the GT spectrogram and the m/z interval selected by the boxes. In the deconvolved mass spectrum, the curved baselines from each included charge state are summed. This composite curved baseline (which often resembles a "pedestal" with rounded sides in the deconvolved mass spectrum) can be directly estimated by applying exactly the same GT resolution and box selection parameters to a constant-valued mass spectrum and computing its "deconvolution." The resulting model baseline is scaled vertically by fitting the rise and fall at its low- and high-mass edges to those of the "raw" deconvolved experimental spectrum. Figure 2 demonstrates this baseline correction applied to the same NIST mAb mass spectrum shown in Figure 1. The baseline correction evidently increases the modulation depth (hence, apparent resolution) of adjacent peaks, especially for lower-abundance peaks.

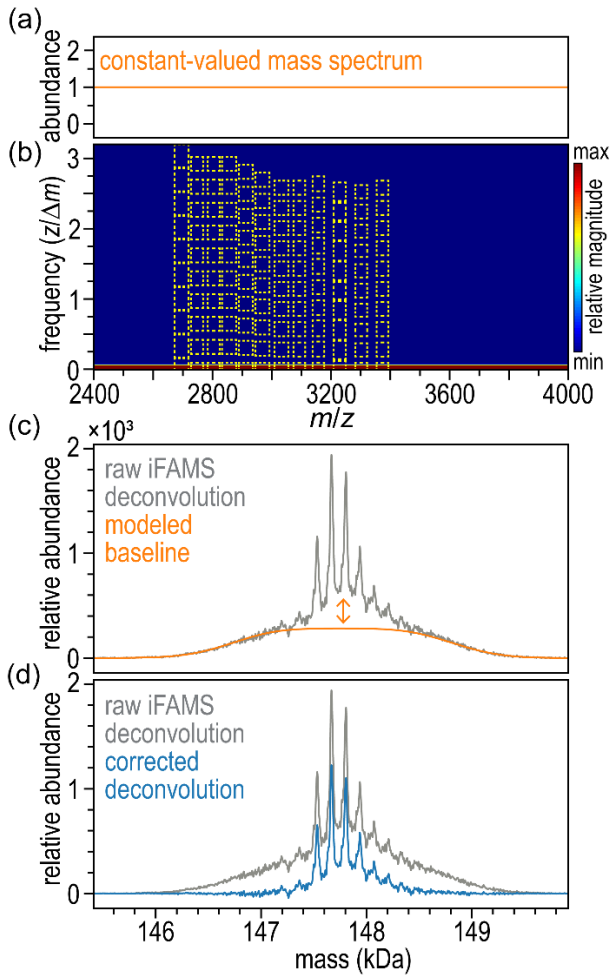


Figure 2. Schematic for automatic baseline correction. (a) Constant-valued mass spectrum used to generate an "empty" Gábor spectrogram. (b) "Empty" Gábor spectrogram with identical selections applied. Selections shown here are based on the 10 ng NIST mAb example in Figure 1. (c) Resulting modeled baseline shape (orange). Baseline edges are then fit to the raw deconvolved mass spectrum (gray) which is the non-baseline corrected deconvolved mass spectrum from the 10 ng NIST mAb example. (d) Deconvolved mass spectrum before (gray) and after (blue) baseline correction.

Mass accuracy. In addition to apparent resolution, another characteristic of interest for mass spectral deconvolution is mass accuracy of the deconvolved peaks. The measured (centroided) accurate masses from the 10 ng NIST mAb deconvolution shown in Figure 1, were compared to the corresponding exact theoretical glycoform masses for NIST mAb (obtained from BioConfirm 10 sp1).[38] For the six most abundant glycoforms, the observed masses were well within United States Food and Drug Administration (FDA)-recommended mass tolerance for intact mAbs (±25 ppm) with errors ranging from 1.1 to 10.1 ppm and an RMSD of 7.3 ppm.[28] Less abundant glycoforms were not further characterized here, because the accurate measured

masses for some pairs of glycoforms, as determined using iFAMS Quant, differ by less than twice the FDA-recommended mass tolerance [i.e., ~7 Da = 48 ppm < 50 ppm] and therefore are not expected to be confidently distinguishable without adjuvant experiments, such as digestion or MS/MS experiments.[38,39] For a different NIST mAb sample at a significantly higher concentration (500 ng/uL, for the same mass on column as in a vendor application note[38]), the RMSD improved to 2.9 ppm for the same six glycoforms, in close agreement with the reported value of 2.2 ppm.[38]

Mass accuracy was also assessed with and without the automatic baseline correction applied (see Table S4), and there was no change to mass accuracy for the peaks in the center of the deconvolved mass distribution (< 0.05 ppm). Since only slight changes (< 1.2 ppm) occurred on either side of the distribution with the greatest effect on the extremes, we determined that the baseline correction had no significant impact on mass accuracy. Accurate masses and mass errors for each glycoform can be found in the Supporting Information (Table S4).

Charge state distributions without resolvable periodic signal can still be deconvolved with GT. Although strong periodic signals often facilitate GT deconvolution, as the periodic glycoform spacing above in the NIST mAb deconvolution, periodical signal is not strictly required for deconvolution. Any sharp feature in the mass spectrum will appear as a broad frequency peak in the GT spectrogram centered about a frequency of zero which can be easily picked out from baseline (which itself yields a very narrow band of signal about zero frequency). However, given user-selected near-zero frequency bands for two adjacent charge states, iFAMS can estimate the positions of the rest of the distributions and automate those selections. This strategy is similar to simpler algorithms in which data for each charge state in the mass spectrum are extracted, rescaled, and summed, but differs from these in that, in iFAMS Quant, each charge state undergoes a low-pass frequency filtration step tailored to that charge state before summation.[26,40]. An example of this strategy for GT deconvolution is shown in Figure 3 for a sample containing both BSA and NIST mAb. The BSA and the NIST mAb signals were easily deconvolved in iFAMS despite the lack of resolved periodic signal from BSA.

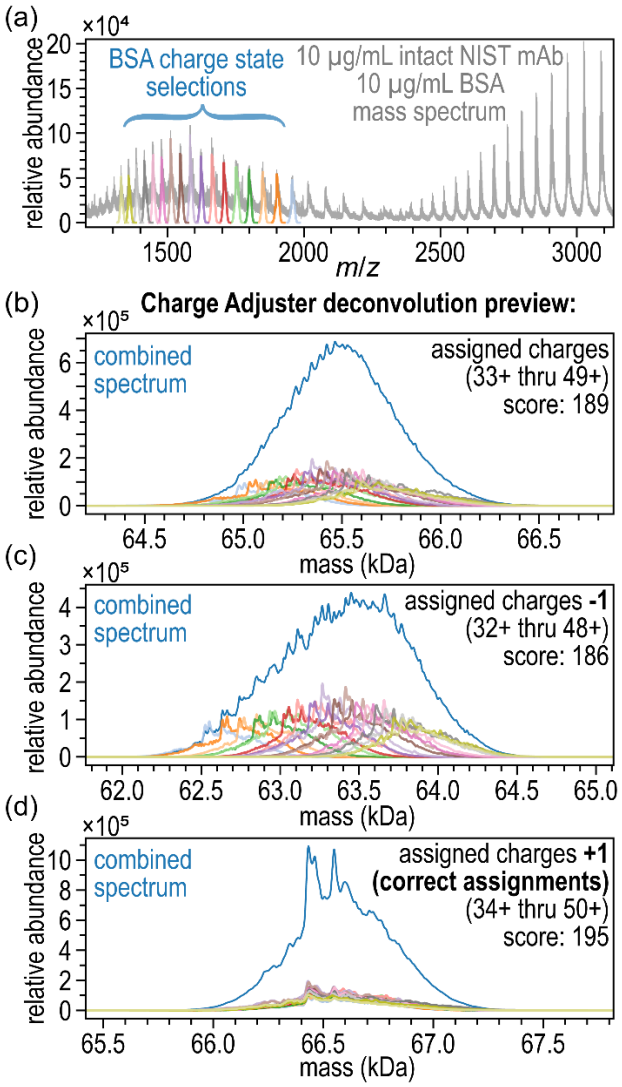


Figure 3. Demonstration of charge adjuster tool. (a) Mass spectrum containing 10 µg/mL intact NIST mAb and 10 µg/mL BSA (gray) with inverse-GT single charge state selections overlayed (multi-colored). (b-d) Charge adjuster deconvolution previews with assigned charges, charges shifted minus one from assigned, and charges shifted plus one from assigned, respectively. Note that individual charge-corrected inverse-GT spectra have doubled abundances for illustrative purposes, and curved baseline correction has not yet been applied (see text). (d) Deconvolution preview of correct charge assignments (charge state score is highest for correct assignments; see Figure S1 for more detail).

Charge assignment and deconvolution confidence. Confident assignment of charge states for large-ion charge state distributions is a common challenge that has prompted development of several experimental solutions, including solution- and gas-phase charge reduction, Charge Detection MS,[41–45] Cation-to-Anion Proton Transfer Reactions,[25,46] and mass shifting by adduction of large ligands.[47] iFAMS Quant offers the user options to adjust automated charge

state assignments for challenging data without the need for instrument or experiment modifications. When using the automated GT filtration tools in iFAMS Quant, charge state assignments can occasionally be offset from the correct charge states, especially in spectra with a high level of noise or for highly charged species. The initial, automated deconvolution of denatured BSA shown in Figure 3b is an example of this phenomenon. The charge state adjustment tool allows this type of error to be quickly corrected. The charge-state correction tool in iFAMS Quant displays each GT-filtered charge state, corrected for charge and overlayed. If the assignments are correct, the major features in most charge states should align. Shifting all charge assignments by one more or less than the correct assignments will misalign the peak spacing and often results in a broad deconvolution with low apparent resolution. By shifting all the assignments one integer at a time with the built-in interface, a user can visually tune the assignments to correct an assignment error. Figures 3c and 3d shows results for the BSA data, in which the automatically assigned charge states have been adjusted up and down by one charge state. Incorrect assignments result in poor peak alignment, whereas the correct adjustment (+1 charge state relative to the automated assignments, for the example shown in Figure 3) clearly results in excellent peak alignment and much higher apparent peak resolution. Occasionally, neighboring unrelated signals are included within the GT-filtration selections and can create artifacts in the deconvolved mass spectrum. Unless the unrelated signal has a nearly identical charge-state distribution to the signal of interest, its contribution to the deconvolved mass spectrum will not appear for all charge states and can thus often be visually identified as an artifact. In cases where periodic signal is well resolved, the charge adjuster deconvolution preview only uses the signal from the fundamentals which creates a deconvolved mass spectrum with signal oscillating about zero. In many cases, this makes visual confirmation of correct assignments much easier due to more destructive interference with erroneous assignments. An example of this is shown in the Supporting Information (see Figure S1). iFAMS Quant has a built-in scoring algorithm within the charge adjuster that is often helpful in determining correct charge state assignments (see Figure S1).

Challenging multi-protein samples. Although LC is routinely available in many types of biomolecular experiments, it is not currently possible to implement it in others, for example, in imaging mass spectrometry, with very small samples, or when directly monitoring a chemical reaction in real time. Nevertheless, by spreading signal out in a 2-dimensional domain in

predictable patterns, Gábor Transform provides a form of signal separation that can help isolate signals belonging to different proteins that strongly overlap in the *m*/*z* spectrum. Figure 4a shows a static nano-electrospray mass spectrum with isotopic resolution for a sample containing a mixture of five common MS standard proteins (bovine ubiquitin, equine cytochrome c, hen egg avidin, equine myoglobin, and bovine β-lactoglobulin). The latter three of these proteins have multiple glycoforms and/or ligand-adduction states, giving rise to further complexity in the resulting mass spectrum. (Note that, although avidin is a native tetramer, avidin monomer was analyzed in this study to represent a protein with a complex glycoform distribution; no tetramer was detected within the mass range measured, so avidin was excluded from relative abundance analysis.) The charge state distributions of the protein ions are strongly interleaved, with some peaks directly overlapping in the *m*/*z* domain. Charge state distributions for each protein can be identified in the corresponding Gábor spectrogram at a glance due to the isotope resolution of the signal (see Figure 4b). That is, because isotopes give rise to signals at frequencies ($z/\Delta m$) arising from ~1 Da isotope spacings, peaks are observed at frequencies equal to the charge state of the corresponding *m*/*z* peak. Signals for individual charges states for each protein can be selected out one at a time or combined to generate a deconvolved mass spectrum with exactly zero signal in between the proteins' mass distributions. The resulting deconvolution is shown in Figure 4c, in which 63 distinct proteoforms or adduct states (with the adducts being sodium, potassium, heme, or lactose) can be assigned. A handful of minor artifact peaks are observed in the deconvolved mass spectrum, but iFAMS has additional tools for removing these; see "Deconvolution refinement" section below.

The same raw mass spectrum was deconvolved using Agilent MassHunter BioConfirm's built-in MaxEnt deconvolution tool as well as the Bayesian deconvolution tool UniDec (see Figure S2 for deconvolved mass spectra). All three deconvolutions reconstructed the same major proteoforms and sodium adduction states but differed greatly on relative abundances. The five protein standards were added to the sample in similar concentrations, best represented by the iFAMS deconvolution. The UniDec deconvolution more closely matches the iFAMS abundances than the MaxEnt abundances, but both MaxEnt and UniDec report a much lower relative abundance of ubiquitin. The deconvolutions also differ in apparent peak resolution. iFAMS was able to reconstruct the mass distribution with baseline-resolved isotope features using the default parameters. Using Agilent MassHunter BioConfirm's default MaxEnt parameters resulted in a

deconvolution with low apparent resolution on the low-mass end and higher resolution at the high-mass end, with only half of the proteins having clear isotope features. Adjusting UniDec's parameters allowed for isotope reconstruction but with lower resolution than iFAMS and a baseline that increases in abundance and noise with higher masses. Although UniDec's isotope distributions had similar peak full width at half-max (FWHM) as iFAMS, the isotope distributions had significant abundance noise that was not typical of iFAMS deconvolutions. MaxEnt and UniDec both generate similar artifacts that do not appear in iFAMS, such as the peak at ~11080 Da that arises from the 9^{th} harmonic of a 1+ contaminating ion of mass 1230 Da, and most of the artifacts in iFAMS are not present in the MaxEnt and UniDec deconvolutions.

In challenging cases where the proteins have similar mass and charge, the Gábor spectrogram may not provide enough separation in the frequency domain to completely separate their signals. In such cases, the frequency resolution of the GT spectrogram can be deliberately decreased to increase mass resolution. In Figure 4d-f, a mass spectrum containing strongly overlapped signals arising from intact trastuzumab and trastuzumab fragment ions from mouse serum was analyzed using the default frequency resolution setting as well as a "high mass resolution" setting in the iFAMS Quant STFT parameters menu. With the default frequency resolution, approximately three-quarters of the trastuzumab fragment ion peaks completely overlap in the Gábor spectrogram with signals from intact trastuzumab ions. Using the high mass resolution setting, more than half of the fragment ion peaks are well-separated in the Gábor spectrogram from neighboring intact trastuzumab ion peaks. By omitting the remaining, still-overlapped charge states from the Gábor spectrogram selection, it was possible to deconvolve both the intact trastuzumab and the major interfering fragment and identify the fragment as trastuzumab minus one light chain based on its deconvolved mass.

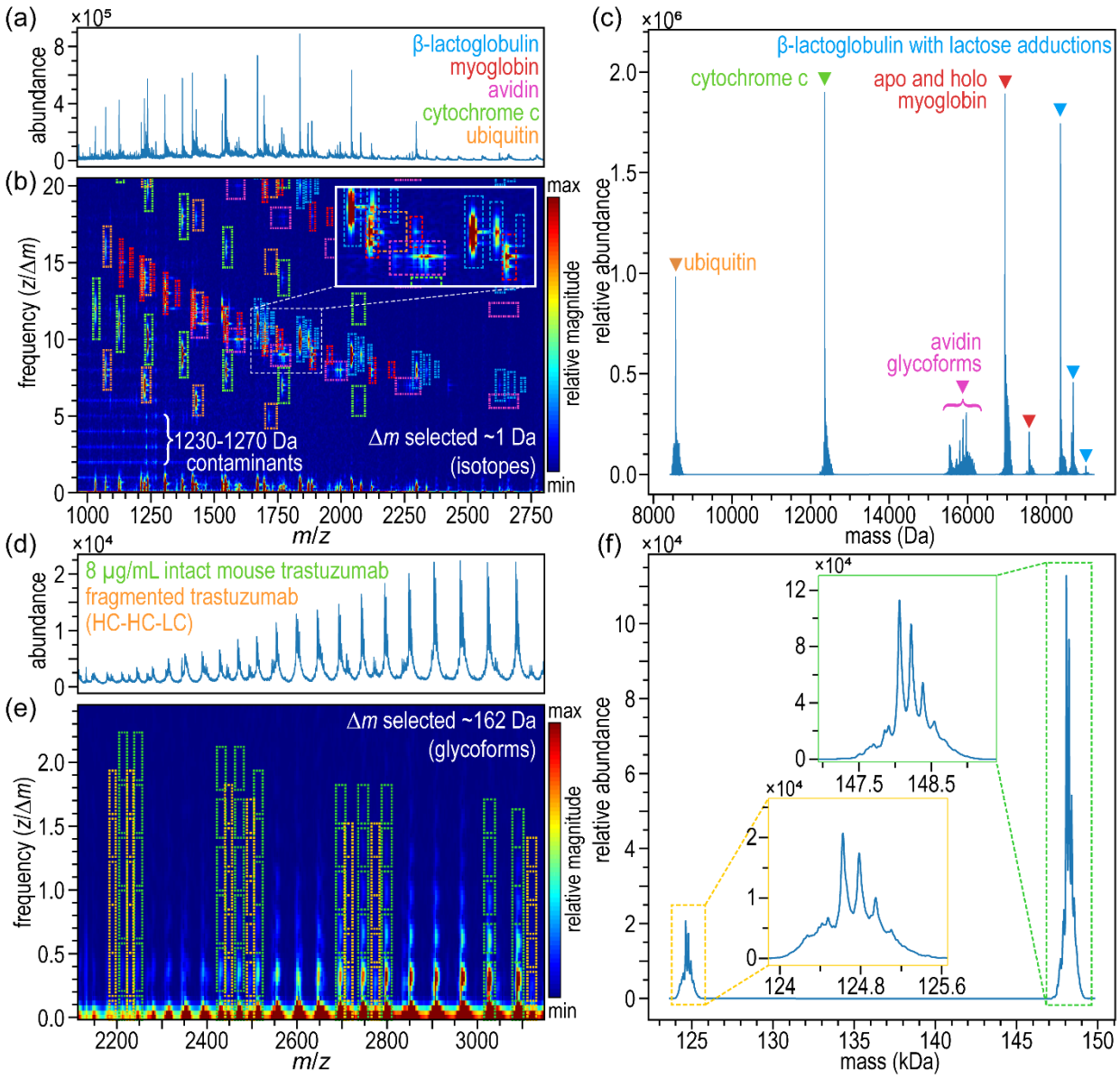


Figure 4. Simultaneous deconvolution of multiple proteins in challenging cases. (a) Mass spectrum of the "multi-protein" sample containing ubiquitin, cytochrome c, avidin, myoglobin, and β-lactoglobulin. (b) Corresponding Gábor spectrogram with signal selections shown with multi-colored dashed boxes. Zoom-in overlayed to emphasize congestion of the mass spectrum and beneficial separation provided by GT. (c) Resulting deconvolved mass spectrum with major proteins/protein-adduct states labeled. (d) Mass spectrum containing 8 µg/mL intact mouse trastuzumab and a trastuzumab fragment missing a single light chain. (e) Corresponding Gábor spectrogram with high mass resolution. Dashed boxes indicate intact trastuzumab selections in green and trastuzumab fragment selections in orange. Note that selections were only made where adequate separation of the two species was achieved in the Gábor Spectrogram. (f) Resulting deconvolved mass spectrum.

Comparison of GT deconvolution to theoretical isotope distributions. To assess the quality of GT deconvolution, theoretical isotope distributions were calculated from the UniProt amino acid sequences for the protein standards in the multi-protein sample. The UniProt amino acid sequences and assumptions made regarding post-translational modifications (PTMs) can be found in the Supporting Information (see Table S5). The built-in isotope distribution calculator (which also uses Fourier Transform methods) in iFAMS was used, and the resulting overlays of the deconvolved and theoretical mass distributions are also shown in the Supporting Information (see Figure S3).

The deconvolved mass spectra for ubiquitin and cytochrome c, which have no PTMs, had excellent agreement in terms of both mass and abundance relative to the highest-abundance isotope peak with the theoretical distributions, including adductions of up to five sodium ions. Avidin, by contrast, is a glycoprotein with numerous glycoforms that can be further obfuscated by overlapping sodium-adduction states, thus this protein represents a considerably more challenging case than ubiquitin or cytochrome c. Although peaks consistent with several previously described major glycoforms of avidin were identified in the deconvolved mass spectrum,[48] isotope distributions could not be as confidently compared with theoretical distributions due to a high degree of overlap of various sodium and potassium adduct glycoforms with unadducted glycoforms (see Figure S4 for avidin glycoform assignments). Apo- and holo-myoglobin were identified in the mass spectrum with a variety of sodium adductions with accurate average masses. Apomyoglobin and associated adduct states had excellent agreement with the theoretical distribution, but, notably, holomyoglobin had a broader distribution than was predicted by the sequence (see Figure S3). β-lactoglobulin appeared with zero, one, and two adductions of lactose (see Figures 4c and S3), and isotope masses for all three adduct states had excellent agreement with the theoretical distributions (see Table S6).

iFAMS-, MaxEnt-, and UniDec-deconvolved mass spectra for these proteins were compared to each other and their theoretical isotope distributions (example results for ubiquitin and β-lactoglobulin shown in Figure 5). It should be noted that the Agilent implementation of MaxEnt in MassHunter BioConfirm software is not designed to model isotope-resolved deconvolved mass spectra of proteins, but it is of interest to compare the width of peaks produced by MaxEnt to isotopically resolved ones. Of the six major protein-adduct peaks that were compared to theoretical isotope distributions (not including the many sodium-adducted

protein peaks or avidin peaks), only holomyoglobin had a significant discrepancy (see Figure S3), with the deconvolved distribution being markedly wider (about 35% wider) than the theoretical one whereas the other peaks had a 5% RMSD. The origin of this discrepancy is unclear.

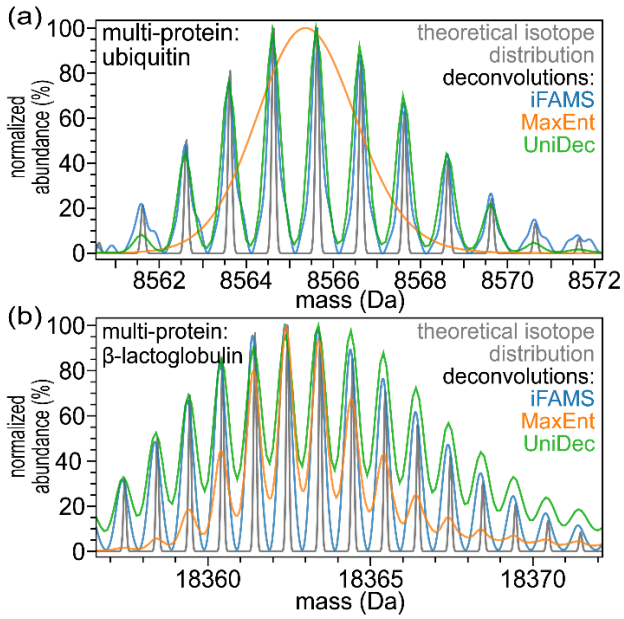


Figure 5. Comparison of deconvolution with iFAMS to UniDec and Agilent MassHunter Bioconfirm's MaxEnt. (a-b) Comparison of deconvolved mass spectra of ubiquitin and β-lactoglobulin, respectively, from the multi-protein sample (same as sample as in Figure 4). The theoretical isotope distribution is also shown in gray with an arbitrary peak width. Protein sequences and assumptions made on structure are detailed in Table S5.

Comparison of GT deconvolution to MaxEnt and UniDec. To compare the quality of multi-protein deconvolution in iFAMS to more conventional deconvolution methods, several of the mass spectra described above were deconvolved in MaxEnt and UniDec (see Figure 5). Mass accuracy was compared between methods using the 10 ng NIST mAb spectrum from Figure 1. Mass accuracy generally appeared to be similar for MaxEnt and even better for UniDec except for the G2F/G2F glycoform, which was outside of FDA tolerance (± 25 ppm) for both MaxEnt and UniDec (see Figure S5 for deconvolved mass spectra and Table S4 for accurate mass assessment of each glycoform). Due to that glycoform, the RMSDs across the six most abundant glycoforms for MaxEnt (13.1 ppm) and UniDec (15.0 ppm) were larger than for iFAMS (8.0 ppm).

Peak resolution was also compared between iFAMS Quant, MaxEnt, and UniDec. Qualitatively, the MaxEnt deconvolution appeared to better resolve many of the less abundant glycoforms, especially minor ones in between the four major glycoforms. For the multi-protein sample, MaxEnt resulted in inconsistent isotope resolution across the deconvolved mass spectrum, ranging from no isotope resolution for ubiquitin, cytochrome c, and apomyoglobin to some isotope resolution for holomyglobin and β-lactoglobulin (see Figure 5, S3). In sharp contrast to results from iFAMS Quant, baseline isotope resolution was not achieved for any of the five proteins using MaxEnt. Isotope resolution was achieved with UniDec after adjusting input parameters to increase the sampling rate of the mass spectrum and deconvolved mass spectrum. Additionally, the deconvolution with UniDec appeared to have increased noise and an elevated baseline that both increased at higher masses which made feature identification more difficult. For these reasons, iFAMS Quant had the most identifiable features (63 compared to 53 with MaxEnt and 50 with UniDec). In addition to inconsistent resolution, the isotope envelope for every MaxEnt peak was approximately 55-65% as wide as its theoretical mass distribution, with the largest discrepancies in the most abundant peaks (42% RMSD compared to 5% RMSD with iFAMS and 6% with UniDec for the top five most abundant features). Together, these results suggest that, while MaxEnt can provide good accurate mass values, resolution of non-isotopically resolved peaks (as in the NIST mAb spectrum) using MaxEnt can be artificially high. By contrast, iFAMS Quant and UniDec achieved both excellent mass accuracy (with RMSDs of 40.8 ppm and 37.9 ppm, respectively, compared to MaxEnt with 39.9 ppm) and isotope profiles (with FWHM RMSDs of 16.4% and 18.2%, respectively, compared to MaxEnt with 32.6%) as compared to theory for 17 of the major proteoforms and adduct states identified. More details on the comparison between the different deconvolution methods applied to the multi-protein sample can be found in the Supporting Information (see Figure S6 and Table S6).

Deconvolution refinement. As described above, GT deconvolution with iFAMS can sometimes result in artifactual signals in the deconvolved mass spectrum due to inclusion of contaminant data during the box selection process. However, many such artifacts can be easily removed using a "refinement" procedure. For signal that has well resolved isotopes, the correctly deconvolved mass spectrum should have a regular spacing of one Dalton. Signals that have improperly assigned charge states will have slightly different spacings that will misalign with the correct signal and cause asymmetric peak broadening or interspersed peaks in the deconvolved

mass spectrum. Performing GT on an already deconvolved mass spectrum with isotope resolution yields a "double GT" spectrogram with signal symmetric about a frequency of one and identically shaped harmonics at integer value frequencies. Artifactual signals with erroneously assigned charges states will appear in the double GT spectrogram as highly asymmetric, offset from an integer frequency, and/or only at integers ≥ 2. Selecting only the signals occurring at integer frequency (and their harmonics) in the double GT spectrogram and then performing inverse GT, artifactual signals can effectively be filtered out of the deconvolved mass spectrum. This refinement process is illustrated for the multi-protein sample in Figure 6, for which all but one artifact could be removed. This also improved the accuracy of isotope envelope widths in the deconvolved mass spectrum from 16.4% RMSD to 13.0% and the mass accuracy RMSD from 40.8 ppm to 36.4 ppm (for the same 17 features used above). While the refinement procedure in this version of iFAMS Quant is manual, a future release with automated refinement is planned.

We anticipate that this procedure could often be useful for detecting higher-order oligomers, because their signal would appear at higher integer-valued frequencies. However, in cases where monomers (or other small oligomers) of the same protein strongly overlap with higher-order oligomer signals at the same m/z, uniquely reconstructing their mass distributions could be challenging without making assumptions about how they contribute to these overlapped signals.[33] However, in native mass spectra, higher-order oligomers often appear at higher m/z than do lower-order oligomers, mitigating this effect.

Some harmonic artifacts can persist through this procedure if proteins with integer multiple masses are present in the sample. This can occur, for example, when one protein in the sample has roughly twice the mass of another protein in the sample, causing the fundamental signal of the heavier protein to align with the second harmonic of the lighter protein. In the double GT spectrogram, the signal will still appear symmetric at a frequency of one, making detection and removal difficult. This can be seen in the multi-protein deconvolution and its refinement (see Figure 6) since ubiquitin is nearly half the mass of apomyoglobin. Interestingly, a similar artifact is also present in the MaxEnt and UniDec deconvolutions which can be seen in the Supporting Information (see Figure S7).

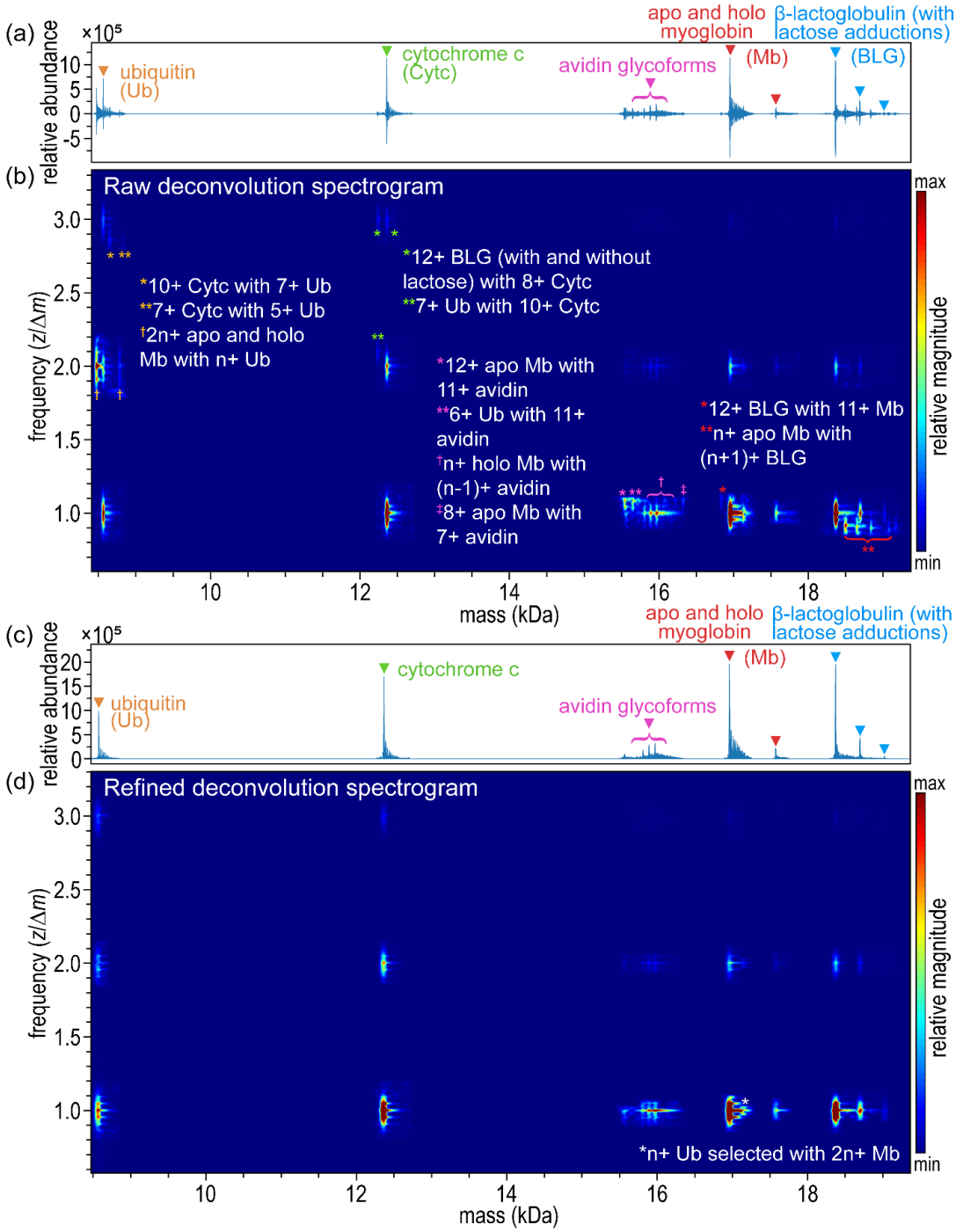


Figure 6. Demonstration of deconvolution refinement on the multi-protein sample. (a) Raw deconvolved mass spectrum of the multi-protein sample from selections prioritizing the inclusion of analyte signal over exclusion of adjacent contaminant signal. Note that the relative abundances oscillate about 0 because near-zero frequencies were excluded, and baseline correction was avoided to reduce estimation errors carrying through. (b) Gábor spectrogram of the raw deconvolved mass spectrum with erroneous signal manually flagged and labeled. (c) Refined deconvolved mass spectrum from excluding flagged signal. (d) Refined Gábor spectrogram with single persistent erroneous signal labeled.

"Quick deconvolution" with iFAMS "Gábor Slicer" tool. Since most protein signals appear within the Gábor spectrogram along predictable curves, analyzing slices of the Gábor spectrogram along these trends of constant $nM/\Delta m$ allows for rapid ion mass estimation. Figure 7 demonstrates how the "Gábor Slicer" tool in iFAMS allows a user to quickly explore distributions of mathematically related signals in the spectrogram. As the user drags the curved trace across the spectrogram using the mouse, the magnitudes of the Gábor spectrogram pixels intersected by the curved trace are instantly plotted against m/z in the Gábor Slicer mass spectrum. Effectively, the Gábor Slicer instantly isolates the mass spectrum signal of a single ion series (albeit with reduced resolution compared to "full" GT deconvolution). The resolution of the Gábor Slicer mass spectrum is determined from the STFT parameters, involving a similar trade-off between frequency resolution and mass spectrum resolution as in "full" GT deconvolution. Once the user releases the curved trace, the ion mass (M) and $\Delta m/n$ (where n is the harmonic number) are instantly estimated from the spacing between the most abundant peaks and the position of the slice within the Gábor spectrogram. (It is important to note that, if the signal arises from different charge states of an analyte in the mass spectrum with no resolvable periodicity, the peaks in the Gábor spectrogram will not be centered along the predicted curve of the slice, and the mass estimation can be highly sensitive to noise and contaminating signal.) The value of $\Delta m/n$ can be useful in determining whether a Gábor Slicer-selected peak series represents fundamentals or a higher harmonic; e.g., if $\Delta m/n \approx 0.5$, the user may have selected the second harmonics of a series corresponding to isotope-resolved charges states, for which $\Delta m/n$ would be ~1. In such a case, the user may wish to find and select the appropriate fundamental peak series for "full" deconvolution.

Occasionally, interfering peaks from other ions may fall along the GT-domain curve selected with the Gábor Slicer. To mitigate the effect of high-magnitude interfering signal (e.g., from a low-mass ion), the Gábor Slicer tool attempts to determine which subset of peaks along the curve most likely represents a charge state distribution for a single mass. In detail, each subset search selects different pairs of local maxima, calculates the mass and charge assuming the two peaks are adjacent charge states, and averages the peak magnitude of data located at the theoretical positions of the six charge states centered about the initial peak pair. The search with the largest average magnitude is used for the final mass estimate. Gábor Slicer mass estimates are often accurate within 10 Da despite being made from a very small subset of the GT

spectrogram data. The Gábor Slicer thus helps the user rapidly determine whether a series of peaks consistent with an expected mass is present, whether it is a fundamental or higher-harmonic series, as well as determine which peaks along a curve of constant $nM/\Delta m$ likely belong to the same mass, which can be especially useful for highly congested GT spectrograms. Information from the Gábor Slicer can also be useful in choosing signals for the higher-accuracy "full" GT deconvolution algorithm described above.

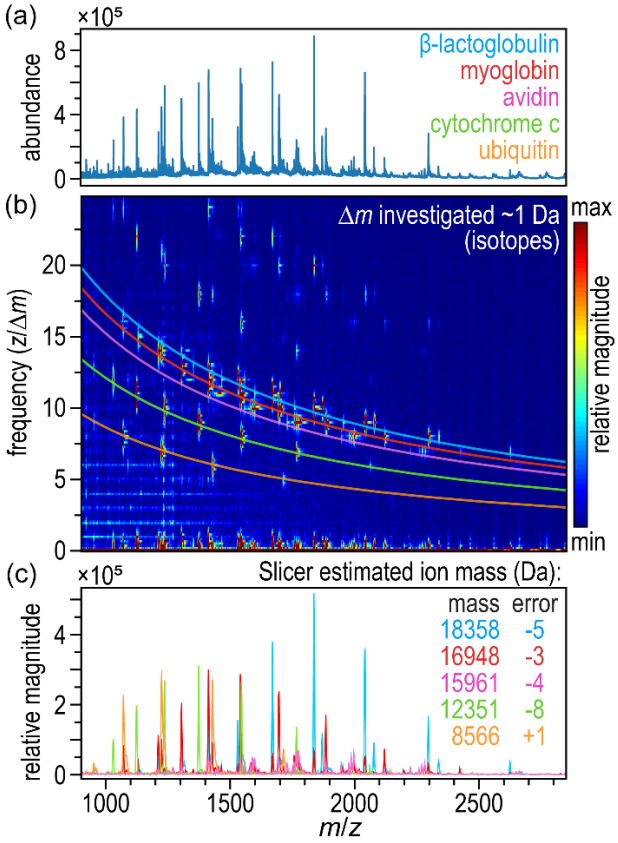


Figure 7. Demonstration of Gábor Slicer on the multi-protein sample. (a) Mass spectrum of the "multi-protein" sample containing ubiquitin, cytochrome c, avidin, myoglobin, and β-lactoglobulin. (b) Gábor spectrogram with "slices" tracing trends corresponding to the signal of each protein. (c) Gábor spectrogram "slices" plot with instant mass estimates and mass errors relative to literature values displayed in Daltons. Note that "slices" appear one at a time in the iFAMS software and multiple have been manually overlayed here for illustrative purposes.

iFAMS has batch processing and calibration capabilities. Having optimized and tested performance characteristics of iFAMS Quant on mass accuracy as well as baseline and artifact removal, we applied it to a state-of-the-art challenge in clinical and biopharmaceutical research: intact mAb quantitation. iFAMS Quant includes options for batch deconvolution and automated

calibration from LC/MS data. Once a spectrum has been processed, the parameters for deconvolution and integration can be saved and loaded into iFAMS at any point for quick and automatic processing of other spectra with the same parameters. (Because identical box selections are used to maximize traceability, we recommend that batch parameters are regenerated for each new analyte.) The time batch processing takes depends mostly on the size of the data files, but, on average, batch processing takes roughly three minutes per ten spectra using a single processor on a desktop computer. Deconvolved mass spectra generated this way can be compared quantitatively and used to generate a calibration curve within iFAMS. Calibration sets can be processed with or without internal standards and fit to linear, quadratic, or logistic curves which can then be used to calculate concentrations from quality control or unknown sample spectra. All user selections are stored in human-readable parameter files for traceability and exact repeatability.

To illustrate the quantitative capabilities of iFAMS Quant, mass spectra for a ten-level calibration series (nominal concentrations from 0.01 to 10 μg/mL) of NIST mAb were acquired with an Agilent 6545XT AdvanceBio LC/Q-TOF instrument in triplicate on three separate days (see Figure 8). Although BSA was spiked into the calibration samples primarily to passivate the LC plumbing, we did also try quantitation using the BSA as an internal standard. Those results are slightly worse, likely due to the BSA not being a protein with mass or structure very similar to NIST mAb and falling in a very different m/z region (see Figure S8). To compare calibration results from iFAMS with results from a more conventional extracted ion chromatogram (EIC) based method, signal response was instead normalized to the average method response per unit concentration of NIST mAb of the top four calibrant levels. Note that, with iFAMS Quant's batch processing option, the selections from the Gábor spectrogram and signal integration parameters are identical for every spectrum processed. Importantly, although the most abundant charge states at low concentrations were barely detectable above the noise, applying the batch selections made at a higher concentration resolved the major glycoforms in the deconvolved mass spectra even for the lowest-concentration data (0.01 μg/mL), which were estimated to be above the lower limit of quantitation (LLOQ) based on a 10:1 signal-to-noise criterion. (Examples of batch deconvolution are shown in Figure S9.) Additional information on how one might determine signal-to-noise and LLOQ for calibration data processed in iFAMS Quant can

be found in the Supporting Information ("*iFAMS Quant signal-to-noise and LLOQ determination*").

Results from this method of calibration were compared to those from an EIC-based method described elsewhere for mAb samples.[16] Briefly, for the EIC-based method, the expected *m*/*z* position of the two most abundant glycoforms (G0F/G1F and G1F/G1F) of the four most abundant charge states (48+ through 51+) were used to generate the chromatograms which were integrated over a fixed retention time interval (3.42-3.64 min). Results from the iFAMS and EIC-based methods are very consistent, but a few differences stand out: the EIC method produced smaller standard deviations between replicates at low levels (0.01-0.4 μg/mL), but the iFAMS calibration had a slightly higher R^2 (0.99986 compared to the EIC calibration R^2 of 0.99969) and smaller standard deviations at high levels (see Figure 8). Since 1/x-weighted quadratic curves are often used in clinical applications, the two methods were also compared with this curve type (see Figure S10) which had poorer fits, as judged by their lower adjusted R^2 values, than the unweighted logistic curves but fit the iFAMS-processed data better than the EIC-processed data. The iFAMS calibration had a higher R^2 (0.99927 vs. 0.99507 for EIC) and a curve that better represented the slight plateau of the data at higher concentrations with a concave-down parabola. The EIC calibration, on the other hand, yielded a concave-up parabola.

Additional advantages of the iFAMS quantitation method include the ability to rapidly change which signals are included for quantitation after processing the mass spectra and increased tolerance for instrumental drift in mass accuracy. Unlike the EIC method, in which any information not included by user selection before the processing step is lost after processing, the entire glycoform distribution is deconvolved in iFAMS, and the desired subset of charge states and glycoforms for quantitation can be easily toggled within the calibration menu. For example, mass spectra of NIST mAb typically exhibit four or five prominent glycoforms that are good candidates for quantitation. However, upon toggling on and off different glycoforms in the iFAMS calibration menu, it was quickly determined that the three most abundant glycoforms yielded the highest-quality calibration. (Effects of different levels of automation in iFAMS calibration are described in greater detail in the Supporting Information and Figure S11.) For these data, omitting the remaining high-abundance glycoforms resulted in a better calibration likely due to an interference of these peaks with others from a 149 kDa contaminant visible in the deconvolved mass spectra at middle to low calibrant levels. This contaminant impacted the

relative abundance of the glycoforms on the high-mass end of the distribution and had little effect on the glycoforms on the low-mass end (including the three most abundant glycoforms).

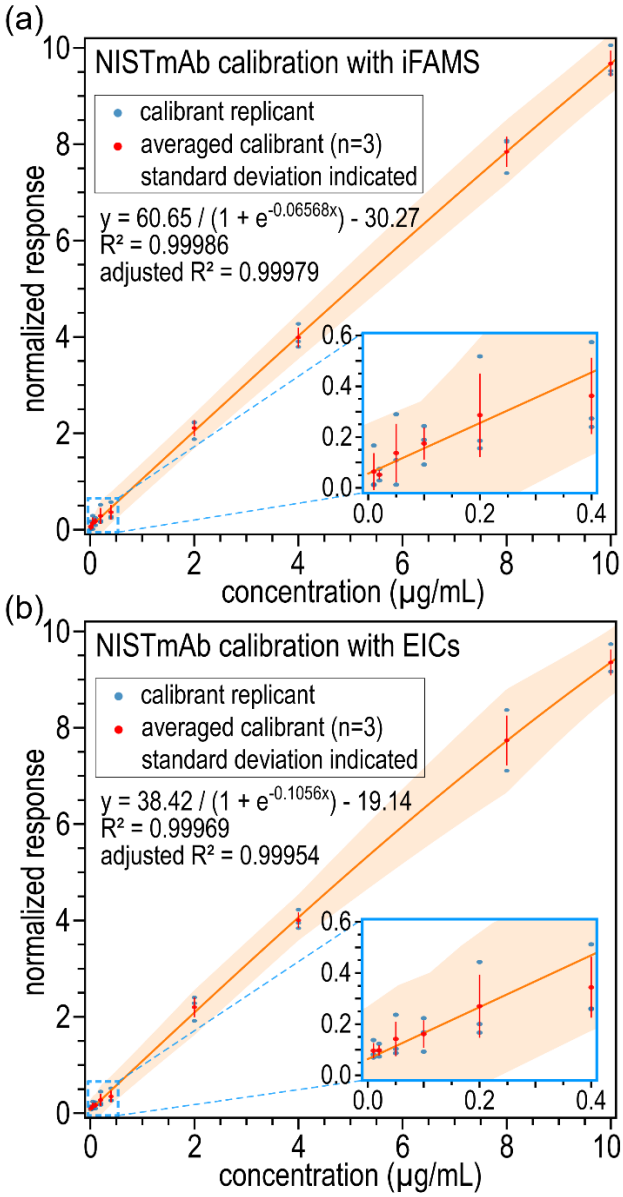


Figure 8. Calibration of intact NIST mAb for quantitative comparison between iFAMS Quant and an EIC-based approach. Concentrations range from 0.01 µg/mL to 10 µg/mL, and each level was averaged across three measurements taken on separate days. Vertical red bars indicate one standard deviation, and orange band indicates the 95% confidence interval. (a) Calibration results using iFAMS Quant. (b) Calibration results using the EIC-based approach.

Conclusions

In this paper, we demonstrated new developments to the open-source Fourier/Gábor Transform-based deconvolution software, iFAMS, including tools for automation and quantitation. The new automated tools not only make the iFAMS deconvolution algorithm easier to apply but also greatly increase the robustness of the deconvolution by improving reproducibility and objectivity and storing all applied parameters. Reducing the amount of user intervention needed and recording parameters used for each step are crucial aspects of a transparent and unbiased workflows often required for clinical applications by oversight organizations such as the U.S. Food and Drug Administration. Additionally, we demonstrated some new tools that make it easier to apply iFAMS deconvolution to a wider range of applications, such as peak selection for non-periodic signal, a charge adjuster for rapid verification of charge state assignments, and the Gábor Slicer tool for "instant" deconvolution in multi-analyte mass spectra. This increased flexibility makes iFAMS a more powerful tool capable of handling complex samples like those encountered in MS-based tissue imaging and pharmaceutical characterization, among many other applications.

The quality of deconvolution from iFAMS was assessed by comparison to other common deconvolution algorithms and expected theoretical values, and iFAMS performed favorably as compared to the widely-used open-source deconvolution software UniDec and Agilent's implementation of MaxEnt. In terms of mass accuracy, iFAMS outperformed MaxEnt and UniDec for the most abundant features but had a similar RMSD for mass accuracy when considering all features identified in the deconvolved mass spectra. Furthermore, iFAMS was able to produce isotopically resolved deconvolved mass spectra that matched sequence-based distributions with greater accuracy than MaxEnt. Generally, the quality of deconvolution generated from iFAMS was comparable to that generated by UniDec despite the two algorithms employing entirely different methods (i.e., based on Gábor spectrogram vs. m/z-domain data). Additionally, we demonstrated a new GT-based deconvolution artifact removal algorithm that further improved the quality of deconvolution.

We also assessed calibration quality by comparing calibration curves generated with iFAMS to one generated using extracted ion chromatograms, which is a currently accepted method of quantitation. iFAMS performed similarly to this accepted method for intact antibody quantitation but with the added benefit of being near-automatic in application and allowing for more insights into analyte composition by also producing a deconvolved mass distribution.

Future studies will more thoroughly assess the application of iFAMS Quant to problems including clinical protein quantitation and tissue imaging with nano-DESI MS, however the usefulness of iFAMS is not limited to intact protein quantitation, as the algorithm can deconvolve many mass spectra containing multiple charge states of one or more ions, such as mass spectra of oligonucleotide samples. Additionally, tools demonstrated here could have further applications to other areas of research in which spectra containing periodic signals are studied. iFAMS Quant software, which is written in Python, is free and publicly available at https://github.com/prellgroup/iFAMS/releases/.

Supporting Information

The Supporting Information is available free of charge at: https://pubs.acs.org

- Detailed instrument acquisition parameters; details on specific iFAMS algorithm parameters; additional deconvolved mass spectra of NIST mAb and the multi-protein sample; additional details on the mass accuracy analyses on NIST mAb and the multi-protein sample; additional example of the charge adjuster tool applied to NIST mAb; sequences and assumptions made for the multi-protein sample theoretical spectra; details on signal to noise and LLOQ determination; and alternative calibration curves to the same NIST mAb calibration series.

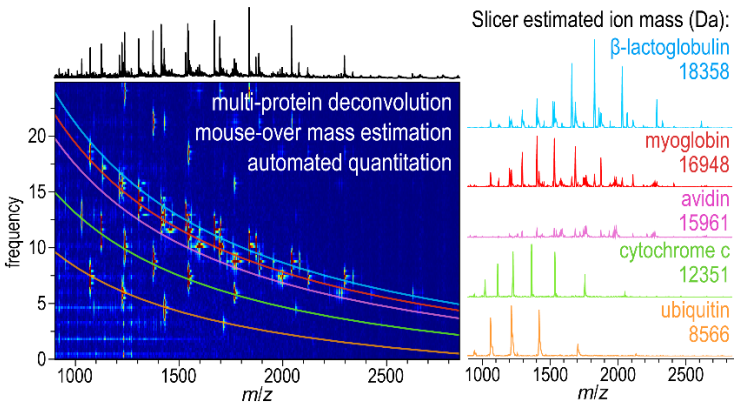
Acknowledgments

Research reported in this publication was supported by the National Science Foundation (award CHE-1752994 to JSP) and Agilent Technologies, Inc. (University Relations award #4524). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation.

Conflict of Interest Disclosure

The Agilent 6545XT AdvanceBio Q-TOF mass spectrometer with a 1290 Infinity II LC used for this research was generously loaned to the University of Oregon by Agilent Technologies, Inc. This research was supported in part by a University Relations award to JSP (#4524).

TOC Graphic.


multi-protein deconvolution
mouse-over mass estimation
automated quantitation
frequency
m/z
Slicer estimated ion mass (Da):
β-lactoglobulin
18358
myoglobin
16948
avidin
15961
cytochrome c
12351
ubiquitin
8566

References

(1) Cradic, K. W.; Ladwig, P. M.; Rivard, A. L.; Katrangi, W.; Wintgens, K. F.; Willrich, M. A. V. Vedolizumab Quantitation Using High-Resolution Accurate Mass-Mass Spectrometry Middle-up Protein Subunit: Method Validation. *Clin. Chem. Lab. Med. CCLM* **2020**, *58* (6), 864–872. https://doi.org/10.1515/cclm-2019-0862.

(2) van der Burg, D.; Wätzig, H.; Sänger-van de Griend, C. E. Method Development for Quantitative Monitoring of Monoclonal Antibodies in Upstream Cell-Culture Process Samples with Limited Sample Preparation – An Evaluation of Various Capillary Coatings. *ELECTROPHORESIS* **2023**, *44* (1–2), 96–106. https://doi.org/10.1002/elps.202200144.

(3) Lázár-Molnár, E.; Delgado, J. C. Immunogenicity Assessment of Tumor Necrosis Factor Antagonists in the Clinical Laboratory. *Clin. Chem.* **2016**, *62* (9), 1186–1198. https://doi.org/10.1373/clinchem.2015.242875.

(4) Mills, J. R.; Cornec, D.; Dasari, S.; Ladwig, P. M.; Hummel, A. M.; Cheu, M.; Murray, D. L.; Willrich, M. A.; Snyder, M. R.; Hoffman, G. S.; Kallenberg, C. G. M.; Langford, C. A.; Merkel, P. A.; Monach, P. A.; Seo, P.; Spiera, R. F.; St. Clair, E. W.; Stone, J. H.; Specks, U.; Barnidge, D. R. Using Mass Spectrometry to Quantify Rituximab and Perform Individualized Immunoglobulin Phenotyping in ANCA-Associated Vasculitis. *Anal. Chem.* **2016**, *88* (12), 6317–6325. https://doi.org/10.1021/acs.analchem.6b00544.

(5) Ji, Q. C.; Rodila, R.; Gage, E. M.; El-Shourbagy, T. A. A Strategy of Plasma Protein Quantitation by Selective Reaction Monitoring of an Intact Protein. *Anal. Chem.* **2003**, *75* (24), 7008–7014. https://doi.org/10.1021/ac034930n.

(6) Wang, S.-L.; Ohrmund, L.; Hauenstein, S.; Salbato, J.; Reddy, R.; Monk, P.; Lockton, S.; Ling, N.; Singh, S. Development and Validation of a Homogeneous Mobility Shift Assay for the Measurement of Infliximab and Antibodies-to-Infliximab Levels in Patient Serum. *J. Immunol. Methods* **2012**, *382* (1), 177–188. https://doi.org/10.1016/j.jim.2012.06.002.

(7) Hickey, J. W.; Neumann, E. K.; Radtke, A. J.; Camarillo, J. M.; Beuschel, R. T.; Albanese, A.; McDonough, E.; Hatler, J.; Wiblin, A. E.; Fisher, J.; Croteau, J.; Small, E. C.; Sood, A.; Caprioli, R. M.; Angelo, R. M.; Nolan, G. P.; Chung, K.; Hewitt, S. M.; Germain, R. N.; Spraggins, J. M.; Lundberg, E.; Snyder, M. P.; Kelleher, N. L.; Saka, S. K. Spatial Mapping of Protein Composition and Tissue Organization: A Primer for Multiplexed Antibody-Based Imaging. *Nat. Methods* **2022**, *19* (3), 284–295. https://doi.org/10.1038/s41592-021-01316-y.

(8) Hu, H.; Helminiak, D.; Yang, M.; Unsihuay, D.; Hilger, R. T.; Ye, D. H.; Laskin, J. High-Throughput Mass Spectrometry Imaging with Dynamic Sparse Sampling. *ACS Meas. Sci. Au* **2022**, *2* (5), 466–474. https://doi.org/10.1021/acsmeasuresciau.2c00031.

(9) Piehowski, P. D.; Zhu, Y.; Bramer, L. M.; Stratton, K. G.; Zhao, R.; Orton, D. J.; Moore, R. J.; Yuan, J.; Mitchell, H. D.; Gao, Y.; Webb-Robertson, B.-J. M.; Dey, S. K.; Kelly, R. T.; Burnum-Johnson, K. E. Automated Mass Spectrometry Imaging of over 2000 Proteins from

Tissue Sections at 100-Mm Spatial Resolution. *Nat. Commun.* **2020**, *11* (1), 8. https://doi.org/10.1038/s41467-019-13858-z.

(10) Hart, M. H.; de Vrieze, H.; Wouters, D.; Wolbink, G.-J.; Killestein, J.; de Groot, E. R.; Aarden, L. A.; Rispens, T. Differential Effect of Drug Interference in Immunogenicity Assays. *J. Immunol. Methods* **2011**, *372* (1), 196–203. https://doi.org/10.1016/j.jim.2011.07.019.

(11) Ladwig, P. M.; Barnidge, D. R.; Willrich, M. A. V. Mass Spectrometry Approaches for Identification and Quantitation of Therapeutic Monoclonal Antibodies in the Clinical Laboratory. *Clin. Vaccine Immunol.* **2017**, *24* (5), e00545-16. https://doi.org/10.1128/CVI.00545-16.

(12) Vasicek, L. A.; Spellman, D. S.; Hsieh, S.; Seghezzi, W.; Zhang, S.; Santostefano, M.; Bateman, K. P. Quantitation of a Therapeutic Antibody in Serum Using Intact Sequential Affinity Capture, Trypsin Digestion, and LC-MS/MS. *Anal. Chem.* **2018**, *90* (1), 866–871. https://doi.org/10.1021/acs.analchem.7b03716.

(13) Schaffer, L. V.; Millikin, R. J.; Shortreed, M. R.; Scalf, M.; Smith, L. M. Improving Proteoform Identifications in Complex Systems Through Integration of Bottom-Up and Top-Down Data. *J. Proteome Res.* **2020**, *19* (8), 3510–3517. https://doi.org/10.1021/acs.jproteome.0c00332.

(14) Jian, W.; Kang, L.; Burton, L.; Weng, N. A Workflow for Absolute Quantitation of Large Therapeutic Proteins in Biological Samples at Intact Level Using LC-HRMS. *Bioanalysis* **2016**, *8* (16), 1679–1691. https://doi.org/10.4155/bio-2016-0096.

(15) Wong, D. L.; Han, M.; Barnaby, O.; Yang, Y. An Integrated Workflow for Sensitive Intact Protein Quantitation of Monoclonal Antibodies from Biological Matrix. *Agilent Technologies application note*, publication number 5994-1249EN, **2019**.

(16) Qiu, X.; Hale, W. A.; Wong, D. Ultra-Sensitive Intact Monoclonal Antibody Quantification Using AssayMAP Bravo Liquid Handling Platform and 6545XT AdvanceBio LC/Q-TOF Mass Spectrometer. *Agilent Technologies application note*, publication number 5994-1984EN, **2020**.

(17) Vasicek, L. A.; Zhu, X.; Spellman, D. S.; Bateman, K. P. Direct Quantitation of Therapeutic Antibodies for Pharmacokinetic Studies Using Immuno-Purification and Intact Mass Analysis. *Bioanalysis* **2019**, *11* (3), 203–213. https://doi.org/10.4155/bio-2018-0240.

(18) Alelyunas, Y. W.; Shion, H.; Edwards, I.; Wrona, M. Intact Trastuzumab Monoclonal Antibody Quantification Using the Xevo G2-XS QTof High Resolution Mass Spectrometer. *Waters application note*, publication number 720006222EN, **2018**.

(19) Kellie, J. F.; Schneck, N. A.; Causon, J. C.; Baba, T.; Mehl, J. T.; Pohl, K. I. Top-Down Characterization and Intact Mass Quantitation of a Monoclonal Antibody Drug from Serum by Use of a Quadrupole TOF MS System Equipped with Electron-Activated Dissociation. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (1), 17–26. https://doi.org/10.1021/jasms.2c00206.

(20) Zhang, L.; Vasicek, L. A.; Hsieh, S.; Zhang, S.; Bateman, K. P.; Henion, J. Top-down LC–MS Quantitation of Intact Denatured and Native Monoclonal Antibodies in Biological Samples. *Bioanalysis* **2018**, *10* (13), 1039–1054. https://doi.org/10.4155/bio-2017-0282.

(21) Ahene, A. B. Application and Interpretation of Free and Total Drug Measurements in the Development of Biologics. *Bioanalysis* **2011**, *3* (11), 1287–1295. https://doi.org/10.4155/bio.11.104.

(22) Hu, H.; Yin, R.; Brown, H. M.; Laskin, J. Spatial Segmentation of Mass Spectrometry Imaging Data by Combining Multivariate Clustering and Univariate Thresholding. *Anal. Chem.* **2021**, *93* (7), 3477–3485. https://doi.org/10.1021/acs.analchem.0c04798.

(23) Hale, O. J.; Cooper, H. J.; Marty, M. T. High-Throughput Deconvolution of Native Protein Mass Spectrometry Imaging Data Sets for Mass Domain Analysis. *Anal. Chem.* **2023**, *95* (37), 14009–14015. https://doi.org/10.1021/acs.analchem.3c02616.

(24) Massonnet, P.; Grifnée, E.; Farré-Segura, J.; Demeuse, J.; Huyghebaert, L.; Dubrowski, T.; Dufour, P.; Schoumacher, M.; Peeters, S.; Goff, C. L.; Cavalier, E. Concise Review on the Combined Use of Immunocapture, Mass Spectrometry and Liquid Chromatography for Clinical Applications. *Clin. Chem. Lab. Med. CCLM* **2023**, *61* (10), 1700–1707. https://doi.org/10.1515/cclm-2023-0253.

(25) Gozzo, T. A.; Bush, M. F. Quantitatively Differentiating Antibodies Using Charge-State Manipulation, Collisional Activation, and Ion Mobility-Mass Spectrometry. *Anal. Chem.* **2024**, *96* (1), 505–513. https://doi.org/10.1021/acs.analchem.3c04638.

(26) Rolland, A. D.; Prell, J. S. Approaches to Heterogeneity in Native Mass Spectrometry. *Chem. Rev.* **2022**, *122* (8), 7909–7951. https://doi.org/10.1021/acs.chemrev.1c00696.

(27) Qiu, X.; Kang, L.; Case, M.; Weng, N.; Jian, W. Quantitation of Intact Monoclonal Antibody in Biological Samples: Comparison of Different Data Processing Strategies. *Bioanalysis* **2018**, *10* (13), 1055–1067. https://doi.org/10.4155/bio-2018-0016.

(28) US Food and Drug Administration. Bioanalytical method validation guidance for industry. *US Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research and Center for Veterinary Medicine*, **2018**.

(29) Kellie, J. F.; Kehler, J. R.; Karlinsey, M. Z.; Summerfield, S. G. Toward Best Practices in Data Processing and Analysis for Intact Biotherapeutics by MS in Quantitative Bioanalysis. *Bioanalysis* **2017**, *9* (23), 1883–1893. https://doi.org/10.4155/bio-2017-0179.

(30) Sawyer, W. S.; Srikumar, N.; Carver, J.; Chu, P. Y.; Shen, A.; Xu, A.; Williams, A. J.; Spiess, C.; Wu, C.; Liu, Y.; Tran, J. C. High-Throughput Antibody Screening from Complex Matrices Using Intact Protein Electrospray Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2020**, *117* (18), 9851–9856. https://doi.org/10.1073/pnas.1917383117.

(31) Phung, W.; Bakalarski, C. E.; Hinkle, T. B.; Sandoval, W.; Marty, M. T. UniDec Processing Pipeline for Rapid Analysis of Biotherapeutic Mass Spectrometry Data. *Anal. Chem.* **2023**, *95* (30), 11491–11498. https://doi.org/10.1021/acs.analchem.3c02010.

(32) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370–4376. https://doi.org/10.1021/acs.analchem.5b00140.(33) Cleary, S. P.; Thompson, A. M.; Prell, J. S. Fourier Analysis Method for Analyzing Highly Congested Mass Spectra of Ion Populations with Repeated Subunits. *Anal. Chem.* **2016**, *88* (12), 6205–6213. https://doi.org/10.1021/acs.analchem.6b01088.

(34) Cleary, S. P.; Li, H.; Bagal, D.; Loo, J. A.; Campuzano, I. D. G.; Prell, J. S. Extracting Charge and Mass Information from Highly Congested Mass Spectra Using Fourier-Domain Harmonics. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (10), 2067–2080. https://doi.org/10.1007/s13361-018-2018-7.

(35) Cleary, S. P.; Prell, J. S. Liberating Native Mass Spectrometry from Dependence on Volatile Salt Buffers by Use of Gábor Transform. *ChemPhysChem* **2019**, *20* (4), 519–523. https://doi.org/10.1002/cphc.201900022.

(36) Swansiger, A. K.; Marty, M. T.; Prell, J. S. Fourier-Transform Approach for Reconstructing Macromolecular Mass Defect Profiles. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (1), 172–180. https://doi.org/10.1021/jasms.1c00317.

(37) Galitzine, C.; Egertson, J. D.; Abbatiello, S.; Henderson, C. M.; Pino, L. K.; MacCoss, M.; Hoofnagle, A. N.; Vitek, O. Nonlinear Regression Improves Accuracy of Characterization of Multiplexed Mass Spectrometric Assays*. *Mol. Cell. Proteomics* **2018**, *17* (5), 913–924. https://doi.org/10.1074/mcp.RA117.000322.

(38) Wong, D. L. Precise Characterization of Intact Monoclonal Antibodies by the Agilent 6545XT AdvanceBio LC/Q-TOF. *Agilent Technologies application note*, publication number 5991-7813EN, **2017**.

(39) Kurzyniec, S.; Johnson, V. Monoclonal Antibody Workflows on the Shimadzu Q-TOF LCMSTM-9030 Using the Protein Metrics Software Suite. *Shimadzu Scientific Instruments application note*, publication number SSI-LCMS-103, **2019**.

(40) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246* (4926), 64–71.

(41) Contino, N. C.; Jarrold, M. F. Charge Detection Mass Spectrometry for Single Ions with a Limit of Detection of 30 Charges. *Int. J. Mass Spectrom.* **2013**, *345–347*, 153–159. https://doi.org/10.1016/j.ijms.2012.07.010.

(42) Goodwin, M. P.; Grinfeld, D.; Yip, P.; Bowen, K. P.; Kafader, J. O.; Kelleher, N. L.; Senko, M. W. Improved Signal Processing for Mass Shifting Ions in Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2024**. https://doi.org/10.1021/jasms.3c00435.

(43) Jarrold, M. F. Applications of Charge Detection Mass Spectrometry in Molecular Biology and Biotechnology. *Chem. Rev.* **2022**, *122* (8), 7415–7441. https://doi.org/10.1021/acs.chemrev.1c00377.

(44) Harper, C. C.; Elliott, A. G.; Oltrogge, L. M.; Savage, D. F.; Williams, E. R. Multiplexed Charge Detection Mass Spectrometry for High-Throughput Single Ion Analysis of Large Molecules. *Anal. Chem.* **2019**, *91* (11), 7458–7465. https://doi.org/10.1021/acs.analchem.9b01669.

(45) Wörner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. Resolving Heterogeneous Macromolecular Assemblies by Orbitrap-Based Single-Particle Charge Detection Mass Spectrometry. *Nat. Methods* **2020**, *17* (4), 395–398. https://doi.org/10.1038/s41592-020-0770-7.

(46) Laszlo, K. J.; Bush, M. F. Analysis of Native-Like Proteins and Protein Complexes Using Cation to Anion Proton Transfer Reactions (CAPTR). *J. Am. Soc. Mass Spectrom.* **2015**, *26* (12), 2152–2161. https://doi.org/10.1007/s13361-015-1245-4.

(47) Abdillahi, A. M.; Lee, K. W.; McLuckey, S. A. Mass Analysis of Macro-Molecular Analytes via Multiply-Charged Ion Attachment. *Anal. Chem.* **2020**, *92* (24), 16301–16306. https://doi.org/10.1021/acs.analchem.0c04335.

(48) Liu, F. C.; Cropley, T. C.; Ridgeway, M. E.; Park, M. A.; Bleiholder, C. Structural Analysis of the Glycoprotein Complex Avidin by Tandem-Trapped Ion Mobility Spectrometry–Mass Spectrometry (Tandem-TIMS/MS). *Anal. Chem.* **2020**, *92* (6), 4459–4467. https://doi.org/10.1021/acs.analchem.9b05481.

Supporting Information for

Gábor Transform-based signal isolation, rapid deconvolution, and quantitation of intact protein ions with mass spectrometry

Kayd L. Meldrum[1], Andrew K. Swansiger[1], Meghan M. Daniels[1], Wendi A. Hale[2], Crystal Kirmiz Cody[2], Xi Qiu[2], Michael Knierman[2], John Sausen[2], James S. Prell[1,3,*]

[1]Department of Chemistry and Biochemistry, 1253 University of Oregon, Eugene, OR, USA 97403-1253

[2]Agilent Technologies, Inc., 5301 Stevens Creek Blvd., Santa Clara, CA, USA, 95051

[3]Materials Science Institute, 1252 University of Oregon, Eugene, OR, USA 97403-1252

[*]Corresponding author email: jprell@uoregon.edu

Table of Contents

Table S1. LC/MS acquisition parameters for NIST mAb sample used in deconvolution mass accuracy assessments.

Agilent 1290 Infinity II LC	
Column	Agilent PLRP-S, 2.1 × 50 mm, 1000 Å, 5 µm (p/n PL 1912-1502)
Autosampler Thermostat	4 °C
Solvent A	0.1% formic acid in 18 MΩ water
Solvent B	0.1% formic acid in 100% acetonitrile
Gradient	0 to 1 minutes, 5 to 20% B 1 to 3 minutes, 20 to 50% B 3 to 4 minutes, 50 to 70 % B
Multicolumn Thermostat Temperature	80 °C
Flow Rate	0.4 mL/min
Injection Volume	1 µL
Agilent 6545XT AdvanceBio Q-TOF	
Source	Dual Agilent Jet Stream
Gas Temperature	350 °C
Gas Flow	13 L/min
Nebulizer	35 psi
Sheath Gas Temperature	380 °C
Sheath Gas Flow	12 L/min
Vcap	5500 V
Nozzle Voltage	2000 V
Fragmentor	380 V
Skimmer	140 V
Quad AMU	*m*/*z* 127
Mass Range	*m*/*z* 500 to 7000
Acquisition Rate	1.0 spectra/s
Reference Mass	922.0098
Acquisition Mode	Positive, extended (m/z 10,000) mass range

Table S2. LC/MS acquisition parameters for NIST mAb calibration series.

Agilent 1290 Infinity II LC	
Column	Agilent PLRP-S, 2.1 × 50 mm, 1000 Å, 5 µm (p/n PL 1912-1502)
Autosampler Thermostat	4 °C
Solvent A	0.1% formic acid in 18 MΩ water
Solvent B	0.1% formic acid in 100% acetonitrile
Gradient	0 to 1 minutes, 5 to 20% B 1 to 6 minutes, 20 to 50% B 6 to 7 minutes, 50 to 70% B 7 to 7.1 minutes, 70 to 90% B 7.1 to 8 minutes, 90% B 8 to 8.1 minutes, 90 to 5% B
Multicolumn Thermostat Temperature	80 °C
Flow Rate	0.4 mL/min
Injection Volume	10 µL
Agilent 6545XT AdvanceBio Q-TOF	
Source	Dual Agilent Jet Stream
Gas Temperature	350 °C
Gas Flow	13 L/min
Nebulizer	35 psi
Sheath Gas Temperature	380 °C
Sheath Gas Flow	12 L/min
Vcap	5500 V
Nozzle Voltage	2000 V
Fragmentor	380 V
Skimmer	140 V
Quad AMU	*m*/*z* 127
Mass Range	*m*/*z* 100 to 3200
Acquisition Rate	1.0 spectra/s
Reference Mass	none
Acquisition Mode	Positive, standard (m/z 3200) mass range

Table S3. MS acquisition parameters for multi-protein sample.

Agilent 6545XT AdvanceBio Q-TOF	
Syringe Pump Flow	400 µL/min
Source	Dual Agilent Jet Stream
Gas Temperature	350 °C
Gas Flow	13 L/min
Nebulizer	35 psi
Sheath Gas Temperature	400 °C
Sheath Gas Flow	12 L/min
Vcap	5000 V
Nozzle Voltage	2000 V
Fragmentor	380 V
Skimmer	140 V
Quad AMU	*m*/*z* 127
Mass Range	*m*/*z* 100 to 3200
Acquisition Rate	1.0 spectra/s
Reference Mass	none
Acquisition Mode	Positive, standard (m/z 3200) mass range

iFAMS Quant default signal selection and integration parameters. The iFAMS Quant parameters that are used for signal detection and selection in the automated tools have been optimized using several types of protein mass spectra. The automated charge state selection feature includes all identified charge states with maximum magnitude at least 6% of the magnitude of the most-abundant charge state in the series. This threshold can be toggled by the user to select all potential charge states within the mass spectrum *m*/*z* range. To optimize the data included for each charge state, GT data with magnitude ≥ 15% of the charge state's peak magnitude are included in reconstructing each GT-selected peak. This threshold can also be toggled by the user to use an average box shape defined by the two user-drawn input boxes. The dimensions of the boxes are defined relative to the peak maximum and are stretched to accommodate peak narrowing and widening that occurs with higher and lower charge states, respectively. Automated harmonic selection first calculates the average of the highest identifiable harmonic (by default, having ≥ 1% of the signal of the fundamental) for all included charge states, then includes all harmonics for each charge state up to this average harmonic number.

This value is recommended to the user, but the user can easily type in an alternative desired value. Every harmonic up to the value entered will be selected for each charge state. A separate option allows the user to include the "near-zero frequency" information which is generally recommended unless three or more harmonics are well resolved or there is heavy overlap in the mass spectrum.

"Edge" cases of mass spectra with signal arising from multiple periodic mass differences (Δm) occasionally require box thresholding toggled off and manual input of two boxes for averaging. When the mass spectral signal is a convolution of two distributions with different Δm's, with one giving rise to much stronger features in the GT spectrogram than the other, the GT spectrogram has the appearance of peaks at harmonics of the "dominant" frequency signal modulated in either the frequency or m/z direction depending on the GT spectrogram resolution. (For examples, see Figure 4b or 6d in the main article: myoglobin signal is dominated by isotope resolution but modulated in the frequency direction from sodiation; avidin signal is dominated by isotope resolution but modulated in the m/z direction from glycoforms). In cases where the modulation has relatively large spacing (as in the avidin example), the automated search thresholds will result in boxes that are too small to include the modulation, so the user will need to toggle the threshold off and draw input boxes large enough to encompass the entire modulated peak.

The automatic baseline correction uses identical box selections to reproduce the baseline shape of the deconvolved mass spectrum that arises from a combination of elevated baseline and gaussian windowing from all charge states used to produce the deconvolved spectrum. This baseline shape is then vertically scaled to fit the deconvolved mass spectrum using the lowest 25% of the baseline to avoid signals from the deconvolved ions.

Due to potential overlap of neighboring peaks, peak centroids are automatically calculated as the weight-average mass of the top 25% highest-abundance data points for each peak, with the high- and low-mass boundaries for each peak determined by a combination of an abundance threshold (default ≥ 3% of the highest feature) and minimum peak spacing (default 0.8 Da). (While a noise tolerance can be user-specified as needed to avoid identifying small local minima as peaks, we found that setting this tolerance to a default value of 0 often yields better results than higher noise tolerance values.)

Table S4. NIST mAb mass accuracy assessment and comparison across methods and concentrations for the top six most abundant glycoforms. Average masses calculated from BioConfirm 10 sp1.[1] Red cells indicate mass errors outside of the FDA recommended tolerance of ±25 ppm.

	Glycoform	G0F/G0F -GlcNAc	G0F/G0F	G0F/G1F	G1F/G1F	G1F/G2F	G2F/G2F	RMSD
	Exact Mass* (Da)	147836.35	148039.43	148201.57	148363.72	148525.86	148688.00	
10 ng NIST mAb (iFAMS)	Observed Mass (Da)	147835.01	148038.17	148200.62	148364.30	148526.03	148686.50	
	Abundance (%)	7	55	100	98	52	21	
	Error (ppm)	**-9.07**	**-8.52**	**-6.44**	**3.93**	**1.12**	**-10.11**	**7.25**
10 ng NIST mAb (iFAMS without baseline correction)	Observed Mass (Da)	147835.03	148038.17	148200.63	148364.30	148526.03	148686.33	
	Abundance (%)	15.8	62.8	99.1	100.0	61.5	30.4	
	Error (ppm)	**-8.93**	**-8.51**	**-6.39**	**3.92**	**1.16**	**-11.28**	**7.49**
500 ng NIST mAb (iFAMS)	Observed Mass (Da)	147835.66	148039.61	148201.90	148363.60	148525.28	148687.64	
	Abundance (%)	10	48	100	100	39	21	
	Error (ppm)	**-4.69**	**1.18**	**2.18**	**-0.78**	**-3.92**	**-2.45**	**2.89**
MaxEnt (10 ng NIST mAb)	Observed Mass (Da)	147835.57	148039.74	148202.11	148364.46	148526.53	148683.36	
	Abundance (%)	4	60	100	85	49	32	
	Error (ppm)	**-5.28**	**2.10**	**3.62**	**5.01**	**4.50**	**-31.26**	**13.34**
UniDec (10 ng NIST mAb)	Observed Mass (Da)	147836.64	148039.88	148201.73	148363.37	148526.35	148682.62	
	Abundance (%)	24	63	100	82	49	27	
	Error (ppm)	**1.98**	**3.07**	**1.03**	**-2.35**	**3.31**	**-36.22**	**14.96**
MaxEnt Literature Reference (500 ng NIST mAb)[1]	Observed Mass (Da)	147836.90	148039.30	148201.80	148363.70	148526.00	148687.50	
	Error (ppm)	**3.54**	**-0.64**	**1.44**	**-0.06**	**0.78**	**-3.72**	**2.22**

*Average masses from BioConfirm 10 sp1

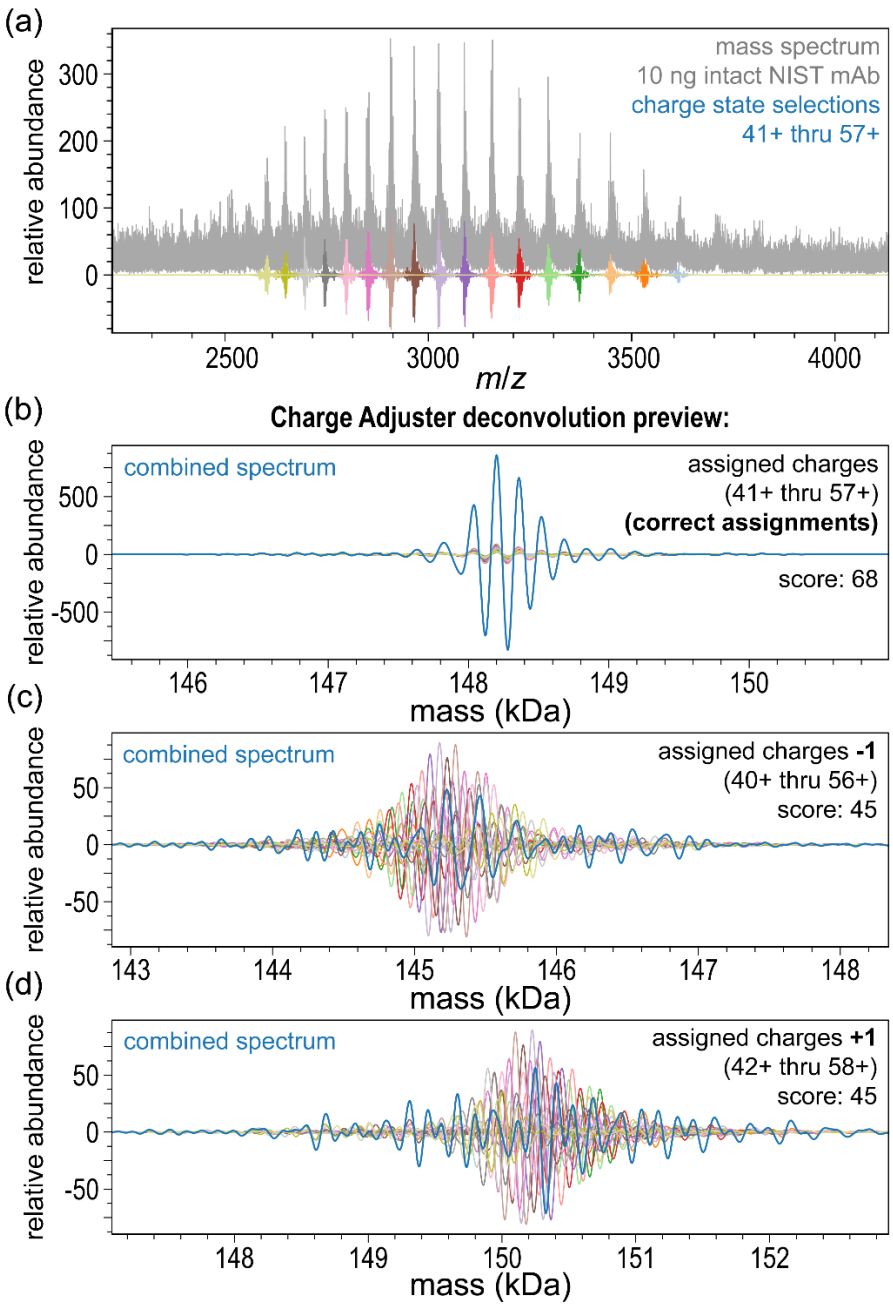


Figure S1. Demonstration of charge adjuster tool on a mass spectrum with well resolved periodic signal. (a) Mass spectrum containing 10 ng intact NIST mAb (gray) with inverse-GT single charge state selections overlayed (multi-colored). (b-d) Charge adjuster deconvolution previews with assigned charges, charges shifted minus one from assigned, and charges shifted plus one from assigned, respectively. Since charge state selections from the Gábor spectrogram used the fundamental of a periodic signal, the deconvolution previews have abundances that oscillate about zero. To help the user identify the correct charge state assignment, iFAMS calculates a score which will typically be largest for correct assignments. The score is calculated as the natural logarithm of the product of the charge-corrected inverse-GT selections (multi-colored spectra in b-d).

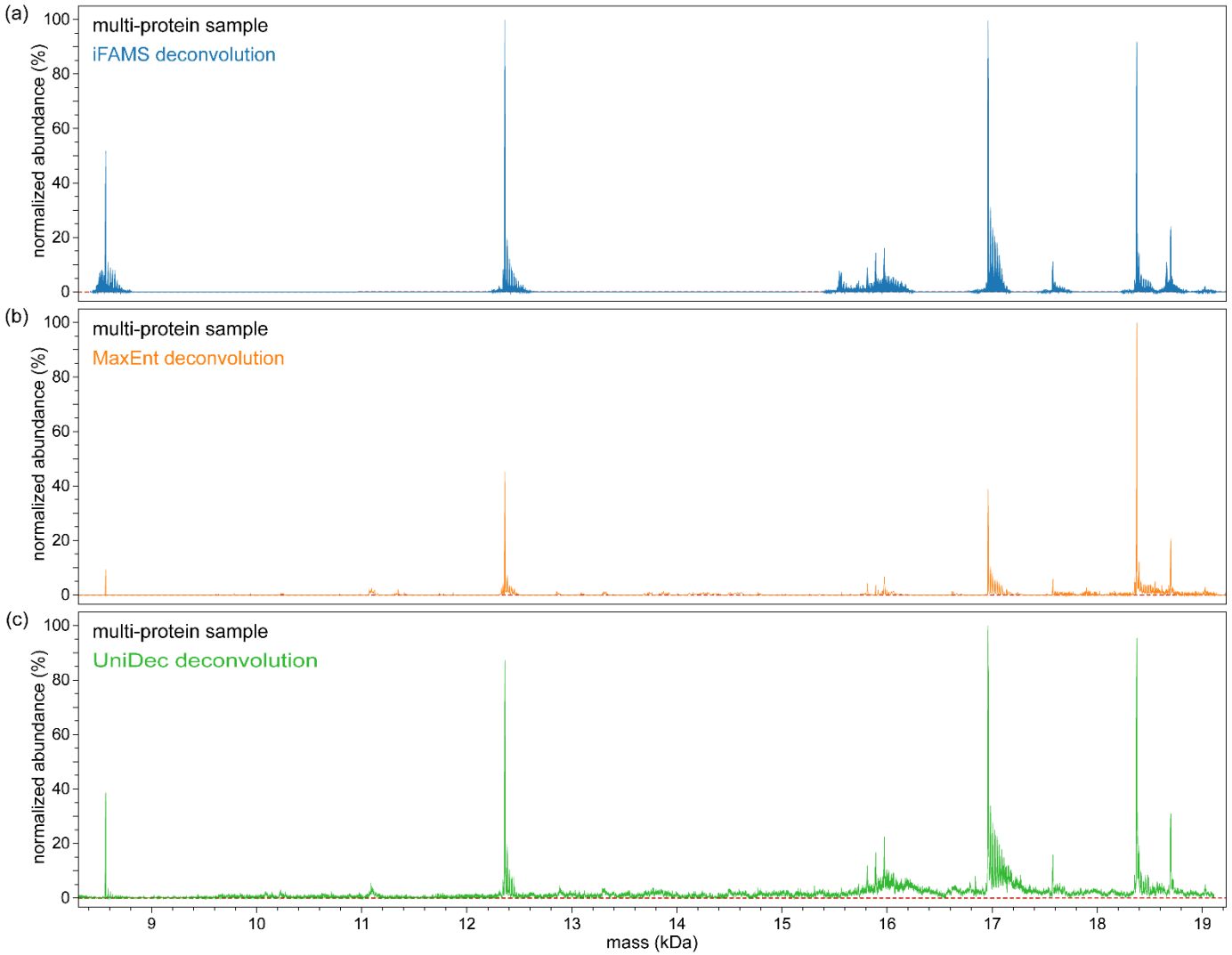


Figure S2. Full deconvolved mass spectrum comparison of different deconvolution methods with the multi-protein sample. (a) iFAMS deconvolution of multi-protein sample. Default automation parameters were used except automatic box resizing was toggled off. (b) Agilent's MaxEnt deconvolution of multi-protein sample using default parameters with an input of charge states 4-16 and mass range of 8.3-22kDa. (c) UniDec deconvolution of multi-protein sample with input parameters of charge states 4-16, mass range of 8.3-22kDa, high artifact suppression, 0.1 Da sampling step, and FWHM of 0.05 Th.

Table S5. Multi-protein amino acid sequences and modification assumptions.[2–7]

Protein	Source	Millipore Sigma SKU	UniProt Entry	Sequence	Assumptions
ubiquitin	bovine erythrocytes	U6253	P0CG53	MQIFVKTLTGKTITLEVEP SDTIENVKAKIQDKEGIPP DQQRLIFAGKQLEDGRTL SDYNIQKESTLHLVLRLR GG	Single chain
cytochrome c	equine heart	C2867	P00004	GDVEKGKKIFVQKCAQC HTVEKGGKHKTGPNLHG LFGRKTGQAPGFTYTDA NKNKGITWKEETLMEYLE NPKKYIPGTKMIFAGIKKK TEREDLIAYLKKATNE	Removed first methionine, added heme C mass minus one hydrogen to account for Fe(III) rather than Fe(II) ($+C_{34}H_{35}O_4N_4S_2Fe$), two thioether bonds (-2S and -4H), added one acetylation ($+C_2H_2O$)
avidin	egg whites	A9275	P02701	ARKCSLTGKWTNDLGSN MTIGAVNSRGEFTGTYITA VTATSNEIKESPLHGTQN TINKRTQPTFGFTVNWKF SESTTVFTGQCFIDRNGK EVLKTMWLLRSSVNDIGD DWKATRVGINIFTRLRTQ KE	Monomer, one disulfide bond (-2H), and most abundant glycoform composition with five mannose and four N-Acetylglucosamine ($+C_{62}H_{102}N_4O_{45}$)
apomyoglobin	equine skeletal muscle	M1882	P68082	GLSDGEWQQVLNVWGK VEADIAGHGQEVLIRLFT GHPETLEKFDKFKHLKTE AEMKASEDLKKHGTVVLT ALGGILKKKGHHEAELKP LAQSHATKHKIPIKYLEFIS DAIIHVLHSKHPGDFGAD AQGAMTKALELFRNDIAA KYKELGFQG	Removed first methionine, (for holomyoglobin) added heme B mass minus one hydrogen to account for Fe(III) rather than Fe(II) ($+C_{34}H_{31}O_4N_4Fe$)
beta-lactoglobulin A	bovine milk	L7880	P02754 (BLG B)	LIVTQTMKGLDIQKVAGT WYSLAMAASDISLLDAQS APLRVYVEELKPTPEGDL EILLQKWEN**D**ECAQKKIIA EKTKIPAVFKIDALNENKV LVLDTDYKKYLLFCMENS AEPEQSL**V**CQCLVRTPEV DDEALEKFDKALKALPMH IRLSFNPTQLEEQCHI	BLG A form differs from BLG B form (BLG B → BLG A) at position 64 (G → D) and position 118 (A → V), two disulfide bonds (-4H), ($+C_{12}H_{20}O_{10}$ for lactose adductions)

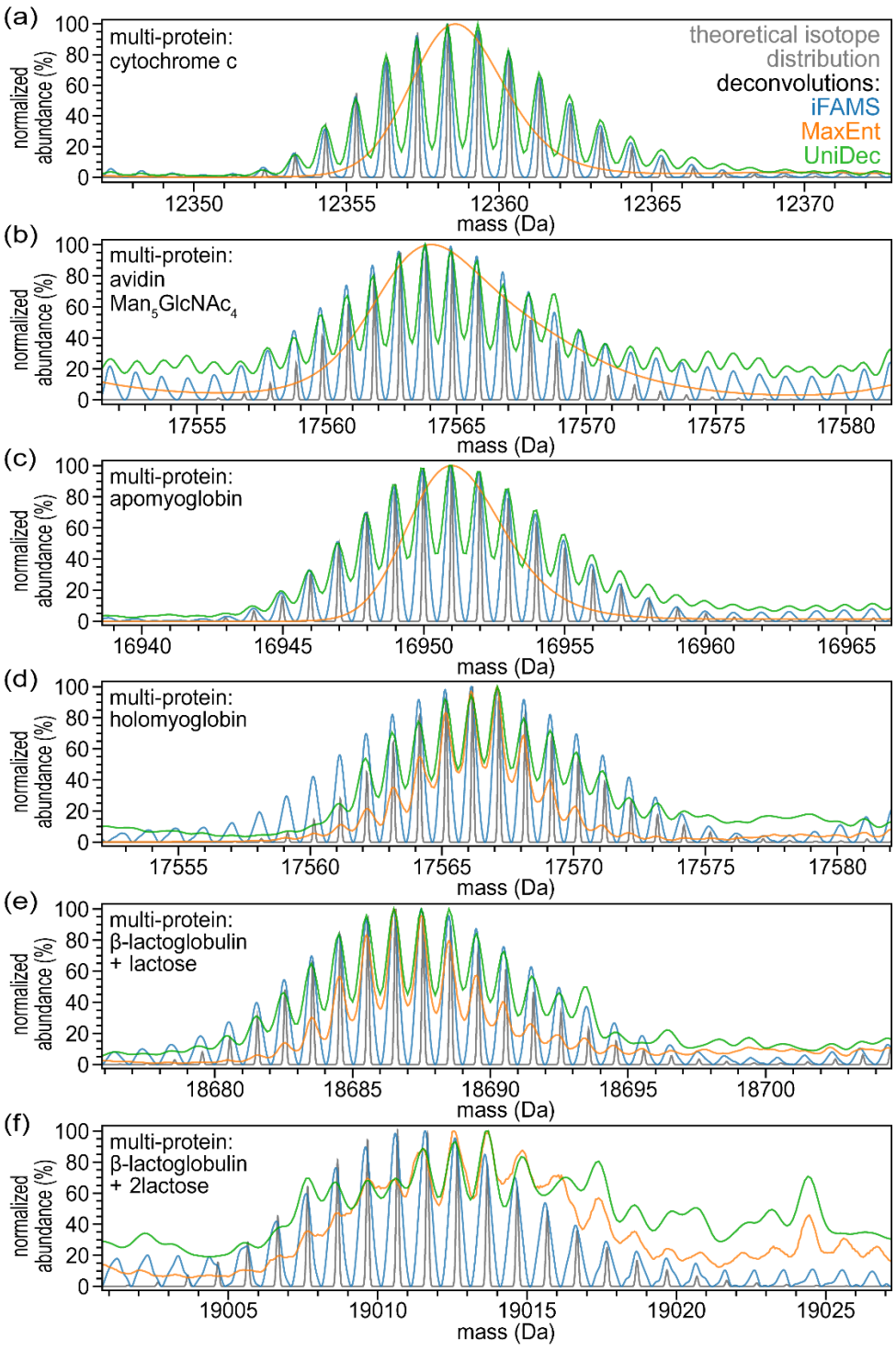


Figure S3. Theoretical isotope distributions compared to different deconvolution methods with the multi-protein sample. iFAMS, MaxEnt, and UniDec deconvolved mass spectra are shown in blue, orange, and pink, respectively with the theoretical isotope distributions shown in red with arbitrary peak thickness. (a-f) Zoom-ins on the most abundant peak for each protein or major protein-adduct state. Ubiquitin and beta-lactoglobulin without lactose are shown in Figure 5 of the main article.

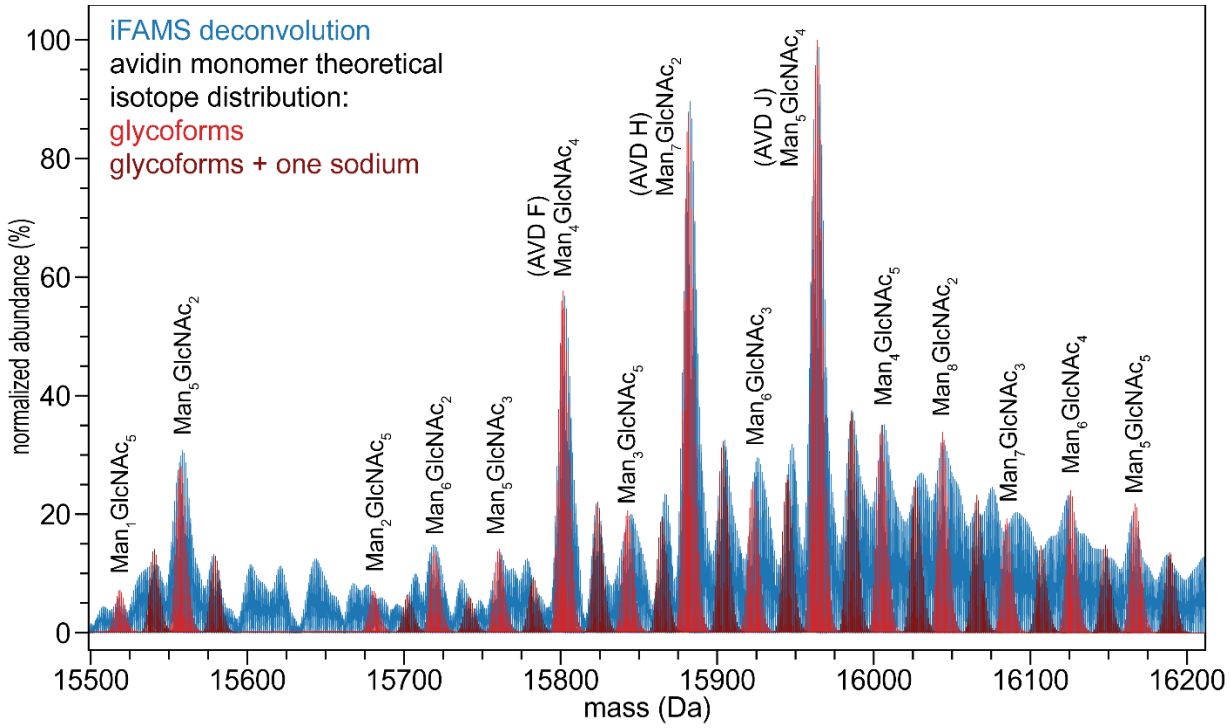


Figure S4. Avidin glycoform assignments in iFAMS deconvolved mass spectrum. The iFAMS deconvolved mass spectrum is shown in blue with the theoretical isotope distribution for major avidin glycoforms in red. Glycoforms with a single adduction of sodium are shown in dark red, and adductions of multiple sodium ions were omitted due to the high degree of overlap that would occur between neighboring glycoforms. The relative abundance of each glycoform and glycoform-adduct state was matched was matched to the abundances in the iFAMS deconvolved mass spectrum to assess alignment.

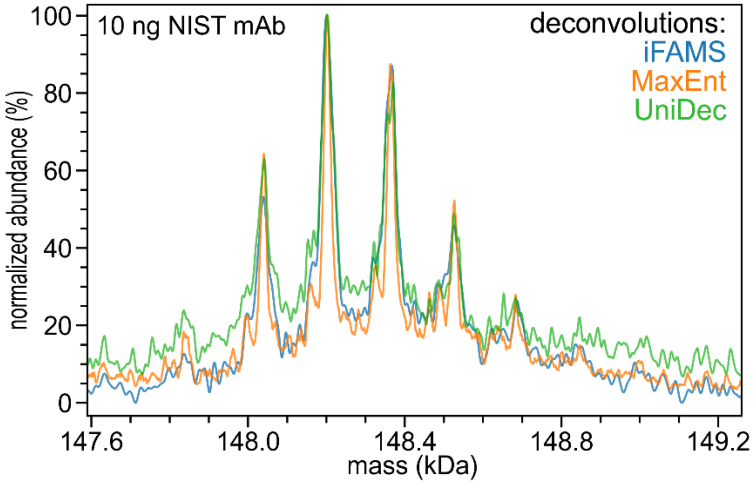


Figure S5. Deconvolution method comparison of 10 ng intact NIST mAb. Deconvolutions were performed using charge states 44-55. Default values were used for all other parameters. See Table S4 for mass accuracy assessment.

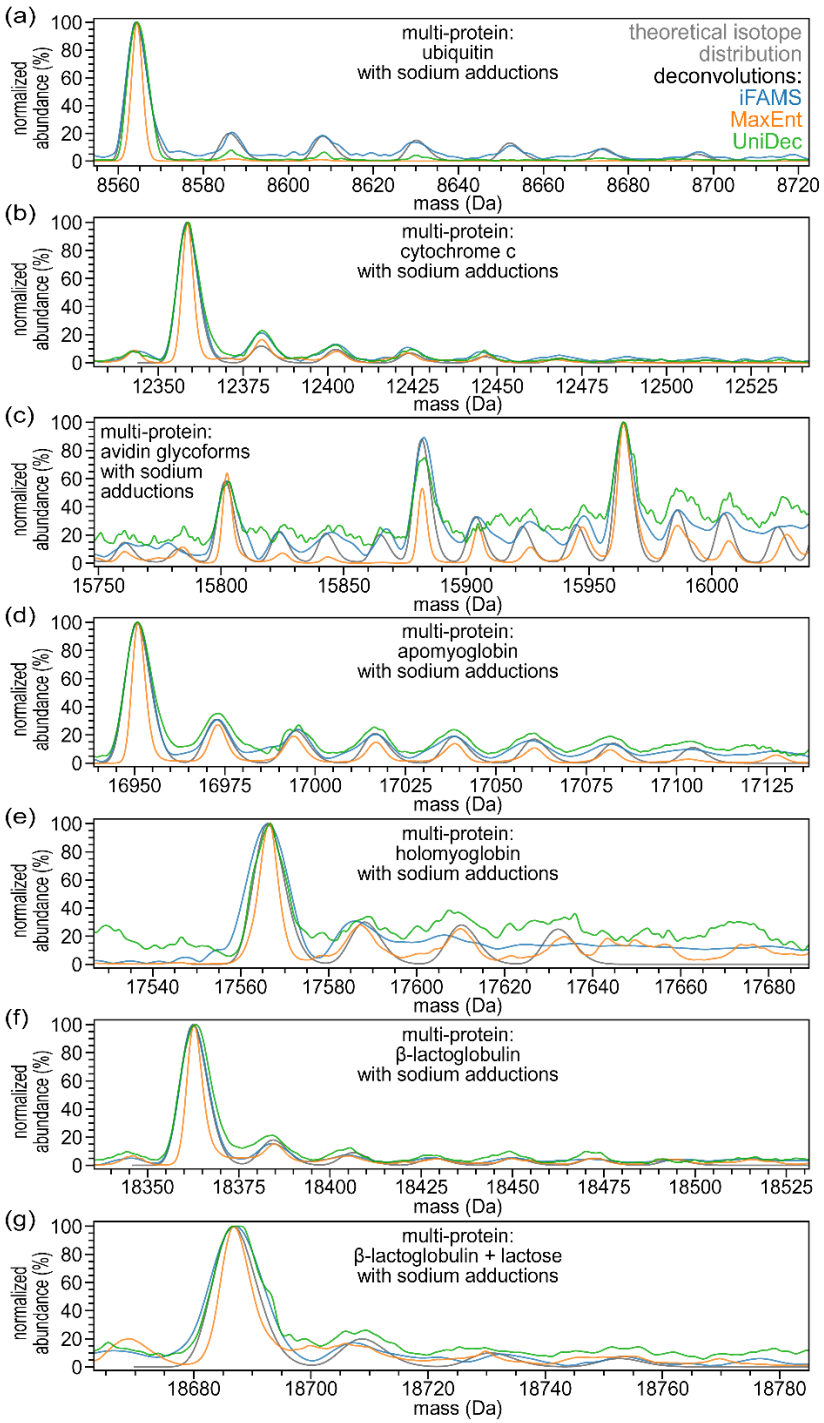


Figure S6. Deconvolution method comparison of multi-protein with multiple sodium adductions. iFAMS, MaxEnt, and UniDec deconvolved mass spectra are shown in blue, orange, and pink, respectively with the theoretical isotope distributions shown in red with multiple sodium adductions. Isotope-resolution was smoothed by Gaussian convolution for clarity and to assess peak FWHM. Relative abundances of each sodium adduction state were arbitrarily assigned roughly based on the abundances in the iFAMS deconvolved mass spectrum. Abundances were not used to assess quality of deconvolution. (a-g) Zoom-ins on the most abundant peaks and respective sodium adduction states for each protein or major protein-adduct state.

Table S6. Deconvolution method accuracy to theoretical distribution based on peak FWHM and mass accuracy measurements. Seventeen of the major peaks from the multi-protein sample were assessed for all three deconvolution methods based on peak FWHM and peak centroided mass compared to theoretical distributions calculated from amino acid sequence and assumptions outlined in Table S5. All measurements were made on isotopically smoothed spectra. The top 5 RMSD only includes the top five most abundant peaks which are highlighted in yellow.

	Relative Deconvolved Abundance (normalized to BLG)				FWHM Relative Error (%)				Mass Error (ppm)			
Feature	**iFAMS**	**Refined iFAMS**	**MaxEnt**	**UniDec**	**iFAMS**	**Refined iFAMS**	**MaxEnt**	**UniDec**	**iFAMS**	**Refined iFAMS**	**MaxEnt**	**UniDec**
BLG + 2Lac	2.6%	2.9%	4.0%	5.8%	11.4%	-11.4%	-2.3%	21.6%	12.8	-10.1	73.4	103.2
BLG + 1Lac	26.4%	22.3%	23.1%	34.8%	-11.4%	15.9%	-34.1%	8.0%	-22.6	16.8	9.6	22.1
BLG + 1Na	15.8%	15.4%	15.4%	21.3%	8.0%	-1.1%	-31.0%	-17.2%	-11.7	-13.5	12.1	-60.2
BLG	100.0%	100.0%	100.0%	100.0%	0.0%	-1.1%	-42.5%	8.0%	8.0	7.5	15.6	29.7
holoMYG + 1Na	4.1%	3.5%	2.0%	6.1%	12.0%	12.0%	-16.9%	-16.9%	-80.4	-104.4	-61.4	14.1
holoMYG	12.1%	11.3%	7.0%	17.9%	34.9%	26.5%	-37.3%	3.6%	-13.7	-11.0	3.8	6.4
apoMYG + 1Na	33.9%	30.9%	17.7%	39.0%	-12.3%	-12.3%	-32.1%	11.1%	18.2	16.4	19.7	13.9
apoMYG	107.6%	99.5%	64.1%	110.2%	3.7%	0.0%	-45.7%	3.7%	11.5	8.6	18.5	12.4
AVD J	17.0%	16.2%	11.2%	23.1%	24.1%	8.9%	-12.7%	-1.3%	16.3	7.6	21.6	11.4
AVD H	15.1%	14.4%	5.8%	17.2%	22.8%	10.1%	-44.3%	12.7%	63.3	51.7	9.0	33.9
AVD F	9.4%	9.4%	7.0%	13.3%	25.6%	23.1%	-43.6%	2.6%	67.4	41.8	42.4	46.2
CytC + 2Na	11.1%	9.2%	6.1%	9.7%	-1.4%	9.7%	-16.7%	9.7%	-0.2	1.2	11.6	-48.5
CytC + 1Na	17.2%	14.8%	12.3%	17.4%	8.3%	6.9%	-27.8%	-5.6%	-9.2	-1.0	7.1	19.3
CytC	85.9%	69.4%	73.0%	76.2%	-1.4%	-2.8%	-43.1%	4.2%	10.3	7.1	12.3	13.8
UBQ + 2Na	7.5%	7.1%	0.2%	1.6%	-25.6%	-5.9%	-21.0%	-51.8%	-44.8	-24.0	-77.4	12.9
UBQ + 1Na	8.8%	8.1%	0.2%	1.9%	-15.0%	-23.9%	-7.1%	-36.3%	96.0	73.3	88.7	27.7
UBQ	36.1%	38.9%	14.1%	25.3%	0.5%	-2.2%	-44.1%	-1.0%	2.0	-3.9	11.1	30.2
Top 5 RMSD					5.4	7.3	42.1	5.7	12.8	9.8	13.8	22.9
Overall RMSD					16.4	13.0	32.6	18.2	40.8	36.4	39.9	37.9

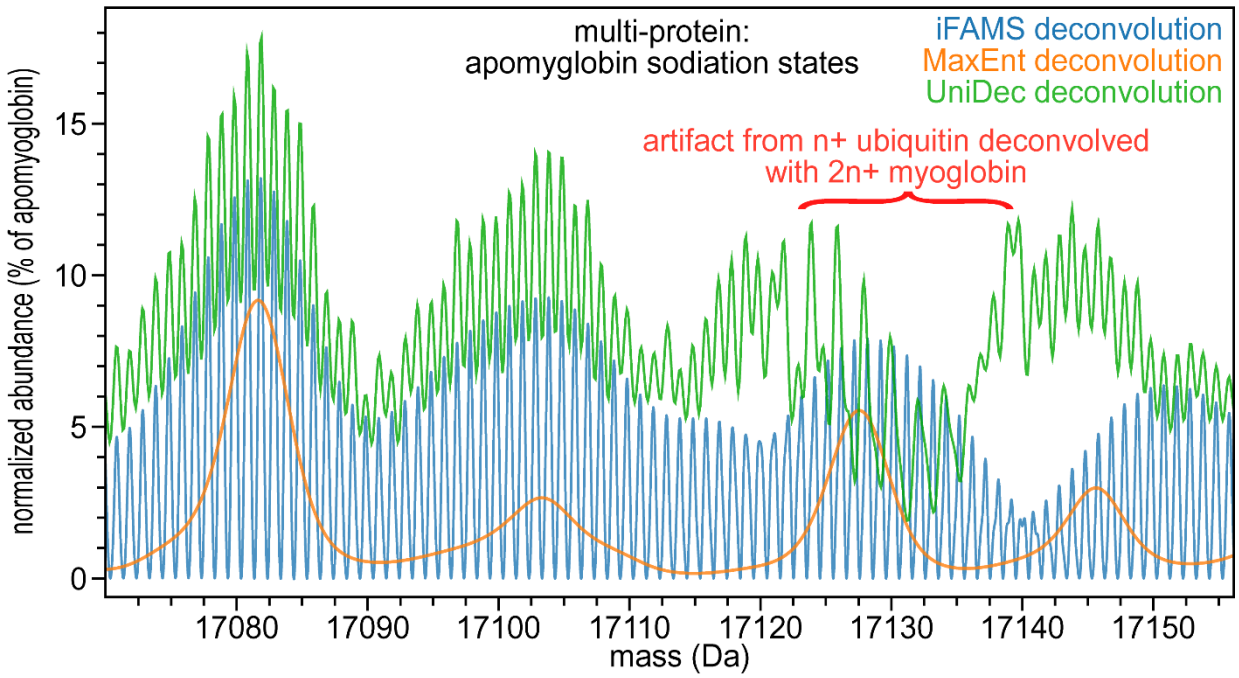


Figure S7. Persistent artifact in all three apomyoglobin deconvolved mass spectra. iFAMS, MaxEnt, and UniDec deconvolved mass spectra are shown in blue, orange, and pink, respectively. Shown here is a zoom-in on the apomyoglobin-sodium adduction states with six, seven, and eight adductions of sodium visible. At the eighth adduction state, an artifact occurs due to the mass of ubiquitin being roughly half that of apomyoglobin with eight sodium adductions and likely affects the relative abundance of that adduct state in all three spectra. This artifact can be identified in the iFAMS and UniDec deconvolved mass spectra due to the irregular peak spacing, but identification in the MaxEnt spectrum would be difficult due to low resolution even though the relative abundance is clearly larger than anticipated given the trend in adduction state abundances. The peak broadening and splitting in the iFAMS deconvolved mass spectrum are more subtle than the doubled peak spacing in the UniDec spectrum because some, but not all, of the artifact's effect on the iFAMS spectrum was filtered out in the deconvolution refinement process.

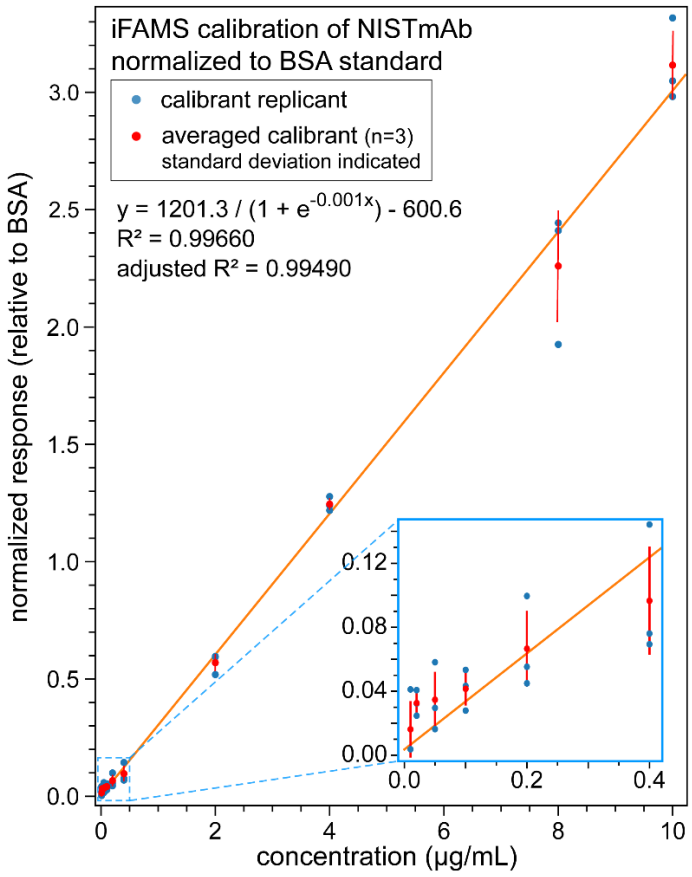


Figure S8. iFAMS calibration of intact NIST mAb using a BSA internal standard. Concentrations range from 0.01 µg/mL to 10 µg/mL, and each level was averaged across three measurements taken on separate days. Vertical red bars indicate one standard deviation.

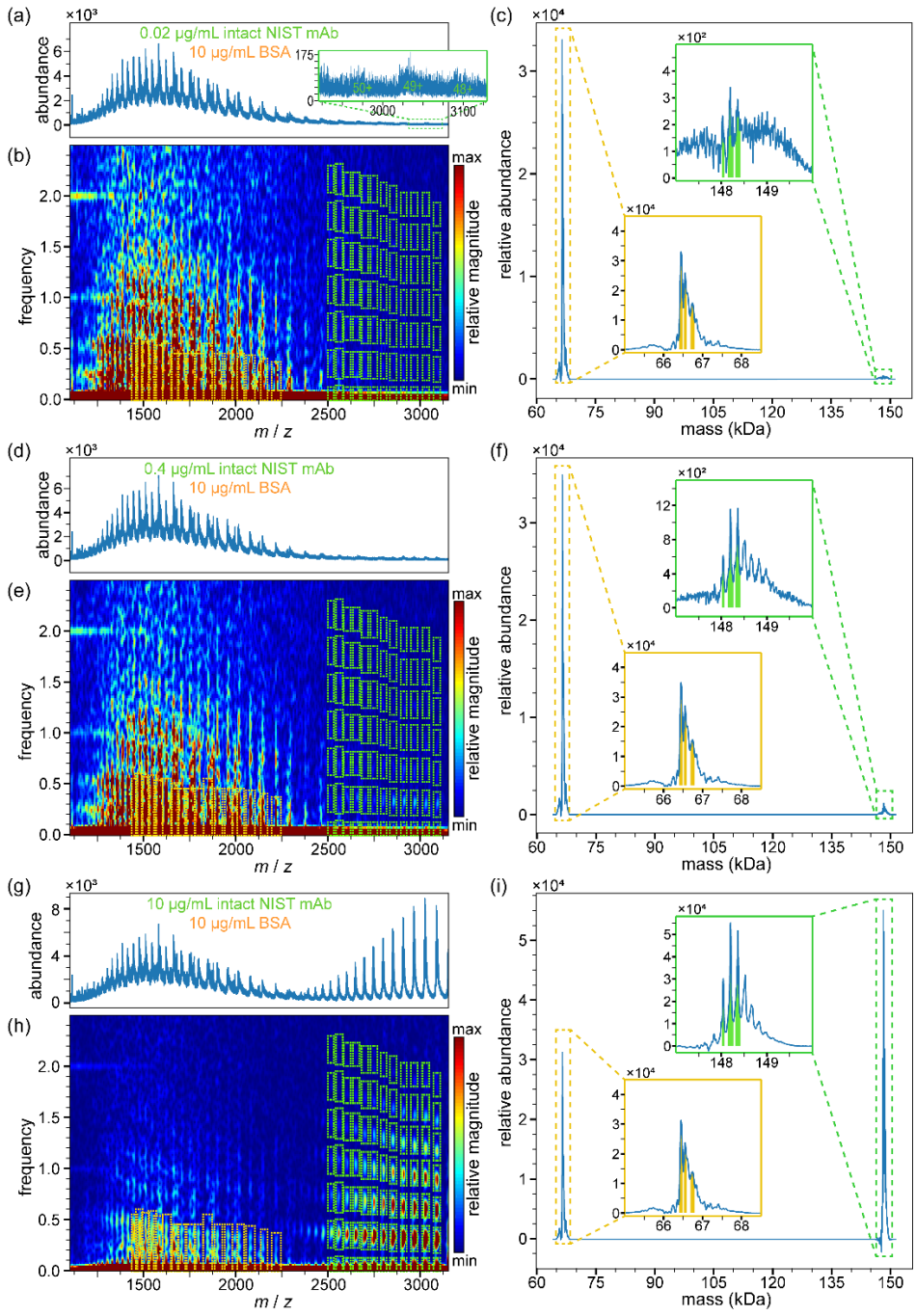


Figure S9. Representative batch deconvolution of intact NIST mAb calibration series. (a-c) Mass spectrum, Gábor spectrogram, and deconvolved mass spectrum of the second smallest NIST mAb concentration in the series (0.02 µg/mL). Zoom-in overlay in (a) shows the three most abundant charge states of NIST mAb in the mass spectrum to emphasize the low signal. It is also worth noting that charge state 49+ is partially overlapped with some BSA signal. (d-f) Mass spectrum, Gábor spectrogram, and deconvolved mass spectrum of a middle NIST mAb concentration in the series (0.4 µg/mL). (g-i) Mass spectrum, Gábor spectrogram, and deconvolved mass spectrum of the largest NIST mAb concentration in the series (10 µg/mL). Gábor spectrograms show identical selections of NIST mAb (green) and BSA (orange) applied from a batch parameter file.

iFAMS Quant signal-to-noise and LLOQ determination. To estimate the lower limit of quantitation (LLOQ), average noise and noise RMSD was calculated from regions free of protein or contaminant signal in the deconvolved mass spectra (between masses 147000-147400 Da, see Figure S9c). Signal was calculated as the sum of the heights of the top three glycoforms above the average noise, and the combined noise was estimated as the square root of three multiplied by the noise RMSD. The average signal-to-noise was calculated for both the 0.01 and 0.02 µg/mL calibrant levels as 19 and 14, respectively. Using a 10:1 signal-to-noise criterion to define LLOQ, we estimated the LLOQ to be < 0.01 µg/mL. This is further supported by the calibration results from removing the 0.01 µg/mL level which lowers the R^2 from 0.99986 to 0.99982.

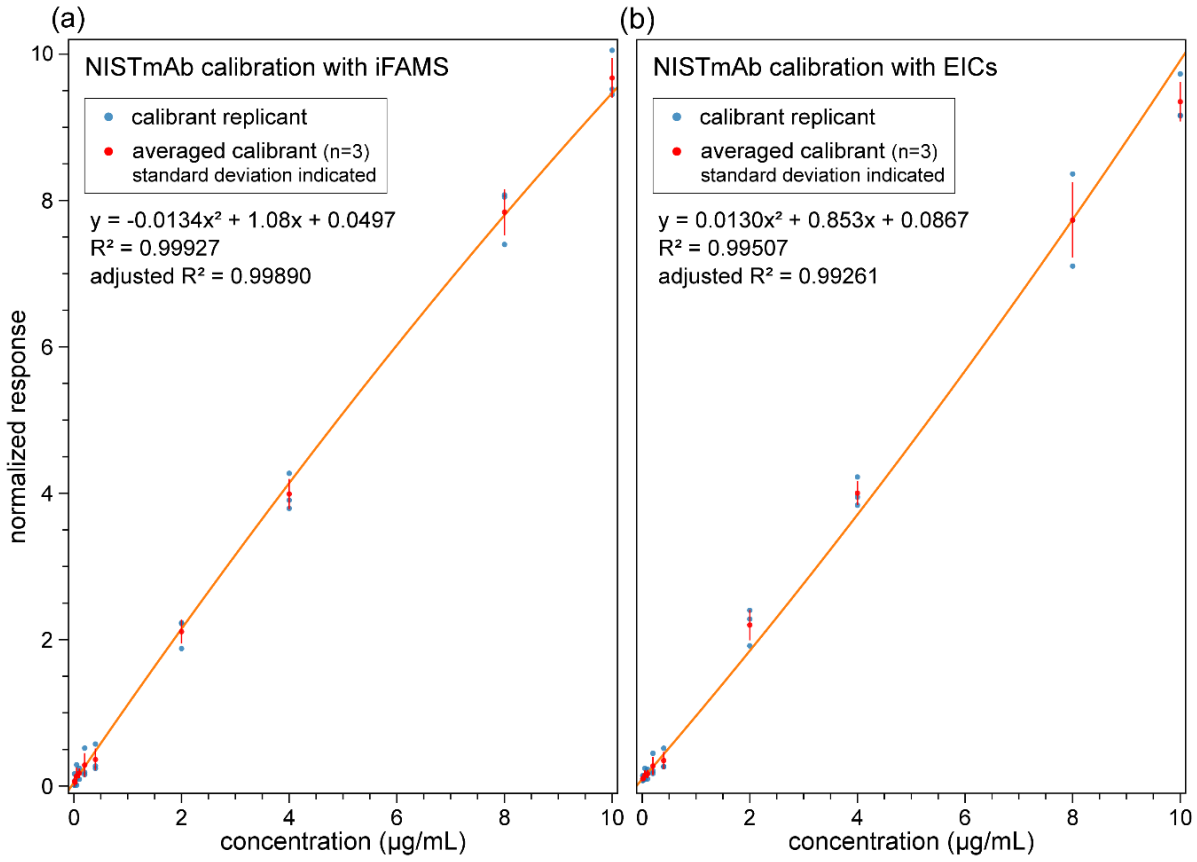


Figure S10. Weighted quadratic calibration of intact NIST mAb for quantitative comparison between iFAMS Quant and an EIC-based approach. A 1/x-weighted quadratic fit was used for both calibrations. Concentrations range from 0.01 µg/mL to 10 µg/mL, and each level was averaged across three measurements taken on separate days. Vertical red bars indicate one standard deviation. (a) Calibration results using iFAMS Quant. (b) Calibration results using the EIC-based approach.

Automation effects on iFAMS calibration. In the iFAMS NIST mAb calibration shown in Figure 8 of the main text, a few manual edits to the default processing parameters were made to improve the overall calibration. Although iFAMS requires, at a minimum, user selection of just two adjacent charge states from the GT spectrogram for automatic deconvolution, some manual intervention can improve the quantitation. For the calibration described above, one charge state (60+) was manually removed, an additional higher harmonic was included, and calibrations using different subsets of glycoforms were compared by toggling to determine the optimal set for calibration. To illustrate that the default parameters in iFAMS still produce high-quality calibration, a maximally automated iFAMS method was applied to the same set of raw data (see Figure S11). Although the calibration was slightly worse than the iFAMS method shown in Figure 8 and the standard deviations were slightly larger than the EIC method at most calibrant levels, the parameters for the resulting logistic fit were very similar for both iFAMS methods, and the adjusted R^2 (0.99978) was still larger than the EIC R^2 (0.99954). These results illustrate that the default parameters in iFAMS can be sufficient to produce calibrations close in quality to those obtained with more extensive optimization.

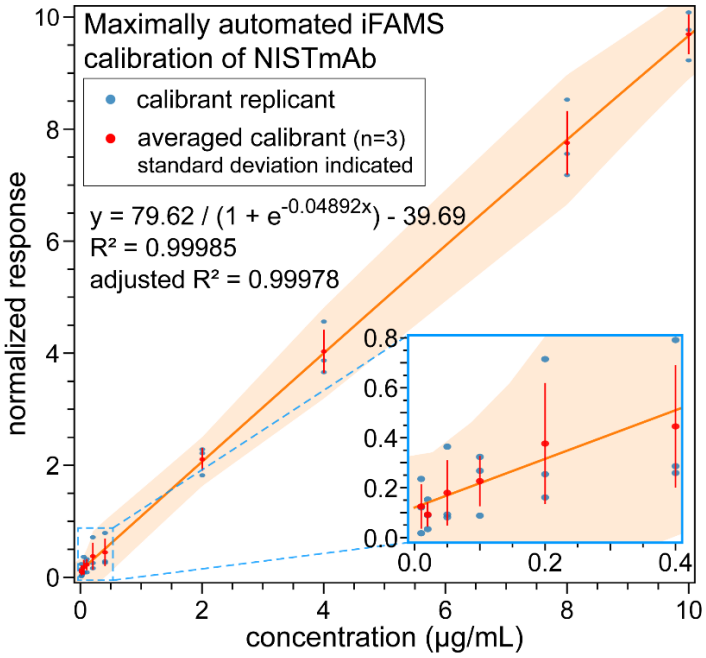


Figure S11. Maximally automated iFAMS calibration of intact NIST mAb. Concentrations range from 0.01 µg/mL to 10 µg/mL, and each level was averaged across three measurements taken on separate days. Vertical red bars indicate one standard deviation, and orange band indicates the 95% confidence interval.

References

(1) Wong, D. L. Precise Characterization of Intact Monoclonal Antibodies by the Agilent 6545XT AdvanceBio LC/Q-TOF. *Agilent Technologies application note*, publication number 5991-7813EN, **2017**.

(2) He, F.; Hendrickson, C. L.; Marshall, A. G. Unequivocal Determination of Metal Atom Oxidation State in Naked Heme Proteins: Fe(III)Myoglobin, Fe(III)Cytochrome c, Fe(III)Cytochrome B5, and Fe(III)Cytochrome B5 L47R. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (2), 120–126. https://doi.org/10.1016/S1044-0305(99)00132-4.

(3) Stevens, J. M.; Mavridou, D. A. I.; Hamer, R.; Kritsiligkou, P.; Goddard, A. D.; Ferguson, S. J. Cytochrome c Biogenesis System I. *FEBS J.* **2011**, *278* (22), 4170–4178. https://doi.org/10.1111/j.1742-4658.2011.08376.x.

(4) Márquez, I.; Pérez-Mejías, G.; Guerra-Castellano, A.; Olloqui-Sariego, J. L.; Andreu, R.; Calvente, J. J.; De la Rosa, M. A.; Díaz-Moreno, I. Structural and Functional Insights into Lysine Acetylation of Cytochrome c Using Mimetic Point Mutants. *FEBS Open Bio* **2021**, *11* (12), 3304–3323. https://doi.org/10.1002/2211-5463.13284.

(5) Liu, F. C.; Cropley, T. C.; Ridgeway, M. E.; Park, M. A.; Bleiholder, C. Structural Analysis of the Glycoprotein Complex Avidin by Tandem-Trapped Ion Mobility Spectrometry–Mass Spectrometry (Tandem-TIMS/MS). *Anal. Chem.* **2020**, *9*^{2} (6), 4459–4467. https://doi.org/10.1021/acs.analchem.9b05481.

(6) Davis, S. R.; Ward, H. E.; Kelly, V.; Palmer, D.; Ankersmit-Udy, A. E.; Lopdell, T. J.; Berry, S. D.; Littlejohn, M. D.; Tiplady, K.; Adams, L. F.; Carnie, K.; Burrett, A.; Thomas, N.; Snell, R. G.; Spelman, R. J.; Lehnert, K. Screening for Phenotypic Outliers Identifies an Unusually Low Concentration of a β-Lactoglobulin B Protein Isoform in Bovine Milk Caused by a Synonymous SNP. *Genet. Sel. Evol.* **2022**, *54* (1), 22. https://doi.org/10.1186/s12711-022-00711-z.

(7) Creamer, L. K.; Bienvenue, A.; Nilsson, H.; Paulsson, M.; van Wanroij, M.; Lowe, E. K.; Anema, S. G.; Boland, M. J.; Jiménez-Flores, R. Heat-Induced Redistribution of Disulfide Bonds in Milk Proteins. 1. Bovine β-Lactoglobulin. *J. Agric. Food Chem.* **2004**, *52* (25), 7660–7668. https://doi.org/10.1021/jf049388y.