Deconvolving Native and Intact Protein Mass Spectra with UniDec

Authors: Marius M. Kostelic,[1] Michael T. Marty[1, 2, *]

[1]Department of Chemistry and Biochemistry and [2]Bio5 Institute, University of Arizona, Tucson, AZ 85721

*mtmarty@arizona.edu

Summary: Intact protein, top-down, and native mass spectrometry (MS) generally require the deconvolution of electrospray ionization (ESI) mass spectra to assign the mass of components from their charge state distribution. For small, well-resolved proteins, the charge can usually be assigned based on the isotope distribution. However, it can be challenging to determine charge states with larger proteins that lack isotopic resolution, in complex mass spectra with overlapping charge states, and in native spectra that show adduction. To overcome these challenges, UniDec uses Bayesian deconvolution to assign charge states and to create a zero-charge mass distribution. UniDec is fast, user-friendly, and includes a range of advanced tools to assist in intact protein, top-down, and native MS data analysis. This chapter provides a step-by-step protocol, an in-depth explanation of the UniDec algorithm, and highlights the parameters that affect the deconvolution. It also covers advanced data analysis tools, such as macromolecular mass defect analysis and tools for assigning potential PTMs and bound ligands. Overall, the chapter provides users with a deeper understanding of UniDec, which will enhance the quality of deconvolutions and allow for more intricate MS experiments.

Keywords: Mass spectrometry data analysis, native mass spectrometry, top-down proteomics, Bayesian deconvolution, electrospray ionization, nanodiscs, intact protein analysis

1 Introduction

One of the breakthroughs of electrospray ionization (ESI) is that it enabled high-resolution analysis of intact proteins. Intact protein analysis has since become useful in the routine characterization of biopharmaceuticals, which enables rapid profiling of the mass distribution and covalent protein modifications.[1] Building on intact protein analysis, top-down proteomics uses gas-phase fragmentation to gain further insights into the protein sequence and sites of covalent modifications, which provides an indispensable tool for characterizing proteoforms.[2-6] A related tool, native MS uses nondenaturing ionization conditions to observe not only covalent modifications but also noncovalent interactions such as bound ligands or complex formation.[7-11] The combination of native MS with top-down provides a powerful tool for the simultaneous characterization of covalent and noncovalent protein interactions.[12-14] For all of these techniques, the initial spectrum of the intact protein has multiple charge states that must be assigned to determine the mass. Charge state assignment can either be done by comparing the spacing between peaks with different charge states or by comparing the spacing of the peaks within an isotope series when high-resolution analysis is available.

However, as intact protein MS analysis moves towards more complex experiments and larger proteins, data analysis becomes challenging due to adduction, loss of isotopic resolution, and overlapping charge state peaks from multiple charge state distributions. These challenges can make it difficult to assign charge states for different *m/z* peaks. Furthermore, there is an increasing demand for automated and higher throughput data analysis approaches that are capable of handling these more complex mass spectra. To overcome these challenges, multiple deconvolution software programs have been created to help assign charge states and deconvolve intact MS data.[15-25] A recent review has covered a wide variety of deconvolution software.[26] UniDec (Universal Deconvolution) software uses Bayesian deconvolution to quickly assign the charge states of *m/z* peaks and thus determine the mass.[27,28] UniDec can deconvolve a wide variety of complex spectra such as native ion-mobility (IM), intact protein spectra, native membrane proteins, nanodiscs, antibodies, and DNA-protein complexes.[29,30,10,31,32,16,33]

Although the UniDec algorithm is open source and has been published [27,28,30,34,35], there is no written step-by-step protocol for using the UniDec software. Moreover, because the default parameters work well for many applications, many users may lack a clear understanding of the algorithm and the advanced features. UniDec is not magic. It will create the most likely deconvolved mass spectrum based on the input parameters and can produce a distorted deconvolution if used incorrectly. Conversely, a deeper understanding of the available parameters provides powerful tools for deconvolving complex spectra. This chapter provides a step-by-step protocol for using UniDec, including comprehensive descriptions of the most common deconvolution parameters and data analysis tools. By understanding these crucial parameters, users will be able to deconvolve a wide variety of complex native and intact protein MS samples.

2 How Does UniDec Work?

Before outlining a step-by-step protocol, we will first provide an overview of the underlying UniDec algorithm, focusing on a typical 1-D mass spectrum (Fig. 1). The goal of UniDec deconvolution is to separate the charge and mass components from an *m/*z spectrum. To accomplish this, UniDec goes through multiple iterations of three steps to create a 2-D charge vs *m/z* matrix that contains the most likely charge state distribution for each *m/z* data point in the spectrum. The 2-D charge vs *m/z* plot is shown directly below the processed *m/z* spectrum after deconvolution in the software.

Initially, UniDec assumes that all charge states are equally likely for any *m/z* point (Fig. 1b). UniDec can determine the most probable charge states in three ways: by using the charge state distribution, the mass distribution, or the isotope distribution. These options are not mutually exclusive and can be used as complimentary filters. Without at least one of these filters, UniDec will not be able to assign a charge state. Here, we will focus on the most common case for native and intact protein spectra, which assumes a smooth charge state distribution common in ESI of proteins.

The first step is to filter the charge state distribution (Fig. 1c). This means that UniDec will adjust the intensity of potential charge states in the 2-D charge vs. *m/z* matrix based on their neighboring charge states. For a point at *z* and *m/z* in the matrix, UniDec will look at the intensities at neighboring pairs at *z+1* with *m/(z+1)* and *z-1* with *m/(z-1)*. If the neighboring charge states have strong intensities, UniDec will increase the intensity of the point at *z* and *m/z*. If not, it will decrease the intensity assigned to that point in the matrix.

Second, UniDec will sum the intensities of all charge assignments in the 2-D charge vs *m/z* matrix into a simulated 1-D *m/z* spectrum that has the same dimensions as the data (Fig. 1d). The simulated spectrum can be convolved with a known peak shape—based on the full width at half maximum (FWHM) and peak shape parameters—to define what the minimum peak size should be. Setting a minimum peak width can help to avoid deconvolution artifacts by telling the algorithm that points within the peak width should have the same charge state.

Third, the simulated mass spectrum intensities are compared to the intensities in the actual mass spectrum. If an *m/z* peak in the simulated spectrum has a higher intensity than its actual counterpart in the original data, UniDec will lower the intensity of all charge states at that *m/z*. If the simulation is lower, UniDec will increase the intensity of all potential charge states at that *m/z* (Fig. 1e). After this, the loop restarts by adjusting the charge state distributions and continues.

The loop ends when the intensities in the simulated mass spectrum match those in the original data or have converged. Finally, the 2-D charge vs. *m/z* matrix is converted into a 2-D charge vs. mass matrix (Fig. 1h). Summing the charge vs. mass matrix across all charge states produces the zero-charge mass spectrum (Fig. 1i). Armed with a basic understanding of the algorithm, we will now outline how to use the software and describe specific parameters.

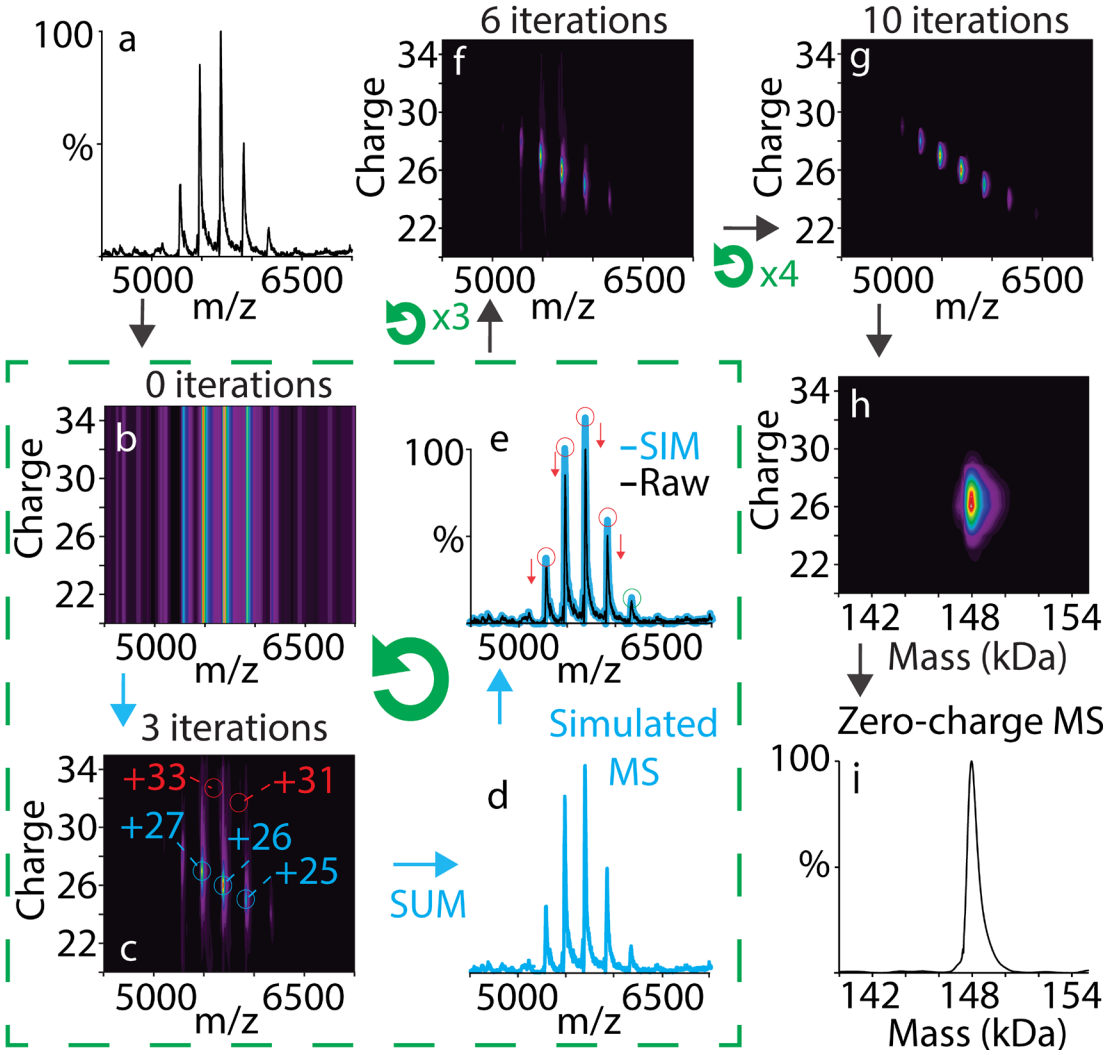

Fig. 1 Schematic of the UniDec algorithm for a mass spectrum of alcohol dehydrogenase (ADH). **a**) Raw mass spectrum of ADH where the charge states and intact mass are unknown. **b**) Initial charge vs *m/z* matrix where all charge states are equally likely. **c**) The charge vs *m/z* plot after 3 iterations, which also conceptually shows the first step of the algorithm filtering the charge state distribution. The +26 charge state is more likely than the +32 charge state because it has support from the neighboring +27 and +25 charge states. **d**) Step 2 of the algorithm, where the charge vs *m/z* plot is summed into a 1-D simulated mass spectrum. **e**) Step 3 of the algorithm, where the simulated mass spectrum is compared to the raw data. Steps **c-e** show one iteration of the algorithm and are symbolized by the green loop. The charge vs *m/z* matrix is then shown at **f)** 6 and **g**) 10 iterations before transformation into the **h**) charge vs mass (kDa) matrix, which is summed to yield the **i**) zero-charge mass spectrum after completing all the iterations.

3 Downloading and Installing UniDec

The latest version of UniDec can be downloaded from GitHub at: (https://github.com/michaelmarty/UniDec/releases). This protocol will cover version 4.3.0 but applies to other versions as well. Unzip and extract the folder into a convenient location such as the desktop. No further installation is necessary because UniDec runs as a standalone, portable program. Then, run the GUI_UniDec.exe file. This compiled version is only available for Windows, but Mac and Linux versions can be built from the source code. UniDec can also be run from the Python source code, which allows command line usage and scripting. Further information on installing the source code can be found here: https://github.com/michaelmarty/UniDec/wiki/.

4 UniDec Data Processing and Deconvolution

4.1 Opening Mass Spectra

The first step in analyzing mass spectra with UniDec is opening the data file. UniDec can read a variety of mass spectra file formats. Where applicable, UniDec will average all spectra in the file into a single mass spectrum. Chromatography data can be parsed externally (see section 4.1.3) or opened with MetaUniDec [28] or UniChrom to parse mass spectra from specific time sections.

4.1.1 Text files, mzML, or Thermo RAW files can be opened in UniDec by going to *File > Open File (text, mzML, or Thermo RAW)*. These files can also be opened by dragging and dropping them into the main window. Text files should be in the form of where each row is a data point with the first column as the *m/z* value and the second column as the intensity. For Bruker data, the spectrum needs to be exported as a "simple x-y ASCII", which can then be opened as a text file in UniDec.

4.1.2 Waters and Agilent data is stored in directories rather than single files and can be opened using *File > Open Waters or Agilent File*. These files can also be opened with drag and drop.

4.1.3 For quick analysis, data can be copied to the clipboard from Thermo QualBrowser or Waters MassLynx. For the Spectrum window in MassLynx, *Edit > Copy Spectrum List* will copy the selected spectrum to the clipboard. For Thermo QualBrowser, right-clicking the mass spectrum and selecting *Export > Clipboard* (*Exact Mass*) will copy the spectrum to the clipboard. To import data from the clipboard into UniDec, go to *File > Get Spectrum from Clipboard* or use the Ctrl+G shortcut. Opening data from the clipboard is a quick way to deconvolve a selected mass spectrum with UniDec, especially during data collection, but it will generally be saved as a temporary file. To save the data in a permanent location, data from the clipboard can be pasted into a text file, saved, and opened as described in section 4.1.1. We recommend saving permanent files so that data analysis can be traced.

4.2 UniDec Presets

After opening a spectrum, a preset can be selected to better tailor the data processing and deconvolution settings. There are multiple fixed presets in *File > Presets*. *High Resolution Native* is a good starting preset for the native protein MS data. For denatured intact protein analysis, the maximum charge state may need to be increased, but this preset should otherwise be a good starting point. Settings are saved as *<filename>_conf.dat* files in the *<filename>_unidecfiles* folder. Past deconvolution settings can be opened using *File > Load External Config File* and selecting any *_conf.dat* file. Adding a config file to the Presents folder in the UniDec directory will make this file available as a custom preset in the Presets menu when the program is reopened. UniDec will

automatically reload settings from previously opened files as long as the *<filename>_unidecfiles* folder and *<filename>_conf.dat* files are still in the same directory as the top file.

4.3 Data Processing

Before deconvolution, the data needs to be processed. In the simplest case, this will simply copy over the raw data and clean up any issues like negative intensities or duplicate *m/z* values. However, additional data processing can be applied to improve and speed up the deconvolution. Data processing parameters can be found in the blue tabs on the control panel to the right of the main UniDec window (Fig. 2a-b).

4.3.1 In the data processing tab, the *m/z* range can be set by entering the minimum and maximum values in the corresponding boxes. Another option to set an *m/z* range is to click and hold the left-click and drag across the *m/z* area of interest in the plot to zoom in then right-click on the plot to automatically save the new minimum and maximum from the selected range. Clicking the "*Full"* button, next to the *m/z* range, will reset the *m/z* range to the lowest and highest *m/z* value in the spectrum.

4.3.2 Below the *m/z* range, clicking the "*Use Background Subtraction"* checkbox will subtract a curved baseline from the spectrum based on smoothing the local minima throughout the spectrum. The type (flat, linear, or curved) and the degree of background subtraction can be further adjusted in the Additional Data Processing Parameters tab (see **Note 1**).

4.3.3 There are multiple additional parameters within the *Additional Data Processing Parameters* tab that are useful for improving the deconvolution. For an in-depth description of each parameter in the *Additional Data Processing Parameters* tab, see **Note 2**.

4.3.4 Clicking "*Process Data*" will apply the parameters within the data processing tabs and write the processed data into the unidecfiles folder (*<filename>_unidecfiles*) in the same directory as the file or embedded in the raw directory for Waters and Agilent data. This folder initially contains the *<filename>_conf.dat, <filename>_rawdata.txt*, and *<filename>_input.dat* files. The *_input.dat* file contains the processed data with *m/z* values in the first column and the corresponding intensities in the second column.

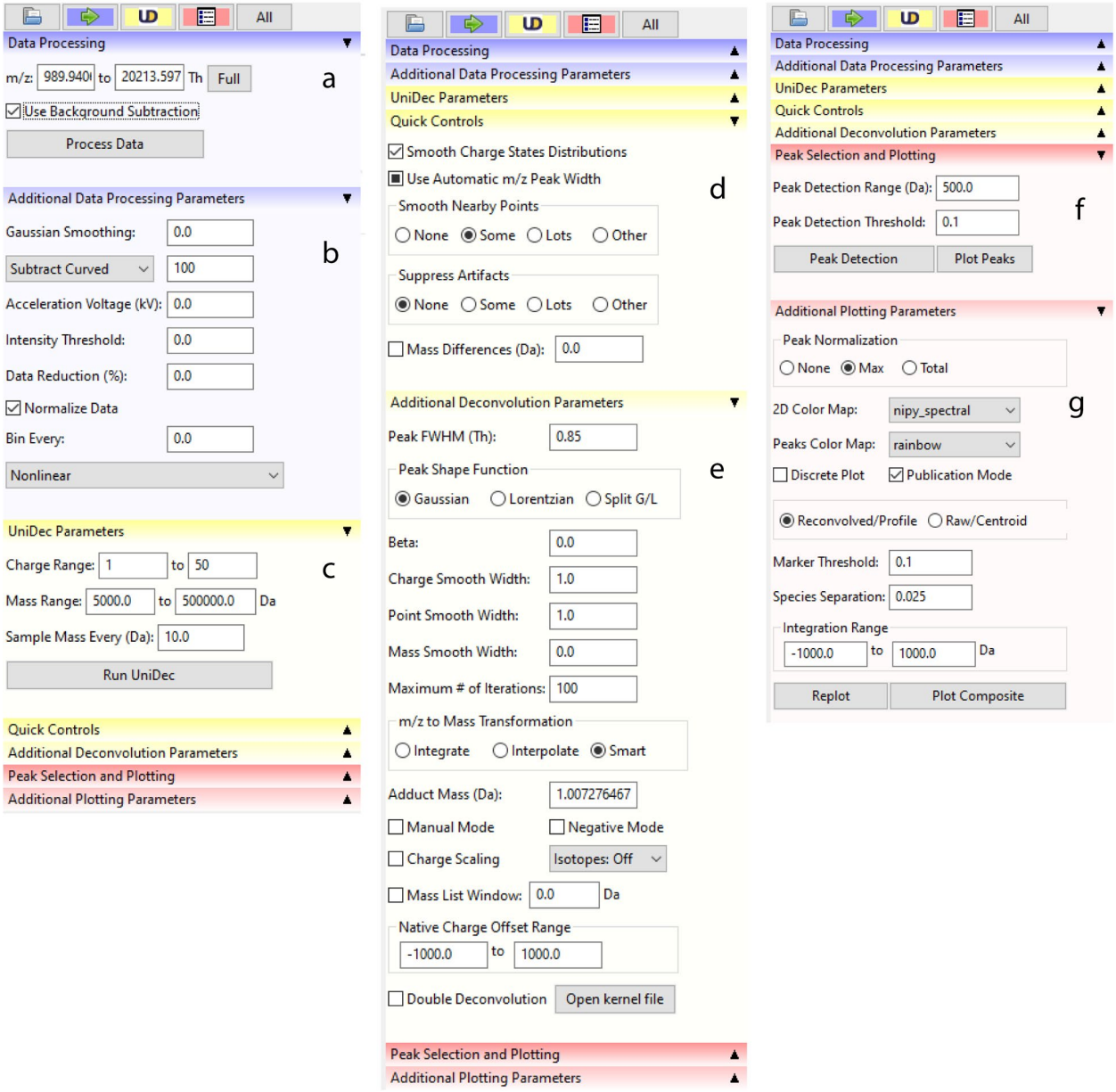

Fig. 2 The UniDec control panel: **a)** Data processing parameters (*blue*), **b**) Advanced data processing parameters (*blue*), **c**) UniDec parameters tab (*yellow*), **d**) Quick control parameters (*yellow*), **e**) Advanced Deconvolution parameters (*yellow*), **f**) Peak selection and plotting parameters (*red*), and **g**) Additional plotting parameters (*red*).

4.4 Deconvolution Parameters & Quick Controls

After the data has been processed, the next step is to deconvolve the data. A few simple deconvolution parameters set the range of potential masses and charge states as well as how precisely to sample the deconvolved mass data. The quick controls then allow the deconvolution to be subtly guided by assumptions about the data. (Fig. 2c-d)

4.4.1 In the *UniDec Parameters* tab, the *Charge Range* and *Mass Range* set the allowed charges and mass values for the deconvolution. The charge range will define the rows in the 2-D

charge vs. *m/z* matrix, and the mass range will forbid points in the matrix outside the mass range. Setting a *Charge Range* or *Mass Range* that is too narrow will cause artifacts in the deconvolution. Thus, it is better to start with a broader range then narrow the charge and mass range to the area of interest. Narrowing the ranges is a useful tool to eliminate artifacts from harmonics at double or half the charge and to speed up the algorithm.

4.4.2 *Sample Mass Every* sets the sample rate for the deconvolved mass spectrum. It is important to set a sample rate with adequate mass resolution. If the sample rate is set to 10 Da, then each mass value will fall on an even 10 Da. Thus, it may be more beneficial to set the sample rate to 1 or 0.1 Da for a better mass accuracy at the expense of speed.

4.4.3 Within the *Quick Controls* tab, the first option is to tell UniDec to expect a Smooth Charge State Distribution by clicking the "*Smooth Charge State Distributions"* check box, which sets the *Charge Smooth Width* in the *Additional Deconvolution Parameters* tab to the default of 1 (see **Note 3**). As described above in the first step of the algorithm (section 2), charge smoothing tells UniDec that the probability of a potential charge state, *z*, for an *m/z* data point in the charge vs. *m/z* matrix is correlated to the intensity of neighboring charge states at *z-1* and *z+1* at data points *m/(z-1)* and *m/(z+1)* respectively.

4.4.4 Checking the "*Use Automatic m/z Peak Width*" box will tell UniDec to automatically determine the peak width and peak shape from the most abundant peak in the processed mass spectrum and apply this for the deconvolution. The peak width and peak shape affect the second step of the algorithm. In this step, the data is convolved with a specified peak shape, which means that each mass peak will have the same minimum width and shape. This convolution can be ignored by unchecking the box or by setting the peak width to 0. The peak width and peak shape parameters can be set manually in the *Additional Deconvolution Parameters* tab. Manually increasing the peak width is useful when the deconvolution is overfitting the data and presenting artifacts. Decreasing the peak width is useful when the deconvolution is underfitting the data.

4.4.5 Setting the *Smooth Nearby Points* option will tell UniDec that nearby data points should have the same charge state.[30] This will smooth nearby data points into the same charge state, to reduce artifacts in the deconvolved mass spectrum, which is usually caused by missassigned charge states. The 4 options in the *Smooth Nearby Points* affect the *Point Smooth Width* parameter in the *Additional Deconvolution Parameters* tab. The "*some*" option works well for most mass spectra, but some mass spectra may require a higher setting (see **Note 4**).

4.4.6 The *Suppress Artifacts* feature tells UniDec that each *m/z* data point should have a single charge state, which helps eliminate artifacts from missassigned charge states. It uses a SoftMax function to push more intense charge states higher for each *m/z*, and the degree to which it pushes towards a single charge state is specified by the *Beta* parameter in the *Additional Deconvolution Parameters* tab. The *Suppress Artifacts* feature is useful when each peak should have a single charge state assignment but harmonics or satellites artifacts are present (see **Note 5**).[35]

4.4.7 *Mass Differences* helps assign charge states based on an expected mass distribution. Mass distribution smoothing is applied in the first step of the algorithm alongside charge state smoothing (see Section 2). It tells UniDec that each *m/z* peak should have neighboring peaks that are *(m–nM)/z* and (*m+nM)/z*, where *M* is a repeating mass in the spectrum that

can be set in the *Mass Differences* box. The repeating mass could be a repeating monomer unit in a polymer/oligomer or a repeating lipid mass in a nanodisc.[36] The *n* value is set by the *Mass Smooth Width* parameter in the *Additional Deconvolution Parameters* tab, which tells UniDec how many masses to consider on either side of the initial *m/z* value. Clicking the *Mass Differences* checkbox will set the *Mass Smooth Width* to 1.

4.5 Advanced Deconvolution Parameters

In addition to the basic deconvolution parameters and the quick controls, there are a number of advanced parameters that are useful for different applications (Fig. 2e). Although the default values are sufficient for most cases, we will highlight a few parameters that users should consider. Other parameters, such as *Charge Scaling, Mass List Window*, and *Native Charge Offset Range*, are described in **Note 6**.

4.5.1 Clicking the *"Negative Mode"* checkbox will switch the *Adduct Mass* to the negative mass of a proton (see **Note 6**). Unchecking the box will return it to the default for positive mode, the positive mass of a proton. These account for the charge carriers typically present in ESI and ensure that UniDec reports the neutral mass of the protein without charge carriers.

4.5.2 *Manual Mode* allows the user to manually specify the charge state for specific regions of the spectrum and is useful when charge state assignments are difficult. For example, *Manual Mode* is essential for MS/MS where a single charge state is isolated because the lack of a charge state distribution makes this isolated peak impossible to identify without isotopic resolution. Clicking the "*Manual Mode*" checkbox tells UniDec to use manual assignments and will automatically open the *Tools > Manual Assignment* window if no assignments are already set. The *Manual Assignment* window allows the user to select specific regions of the spectrum—defined by a central *m/z* value and an *m/z* range around this value—and set the charge state of this region.

5 Analyzing Deconvolution Data

5.1 Peak Selection and Plotting

After the deconvolution, UniDec has a wide variety of tools to analyze and visualize the data. We will first discuss tools for peak selection and plotting (Fig. 2f-g).

5.1.1 The parameters in the *Peak Selection and Plotting* tab allow the user to automatically detect peaks within the deconvolved mass spectrum. The *Peak Detection Range* tells UniDec to select peaks in the deconvolved mass spectrum that are local maxima within a certain window. Thus, to detect two peaks that are 500 Da apart, a *Peak Detection Range* of less than 500 would be required.

5.1.2 The *Peak Detection Threshold* tells UniDec to only pick peaks that are above a set intensity threshold. The threshold is the decimal value entered in the box multiplied by the intensity of the most abundant peak. For example, a value of 0.1 will set a peak threshold at 10% of the maximum intensity.

5.1.3 Clicking "*Peak Detection"* will detect peaks in the deconvolved mass spectrum, label them with symbols on both the deconvolved mass and the processed data below, and display the *Mass*, *Intensity*, *DScore*, and *Name* for each peak directly to the right of the deconvolved mass spectrum. The *DScore* is part of the UniDec scoring function and a user would usually want a DScore above 60. For a further explanation of the *DScore* function, see **Note 7**.

5.1.4 Clicking "*Plot Peaks"* next to the *Peak Detection* button will plot each mass species charge state distribution independently. Note that each peak will have the same simulated minimum peak width and shape and may not reflect the peak shapes real data. *Species Separation*, in the *Additional Plotting Parameters*, can be used to increase the distance between the mass species within the same plot (see **Note 8**).

5.1.5 UniDec can integrate the selected peaks in the deconvolved mass spectrum by clicking *Analysis > Integrate Peaks*. The range of the integration will be highlighted for each selected peak in the deconvolved mass spectrum and is determined by the *Integration Range* in the *Additional Plotting Parameters* (see **Note 8**). An asymmetric integration range can be set by adjusting these parameters. The area calculated from the integration for each peak is shown in the peak list in the new *Area* column.

5.1.6 Right-clicking a peak in the list will provide a drop-down menu with options on how to isolate, display, and export the data.

5.1.7 To label the charge states of a specific mass species, right-click the species in the peak list and select "*Label Charge States*". This will plot all possible charge states for that mass, which helps visualize and confirm the charge state assignment for a deconvolved mass peak.

5.1.8 A useful method to display the mass differences between mass peaks is to right-click any mass species and select "*Display Differences*". With *Display Differences* selected, the mass peaks will be labeled as their mass difference from the selected peak. This plotting is useful for quickly identifying post-translational modifications (PTMs), oligomers, and bound ligands in the mass spectrum.

5.2 Mass and Oligomer Tools for Peak Assignment

The *Mass and Oligomer Tools* is useful for quickly determining possible PTMs or bound ligands based on the masses of each peak. The overall workflow is to first define potential mass species in the *Oligomer Maker* and then match them to measured peaks (Fig. 3).

5.2.1 After selecting peaks in the main UniDec window, go to *Tools > Mass and Oligomer Tools* or use the shortcut Ctrl+T. To the left is the *Mass List*, which is discussed in **Note 6**. The *Oligomer Maker* is in the center of the window (Fig. 3b).

5.2.2 UniDec defines potential species as oligomers with 5 parameters: a *Base Offset* (Da), *Monomer Mass* (Da), *Min # of Oligomers*, *Max # of Oligomers*, and *Name*. Potential masses will be calculated as *Base Offset* + *Monomer Mass* * *n,* where *n* is an integer from *Min# of Oligomers* to *Max# of Oligomers*. Thus, a potential monomeric protein can be input using a base mass of 0, the protein mass for the *Monomer Mass*, a minimum of 0, and a maximum of 1. Setting the minimum to 1 will force all potential species to have that protein mass as a fixed component.

5.2.3 Within the *Oligomer Maker*, potential species can be manually added by selecting "*Add Oligomer Species"*. Once a new oligomer species has been added, each parameter can be edited. UniDec will automatically export the oligomers to a file in the unidecfiles folder called <*filename*>*_ofile.dat*. These files can be imported for other analysis and are simply text files with columns and rows that match the list in the window. Masses can also be calculated from the sequence of a protein by right clicking and opening the Biopolymer Calculator window.

5.2.4 The *Common Masses List* presents a variety of common PTMs, lipids, ions, and detergents that may be added or removed from the analyte. To add an item from the *Common Mass List* into the *Oligomer Maker*, right-click an item and select "*Add to Oligomer Builder*". The *Common Mass List* is automatically imported from a CSV file and can be edited to incorporate additional species.

5.2.5 After specifying potential oligomers, the *Mass and Oligomer Tools* can match measured peaks to potential oligomer masses by selecting either "*Match to Isolated Oligomers*" or "*Match to Mixed Oligomers*" in the *Match Peaks to Oligomers* section. The *Match Peaks to Oligomer* section presents each match with the measured *Peak Mass* (Da), the theoretical mass of the *Match*, the *Error* (Da) between the two, and the *Name.* The *Error* is the mass difference, positive or negative, between the measured mass and the theoretical mass. The *Name* is taken from the name in the *Oligomer List*; a monomer match would be labeled [*Name*] and a dimer would be labeled *2*[*Name*]. A match will only be displayed if the absolute value of the error is lower than value set in the *Error Tolerance for Matching* box.

5.2.6 "*Match to Isolated Oligomers"* will only allow isolated rows when calculating potential masses, which is useful for matching multiple protein species (each row) that should not combine to form complexes.

5.2.7 Selecting "*Match to Mixed Oligomers*" will allow all possible combinations of rows, which is useful for matching potential combinations of PTMs on a protein or for matching combinations of oligomers.

5.2.8 Clicking *Ok* will save the matches and label each peak in the main panel with the matched name.

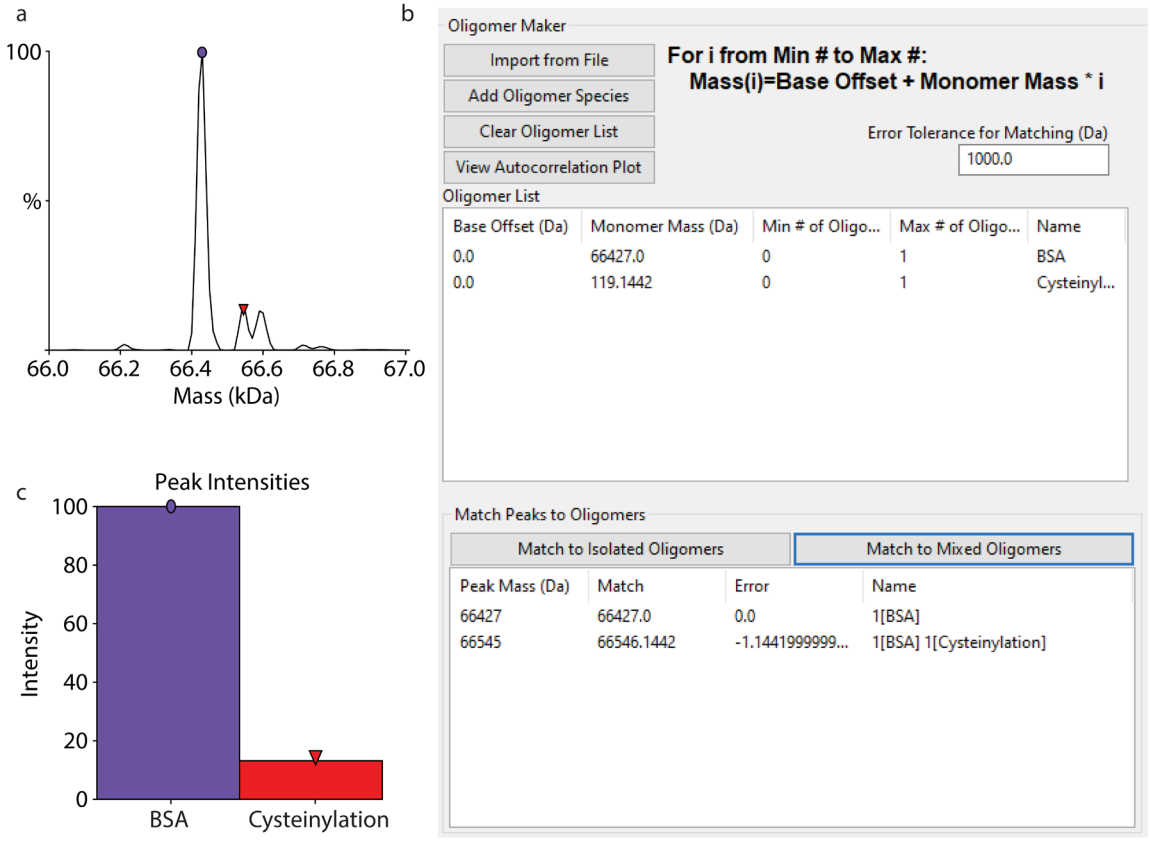

Fig. 3 The *Mass and Oligomer Tools* used to identify a cysteinylation of BSA.*[37]* **a**) Deconvolved mass spectrum of BSA. **b**) *Mass and Oligomer Tools* window, the mass of BSA was manually entered, and the mass of the cysteinylation was imported from the *Common Masses List*. **c**) Relative peak intensities for BSA and BSA with a cysteinylation created in the main panel.

5.3 Macromolecular Mass Defect Analysis

Macromolecular mass defect analysis can be highly useful for visualizing and quantifying complex mass spectra, especially when there are repeating mass units. The principle behind macromolecular mass defect analysis is that small shifts in mass relative to a repeating multiple of a reference mass can be used to assign features of the spectrum. Kendrick mass defect analysis was originally used in hydrocarbon analysis, where the reference mass of CH_2 was used to differentiate a wide diversity of hydrocarbons.[38,39] Recently, mass defect has been used to determine the stoichiometry of membrane proteins and antimicrobial peptides in lipid nanodiscs.[40,31,29,10] Nanodiscs have a repeating pattern of lipids, and subtle shifts in this pattern can reveal the contents of the nanodiscs even when the overall mass distribution does not shift appreciably, as shown in Fig. 4. The key advantage of mass defect analysis is that polydisperse spectra that differ only in multiples of a reference mass will have the same mass defect, so common features in complex spectra can be visualized.

The mass defect is calculated by dividing the measured masses by the reference mass. The remainder of this division is the mass defect. It can be expressed as a normalized mass defect with a decimal value between 0 and 1 or as an absolute mass defect in units of Da between 0 and the reference mass. Combining mass defects follows modular arithmetic such that a normalized mass

defect larger than 1 would simply drop the integer part and retain the decimal place. For example, an exact multiple of the reference mass would have a mass defect of 0, which is also equivalent to 1. A component that is 0.4 times the reference mass would have a mass defect of 0.4. Two such components would cause a mass defect of 0.8, and three would cause a mass defect of 0.2. Importantly, these mass defect values are independent of the absolute mass and are the same across any number of repeating reference mass units.

5.3.1 After deconvolving the spectrum, go to *Analysis* and select "*Mass Defect Tools"* or use the shortcut Ctrl+K. The reference mass is automatically imported from the *Mass Difference* in the main panel. If no *Mass Difference* value is set, a reference mass will need to be manually entered in the *Mass Defect Tools* window and replotted.

5.3.2 Within the *Mass Defect* window, the zero-charge mass spectrum in the top left is colored with mass defect, which shows the mass defect values for each point on the zero-charge mass spectrum (Fig. 4c). Below the zero-charge mass spectrum, a 2-D plot shows mass defect on the y-axis and mass (kDa) on the x-axis. This plot is useful for also visualizing the mass defect of different mass distributions in the spectrum. To the right, the two plots show the total relative intensity summed across all masses on the y-axis and mass defect on the x-axis (Fig. 4d). This plot is useful for identifying different mass defect species in the zero-charge mass spectrum. The apex of each peak can be detected by clicking *Plot* > *Label Peaks*.

5.3.3 For higher resolution data, increasing the number of mass defect bins may be necessary to capture all relevant features. Clicking "*Replot*" will reset the plots with the new parameters. Additional options allow the user to specify the range and whether to use normalized or absolute mass defect units.

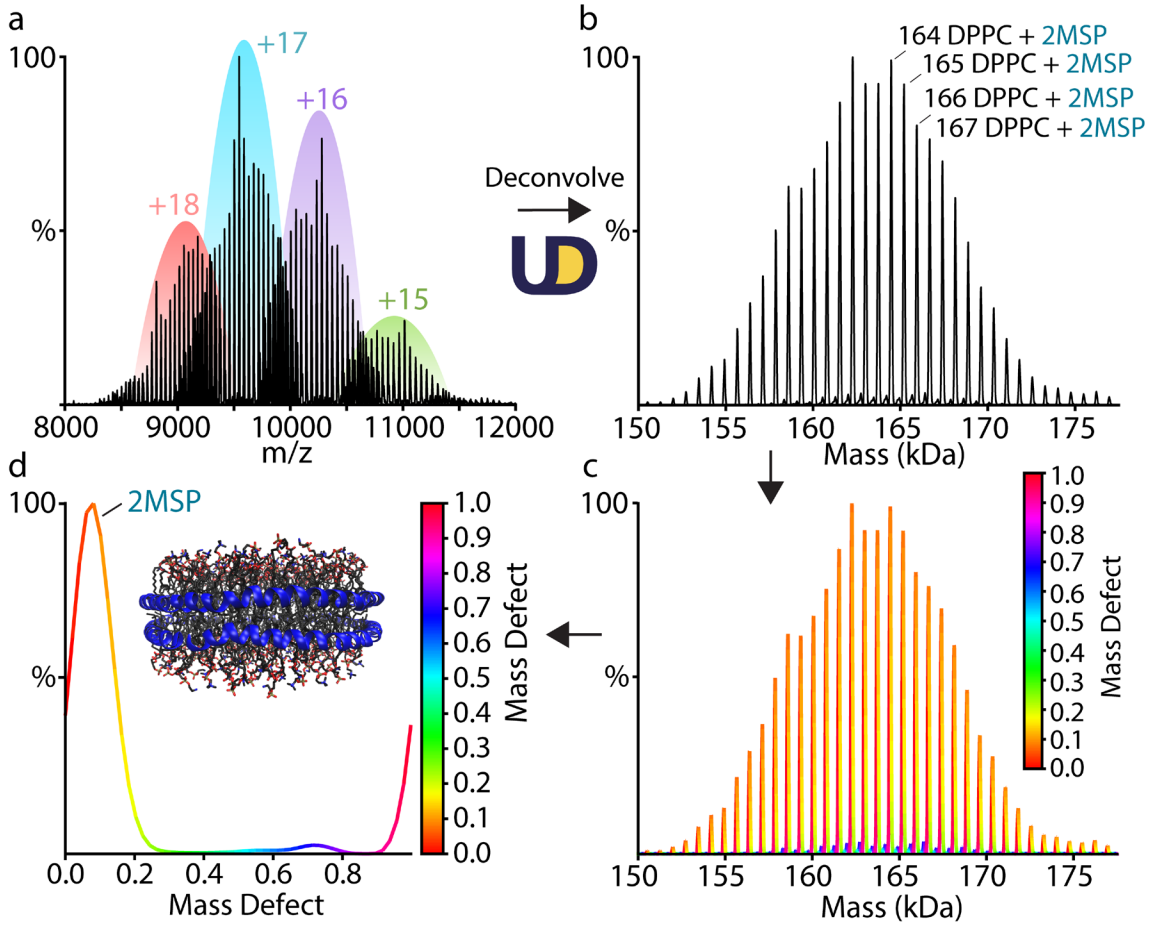

Fig. 4 Native mass spectrum of DPPC nanodiscs with the corresponding mass defect analysis. **a**) Mass spectrum of DPPC nanodiscs with charge states highlighted. **b**) Deconvolved mass distribution of DPPC nanodiscs with the number of lipids annotated for select peaks. **c**) The deconvolved mass distribution colored by mass defect. **d**) Summed mass defects from the deconvolved mass distribution. The peak around 0.1 corresponds to a nanodisc with 2 MSPs.

6 Exporting Plots and Mass Spectra

Now that the data has been analyzed and plotted, UniDec has multiple tools to export any plot or mass spectra as figures for publication. Many windows have additional save figure tools in the file menu, but we will cover basic exporting tools that will be useful for any user.

6.1.1 All plots in the main UniDec window can be automatically exported as a range of file types, including PDF or PNG, by selecting the *File > Save Figures As* menu or selecting one of the *File > Save Figure* presets.

6.1.2 Additionally, users can middle-click any plot within UniDec, including any analysis window, to open a save figure dialog and specify a save location. Ending the file name with .png or .pdf will set the file type, and any file types supported by MatPlotLib can be used.[41]

7 Conclusion

Here, we covered how the UniDec algorithm works, the deconvolution parameters, mass and oligomer tools, and mass defect analysis. Although this chapter provides an explanation of the key features of UniDec, we omitted some tools that may be useful for different types of analysis, including K_D fitting, native charge state analysis tools, and Fourier analysis. Additionally, MetaUniDec is useful for analyzing collections of MS data sets [28], and UniChrom adapts UniDec to time-resolved MS data, such as from LC/MS. Additional information on some of these tools, future features, and more can be found at https://github.com/michaelmarty/UniDec/wiki. Overall, with the deeper understanding of UniDec and associated analysis tools, we hope users will achieve the best quality deconvolution and analysis for intact and native MS spectra.

8 Notes

1. There are three ways to subtract the background in UniDec: *minimum*, *line*, and *curved*. *Subtract minimum* will find the lowest intensity data point and subtract the whole spectrum by that intensity. The number in the box, next to the subtract setting, does not matter for *Subtract minimum*, any number other than zero will turn it on. *Subtract minimum* is useful for a constant raised baseline. *Subtract line* will create a sloped line from the first and last *m/z* data points in the spectrum. The resulting line will be subtracted from the whole spectrum. The number in the box, *n*, sets how many data points to average at the front and end of the spectrum. *Subtract line* is most useful for mass spectra with linearly sloping baselines. *Subtract curved* will create a curved baseline based on local minima throughout the spectrum. The number in the box, *n*, will set how many data points are considered while creating each local minimum. A smaller *n* will create a rough baseline based on many local minima, and a larger *n* will create a smoother baseline based on fewer local minima. The default of *Subtract Curved* 100 works well for most spectra.
2. This note will describe some of the advanced data processing parameters. *Gaussian Smoothing* can be useful for smoothing noisy data but may broaden the peaks in the mass spectrum. UniDec will smooth the data with a Gaussian that has a width, in data points, set to the number in the box. However, nonlinear resampling (described below) is often a more effective way to smooth data while also speeding up the algorithm.

   Acceleration Voltage is used to correct for the detector efficiency between big ions and small ions in time-of-flight (ToF) mass spectrometers with multichannel plate (MCP) detectors.[42]

   Intensity Threshold is useful for removing data points that are under a fixed relative intensity. For example, if the threshold is set at 0.1, all data points below 10% of the most abundant peak will be removed from the spectrum.

   Data Reduction, like *Intensity Threshold*, is useful for removing less abundant data points from the spectrum. However, *Data Reduction* ranks all the data points from least abundant to most abundant and removes a fixed percentage of the least intense based on the percentage entered in the box.

   To normalize the data, click the box labeled "*Normalize Data"*. Normalizing the data will set the y-axis of all mass and deconvolved mass spectra in relative intensity units such that the maximum is 1. Turning normalization off will leave the y-axis in absolute signal intensity.

   At the bottom of the Additional Data Processing Parameters tab, a dropdown menu presents options on how to resample the data: *Linear*, *Linear Resolution*, *Nonlinear*, *Linear*

Interpolated, and *Linear Resolution Interpolated*. Entering a nonzero value in the Bin Every option will turn the resampling on.

Linear will linearize the spectrum such that the intensities of all data points within a fixed *m/z* range will be summed together into one data point. The value entered in the *Bin Every* option specifies the spacing between data points in the linearized data. For example, a bin value of 5 will create a data point every 5 *m/z*, and all intensities will be summed into the nearest 5 *m/z* bin.

Linear Resolution works similarly to *Linear*, but it linearizes based on a constant resolution rather than a constant *m/z*. If the bin value is set to 5, then the first two data points would be separated by 5 *m/z*, but the next two would be separated by a slightly larger value to keep a constant resolution.

Nonlinear resamples the data differently than linear. *Nonlinear* will tell UniDec to average every *n* number of data points together across the mass spectrum. Both the *m/z* and intensity values are averaged. The value, *n,* is set by the *Bin Every* parameter, and this value specifies the number of data points averaged instead of the *m/z* spacing.

Linear Interpolated works similarly to *Linear*. However, *Linear Interpolated* interpolates the intensity at each linearized data point rather than simply summing everything into fixed bins. *Linear Interpolated* should only be used for oversampling the data.

Linear Resolution Interpolated works similarly to *Linear Interpolated*, but it linearizes with a constant resolution rather than a constant *m/z*.

The *Nonlinear* option is usually the best for reducing the data. It has the least potential for artifacts, resists mass peak broadening, and matches the nonlinear nature of most mass spectra. It is an effective way to smooth the data and speed up the algorithm without introducing artifacts.

3. With a *Charge Smooth Width* of 1, UniDec will use a log mean filter to increase the intensity of potential charge states with neighboring charge states and decrease the intensity of those without neighboring charge states. The *Charge Smooth Width* value tells UniDec how many neighboring charge states to expect on either side. The default value of 1 is usually optimal.
4. The *Smooth Nearby Points* feature will smooth nearby points within a charge state using a mean filter to reduce artifacts in the deconvolved mass spectrum with a width set by the value in the *Point Smooth Width* box. There are four options for this function, *None*, *Some*, *Lots*, and *Other*. Clicking "*Some"* will set the *Point Smooth Width* to 1, which means a mean filter of width ±1 will be used to smooth each row of the charge vs. *m/z* matrix. Clicking "*Lots"* will set the point smooth filter width to 10 and may broaden the mass peaks. Inputting any value other than 1 or 10 in the *Point Smooth Width*, will set the *Smooth Nearby Points* to *Other*.
5. There are four options for *Suppress Artifacts*: *None*, *Some*, *Lots*, and *Other*. Clicking "*Some*" will set the *Beta* parameter to 50, which is a medium setting. Clicking "*Lots*" will set the *Beta* parameter to 500, which is a high setting that will reduce lots of artifacts; this setting may distort the deconvolved mass spectrum, especially with mass spectra that have overlapping mass peaks. In cases with many overlapping peaks, a *Beta* value of 5-15 may be more appropriate. Manually setting a *Beta* value other than 50 or 500 will switch the *Suppress Artifacts* setting to *Other*.

6. This note describes several advanced deconvolution parameters. *Charge Scaling* divides the intensity of each ion by its charge state to correct for different signal responses for different charge states, which is usually the case in FTICR and Orbitrap mass spectra. For example, an ion with a +2 charge state would have double the signal intensity of the same ion with a +1 charge state because it induces double the current in an FTICR or Orbitrap detector.

   Checking the *Mass List Window* option enforces a list of allowed masses that are entered in the *Mass List* section of the *Mass and Oligomer Tools*, which can be opened by going to *Tools* > *Mass and Oligomer Tools* or by using the shortcut Ctrl+T. The value in the box specifies the range around each entered mass value, in Da, that is allowed. For example, setting the value to 1000 would allow a ±1000 Da window around each mass entered in the *Mass List* but would exclude any masses not within that window from one of the defined masses.

   The *Native Charge Offset Range* sets how far an allowed charge state can be from the predicted native charge state of a globular protein of the same mass.[43] This feature is useful for clipping off high or low charge states that are outside of the expected range for a given mass without having to limit the global charge range.

   In the *Additional Deconvolution Parameters* tab, there is a *Negative Mode* check box that will automatically switch the adduct mass to a negative value when turned on. By default, the *Adduct Mass* is set to the mass of a proton. If the charge carrier for the ions were sodium, then the mass of sodium would be used as the *Adduct Mass*. However, if negative ionization mode is being used, then the *Adduct Mass* will have to be switched to a negative value because the charge is caused by the loss of a proton.
7. The *DScore* value is a part of the UniScore tool within UniDec. The UniScore rates the quality of the deconvolution with a range of 0–100, with 0 or less being completely wrong and 100 being perfect. The *DScore* scores each mass peak on the same scale. Briefly, the *DScore* is based on the presence of unique charge state peaks for a mass species, the fit of the peak shape across all charge states, the smoothness of the charge state distribution, and the asymmetry of the mass peak shape. Normally, a *DScore* above 60 indicates a confident deconvolution, and peaks with lower *DScore* values should be checked. Lowering the FWHM parameter can help increase the *DScore*. To learn more about the UniScore function see [34].
8. Other parameters in the *Additional Plotting Parameters* tab include *Peak Normalization*, *Color Maps*, and *Publication Mode. Peak Normalization* will tell UniDec how to normalize the intensity of each mass peak. Setting this parameter to "*None*" will turn off the normalization. "*Max*" will normalize the intensities to a percent relative intensity with the most abundant peak having an intensity of 100%. *"Total"* will set the normalization such that the intensity of all peaks added together will be 100%.

   The *Integration Range* parameter is important for integrating the mass peaks in the deconvolved mass spectrum. UniDec will integrate the area under the selected mass peak with a range specified in the minimum and maximum boxes of the *Integration Range.* For example, the range can be asymmetric such that a range of -500 to 1000 Da could be used.

   Clicking the "*Publication Mode*" checkbox will plot the raw *m/z* and deconvolved mass spectra in a simpler plot that is more amenable for publication.

9 Acknowledgements

The authors thank James Keener for providing the ADH example spectrum. This work was funded by the National Science Foundation (CHE-1845230). The authors thank the broad and diverse community of UniDec users for their helpful suggestions and continued support.

10 References

1. Beck A, Wagner-Rousset E, Ayoub D, Van Dorsselaer A, Sanglier-Cianferani S (2013) Characterization of therapeutic antibodies and related products. Anal Chem 85 (2):715-736. doi:10.1021/ac3032355
2. Gregorich ZR, Ge Y (2014) Top-down proteomics in health and disease: challenges and opportunities. Proteomics 14 (10):1195-1210. doi:10.1002/pmic.201300432
3. Cai W, Tucholski TM, Gregorich ZR, Ge Y (2016) Top-down Proteomics: Technology Advancements and Applications to Heart Diseases. Expert Rev Proteomics 13 (8):717-730. doi:10.1080/14789450.2016.1209414
4. Toby TK, Fornelli L, Kelleher NL (2016) Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annu Rev Anal Chem (Palo Alto Calif) 9 (1):499-519. doi:10.1146/annurev-anchem-071015-041550
5. Smith LM, Kelleher NL, Consortium for Top Down P (2013) Proteoform: a single term describing protein complexity. Nat Methods 10 (3):186-187. doi:10.1038/nmeth.2369
6. Schaffer LV, Millikin RJ, Miller RM, Anderson LC, Fellers RT, Ge Y, Kelleher NL, LeDuc RD, Liu X, Payne SH, Sun L, Thomas PM, Tucholski T, Wang Z, Wu S, Wu Z, Yu D, Shortreed MR, Smith LM (2019) Identification and Quantification of Proteoforms by Mass Spectrometry. Proteomics 19 (10):e1800361. doi:10.1002/pmic.201800361
7. Gault J, Liko I, Landreh M, Shutin D, Bolla JR, Jefferies D, Agasid M, Yen HY, Ladds M, Lane DP, Khalid S, Mullen C, Remes PM, Huguet R, McAlister G, Goodwin M, Viner R, Syka JEP, Robinson CV (2020) Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. Nat Methods 17 (5):505-508. doi:10.1038/s41592-020-0821-0
8. Bolla JR, Agasid MT, Mehmood S, Robinson CV (2019) Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. Annu Rev Biochem 88:85-111. doi:10.1146/annurev-biochem-013118-111508
9. Calabrese AN, Radford SE (2018) Mass spectrometry-enabled structural biology of membrane proteins. Methods 147:187-205. doi:10.1016/j.ymeth.2018.02.020
10. Keener JE, Zambrano DE, Zhang G, Zak CK, Reid DJ, Deodhar BS, Pemberton JE, Prell JS, Marty MT (2019) Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. J Am Chem Soc 141 (2):1054-1061. doi:10.1021/jacs.8b11529
11. Eschweiler JD, Kerr R, Rabuck-Gibbons J, Ruotolo BT (2017) Sizing Up Protein-Ligand Complexes: The Rise of Structural Mass Spectrometry Approaches in the Pharmaceutical Sciences. Annu Rev Anal Chem (Palo Alto Calif) 10 (1):25-44. doi:10.1146/annurev-anchem-061516-045414
12. Zhou M, Lantz C, Brown KA, Ge Y, Paša-Tolić L, Loo JA, Lermyte F (2020) Higher-order structural characterisation of native proteins and complexes by top-down mass spectrometry. Chem Sci. doi:10.1039/d0sc04392c
13. Peetz O, Hellwig N, Henrich E, Mezhyrova J, Dotsch V, Bernhard F, Morgner N (2019) LILBID and nESI: Different Native Mass Spectrometry Techniques as Tools in Structural Biology. J Am Soc Mass Spectrom 30 (1):181-191. doi:10.1007/s13361-018-2061-4
14. Allison TM, Reading E, Liko I, Baldwin AJ, Laganowsky A, Robinson CV (2015) Quantifying the stabilizing effects of protein-ligand interactions in the gas phase. Nat Commun 6 (1):8551. doi:10.1038/ncomms9551
15. Cleary SP, Prell JS (2019) Liberating Native Mass Spectrometry from Dependence on Volatile Salt Buffers by Use of Gabor Transform. Chemphyschem 20 (4):519-523. doi:10.1002/cphc.201900022
16. Campuzano IDG, Robinson JH, Hui JO, Shi SD, Netirojjanakul C, Nshanian M, Egea PF, Lippens JL, Bagal D, Loo JA, Bern M (2019) Native and Denaturing MS Protein Deconvolution for Biopharma: Monoclonal

Antibodies and Antibody-Drug Conjugates to Polydisperse Membrane Proteins and Beyond. Anal Chem 91 (15):9472-9480. doi:10.1021/acs.analchem.9b00062
17. Cleary SP, Li H, Bagal D, Loo JA, Campuzano IDG, Prell JS (2018) Extracting Charge and Mass Information from Highly Congested Mass Spectra Using Fourier-Domain Harmonics. J Am Soc Mass Spectrom 29 (10):2067-2080. doi:10.1007/s13361-018-2018-7
18. Bern M, Caval T, Kil YJ, Tang W, Becker C, Carlson E, Kletter D, Sen KI, Galy N, Hagemans D, Franc V, Heck AJR (2018) Parsimonious Charge Deconvolution for Native Mass Spectrometry. J Proteome Res 17 (3):1216-1226. doi:10.1021/acs.jproteome.7b00839
19. Cleary SP, Thompson AM, Prell JS (2016) Fourier Analysis Method for Analyzing Highly Congested Mass Spectra of Ion Populations with Repeated Subunits. Anal Chem 88 (12):6205-6213. doi:10.1021/acs.analchem.6b01088
20. Tseng YH, Uetrecht C, Yang SC, Barendregt A, Heck AJ, Peng WP (2013) Game-theory-based search engine to automate the mass assignment in complex native electrospray mass spectra. Anal Chem 85 (23):11275-11283. doi:10.1021/ac401940e
21. Sivalingam GN, Yan J, Sahota H, Thalassinos K (2013) Amphitrite: A program for processing travelling wave ion mobility mass spectrometry data. Int J Mass spectrom 345-347:54-62. doi:10.1016/j.ijms.2012.09.005
22. Morgner N, Robinson CV (2012) Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. Anal Chem 84 (6):2939-2948. doi:10.1021/ac300056a
23. Ferrige AG, Seddon MJ, Green BN, Jarvis SA, Skilling J (1992) Disentangling Electrospray Spectra with Maximum-Entropy. Rapid Commun Mass Spectrom 6 (11):707-711. doi:DOI 10.1002/rcm.1290061115
24. Lu J, Trnka MJ, Roh SH, Robinson PJ, Shiau C, Fujimori DG, Chiu W, Burlingame AL, Guan S (2015) Improved Peak Detection and Deconvolution of Native Electrospray Mass Spectra from Large Protein Complexes. J Am Soc Mass Spectrom 26 (12):2141-2151. doi:10.1007/s13361-015-1235-6
25. Silvia MM, Sara C, Florian F, Jennifer S, Paul G, Ken C, Kai S, Jonathan B (2020) Optimisation of the use of sliding window deconvolution for comprehensive characterisation of trastuzumab and adalimumab charge variants by native high resolution mass spectrometry. Eur J Pharm Biopharm 158:83-95. doi:10.1016/j.ejpb.2020.11.006
26. Allison TM, Barran P, Benesch JLP, Cianferani S, Degiacomi MT, Gabelica V, Grandori R, Marklund EG, Menneteau T, Migas LG, Politis A, Sharon M, Sobott F, Thalassinos K (2020) Software Requirements for the Analysis and Interpretation of Native Ion Mobility Mass Spectrometry Data. Anal Chem 92 (16):10881-10890. doi:10.1021/acs.analchem.9b05792
27. Marty MT, Baldwin AJ, Marklund EG, Hochberg GK, Benesch JL, Robinson CV (2015) Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Anal Chem 87 (8):4370-4376. doi:10.1021/acs.analchem.5b00140
28. Reid DJ, Diesing JM, Miller MA, Perry SM, Wales JA, Montfort WR, Marty MT (2019) MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra. J Am Soc Mass Spectrom 30 (1):118-127. doi:10.1007/s13361-018-1951-9
29. Walker LR, Marty MT (2020) Revealing the Specificity of a Range of Antimicrobial Peptides in Lipid Nanodiscs by Native Mass Spectrometry. Biochemistry 59 (23):2135-2142. doi:10.1021/acs.biochem.0c00335
30. Kostelic MM, Ryan AM, Reid DJ, Noun JM, Marty MT (2019) Expanding the Types of Lipids Amenable to Native Mass Spectrometry of Lipoprotein Complexes. J Am Soc Mass Spectrom 30 (8):1416-1425. doi:10.1007/s13361-019-02174-x
31. Walker LR, Marzluff EM, Townsend JA, Resager WC, Marty MT (2019) Native Mass Spectrometry of Antimicrobial Peptides in Lipid Nanodiscs Elucidates Complex Assembly. Anal Chem 91 (14):9284-9291. doi:10.1021/acs.analchem.9b02261

32. Campuzano ID, Li H, Bagal D, Lippens JL, Svitel J, Kurzeja RJ, Xu H, Schnier PD, Loo JA (2016) Native MS Analysis of Bacteriorhodopsin and an Empty Nanodisc by Orthogonal Acceleration Time-of-Flight, Orbitrap and Ion Cyclotron Resonance. Anal Chem 88 (24):12427-12436. doi:10.1021/acs.analchem.6b03762
33. Ahdash Z, Lau AM, Martens C, Politis A (2018) Analyzing Protein Architectures and Protein-Ligand Complexes by Integrative Structural Mass Spectrometry. J Vis Exp (140):e57966. doi:10.3791/57966
34. Marty MT (2020) A Universal Score for Deconvolution of Intact Protein and Native Electrospray Mass Spectra. Anal Chem 92 (6):4395-4401. doi:10.1021/acs.analchem.9b05272
35. Marty MT (2019) Eliminating Artifacts in Electrospray Deconvolution with a SoftMax Function. J Am Soc Mass Spectrom 30 (10):2174-2177. doi:10.1007/s13361-019-02286-4
36. Marty MT, Hoi KK, Robinson CV (2016) Interfacing Membrane Mimetics with Mass Spectrometry. Acc Chem Res 49 (11):2459-2467. doi:10.1021/acs.accounts.6b00379
37. Pratesi A, Cirri D, Ciofi L, Messori L (2018) Reactions of Auranofin and Its Pseudohalide Derivatives with Serum Albumin Investigated through ESI-Q-TOF MS. Inorg Chem 57 (17):10507-10510. doi:10.1021/acs.inorgchem.8b02177
38. Kendrick E (1963) A Mass Scale Based on CH2= 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. Anal Chem 35 (13):2146-2154. doi:10.1021/ac60206a048
39. Marshall AG, Rodgers RP (2004) Petroleomics: the next grand challenge for chemical analysis. Acc Chem Res 37 (1):53-59. doi:10.1021/ar020177t
40. Marty MT, Hoi KK, Gault J, Robinson CV (2016) Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs. Angew Chem Int Ed Engl 55 (2):550-554. doi:10.1002/anie.201508289
41. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. Comput Sci Eng 9 (3):90-95. doi:10.1109/mcse.2007.55
42. Fraser GW (2002) The ion detection efficiency of microchannel plates (MCPs). Int J Mass spectrom 215 (1-3):13-30. doi:10.1016/s1387-3806(01)00553-x
43. Stengel F, Baldwin AJ, Painter AJ, Jaya N, Basha E, Kay LE, Vierling E, Robinson CV, Benesch JL (2010) Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. Proc Natl Acad Sci U S A 107 (5):2007-2012. doi:10.1073/pnas.0910126107