Supporting information

Mass spectrometry-based structural analysis of cysteine-rich metal-binding sites in proteins with MetaOdysseus R software

Manuel David Peris-Díaz,[a] Roman Guran,[b,c] Ondrej Zitka,[b,c] Vojtech Adam,[b,c]

and Artur Krężel[a]*

[a] *Department of Chemical Biology, Faculty of Biotechnology, University of Wrocław, F. Joliot-Curie 14a, 50-383 Wrocław, Poland*

[b] *Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University in Brno, Zemedelska 1, 613 00 Brno, Czech Republic*

[c] *Central European Institute of Technology, Brno University of Technology, Purkynova 656/123, 612 00 Brno, Czech Republic*

Keywords: metal ions, metallothionein, mass spectrometry, (zinc buffering, metalation)

* corresponding author, E-mail: artur.krezel@uwr.edu.pl

Table of contents

Experimental section

Materials. The following reagents were purchased from Sigma-Aldrich: $ZnSO_4 \cdot 7H_2O$, 4-(2-pyridylazo)resorcinol (PAR), $(NH_4)_2CO_3$, tris(hydroxymethyl)aminomethane (Tris base) and 4-(2-hydroxyethyl)-1 piperazineethanesulfonic acid (HEPES), iodoacetamide (IAM), iodoacetic acid (IAA), ethyl iodoacetate (ET), *N*-ethylmaleimide (NEM), tris(2-carboxyethyl)phosphine hydrochloride (TCEP), ethylenediamine-tetraacetic acid (EDTA), proteomics grade trypsin, mass spectrometry grade methanol, a commercial ESI-TOF Tuning mix, mass spectrometry grade ammonium acetate and mass spectrometry grade acetonitrile (ACN). The metal-chelating resin Chelex 100 was acquired from Bio-Rad and 98% hydrochloric acid (HCl) was purchased from VWR Chemicals. Trypton, yeast extract, LB broth, agar, agarose, isopropyl-β-D-1-thiogalactopyranoside (IPTG), and SDS were from Lab Empire, NaCl, NaOH, glycerol, $KH_2PO_4 \cdot H2O$, K_2HPO_4 from POCH (Gliwice Poland), pTYB21 vector and chitin resin were from New England BioLabs, 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) from TCI Europe N.V. was purchased from Sigma-Aldrich. Dithiothreitol (DTT) was purchased from Iris Biotech GmbH.

External native MS data set. External native MS spectra of soluble proteins (myoglobin, BSA, ADH, and Herceptin) analyzed with an Orbitrap Eclipse tribrid published by Robinson group[1].

Expression and purification of metallothioneins. Expression vectors of MT2, MT3 and MT1e were transformed into BL21(DE3)pLysS *E. coli* cells and cultured in a rich full culture medium (1.1% tryptone, 2.2% yeast extract, 0.45% glycerol, 1.3% K_2HPO_4, 0.38% KH_2PO_4) at 37°C until OD_{600} = 0.5. Cells were induced with 0.1 mM IPTG and incubated overnight at 20°C with vigorous shaking. The next purification steps were conducted at 4°C. Cells were collected by centrifugation (4,000 × *g* for 10 min), resuspended in 50 mL of cold buffer A (20 mM HEPES, pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM TCEP) and sonicated for 30 min (1 min "on" and 1 min "off") followed by centrifugation (20,000 × *g* for 15 min). The supernatant was incubated with 20 mL of chitin resin in buffer A and kept overnight with mild shaking. After the incubation resin was washed 4-5 times with 50 mL of buffer A and cleavage was induced by the addition of 100 mM DTT and incubation for 48 h at room temperature on a rocking bed[2]. Eluted solution was acidified to pH ~ 2.5 with 7% HCl and concentrated using Amicon Ultra-4 Centrifugal Filter Units with a membrane cut-off of 3 kDa (Merck Millipore, USA) and purified on a size exclusion chromatography SEC-70 gel filtration column (Bio-Rad) equilibrated with 10 mM HCl[3]. The identity of apo-MT2 and apo-MT3 protein (thionein) was

confirmed by ESI-MS (API 2000 instrument, Applied Biosystems, USA). The concentration of thiols was determined spectrophotometrically using DTNB assay[4]. Protein binding capacity was confirmed spectrophotometrically by Zn(II) and Cd(II) titrations[5]. Purified thioneins (apo-forms of MTs) were mixed with 10 molar excess of $ZnSO_4$ under a nitrogen blanket and pH adjusted to 8.6 with 1 M solution of Tris base. Samples were concentrated with Amicon Ultra-4 Centrifugal Filter Units with a membrane cut-off of 3 kDa (Merck Millipore, USA) and purified on a size exclusion chromatography SEC-70 gel filtration column (Bio-Rad) equilibrated with 20 Tris-HCl buffer, pH 8.6. The concentration of thiols and Zn^{2+} was determined spectrophotometrically using DTNB and PAR assay, respectively.

R packages embedded in MetaOdysseus. The R package includes several algorithms for pre-processing mass spectrometry data, aimed at correcting the baseline drift and remove the background noise, which are common steps for pre-processing MS data[7]. For baseline removal, there was incorporated the asymmetric least squares algorithm implemented from the ptw R package[8]. Removing the background noise is accomplished with any of the smoothing algorithms embedded, namely, the Savitzky-Golay filter from the MassSpecWavelet R package[9], finite difference penalty from the ptw R package[8] and discrete wavelet transform from the MassSpecWavelet R package[9]. Plots were implemented using the ggplot2 R package[10] and the R package OrgMassSpec[11], which was embedded and modified.

Charge deconvolution with the simulation method. Once the experimental spectrum has been preprocessed and a list of monoisotopic peaks obtained, the first step is to find the most intensive unprocessed peak $h_{m/z}$ above a defined signal-to-noise ratio (SNR). The optimum SNR is automatically calculated and updated for each peak $h_{m/z}$, starting with an initial SNR value usually set up to 2. From the user-defined maximum mass and charge state, z_{max}, $h_{m/z}$ take a charge state of z_i with $z_i = \{z_i, z_{i+1}, \ldots, z_{max}\}$. After, for each assigned z_i it calculates the *m/z* value for a set of ions with different charge states surrounding z_i (i.e., z_i +1, z_i +2, z_i +3, …, z_i -1, z_i -2) and with a peak finding routine search in the experimental spectrum within a m/z tolerance window. For each found ion *m/z*, a score $s_{m/z}$ *is* defined as the $(\log(SNR))_{m/z}$. The total score *S* for each charge state z_i is simply $\sum s_{m/z}$. A first filter assigns that the charge state *z* is determined as the max(*S*). The charge state distributions should be continuous, however, it might be the case that one charge state deviates from the tolerance set up and it is excluded from the peak series. Thus first, we allowed a maximum difference in the charge state *z* between the peaks is 2. The peak series that has the maximum of

consecutive charges determines the charge state *z* for $h_{m/z}$ and the neutral mass *M* is calculated for *h* and all peak series. For every peak in the charge state series, we extract a mass spectrum in the m/z domain determined by FWHM (default to 2). The FWHM is optimized to extract at least five m/z signals. Then, every m/z domain extracted is fitted to a Gaussian distribution with a nonlinear least squares with the Gauss-Newton algorithm. From the fitting, we obtain the mean, sigma, and height for every peak that corresponds to the mean *m/z*, the broadness, and the peak height, respectively. After we simulate a Gaussian distribution using the parameters obtained for each peak, transformed the zero-charge mass domain and their intensities summed. A difference with Massign[12], we do not fit the intensity *vs* charge state that requires several charge states. For small proteins or cases where no more than 3 charge states are observed, this premise might not work. This approach leads to an overestimation of the peak intensities in the next rounds if the part of the spectrum in the m/z domain that was extracted is shared with another peak series. To avoid this, in the next round we are starting the difference spectrum and not with the experimental spectrum. The difference spectrum is calculated as the subtracted between the simulated and the original and the algorithm goes back to the starting point. The overall simulation is obtained by the sum of the simulations. The approach presented here combines ideas from the Z-Score[13] and the Massign[12] software.

Charge deconvolution with peak assignment method. Here, we adapted and slightly modified the Z-Score[13] algorithm for high-resolution and low-charge ESI spectra. This method allows to determine the charge state from single isotopic multiplet, and thus species that only carry one charge. Briefly, as above the first step is to find the most intensive unprocessed peak $h_{m/z}$ above an SNR threshold. The maximum possible charge state z_{max} is calculated from the separation between $h_{m/z}$ and the nearest-neighbor peak. Then, $h_{m/z}$ take a charge state of z_i with $z_i = \{z_i, z_{i+1}, \ldots, z_{max}\}$. After, for each assigned z_i it calculates the *m/z* value for a set of isotopic ions lighter and heavier than $h_{m/z}$ (i.e., $h_{m/z}$ + 1/ z_i, $h_{m/z}$ + 2/ z_i, …, $h_{m/z}$ - 1/ z_i, $h_{m/z}$ - 2/ z_i) and with a peak finding routine search in the experimental spectrum within a m/z tolerance window. For each found ion *m/z*, a score $s_{m/z}$ *is* defined as the $(\log(SNR))_{m/z}$. The total score *S* for each charge state z_i is simply $\sum s_{m/z}$. The charge state *z* is determined as the max(*S*). In the original implementation of the Zscore, the charge state *z* was determined considering only which z_i gives the max($s_{m/z}$) for $h_{m/z}$, and not the all isotopic multiplet familiy of $h_{m/z}$. The found isotopic multiplet is transformed to the zero-charge mass spectrum with the original intensities and these peaks denoted as processed. For this, the original m/z spectrum is subtracted from the m/z signals from the multiplet assigned. Then, the algorithm goes back to the initial step. The zero-charge mass spectrum is constructed by summing the individual multiplets.

Because not all of the m/z domain is sampled, a linear interpolation is performed over transformed data.

Implementation of the universal deconvolution score UniScore. The UniScore incorporated in this software has been explained in detailed elsewhere[14]. Briefly, the UniScore scores a deconvolved mass distribution based on four components: 1) The UScore that capture the uniqueness and fit of each charge state; 2) The MScore that evaluates the changes in the peak shape for all charge states from the same mas; 3) The CSScore that scores how distributes the charge state intensity to a Gaussian distribution and; 4) the Fscore that penalize peaks that are not good separated or exhibits high peak asymmetry. The deconvolution score (DScore) or a particular peak is the product of these four components. The total UScore is the R^2 obtained from the fitting of the deconvolution multiplied by the average DScore of each peak weighted by its squared intensity.

PMF score calculation and validation by permutation tests. To assess the quality of the deconvoluted and mass assigned mass spectrum we incorporated a global PMF score[15] in a function named *ScoresPMF.R*. It is based on the combination between the hit ratio and the mass coverage, defined as follows: Hit ratio = a number of masses matched/number of masses of the theoretical digested protein. Mass coverage was defined as (sequence coverage/100) × molecular weight protein (kDa). The PMF score was computed combining these scores as follows. PMF score = (hit ratio × 100) + mass coverage. The confidence interval of the calculated PMF score was estimated with nonparametric bootstrapping with 1000 bootstrap replicates. The PMF score as defined herein is a relative metric and thus validation is required in order to judge its quality. We included a statistical method to assign an empirical *p*-value to the calculated PMF score based on permutation tests. A list of decoy sequences is constructed by randomizing the known and target protein sequence, keeping the total amino acid composition[16]. Then, a null (random) distribution of PMF scores is determined by constructing a histogram of the PMF scores for the decoy list that was scored against the experimental spectrum. Once it is obtained a vector s that contains the PMF scores from the null distribution $s = \{s_1, s_2, \ldots, s_N\}$ where s_i is the PMF score computed on the ith randomized sequence and Then a one-sided empirical *p*-value for s_0 can be easily computed as *p*-value = (1+sum (s_i> = s_0))/(N+1), where s_0 is the PMF score obtained for the alternative hypothesis. These represent the classical *p*-value computation as the tail probability in a null distribution.

Comparison between MetaOdysseus and Bruker Daltonics for MALDI-MS. The peak picking parameters used in the SNAP algorithm included in BioTools (Bruker Daltonics) were as S/N threshold 2, quality factor threshold, 20 and relative intensity threshold 0. To calculate the confusion matrix for each dataset, first, we calculated the intersection and excluded annotated peptides by using a feature finding algorithm between MetaOdysseus and BioTools. Afterward, those that are unique features for each software were manually inspected to assign the number of FP and TP. The number of FN for a particular software was determined as the TP of the opposite software. We calculated the empirical FDR as FP/TP. The precision was calculated as TP/(TP+FP). The Recall was calculated as TP/(TP+FN). The TN was determined as the subtraction between the number of peaks picked and the sum of the peptides annotated and the FN of the opposite software. The accuracy was calculated as (TP+TN)/(TN+TP+FP+FN). The F1 score was calculated as (2 × Precision × Recall)/(Precision + Recall).

Scoring PSMs and filtering by FDR strategy in a bottom-up MS. In order to score the peptide-spectrum matches (PSMs), we incorporated a simple peak match probability scored that is based on computing probabilities from a binomial distribution. That is:

$$P(X) = \sum_{k=n}^{N} \binom{N}{k} p^k \, (1-p)^{N-k}$$

Where the cumulative binomial probability *P* was calculated using the number of trials *N*, the number of successes *n*, and the probability of success p. The p was calculated $p = (2\,x\,\text{mass window})/(\text{mass range})$. The total *N* is the total number of fragment ions for a given peptide. The *n* equals the number of ions matched into the experimental spectrum. Then the score was calculated as the $-\log(P)$. MetaOdysseus estimates FDRs using the target-decoy approach (TDA)[16,] constructing the decoy database with a randomized target protein sequence. We used a conservative approach with constructing a separate target-decoy database search and assessing the quality of matching individual spectra to peptide sequences. (Fig Score vs FP, Score threshold found 122). The TDA searching provided an estimate of the false-positive rate (FPR) in order to compare the different algorithms.

Comparison between MetaOdysseus, Mascot, and MS-GF+ for a bottom-up MS. We compared the different software with a target-decoy approach in order to deliver false-positive estimations. As commented, we used a more conservative approach with a separate target and decoy sequence instead

of a concatenated target-decoy sequence list[16]. Here, first, we search MS/MS spectra against decoy sequences, obtaining a null distribution of scores. On the other hand, we populate the target database with only target sequences, and thus the scores that are dependent on the search space will be mostly greater for the correct identifications, than if using concatenated target-decoy search. This no competition between target and decoy PSMs results in that the fraction of false PSMs among all target PSMs is not summed. For small databases, like herein, it is more recommended to use a more conservative approach to estimate the FDR[17]. Once we obtained the peptide hits for the target and decoy separately, we estimated how many targets and how many decoys exhibit a score S greater than or equal than and lower than a score threshold St_i with $St_i = \{St_0, St_1, \ldots, S_{N-1}\}$ where N is the score maximum. The number of FN was calculated as the sum of target hits with an $S < St_i$. The number of TN was calculated as the sum of decoy hits with an $S < St_i$. The number of FP was calculated as the sum of decoy hits with an $S \geq St_i$. The number of TP was calculated as the sum of target hits with an $S \geq St_i$. The FDR was estimated as FP/TP[18]. Sensitivity, precision, and specificity were calculated as above described.

Comparison between MetaOdysseus and MASH suite Pro for mass deconvolution of top-down MS.

To do the fairer comparison possible, we set up the deconvolution parameters in the MASH Suite Pro[20] and MetaOdysseus to obtain a similar number of deconvoluted masses. A reference list of charge state peak series was determined manually. Then, we inspected the peak charge state series assigned to determine TP and FP. If the charge state was correctly assigned, a TP was counted. If the charge state was incorrectly assigned, the peak was counted as FP. Then, to calculate the FN and TN we followed a simple rule: If MetaOdysseus assigned a peak that it is not assigned in MASH Suite pro, the peak could be a TP or FP. Once it is determined manually if it is TP or FP, the TN of FN is assigned. If it is an FP and it is not a duplicated peak, then it is a TN for MASH Suite pro. If it is a TP, then it is an FN for MASH Suite pro. Then we calculated the true positive rate (TPR) as TP/(TP+FN), and the FPR as FP/(FP+TN).

Table S1. Deconvolution results obtained with MetaOdysseus for a set of ESI-MS spectra. The R^2 reported the correlation coefficient for the fitted simulated component spectra to the experimental one. The UniScore stands for the recently presented universal score, defined in the main text of this manuscript (Case 1) and more details elsewhere.[14]

Sample	R^2	UniScore
Zn_0MT2 low resolution	0.96	69
Zn_0MT2 high resolution	0.97	51
Zn_7MT2 low resolution	0.95	38
Zn_7MT2 high resolution	0.85	43
Zn_7MT2 + IAM high resolution	0.98	62
Zn_7MT2 + IAM low resolution	0.90	47
Zn_7MT2 + NEM low resolution	0.86	48
Zn_7MT2 + NEM high resolution	0.89	53
BSA	0.98	80
Myoglobin	0.92	97
ADH	0.98	66
Herceptin	0.87	72

Table S2. The software parameters used in the MetaOdysseus for deconvolution results in Table S1. SNR.th stands for the signal-to-noise threshold; Wa stands for the mass error window; FWHM stands for the full width at half maximum; top stands for the top N peaks retained in percentage. Interval stands for the mass interval extracted from the m/z domain to the zero-charge mass spectrum. IntervaLo2 and intervaLo are used for the calculation of the UniScore.

Sample	SNR.th	Wa	FWHM	top	interval	intervaLo2	intervaLo
Zn_0MT2_low resolution	5	0.8	0.8	100	20	10	200
Zn_0MT2_high resolution	5	0.8	0.5	100	20	10	500
Zn_7MT2_high resolution	5	0.5	0.5	100	20	10	500
Zn_7MT2_low resolution	5	0.8	1.5	100	20	10	200
Zn_7MT2+IAM_highresolution	5	0.5	0.5	100	20	10	500
Zn_7MT2+IAM_lowresolution	5	0.8	1.1	100	20	10	200
Zn_7MT2 +NEM low resolution	5	0.5	1.1	20	20	10	500
Zn_7MT2 +NEM high resolution	5	0.5	0.5	20	20	10	500
Myoglobin	1	1.5	20	100	50	100	3000
BSA	1	1.5	20	100	50	200	3000
ADH	1	1.5	30	100	500	400	3000
Herceptin	1	1.5	50	100	1000	400	3000

Table S3. Mass assignment of the deconvolved spectrum obtained with MetaOdysseus in Table S1-S2. The mass error is calculated as the difference between the experimental and theoretical mass. The DScore, UScore, FScore, CCScore, and MScore are calculated as detailed elsewhere.[14]

Protein	Resolution	Assigned	Mass[a]	Error (Da)	DScore	UScore	FScore	CCScore	MScore
Apo-MT2 + 0 Zn^{2+} Eq	Low	Metal 0	6041.16	-1.04	0.75	0.90	1	1	0.84
Apo-MT2 + 0 Zn^{2+} Eq	High	Metal 0	6039.78	0.34	0.55	0.82	1	1	0.66
Apo-MT2 + 7 Zn^{2+} Eq	High	Metal 6	6419.39	-1.36	0.42	0.82	1	1	0.52
Apo-MT2 + 7 Zn^{2+} Eq	High	Metal 7	6485.38	0.68	0.53	0.87	1	1	0.60
Apo-MT2 + 7 Zn^{2+} Eq	Low	Metal 6	6419.57	-1.19	0.13	0.69	1	1	0.19
Apo-MT2 + 7 Zn^{2+} Eq	Low	Metal 7	6485.38	0.69	0.38	0.58	1	1	0.66
Apo-MT2 + 7 Zn^{2+} Eq + IAM	Low	$Zn_0IAM_{19}MT2$	7183.32	-1.22	0.20	0.86	1	1	0.23
Apo-MT2 + 7 Zn^{2+} Eq + IAM	Low	$Zn_0IAM_{20}MT2$	7240.30	-1.24	0.14	0.60	1	1	0.23
Apo-MT2 + 7 Zn^{2+} Eq + IAM	High	$Zn_0IAM_{19}MT2$	7183.52	-1.01	0.63	0.93	1	1	0.68
Apo-MT2 + 7 Zn^{2+} Eq + IAM	High	$Zn_0IAM_{20}MT2$	7240.45	-1.10	0.65	0.82	1	1	0.79
Apo-MT2 + 7 Zn^{2+} Eq + NEM	Low	$Zn_0NEM_{20}MT2$	8543.60	-1.10	0.71	0.90	1	1	0.79
Apo-MT2 + 7 Zn^{2+} Eq + NEM	High	$Zn_0NEM_{20}MT2$	8544.96	0.24	0.66	0.91	1	1	0.73
Myoglobin	High	-	17584.1	-	0.73	0.95	1	1	0.76
BSA	High	-	66464.8	-	0.83	0.94	1	1	0.88
ADH	High	-	174610	-	0.68	0.75	1	1	0.91
Herceptin	High	-	148213	-	0.83	0.91	1	1	0.91

[a]Assigned masses: average mass. Samples myoglobin, BSA, ADH and herceptin were obtained from external data set and were not mass assigned.

Table S4. Deconvolution results obtained with UniDec for the samples assayed. The footnotes present the preset used in UniDec to process particular sample. More information about the significance of the output can be found in UniDec or UniScore.[14]

Sample	R^2	DScore	Uniqueness/Fit	Peak shape	Charge distribution	FWHM penalty	UniScore
Zn_0MT2 low resolution[a]	0.99	93	95	98	100	100	92
Zn_0MT2 high resolution[b]	0.74	42	42	100	100	100	42
Zn_7MT2 low resolution[a]	0.99	50	55	93	99	100	33
Zn_7MT2 high resolution[b]	0.87	55	57	97	100	100	39
Zn_7MT2 +IAM high resolution[c]	0.55	63	65	96	100	100	34
Zn_7MT2 +IAM low resolution[a]	0.99	86	90	96	100	100	85
Zn_7MT2 +NEM low resolution[a]	0.96	41	46	88	100	100	55
Zn_7MT2 +NEM high resolution[c]	0.68	49	55	89	100	100	32
BSA[c]	0.98	81	89	96	100	95	81
Myoglobin[c]	0.95	78	82	96	100	100	77
ADH[c]	0.99	79	90	88	100	100	78
Herceptin[c]	0.99	90	96	94	100	100	90

[a]Presets: default.

[b]Presets: isotopic resolution.

[c]Presets: high-resolution native

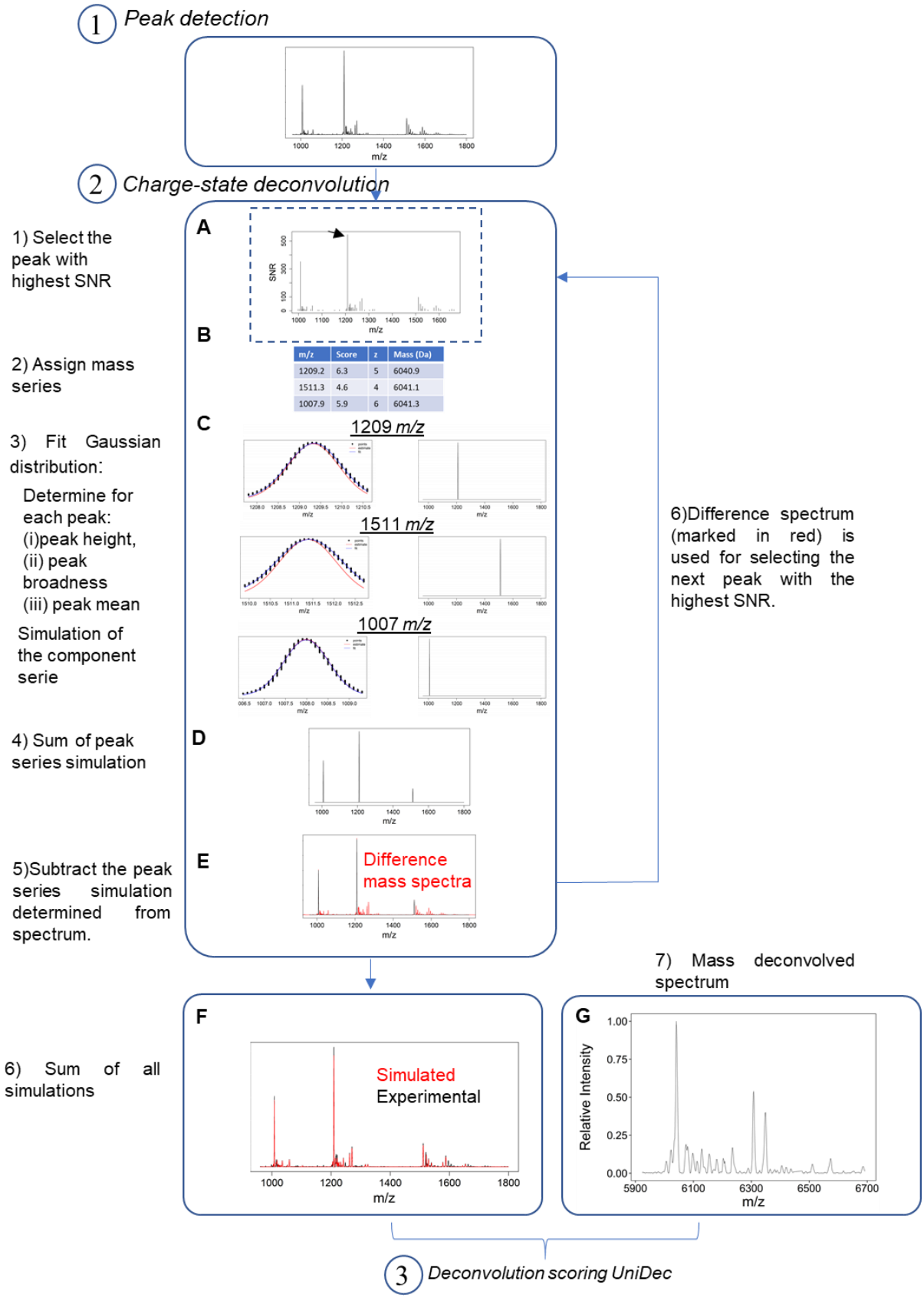

Figure S1. The charge-deconvolution overview explains above in the page S5-S6.

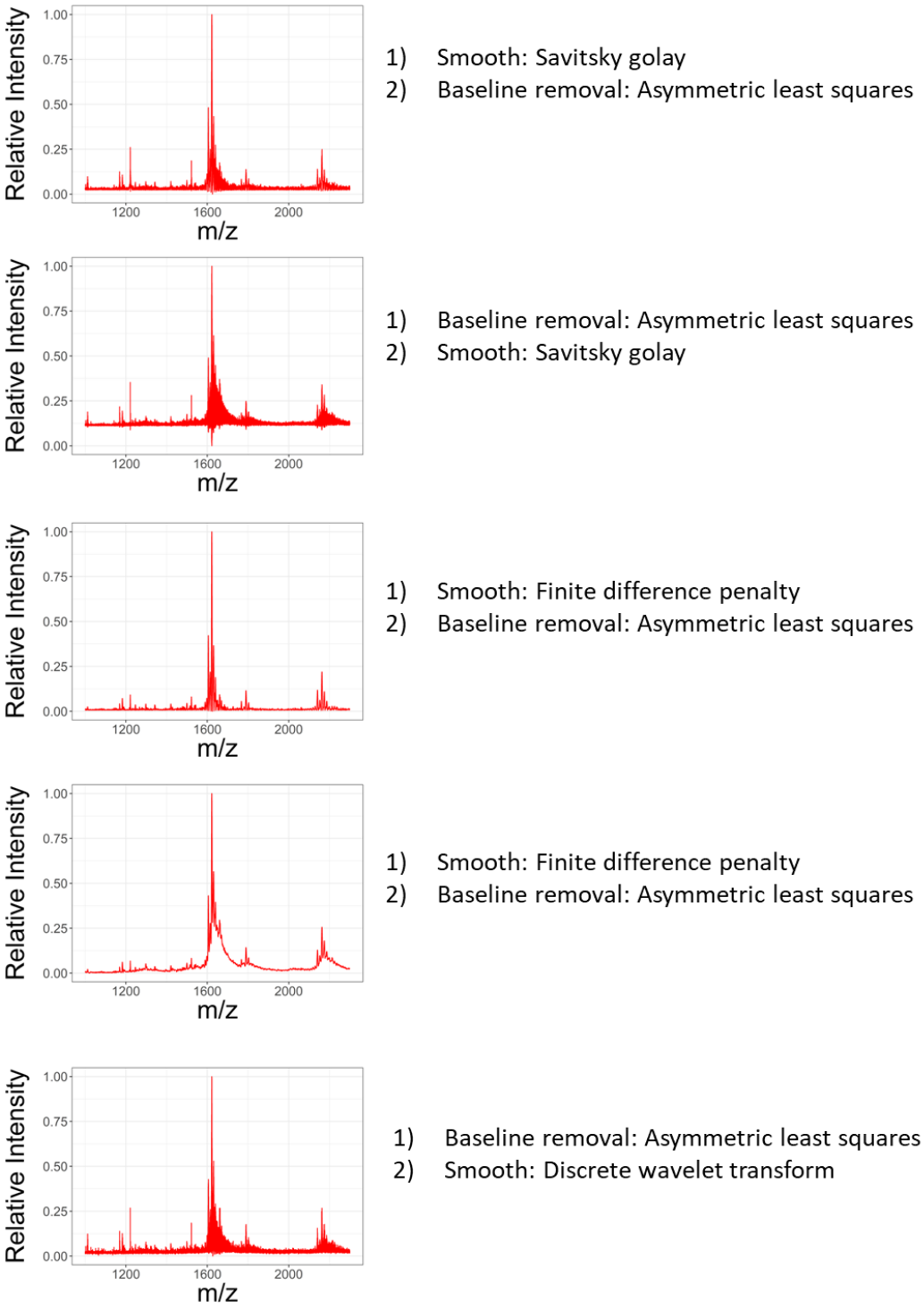

Figure S2. Comparison of combined algorithms for ESI-MS spectrum processing performed over apo-MT2.

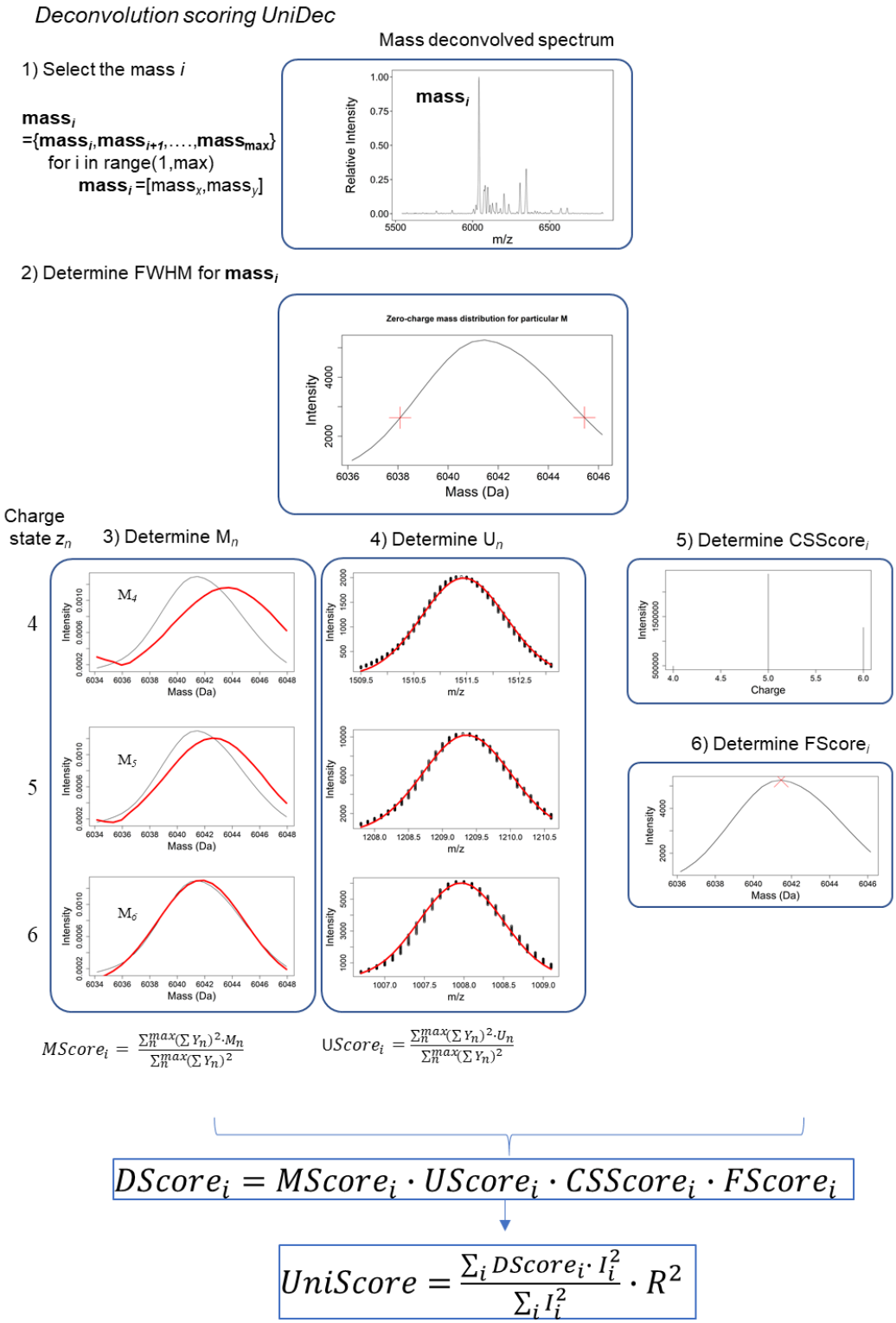

Figure S3. Deconvolution scoring scheme based on UniDec.[14]

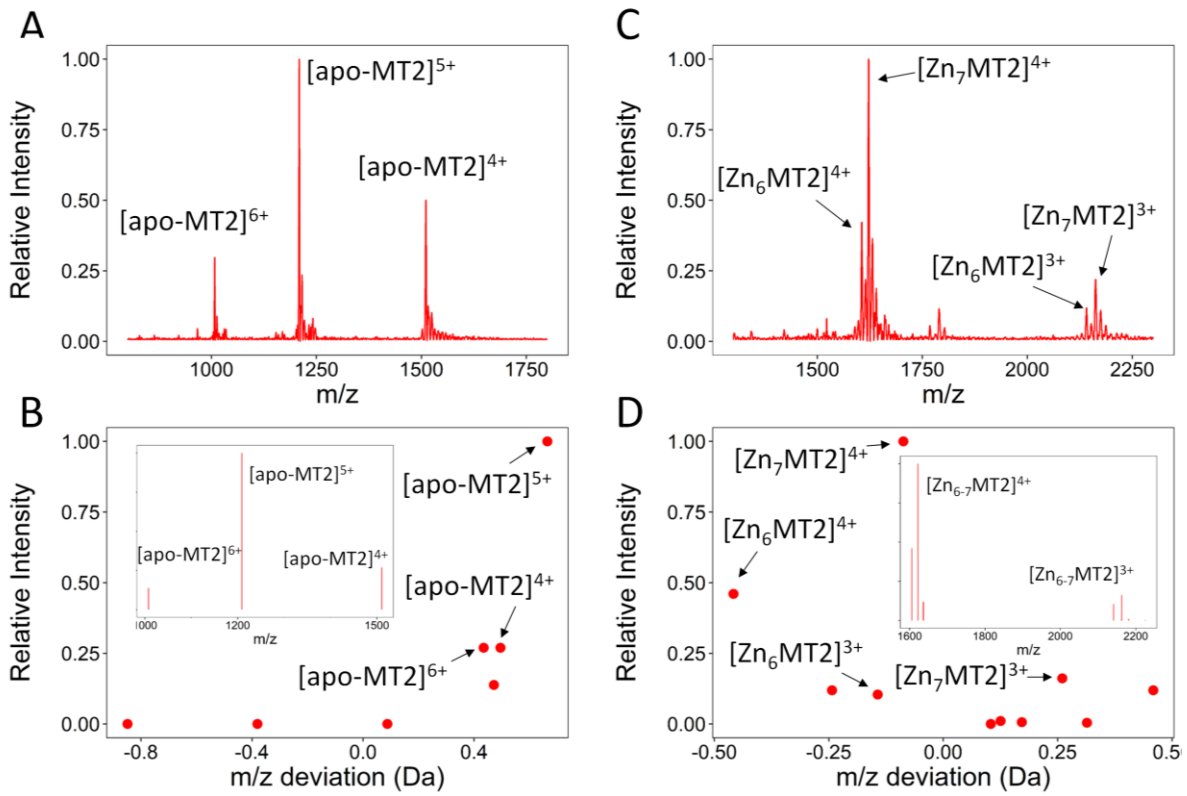

Figure S4. Analysis of ESI-MS spectra for apo-MT2 and Zn7MT2 acquired in native conditions (ammonium acetate 20 mM). A) ESI-MS spectrum for apo-MT2 displays the ions observed with charge state ranging from $[M+ nH]^{n+}$ where n = 4-6, with n = 5 the most abundant charge state. B) Error control plot that shows the mass accuracy deviation achieved for each of the peaks detected and filtered for apo-MT2. Visualization of the detected and filtered peaks by global peak intensity (10%) is shown in the inset. C) ESI-MS spectrum for 7 Zn^{2+} Eq added to apo-MT2 that displays Zn_7MT2 and Zn_6MT2 observed with charge state ranging from $[M+ nH]^{n+}$ where n = 3-4, with n = 4 the most abundant charge state. D) Error control plot that shows the mass accuracy deviation achieved for each of the peaks detected and filtered in the ESI-MS spectrum for 7 Zn^{2+} eq. added to apo-MT2. Visualization of the detected and filtered peaks by global peak intensity (10%) is shown in the inset.

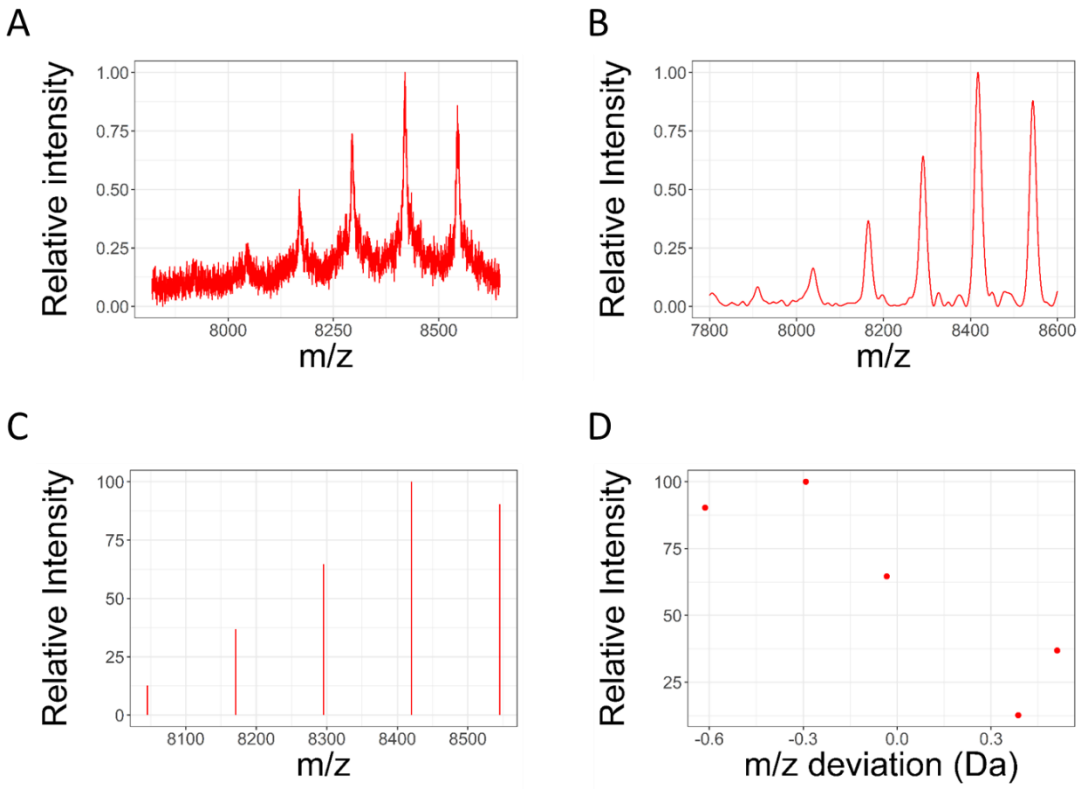

Figure S5. MALDI-TOF-MS analysis of NEM-labelled apo-MT2. A) Experimental MS spectrum acquired. B) Signal processed by baseline correction and smoothing by algorithm included in the R package. C) Visualization of the detected and filtered peaks from the MALDI-MS spectrum processed. D) Quality control plot that shows the mass accuracy deviation achieved for each of the peaks detected and filtered by global peak intensity (10%) for NEM-alkylated apo-MT2.

Table S5. Annotated MALDI-TOF-MS spectra of chemically labeled apo- and Zn7MT2 through the use of the R package. Results indicate the number of peaks detected with the corresponding mass deviation error, maximum and relative intensity. ET, IAA, NEM, and IAM refer to ethyl iodoacetate, iodoacetic acid, N-ethylmaleimide, and iodoacetamide, respectively.

Protein sample	Experimental m/z	Maximum intensity	Error (Da)	Number of modifications	Relative intensity (%)	Alkylating reagent
7 Zn^{2+} Eq	6472.7114	9374	-0.95	5	22	ET
7 Zn^{2+} Eq	6387.3096	20057	-0.26	4	47	ET
7 Zn^{2+} Eq	6301.2495	35931	-0.23	3	85	ET
7 Zn^{2+} Eq	6215.0581	42003	-0.34	2	100	ET
7 Zn^{2+} Eq	6129.0347	26933	-0.27	1	64	ET
7 Zn^{2+} Eq	6042.6943	6340	-0.52	0	14	ET
0 Zn^{2+} Eq	6815.375	14193	-2.64	9	16	ET
0 Zn^{2+} Eq	6730.147	31631	-1.78	8	36	ET
0 Zn^{2+} Eq	6644.644	57453	-1.2	7	67	ET
0 Zn^{2+} Eq	6558.9551	80243	-0.8	6	94	ET
0 Zn^{2+} Eq	6473.1675	85625	-0.49	5	100	ET
0 Zn^{2+} Eq	6387.2217	59228	-0.35	4	69	ET
0 Zn^{2+} Eq	6301.4204	29632	-0.06	3	34	ET
0 Zn^{2+} Eq	6215.415	9796	0.02	2	10	ET
7 Zn^{2+} Eq	6273.5405	3319	-1.82	4	17	IAA
7 Zn^{2+} Eq	6216.895	10468	-0.43	3	56	IAA
7 Zn^{2+} Eq	6158.9429	18045	-0.34	2	97	IAA
7 Zn^{2+} Eq	6101.2627	18551	0.01	1	100	IAA
7 Zn^{2+} Eq	6043.0093	8227	-0.21	0	43	IAA
0 Zn^{2+} Eq	6274.3398	22904	-1.02	4	21	IAA
0 Zn^{2+} Eq	6217.3774	61781	0.05	3	59	IAA
0 Zn^{2+} Eq	6159.4761	104533	0.19	2	100	IAA
0 Zn^{2+} Eq	6101.6362	101974	0.39	1	98	IAA
0 Zn^{2+} Eq	6043.791	44226	0.58	0	42	IAA
7 Zn^{2+} Eq	8544.7021	1324	-1.02	20	34	NEM
7 Zn^{2+} Eq	8420.2207	3529	-0.38	19	93	NEM
7 Zn^{2+} Eq	8295.5098	3802	0.04	18	100	NEM
7 Zn^{2+} Eq	8171.165	2518	0.82	17	66	NEM
7 Zn^{2+} Eq	8048.1655	1449	2.94	16	37	NEM

0 Zn^{2+} Eq	8545.1113	7097	-0.61	20	90	NEM
0 Zn^{2+} Eq	8420.3076	7848	-0.29	19	100	NEM
0 Zn^{2+} Eq	8295.4404	5100	-0.03	18	65	NEM
0 Zn^{2+} Eq	8170.8594	2945	0.51	17	37	NEM
0 Zn^{2+} Eq	8045.6094	1064	0.39	16	13	NEM
0 Zn^{2+} Eq	7185.0093	21682	0.77	20	70	IAM
0 Zn^{2+} Eq	7127.4541	30837	0.26	19	100	IAM
0 Zn^{2+} Eq	7070.6514	13928	0.51	18	45	IAM
0 Zn^{2+} Eq	7013.9287	3369	0.84	17	10	IAM
7 Zn^{2+} Eq	6443.2295	2696	0.66	7	19	IAM
7 Zn^{2+} Eq	6385.9780	3065	0.46	6	22	IAM
7 Zn^{2+} Eq	6329.6211	3517	1.15	5	26	IAM
7 Zn^{2+} Eq	6271.4678	6297	0.05	4	48	IAM
7 Zn^{2+} Eq	6216.0430	8567	1.67	3	66	IAM
7 Zn^{2+} Eq	6158.9063	12903	1.59	2	100	IAM
7 Zn^{2+} Eq	6101.5469	11782	1.28	1	91	IAM
7 Zn^{2+} Eq	6044.2485	6413	1.03	0	49	IAM

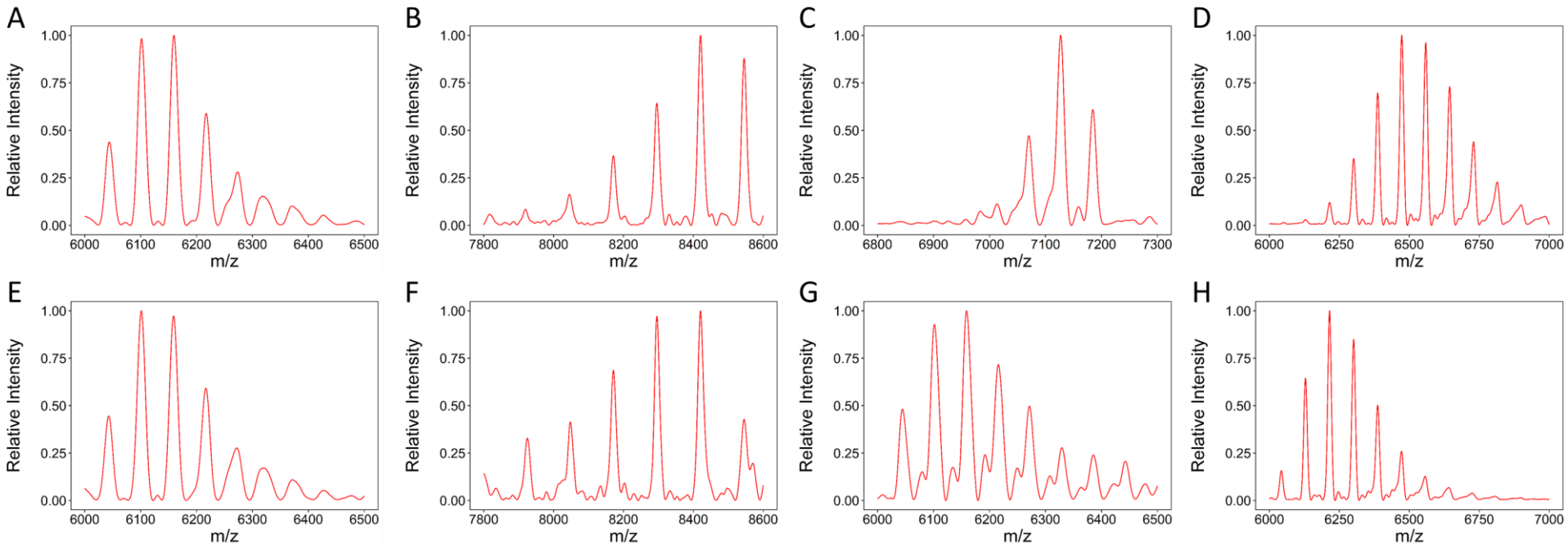

Figure S6. Processed MALDI-TOF-MS spectrum for apo-MT2 and Zn7MT2 chemically labeled by set of alkylation reagents. A) apo-MT2 labeled by IAA. B) apo-MT2 labeled by NEM. C) apo-MT2 labeled by IAM. D) apo-MT2 labeled by ethyl iodoacetate. E) Zn_7MT2 labeled by IAA. F) Zn_7MT2 labeled by NEM. G) Zn_7MT2 labeled by IAM. H) Zn_7MT2 labeled by ethyl iodoacetate.

Table S6. Annotated peptide-mass fingerprint for IAM-labeled apo-MT2. Start and stop denote the amino acid positions for the peptide found.

Experimental m/z	Maximum intensity	Error (Da)	Number of modifications	Relative intensity	Sequence	Start	Stop
2209.53	294	0.77	5	6	MDPNCSCAAGDSCTCAGSCK	1	20
2152.41	54	0.68	4	1	MDPNCSCAAGDSCTCAGSCK	1	20
2093.30	55	-1.41	3	1	MDPNCSCAAGDSCTCAGSCK	1	20
1445.72	1691	0.21	5	33	SCCSCCPVGCAK	32	43
1388.74	163	0.26	4	3	SCCSCCPVGCAK	32	43
1275.28	274	0.84	2	5	SCCSCCPVGCAK	32	43
996.60	5085	0.19	3	100	CAQGCICK	44	51
939.57	438	0.18	2	9	CAQGCICK	44	51
882.52	115	0.16	1	2	CAQGCICK	44	51
2212.35	85	0.58	1	2	MDPNCSCAAGDSCTCAGSCKCK	1	20
2152.41	54	-1.33	0	1	MDPNCSCAAGDSCTCAGSCKCK	1	20
1073.70	78	1.28	3	2	ECKCTSCK	23	30
1015.67	53	0.26	2	1	ECKCTSCK	23	30
958.62	57	0.24	1	1	ECKCTSCK	23	30
898.51	58	-1.84	0	1	ECKCTSCK	23	30
901.55	51	1.20	0	1	ECKCTSCK	23	30
1573.85	177	0.24	5	3	KSCCSCCPVGCAK	32	43
1574.91	109	1.31	5	2	KSCCSCCPVGCAK	32	43

1457.85	59	-1.71	3	1	KSCCSCCPVGCAK	32	43
1402.38	99	-0.16	2	2	KSCCSCCPVGCAK	32	43
1343.58	123	-1.93	1	2	KSCCSCCPVGCAK	32	43
2423.12	74	0.22	8	1	SCCSCCPVGCAKCAQGCICK	33	51
1339.14	74	-1.43	1	1	CAQGCICKGASDK	44	56
1115.62	117	0.22	3	2	GASDKCSCCA	52	61
1058.64	76	0.27	2	1	GASDKCSCCA	52	61
999.64	368	-1.71	1	7	GASDKCSCCA	52	61
1000.62	124	-0.73	1	2	GASDKCSCCA	52	61
943.56	67	0.24	0	1	GASDKCSCCA	52	61
2801.25	73	1.25	5	1	MDPNCSCAAGDSCTCAGSCKCKECK	1	25
2685.50	359	-0.47	3	7	MDPNCSCAAGDSCTCAGSCKCKECK	1	25
1142.71	54	-0.78	2	1	ECKCTSCKK	23	30
2209.53	294	-0.32	7	6	CTSCKKSCCSCCPVGCAK	26	43
2152.41	54	-0.41	6	1	CTSCKKSCCSCCPVGCAK	26	43
2550.74	268	-0.25	8	5	KSCCSCCPVGCAKCAQGCICK	31	51
2209.53	294	0.66	2	6	KSCCSCCPVGCAKCAQGCICK	31	51
2152.41	54	0.57	1	1	KSCCSCCPVGCAKCAQGCICK	31	51
2093.30	55	-0.51	0	1	KSCCSCCPVGCAKCAQGCICK	31	51
2423.12	74	-0.81	0	1	SCCSCCPVGCAKCAQGCICKGASDK	32	56
2424.16	141	0.23	0	3	SCCSCCPVGCAKCAQGCICKGASDK	32	56

2093.30	55	0.52	6	1	CAQGCICKGASDKCSCCA	44	61
2673.79	145	-0.25	1	3	CTSCKKSCCSCCPVGCAKCAQGCICK	26	51
2551.73	294	-0.29	0	6	KSCCSCCPVGCAKCAQGCICKGASDK	31	56

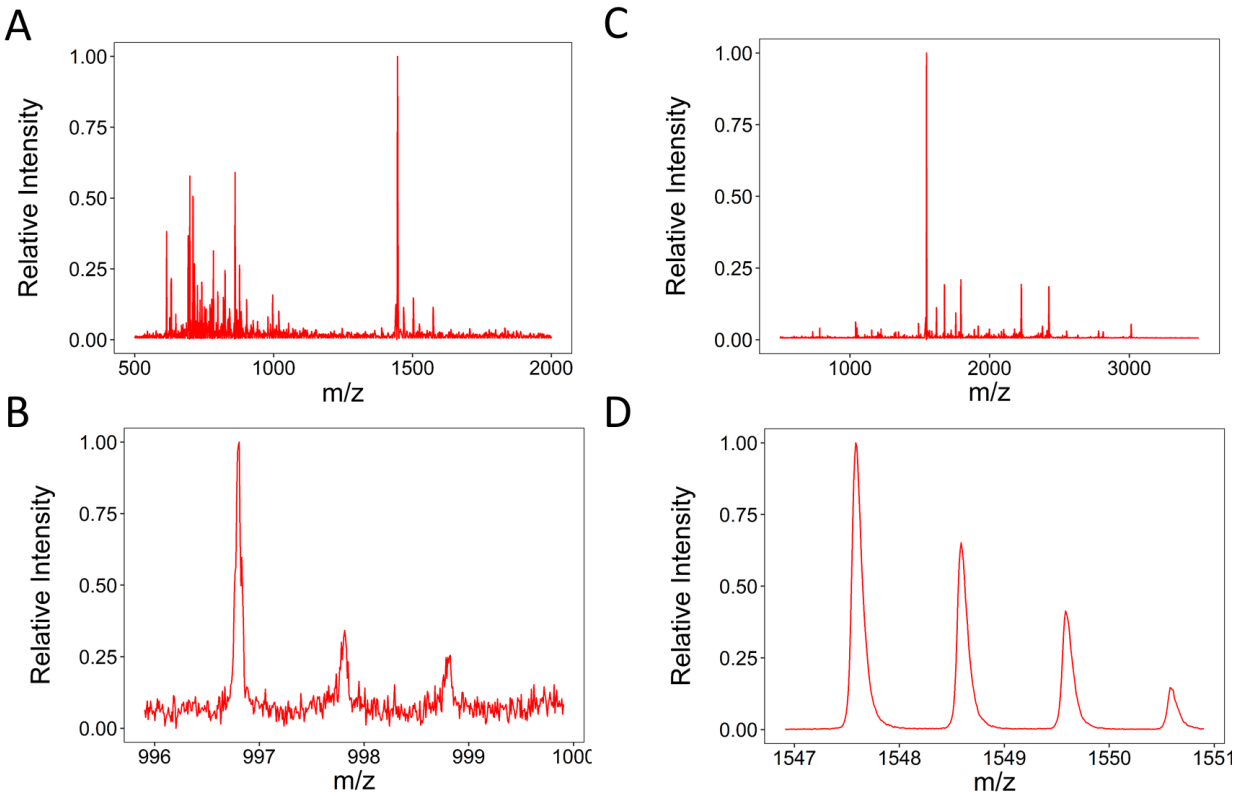

Figure S7. Peptide-mass fingerprint (PMF) of chemically labeled apo-MT2 and apo-MT3 obtained by MALDI-TOF-MS. A) Signal processed (smoothed and baseline corrected) for the MALDI-MS spectrum for the IAM-labelled apo-MT2 trypsin digested. B) Annotated peptide CAQGCICK harboring 3 IAM modifications. C) Signal processed (smoothed and baseline corrected) for the MALDI-MS spectrum for the IAM-labeled apo-MT3 trypsin digested. D) Annotated peptide SCCSCCPAECEK with 5 Cys-IAM labeled residues. We observed how the software identified peaks that are isotopically well resolved but also detected true peaks with low signals.

Table S7. Scores obtained for annotated peptide-mass fingerprint for apo-MT2 and apo-MT3.

Protein	Sequence coverage (%)	Hit ratio	Mass coverage (kDa)	ELDP	PMF score ± sd[a]
Zn_0MT2	100	0.50	6.042	9	56.45 ± 0.69
Zn_4MT2	100	1.27	6.042	-34	133.32 ± 0.95
Zn_5MT2	67.21	0.43	4.061	8	47.03 ± 0.62
Zn_6MT2	67.21	0.14	4.06	-3	18.11 ± 0.97
Zn_7MT2	67.21	0.50	4.06	4	53.65 ± 0.41
Zn_0MT3	100.00	0.19	6.927	-13	26.09 ± 0.28
Zn_4MT3	61.76	0.13	4.28	-2	17.61 ± 0.60
Zn_5MT3	47.06	0.16	3.26	-6	19.93 ± 0.45
Zn_6MT3	69.12	0.13	4.78	1	17.29 ± 0.70
Zn_7MT3	45.59	0.09	3.16	1	12.32 ± 0.43

[a]Sd stands for the standard deviation and it was computed by 1000 bootstrapping.

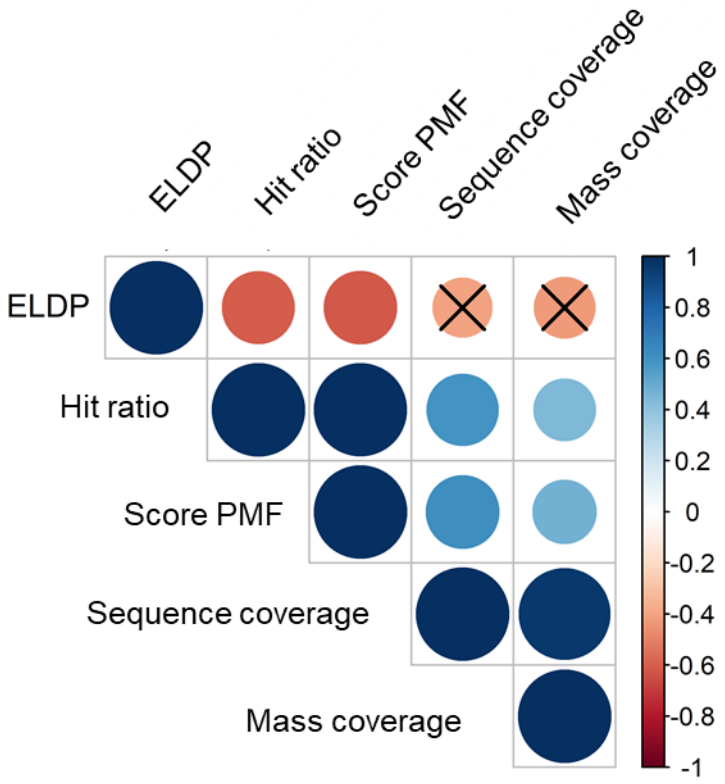

Figure S8. Correlogram for the PMF score, hit ratio, sequence coverage, and mass coverage for the ten datasets analyzed (Zn0MT2…Zn7MT2, Zn0MT3… Zn7MT2) with MALDI-MS. The correlation coefficient R was calculated based on Pearson's product-moment correlation coefficient. An exact *p*-value was computed with a two-sided alternative hypothesis. The crossed circles indicate no-significant correlations (significance level = 0.01).

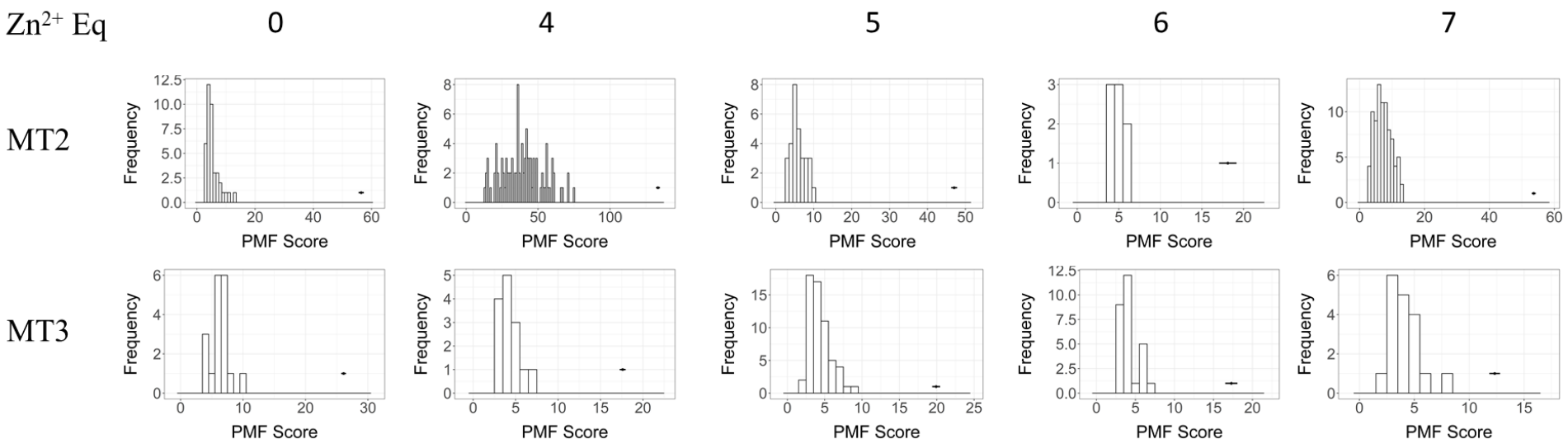

Figure S9. Permutation tests to calculate the p-value from the null distributions for the PMF scores obtained for the decoy list scored against the experimental MALDI-MS spectrum for Zn0MT2…Zn7MT2 and Zn0MT3…Zn7MT3 proteins. The true PMF score obtained for the target protein is indicated at the frequency value of 1 with a point and a segment line that represents its confidence interval determine by bootstrapping.

Table S8. Peptide-spectrum matches (PSMs) for the dataset of MS/MS spectra collected for enzymatically digested Zn0-7MT3 proteins.

Sequence	Mean error (Da)	Number of ions matched	Peaks matched (%)	Score	Parent ion	Label
GGEAAEAEAEKCSCCQ	-1.16	1	2	2	1585	target
CGCEAPSECCAKCKCDSK	-1.99	1	1	2	1862	decoy
CTSXKKSCCSCCPAECEK	-2.00	1	1	2	1968	target
DXVXKGGEAAEAEAEKXSXXQ	-1.47	1	1	2	2419	target
XGXEAPSECCAKCK	0.22	2	4	5	1547	decoy
XTSCKK	-1.71	2	10	5	725	target
CKCDSKAKEECCSK	-0.17	4	8	12	1561	decoy
SXXSXXPAEXEK	-0.17	5	11	17	1547	target
KSXXSCCPAECEKXAK	-0.88	7	12	24	1862	target
KSCCSXXPAEXEK	1.18	7	15	25	1561	target
CGCEAPSECCAKXK	0.45	9	17	33	1490	decoy
XTSXKK	0.47	9	45	37	783	target
XTSCKKSCCSXCPAECEKCAK	0.24	10	13	37	2328	target
KSXCSCXPAECEK	0.40	11	23	43	1504	target
XKXDSKAKEECCSK	-0.28	12	23	48	1675	decoy
GGEAAEAEAEK	-0.13	14	35	57	1061	target
MDPETXPXPSGGSXTXADSXK	-0.26	15	19	62	2377	target
XAKDXVXK	-0.03	14	50	63	1040	target

CKCDSKAKEEXCSK	-0.08	17	33	74	1617	decoy
CGCEAPSECCAKCK	0.00	18	35	77	1432	decoy
SXXSXXPAECEK	-0.43	18	41	82	1490	target
CEGCKCTSCKK	-0.13	20	50	91	1188	target
SXXSCCPAEXEK	-0.09	20	45	93	1432	target
KSCXSXXPAEXEK	-0.98	20	42	93	1617	target
KSXXSXXPAEXEK	-0.30	20	42	94	1675	target
GGEAAEAEAEKXSXXQ	-0.12	22	37	100	1756	target
SXXSCCPAECEK	0.07	22	50	104	1375	target
KSXXSXXPAEXEK	0.08	22	46	106	1675	target
KSXXSCCPAECEK	0.11	23	48	109	1504	target
GGEAAEAEAEKXSXXQ	0.01	25	42	118	1756	target
XKCDSKAKEECXSK	-0.27	25	48	121	1675	decoy
MDPETXPXPSGGSXTXADSXKXEGXK	-0.13	26	26	122	3012	target
DXVXKGGEAAEAEAEK	0.25	30	50	148	1724	target
SXXSCCPAECEKCAKDCVCK	0.00	38	50	195	2225	target
SXXSCCPAECEKCAKDCVCK	-0.07	38	50	196	2225	target
DXVXKGGEAAEAEAEKCSCCQ	-0.12	40	50	207	2247	target
DXVXKGGEAAEAEAEKXSCCQ	-0.28	40	50	208	2304	target
XEGXKXTSCKKSCCSCCPAECEK	-0.23	44	50	234	2605	target
XAKDXVXKGGEAAEAEAEKCSCCQ	-0.17	46	50	245	2606	target

XAKDXVXKGGEAAEAEAEKXSCCQ	0.07	46	50	246	2664	target
XAKDXVXKGGEAAEAEAEKXSXCQ	0.01	46	50	247	2720	target
XAKDXVXKGGEAAEAEAEKXSXCQ	-0.25	46	50	247	2720	target
XAKDXVXKGGEAAEAEAEKXSXXQ	0.02	46	50	248	2778	target
MDPETXPXPSGGSXTXADSCKCEGCK	0.26	50	50	270	2840	target

Table S9. Deconvolution and assignment of the bottom-up MS/MS spectra with MetaOdysseus. The identifications accounted for false positives filtering by false discovery rate threshold at 1 % estimated with a target-decoy approach.

Sequence	Mean Error (Da)	N. ions matched	Peaks matched (%)	Score	Parent ion	Label
DXVXKGGEAAEAEAEK	0.25	30	50	148	1724	target
SXXSCCPAECEKCAKDCVCK	0.00	38	50	195	2225	target
SXXSCCPAECEKCAKDCVCK	-0.07	38	50	196	2225	target
DXVXKGGEAAEAEAEKCSCCQ	-0.12	40	50	207	2247	target
DXVXKGGEAAEAEAEKXSCCQ	-0.28	40	50	208	2304	target
XEGXKXTSCKKSCCSCCPAECEK	-0.23	44	50	234	2605	target
XAKDXVXKGGEAAEAEAEKCSCCQ	-0.17	46	50	245	2606	target
XAKDXVXKGGEAAEAEAEKXSCCQ	0.07	46	50	246	2664	target
XAKDXVXKGGEAAEAEAEKXSXCQ	0.01	46	50	247	2720	target
XAKDXVXKGGEAAEAEAEKXSXCQ	-0.25	46	50	247	2720	target
XAKDXVXKGGEAAEAEAEKXSXXQ	0.02	46	50	248	2778	target
MDPETXPXPSGGSXTXADSCKCEGCK	0.26	50	50	270	2840	target

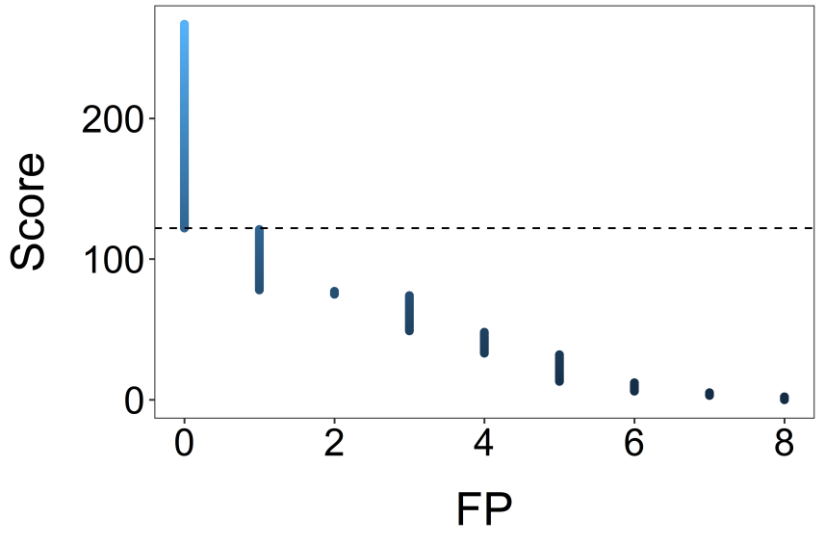

Figure S10. Evaluation of the relationship between the false positives (FP) and the Score achieved for the peptide-spectrum matches computes as a peak match probability scored based on the probabilities obtained from a binomial distribution. The horizontal dashed line indicates the false discovery rate (FDR) threshold of 1 %.

Table S10. Peptide-spectrum matches (PSM) obtained with MS-GF+ for the set of enzymatically digested Zn0-7MT3 proteins. In bold are marked those features that were identified with a 0.1 false discovery rate. The output (ms-gf:denovo score, evalue, rawscore, specevalue, or isotope error) were obtained directly from the MS-GF+ software and they are explained elsewhere.[20] In the brackets is indicated the position of the modification. The mass of the modification is indicated right before the bracket (e.g. 57 stands for IAM).

Calculated (m/z)	Experimental (m/z)	ms-gf: denovo score	ms-gf: evalue	ms-gf: rawscore	ms-gf: specevalue	Isotope error	End	Start	Sequence	Modification
1586.64	1585.00	11	8.5E-02	-36	1.2E-03	-2	52	41	ECEKCAKDCVCK	57 (2), 57 (5), 57 (9), 57 (11)
2420.82	2418.91	66	3.8E-01	-22	5.6E-03	-2	23	2	DPETCPCPSGGSCTCADSCKCE	57 (7), 57 (13), 57 (15), 57 (21)
2417.84	2418.91	66	3.8E-01	-22	5.6E-03	1	31	8	CPSGGSCTCADSCKCEGCKCTSCK	57 (15)
2418.81	2418.91	66	3.8E-01	-22	5.6E-03	0	43	22	CEGCKCTSCKKSCCSCCPAECE	57 (1), 57 (4)
1568.65	1567.00	47	1.3E+00	-49	1.9E-02	-2	59	45	CAKDCVCKGGEAAEA	57 (1), 57 (5)
1756.66	1756.00	144	2.8E-06	39	4.1E-08	-1	68	53	GGEAAEAEAEKCSCCQ	57 (12), 57 (14), 57 (15)
1723.73	1724.00	46	1.7E-02	-19	2.5E-04	0	63	48	DCVCKGGEAAEAEAEK	57 (2), 57 (4)
1040.43	1040.00	22	1.0E-04	-5	1.5E-06	0	52	45	CAKDCVCK	57 (1), 57 (5), 57 (7)
1547.51	1547.00	83	5.0E-08	38	7.3E-10	-1	44	33	SCCSCCPAECEK	57 (2), 57 (3), 57 (5), 57 (6), 57 (10)

1675.60	1675.00	55	5.1E-06	9	7.4E-08	-1	44	32	KSCCSCCPAECEK	57 (3), 57 (4), 57 (6), 57 (7), 57 (11)
1061.47	1061.00	74	1.5E-10	66	2.3E-12	0	63	53	GGEAAEAEAEK	NA
2328.87	2328.00	58	2.4E-02	-11	3.6E-04	-1	47	27	CTSCKKSCCSCCPAECEKCAK	57 (4), 57 (12)
2328.87	2328.00	58	2.4E-02	-11	3.6E-04	-1	47	27	CTSCKKSCCSCCPAECEKCAK	57 (4), 57 (16)
1863.70	1862.00	46	1.7E-03	-9	2.4E-05	-2	47	32	KSCCSCCPAECEKCAK	57 (7), 57 (11), 57 (14)
2599.94	2600.00	59	2.7E-01	-26	4.0E-03	0	31	9	PSGGSCTCADSCKCEGCKCTSCK	57 (6), 57 (8), 57 (14), 57 (17), 57 (19), 57 (22)
1969.71	1968.00	45	1.5E-02	-7	2.2E-04	-2	44	27	CTSCKKSCCSCCPAECEK	57 (1)
1561.56	1561.00	68	5.9E-03	-13	8.7E-05	-1	44	32	KSCCSCCPAECEK	57 (6), 57 (7), 57 (11)
2376.83	2377.00	122	8.0E-17	91	1.2E-18	0	21	1	MDPETCPCPSGGSCTCADSCK	57 (6), 57 (8), 57 (14), 57 (16), 57 (20)
1188.92	1188.00	17	3.8E-03	-27	5.6E-05	-2	21	1	MDPETCPCPSGGSCTCADSCK	57 (6), 57 (8), 57 (14), 57 (16), 57 (20)
783.35	783.35	1	3.4E-05	-7	4.9E-07	0	32	27	CTSCKK	57 (1), 57 (4)
1504.54	1504.00	30	3.5E-02	-24	5.2E-04	-1	44	32	KSCCSCCPAECEK	57 (4), 57 (11)
1433.46	1432.00	12	3.6E-04	-19	5.2E-06	-1	44	33	SCCSCCPAECEK	57 (3), 57 (5), 57 (10)
1490.48	1490.00	14	4.7E-03	-25	6.9E-05	0	44	33	SCCSCCPAECEK	57 (2), 57 (3), 57 (5), 57 (6)

2778.07	2778.00	65	1.5E-01	-23	2.2E-03	0	68	45	CAKDCVCKGGEAAEAEAEKCSCCQ	57 (1), 57 (5), 57 (7), 57 (20), 57 (22), 57 (23)
1618.58	1617.00	19	6.0E-04	-14	8.8E-06	-2	44	32	KSCCSCCPAECEK	57 (3), 57 (6), 57 (7), 57 (11)
2226.79	2225.00	63	4.2E-02	-6	6.2E-04	-2	52	33	SCCSCCPAECEKCAKDCVCK	57 (2), 57 (13)
2226.79	2225.00	63	4.2E-02	-6	6.2E-04	-2	52	33	SCCSCCPAECEKCAKDCVCK	57 (2), 57 (5)
2105.84	2106.00	65	5.1E-02	-8	7.5E-04	0	47	31	KKSCCSCCPAECEKCAK	57 (4), 57 (5), 57 (7), 57 (12), 57 (15)
1376.44	1375.00	25	5.3E-03	-19	7.8E-05	-1	44	33	SCCSCCPAECEK	57 (3), 57 (10)
2245.79	2247.00	66	2.4E-02	1	3.6E-04	1	21	2	DPETCPCPSGGSCTCADSCK	57 (5), 57 (7), 57 (13), 57 (15), 57 (19)
2664.03	2663.00	75	1.8E-01	-18	2.7E-03	-1	68	45	CAKDCVCKGGEAAEAEAEKCSCCQ	57 (1), 57 (5), 57 (20), 57 (22)
2304.87	2304.00	75	7.7E-01	-18	1.1E-02	-1	68	48	DCVCKGGEAAEAEAEKCSCCQ	57 (17), 57 (19), 57 (20)
2606.91	2606.00	66	7.8E-01	-24	1.1E-02	-1	46	22	CEGCKCTSCKKSCCSCCPAECEKCA	NA

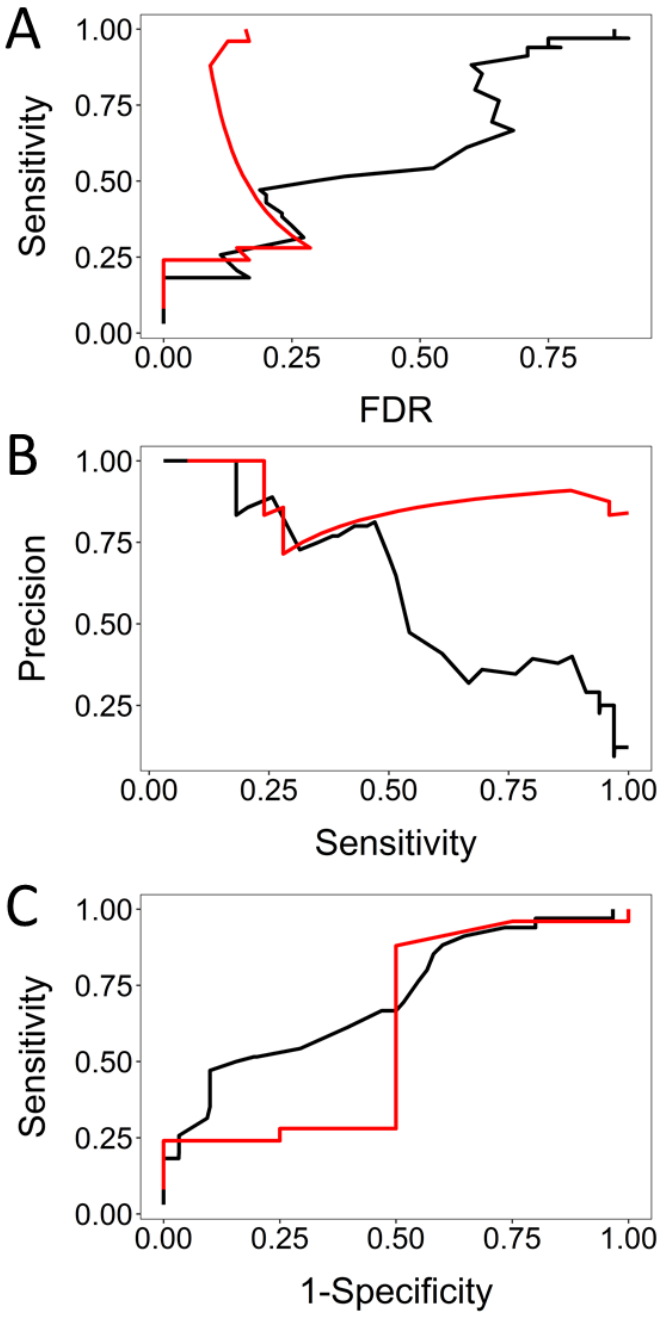

Figure S11. Comparison between MetaOdysseus (red line) and MS-GF+ (black line) for the peptide-spectrum matches (PSMs) results obtained in terms of their sensitivity, specificity, and precision. A) Analysis of the sensitivity achieved at different FDR cut-offs obtained. B) Precision-recall curves for the different FDR cut-offs. C) Receiver-operating characteristic (ROC) analysis at different FDR cut-offs.

Table S11. Mass deconvolution for the native top-down MS/MS obtained with MetaOdysseus using the peak assignment algorithm. The class indicates if the mass was as a true positive (TP) or false positive (FP). The calculation of score and the Total score *S* is explained in the page S6.

Class	Acquisition[a]	Monoisotopic mass (m/z)	Score	Intensity	Charge	Neutral mass (Da)	Total score *S*
TP	DDA	1602.85	6	567	1	1601.84	22
TP	DDA	1624.83	5	215	1	1623.82	10
TP	DDA	1612.87	5	139	4	6447.45	49
FP	DDA	1605.63	5	138	2	3209.24	10
TP	DIA	1624.80	11	44278	1	1623.79	20
TP	DIA	1602.83	11	42568	1	1601.82	29
TP	DIA	1028.55	9	7462	1	1027.54	16
TP	DIA	929.49	9	5784	1	928.48	8
TP	DIA	1252.60	9	5436	1	1251.59	15
TP	DIA	1108.38	9	4956	1	1107.37	16
TP	DIA	994.52	8	4722	1	993.51	8
TP	DIA	1365.67	8	3150	1	1364.66	8
FP	DIA	1628.80	8	3108	1	1627.79	7
TP	DIA	911.47	8	2980	1	910.46	8
TP	DIA	1542.42	8	2498	1	1541.41	8
TP	DIA	1230.62	8	2476	1	1229.61	7
TP	DIA	1646.77	8	1882	1	1645.76	7
TP	DIA	1760.79	7	1438	1	1759.78	14
TP	DIA	1738.82	7	1324	1	1737.81	14

[a]DDA and DIA refer to data-dependent and data-independent acquisition, respectively.

Table S12. Mass deconvolution for the native top-down MS/MS obtained with MASH Suite Pro using the eTRASH algorithm.

Class	Acquisition[a]	Monoisotopic mass (m/z)	Abundance	MonoMZ	Most abundant m/z	Charge	Score	Most abundant mass
TP	DDA	1602.83	566	1603.84	1603.84	1	0.84	1602.83
TP	DDA	1623.82	276	1624.83	1624.83	1	0.79	1623.82
FP	DDA	3192.75	216	1597.38	1597.88	2	0.66	3193.76
TP	DIA	1623.79	44278	1624.80	1624.80	1	0.99	1623.79
TP	DIA	1601.82	42568	1602.83	1602.83	1	0.99	1601.82
TP	DIA	1027.54	7462	1028.55	1028.55	1	0.99	1027.54
TP	DIA	1140.62	6452	1141.62	1141.62	1	0.97	1140.62
TP	DIA	928.48	5784	929.48	929.48	1	0.99	928.48
TP	DIA	1251.59	5436	1252.60	1252.60	1	0.99	1251.59
TP	DIA	1085.40	5426	1086.40	1086.40	1	0.95	1085.40
TP	DIA	1107.37	4956	1108.38	1108.38	1	0.95	1107.37
TP	DIA	993.51	4722	994.52	994.52	1	0.99	993.51
FP	DIA	1252.61	3246	1253.61	1253.61	1	0.68	1252.61
TP	DIA	1364.67	3150	1365.68	1365.68	1	0.98	1364.67
FP	DIA	1627.80	3108	1628.81	1628.81	1	0.83	1627.80
TP	DIA	910.46	2980	911.47	911.47	1	0.91	910.46
FP	DIA	1605.81	2728	1606.82	1606.81	1	0.91	1605.81
TP	DIA	1541.42	2498	1542.43	1542.43	1	0.94	1541.42

FP	DIA	1365.70	2262	1366.70	1366.70	1	0.68	1365.70

[a]DDA and DIA refers to data-dependent and data-independent acquisition, respectively.

Table S13. Mass assignment for the deconvolved masses from the native top-down MS/MS obtained with MetaOdysseus.

Acquisition[a]	Experimental (m/z)	Theoretical (m/z)	Intensity (%)	MS^2 sequence	MS^2 ion type	No. metal	Error (Da)	Score
DDA	1624.83	1623.43	23	MDPNCSCATGGSCTCAG	$[b17]^{1+}$	1	1.39	9.93
DDA	1605.63	1606.46	6	KSCCSCCPVGCAKCAQGCVCKGASEKCSCCA	$[y31]^{2+}$	3	-0.83	9.71
DIA	1624.79	1623.43	100	MDPNCSCATGGSCTCAG	$[b17]^{1+}$	1	1.36	20.31
DIA	1252.59	1252.23	10	MDPNCSCATGGS	$[b12]^{1+}$	2	0.36	15.27
DIA	1108.37	1108.18	8	MDPNCSCATG	$[b10]^{1+}$	2	0.19	7.61
DIA	1230.61	1229.13	3	MDPNCSCATGG	$[b11]^{1+}$	3	1.48	7.49
DIA	1542.42	1544.22	3	VCKGASEKCSCCA	$[y13]^{1+}$	4	-1.80	7.58
DIA	1646.76	1648.07	1	MDPNCSCATGGSCT	$[b14]^{1+}$	5	-1.31	7.40

[a]DDA and DIA refers to data-dependent and data-independent acquisition, respectively.

References

1. Gault, J.; Liko, I.; Landreh, M.; Shutin, D.; Bolla, J. R.; Jefferies, D.; Agasid, M.; Yen, H.Y.; Ladds, M. J. G. W.; Lane, D. P.; Khalid, S.; Mullen, C.; Remes, P.M.; Huguet, R.; McAlister, G.; Goodwin, M.; Viner, R.; Syka, J. E. P.; Robinson, C.V. *Nat. Methods,* **2020**, *17*, 505–508.
2. Drozd, A.; Wojewska, D.; Peris–Díaz, M. D.; Jakimowicz, P.; Krężel, A. Crosstalk of the structural and zinc buffering properties of mammalian metallothionein–2. *Metallomics* **2018**, 10, 595–613-
3. Krężel, A.; Maret, W. Dual nanomolar and picomolar Zn(II) binding properties of metallothionein. *J. Am. Chem. Soc.* **2007**, *129*, 10911–10921.
4. Eyer, P.; Worek, F.; Kiderlen, D.; Sinko, G.; Stuglin, A.; Simeon-Rudolf, V.; Reiner, E. Molar absorption coefficients for the reduced Ellman reagent: reassessment. *Anal. Biochem.* **2003**, *312*, 224–227.
5. Krężel, A.; Maret, W. Thionein/metallothionein control Zn(II) availability and the activity of enzymes. *J. Biol. Inorg. Chem.* **2008**, *13*, 401–409.
6. Wright, P. C.; Evans, C. A.; Davlyatova, L.; Couto, N. Application of the broadband collision-induced dissociation (bbCID) mass spectrometry approach for protein glycosylation and phosphorylation analysis. *Rapid Commun. Mass Spectrom.* **2017**, *32*, 75-85.
7. Satten, G. A., Datta, S., Moura, H., Woolfitt, A. R.; da G. Carvalho, M.; Carlone, G. M.; De, B.K.; Pavlopoulos, A.; Barr, J. R. Standardization and denoising algorithms for mass spectra to classify whole–organism bacterial specimens. *Bioinformatics* **2004**, *20*, 3128–3136.
8. Bloemberg, T. G.; Gerretzen, J.; Wouters, H. J. P.; Gloerich, J.; van Deal, M.; Wessels, H.; Eilers, P. H. C.; van den Heuvel, L.; Buydens, L.; Werhrens, R. Improved parametric time warping for proteomics. *Chemom. Intell. Lab. Syst.* **2010**, *104*, 65–74.
9. Du, P.; Kibbe, W. A.; Lin, S. M. Improved peak detection in mass spectrum by incorporating continuous wavelet transform–based pattern matching. *Bioinformatics* **2006**, *22*, 2059–2065.
10. Wickham, H. ggplot2: Elegant graphics for data analysis. Springer–Verlag New York, **2016**.
11. Dodder, N. OrgMassSpecR: Organic Mass Spectrometry, https://CRAN.R-project.org/package=OrgMassSpecR, **2017**.
12. Morgner, N.; Robinson, C. V. Massign: An assignment strategy for maximizing information from the mass spectra of heterogeneous protein sssemblies. *Anal. Chem.* **2012**, *84*, 2939−2948.
13. Zhang, Z.; Marshall, A. G. A Universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass. Spectrom.* **1998**, 9, 225−233.
14. Marty, M. T. A Universal score for deconvolution of intact protein and native electrospray mass spectra. *Anal. Chem.* **2020**, *92*, 4395−4401
15. Stead, D. A.; Preece, A.; Brown, A. J. P. Universal metrics for quality assessment of protein identifications by mass spectrometry. *Mol. Cell. Proteomics.* **2006**, 5, 1205–1211.
16. Elias, J. E.; Gygi, S. P. Target-decoy search strategy for mass spectrometry-based proteomics, *Methods Mol. Biol.* **2010**, *604*, 55–71.
17. Bandeira, N., Kim, S., Jeong, K., False discovery rates in spectral identification. *BMC Bioinformatics* **2012**, *13(Suppl 16)*: S2.
18. Paton, N.W.; Hubbard, S.J.; Siepen, J.A.; Jones, A.R. Improving sensitivity in proteome studies by analysis of false discovery rates for multiple search engines. *Proteomics*, **2009**, *9*, 1220–1229.
19. Cai, W. X.; Guner, H.; Gregorich, Z. R.; Chen, A. J.; Ayaz-Guner, S.; Peng, Y.; Valeja, S.G.; Liu, X. W.; Ge, Y. MASH suite Pro: A comprehensive software tool for top-down proteomics. *Mol. Cell Proteomics* **2016**, *15*, 703–714.

20. Kim, S.; Pevzner, P. A. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat. Commun*. **2014**, 5: 5277.