Electronic Supplementary Material (ESI) for Metallomics.

Supplementary Information

A mass spectrometry-based approach gives new insight to organotin-protein interactions

Jonas M. Will[a], Catharina Erbacher[a], Michael Sperling[a], and Uwe Karst[a,*]

Affiliations:

[a] J. M. Will, C. Erbacher, Dr. M. Sperling, Prof. Dr. U. Karst*

Institute of Inorganic and Analytical Chemistry

University of Muenster

Corrensstr. 28/30, 48149 Münster (Germany)

Corresponding author e-mail: uk@uni-muenster.de

Table of contents

Experimental details

Enzymatic digestion of LGA-Ph_3Sn adducts

Stock solutions of trypsin and chymotrypsin (1 mg/mL) were prepared in ABC buffer (50 mM, pH 8.0). The digestion experiments were carried out in quadruplicate. First, LGA (54 µM) was incubated with Ph_3Sn (272 µM) in PBS at 37 °C overnight. As the excess Ph_3Sn in the incubation solution could bind to the peptides after enzymatic digestion and therefore give false-positive results, it had to be removed before the addition of the respective enzyme. This desalination was carried out using ultrafiltration devices with a mass weight cut-off of 3 kDa (Amicon Ultra-0.5 Centrifugal Filters, Merck KGaA, Darmstadt, Germany) in a centrifuge at 14 000 rpm for 30 min. The sample was washed three times by diluting the residue back to a volume 500 µL using ABC buffer (50 mM, pH 8.0) and subsequent centrifugation at 14 000 rpm for 30 min. After the last washing step, trypsin, chymotrypsin, or a combination of both was added. The mixture was diluted back to 500 µL using ABC buffer, which led to final concentrations of 1 mg/mL for the protein fraction (equal to the initial 54 µM LGA) and 0.02 mg/mL for the enzyme. The mixture was then incubated at 37 °C overnight. With four dilutions with a factor of at least 15 per step, the theoretical total Ph_3Sn concentration after the washing procedure was approx. 5 nM or lower, and thus negligible compared to final peptide concentrations in the micromolar concentration range.

Protein labeling experiments

The protein labeling experiments were carried out in duplicate. The IAA stock solution (375 mM) was prepared using ABC buffer (50 mM, pH 8.0) in a light-protected vial. For the protein alkylation, 100 µL of LGA solution (50 µM, dissolved in ABC buffer) were then mixed with 20 µL IAA solution (1500 equivalents) and incubated in a light-protected vial at room temperature for 30 min. The reaction was quenched by adding ABC buffer to a total volume of 500 µL. Reaction control was carried out using direct injection ESI-MS analysis. To investigate the interactions of the *S*-alkylated LGA with OTCs, the excess of IAA was removed first by desalination using ultrafiltration devices with a mass weight cut-off of 3 kDa in a centrifuge at

14 000 rpm for 30 min. After three washing cycles as described above, PBS and the respective OTC stock solution were added to yield final concentrations of 10 µM LGA (sum of modified and residual unmodified LGA) and 50 µM OTC. The mixture was then incubated at 37 °C overnight.

The ebselen stock solution (1 mM) was prepared using bd. water. Due to the high selectivity and reactivity towards reduced cysteines, only a slight excess of ebselen was necessary during the reaction with the protein. For this, LGA and ebselen were mixed and diluted with PBS buffer to yield final concentrations of approximately 20 µM ebselen and 10 µM LGA. After just two minutes, the reaction was complete, and a reaction control was carried out using direct injection ESI-MS analysis after a desalination step. To investigate the interactions of the ebselen-labeled LGA with OTCs, the respective OTC stock solution was added to the protein to yield final concentrations of 10 µM modified LGA and 50 µM OTC. The mixture was then incubated at 37 °C overnight.

Detailed ICP-MS parameters

The dwell times for each isotope had to be optimized with regard to the experiment and the peak widths in the respective LC analysis run. For general protein and adduct analysis in TQ mode, the following isotopes were recorded in a transient measurement: $^{118}Sn^{16}O$ (500 ms dwell time), $^{120}Sn^{16}O$ (500 ms), $^{32}S^{16}O$ (500 ms), $^{34}S^{16}O$ (500 ms). When ebselen was used for labeling, the following isotopes were recorded: $^{118}Sn^{16}O$ (400 ms), $^{120}Sn^{16}O$ (400 ms), $^{32}S^{16}O$ (300 ms), $^{34}S^{16}O$ (300 ms), $^{78}Se^{16}O$ (300 ms), $^{80}Se^{16}O$ (300 ms). For peptide analysis in TQ mode, where narrower peaks occur, the following isotopes with shorter dwell times were recorded: $^{118}Sn^{16}O$ (200 ms), $^{120}Sn^{16}O$ (200 ms), $^{32}S^{16}O$ (300 ms), $^{34}S^{16}O$ (300 ms).

Detailed ESI-MS parameters

The micrOTOF was operated with the following parameters: Ion Polarity positive; Mass Range *m*/*z* 800-6000; End Plate Offset = -500 V; Capillary = 4000 V; Nebulizer Gas = 1.2 bar; Dry Gas = 9.0 L/min; Dry Temp. = 200 °C; Hexapole RF = 600 Vpp; Pre Pulse Storage = 30 µs;

Transfer Time = 95 µs. An isCID was applied by setting Capillary Exit to 300 V and Skimmer 1 to 75 V.

The timsTOF fleX parameters were optimized for protein-OTC adducts as well as for tryptic peptides, both under non-denaturing, mild conditions. For protein analysis, the following parameters were applied: Ion Polarity positive; Mass Range *m*/*z* 200-6000; End Plate Offset = -500 V; Capillary = 3500 V; Nebulizer Gas = 1.6 bar; Dry Gas = 9.0 L/min; Dry Temp. = 200 °C; Funnel 1 RF = 500 Vpp; Multipole RF = 1200 Vpp; Collision RF = 2000 Vpp; Pre Pulse Storage = 5 µs; Transfer Time = 115 µs. The isCID energy was set to 100 eV.

For peptide analysis, the timsTOF fleX was operated using the following parameters: Mass Range *m*/*z* 300-3000; End Plate Offset = -500 V; Capillary = 3500 V; Nebulizer Gas = 1.6 bar; Dry Gas = 9.0 L/min; Dry Temp. = 200 °C; Funnel 1 RF = 500 Vpp; Multipole RF = 1200 Vpp; Collision RF = 2500 Vpp; Pre Pulse Storage = 17 µs; Transfer Time = 80 µs. The isCID energy was set to 50 eV.

Supplementary figures and tables

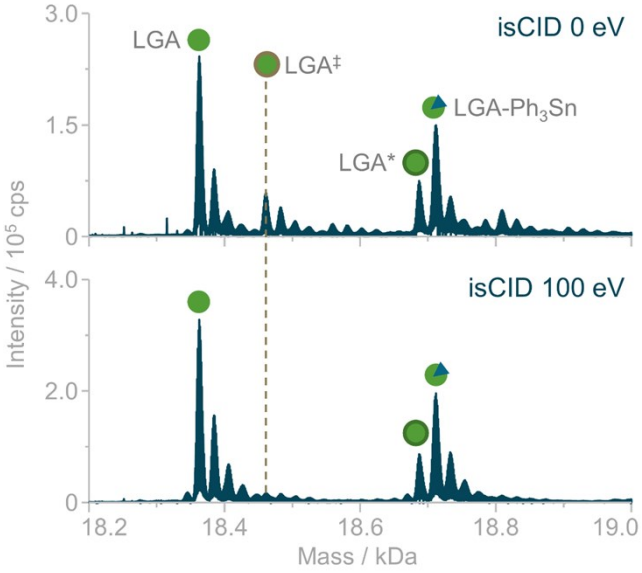

Figure S1. Stability assessment of LGA-Ph_3Sn adducts by isCID experiments. Deconvoluted mass spectra of LGA-Ph_3Sn adducts after analysis via SEC-ESI-HRMS with different applied isCID energies are shown. An isCID energy of 100 eV leads to the neutral loss of phosphate as H_3PO_4, which originates from the PBS incubation and binds to LGA via electrostatic interactions (species marked with ‡). However, decomposition of the LGA-Ph_3Sn adduct is not observed, which indicates high binding energy of the potentially covalently bound Ph_3Sn moiety.

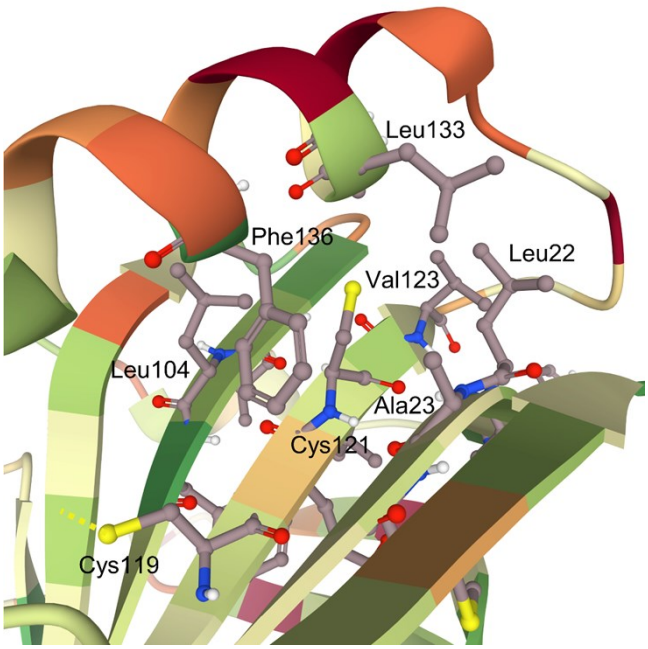

Figure S2. Spatial arrangement of amino acid residues in the Ph_3Sn binding region. An enlarged section of a crystal structure of LGA (PDB ID 1BEB, www.rcsb.org)[1] is shown, as published by Brownlow et al.[2] The structure is presented as the trace of C^{α} atoms and the backbone colors represent the hydrophobicity of the residues (green is hydrophobic, red is hydrophilic). To reduce complexity, the backbone in the background is transparent. The surrounding amino acids (5 Å) of Cys121 are presented as labeled ball&stick structures. The figure was drawn using Mol*.[3] As can be seen, the binding pocket around Cys121 on the solvent-accessible outside of the β-barrel structure is highly hydrophobic, and therefore tolerates the binding of organometallics with lipophilic characters, for instance Ph_3Sn.

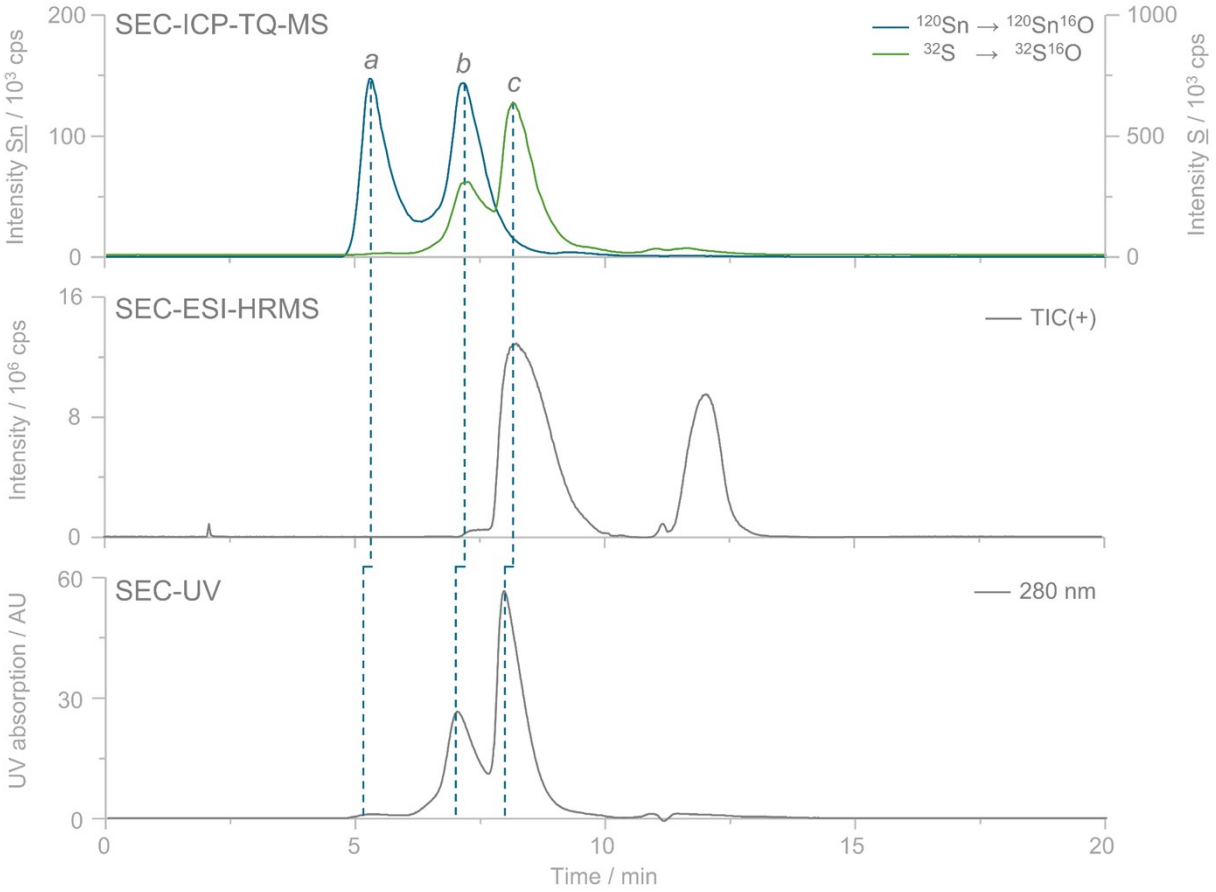

Figure S3. Adduct formation of lactoglobulin and dibutyltin. SEC-ICP-TQ-MS, SEC-ESI-HRMS, and SEC-UV chromatograms of LGA after incubation with dibutyltin are shown. The protein fraction was up-concentrated after incubation using ultracentrifugation filters. Two intense Sn signals show that dibutyltin is not just bound to LGA (peak *c*) but leads to the formation of two new species with shorter retention times, which indicate a higher molecular weight. While the S signal and the t_R of peak *b* imply the formation of a dimeric LGA-dibutyltin species (for mass spectra, see Fig. 4), the composition of the tin species of peak *a* remains unclear.

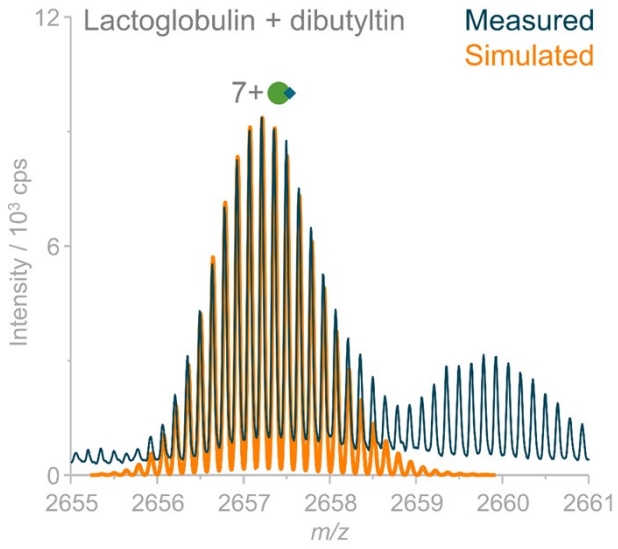

Figure S4. Mass spectrum of the newly identified lactoglobulin-dibutyltin species. Enlarged section of the mass spectrum of LGA after incubation with Bu_2Sn and analysis via SEC-ESI-HRMS. The characteristic isotopic pattern of tin is not visible due to the high molecular weight of the protein, but the high number of naturally stable tin isotopes leads to a significant increase from 27 to 32 isotopes after the adduct formation.

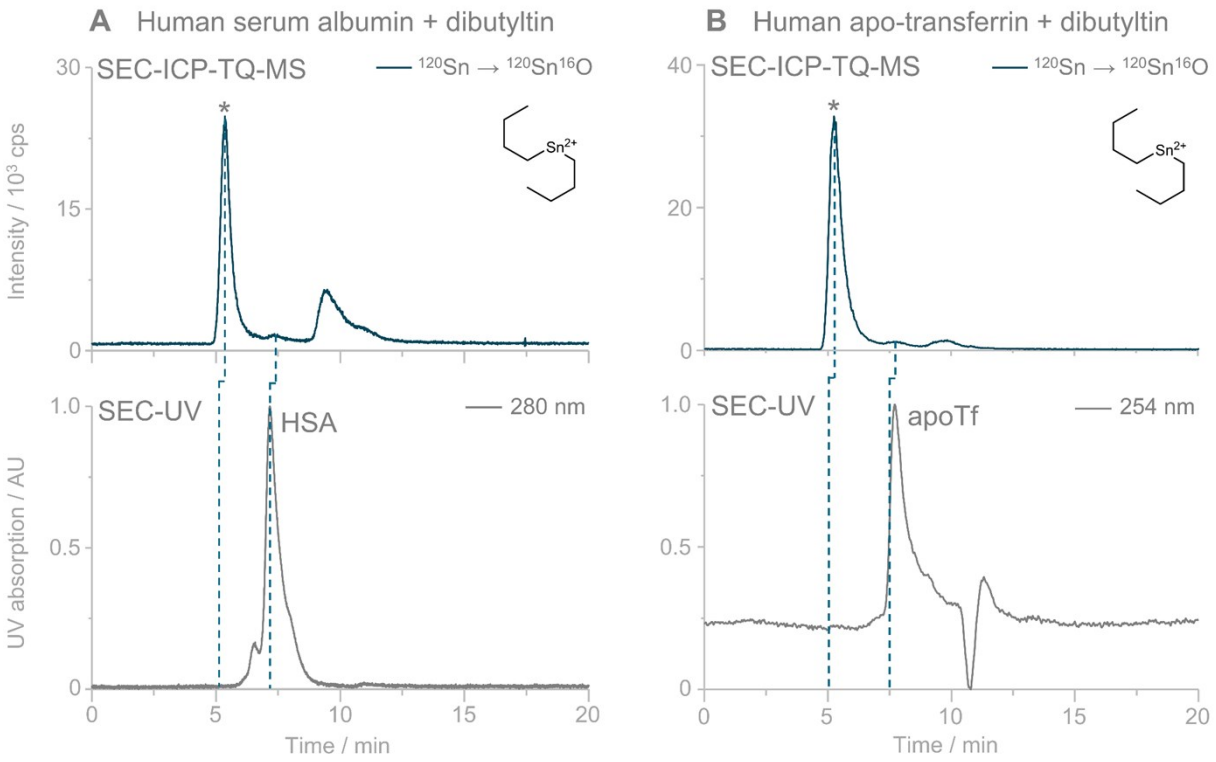

Figure S5. Adduct formation of dibutyltin with toxicological relevant proteins. SEC-ICP-TQ-MS and SEC-UV chromatograms of proteins after incubation with dibutyltin are shown. For both human serum albumin (**A**) and human apo-transferrin (**B**), an intense tin peak (*) with a t_R of 5 min is observed, which is in analogy with LGA and indicates the formation of high molecular weight protein-Bu_2Sn oligomers. With respect to the similar t_R of the species, it can be concluded that the unknown species elute in the exclusion volume of the column.

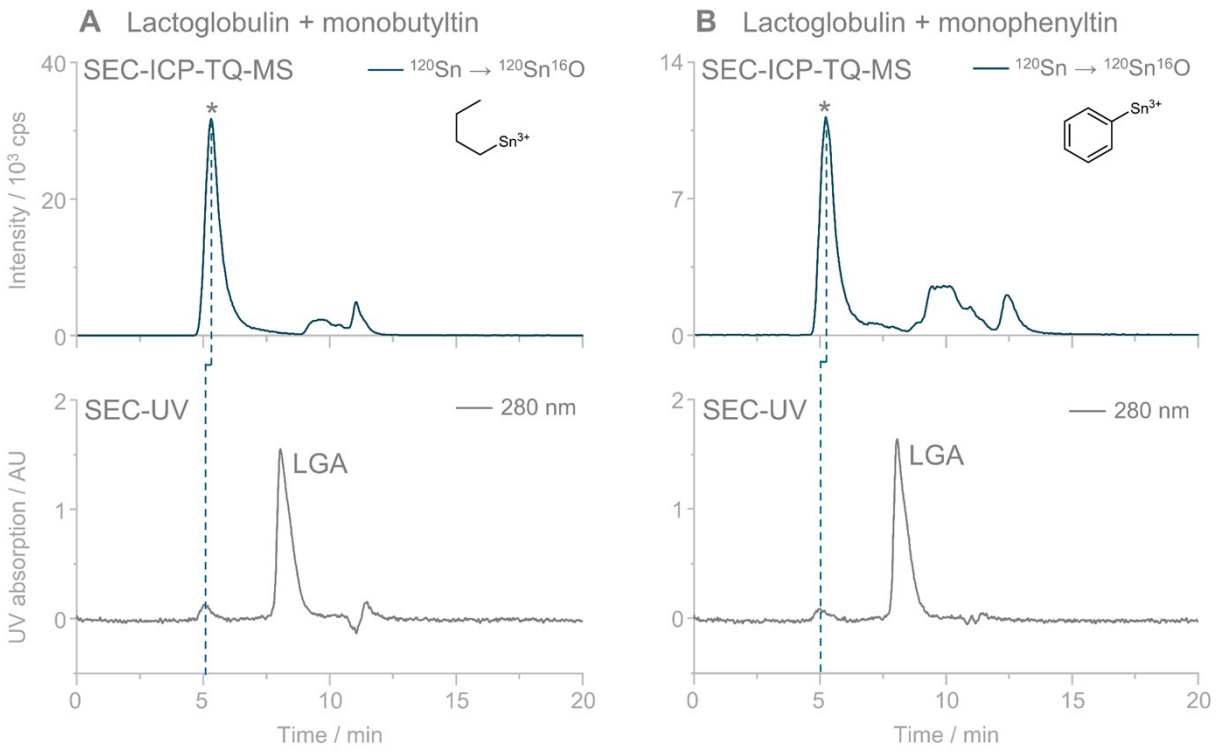

Figure S6. Adduct formation of lactoglobulin with monovalent organotin compounds. SEC-ICP-TQ-MS and SEC-UV chromatograms of LGA after incubation with monobutyltin (**A**) and monophenyltin (**B**) are shown. For both MOTCs, an intense tin peak (*) with a t_R of 5 min is observed, which is in analogy with the binding behavior of DOTCs and indicates the formation of high molecular weight protein-MOTC oligomers.

Table S1. Relevant detected peptides and peptide-Ph_3Sn species. The table gives an overview of the detected peptide species, including their sum formulas, their measured and calculated *m*/*z*, and the resulting relative mass accuracies. Analysis was carried out on the timsTOF fleX instrument.

Pep.[a]	Ion	Seq.	Sum formula	*m*/*z* (meas.)[b]	*m*/*z* (calc.)[b]	Δm / ppm[c]
1	$[M+4H]^{4+}$	[92-135]	$C_{222}H_{357}N_{53}O_{74}S_4$	1,270.1191	1,270.1184	-0.55
			$C_{222}H_{356}N_{53}O_{74}S_4\mathbf{Sn(C_6H_5)_3}$	1,357.3725	1,357.3715	-0.74
2	$[M+3H]^{3+}$	[92-124]	$C_{170}H_{274}N_{41}O_{52}S_4$	1,283.9659	1,283.9658	-0.08
			$C_{170}H_{273}N_{41}O_{52}S_4\mathbf{Sn(C_6H_5)_3}$	1,400.3038	1,400.3032	-0.43
3	$[M+5H]^{5+}$	[92-138]	$C_{241}H_{384}N_{57}O_{79}S_4$	1,094.3353	1,094.3343	-0.91
			$C_{241}H_{383}N_{57}O_{79}S_4\mathbf{Sn(C_6H_5)_3}$	1,164.1380	1,164.1368	-1.0
4	$[M+2H]^{2+}$	[102-124]	$C_{115}H_{183}N_{29}O_{36}S_5$	1,337.6142	1,337.6140	-0.15
			$C_{115}H_{182}N_{29}O_{36}S_5\mathbf{Sn(C_6H_5)_3}$	1,512.1201	1,512.1201	0.00

[a] Peptide number with regard to the paper [b] The highest abundant peaks of the respective isotope patterns were used for evaluation, as the monoisotopic peaks could not be determined for all peptides [c] Relative mass accuracy.

Table S2. Detected LGA-Bu_2Sn adduct species. The table gives on overview of the relevant adduct species and their respective sum formulas, measured and calculated *m*/*z*, and the resulting relative mass accuracy. Analysis was carried out on the timsTOF fleX instrument.

Species	Ion	Sum formula	*m*/*z* (meas.)[a]	*m*/*z* (calc.)[a]	Δm / ppm[b]
LGA	$[M+7H]^{7+}$	$C_{821}H_{1325}N_{206}O_{250}S_9$	2624.2108	2624.2147	1.5
LGA-Bu_2Sn	$[M+7H]^{7+}$	$C_{821}H_{1323}N_{206}O_{250}S_9\mathbf{Sn(C_4H_9)_2}$	2657.2121	2657.2187	2.5
LGA-$(Bu_2Sn)_2$	$[M+7H]^{7+}$	$C_{821}H_{1321}N_{206}O_{250}S_9\mathbf{Sn_2(C_4H_9)_4}$	2690.2174	2690.2227	2.0
LGA-$(Bu_2Sn)_3$	$[M+7H]^{7+}$	$C_{821}H_{1319}N_{206}O_{250}S_9\mathbf{Sn_3(C_4H_9)_6}$	2723.2203	2723.2267	2.4
fGLA	$[M+7H]^{7+}$	$C_{780}H_{1251}N_{196}O_{239}S_8$	2493.4208	2493.4249	1.6
fLGA-Bu_2Sn	$[M+7H]^{7+}$	$C_{780}H_{1249}N_{196}O_{239}S_8\mathbf{Sn(C_4H_9)_2}$	2526.4243	2526.4290	1.9
fLGA-$(Bu_2Sn)_2$	$[M+7H]^{7+}$	$C_{780}H_{1247}N_{196}O_{239}S_8\mathbf{Sn_2(C_4H_9)_4}$	2559.4293	2559.4330	1.4
fLGA-$(Bu_2Sn)_3$	$[M+7H]^{7+}$	$C_{780}H_{1245}N_{196}O_{239}S_8\mathbf{Sn_3(C_4H_9)_6}$	2592.4310	2592.4370	2.3
LGA_2-Bu_2Sn	$[M+13H]^{13+}$	$C_{1642}H_{2647}N_{412}O_{500}S_{18}\mathbf{Sn(C_4H_9)_2}$	2843.8384	2843.8483	3.5
LGA_2-$(Bu_2Sn)_2$	$[M+13H]^{13+}$	$C_{1642}H_{2645}N_{412}O_{500}S_{18}\mathbf{Sn_2(C_4H_9)_4}$	2861.6085	2861.6197	3.9
LGA_2-$(Bu_2Sn)_3$	$[M+13H]^{13+}$	$C_{1642}H_{2643}N_{412}O_{500}S_{18}\mathbf{Sn_3(C_4H_9)_6}$	2879.2998	2879.3141	5.0
fLGA-LGA- Bu_2Sn	$[M+13H]^{13+}$	$C_{1601}H_{2573}N_{402}O_{489}S_{17}\mathbf{Sn(C_4H_9)_2}$	2773.4102	2773.4231	4.7
fLGA-LGA-$(Bu_2Sn)_2$	$[M+13H]^{13+}$	$C_{1601}H_{2571}N_{402}O_{489}S_{17}\mathbf{Sn_2(C_4H_9)_4}$	2791.1857	2791.1945	3.2
fLGA-LGA-$(Bu_2Sn)_3$	$[M+13H]^{13+}$	$C_{1601}H_{2569}N_{402}O_{489}S_{17}\mathbf{Sn_3(C_4H_9)_6}$	2808.9546	2808.9659	4.0

[a] The highest abundant peaks of the respective isotope patterns were used for evaluation, as the monoisotopic peaks could not be determined for all peptides [b] Relative mass accuracy.

References

1 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, The Protein Data Bank, *Nucleic Acids Res.*, 2000, **28**, 235–242.

2 S. Brownlow, J. H. Morais Cabral, R. Cooper, D. R. Flower, S. J. Yewdall, I. Polikarpov, A. C. T. North and L. Sawyer, Bovine β-lactoglobulin at 1.8 Å resolution - Still an enigmatic lipocalin, *Structure*, 1997, **5**, 481–495.

3 D. Sehnal, A. Rose, J. Koča, S. Burley and S. Velankar, Mol*: towards a common library and tools for web molecular graphics, *Proc. Work. Mol. Graph. Vis. Anal. Mol. data*, 2018, 29–33.