

Expanding the mass range for UVPD-based native top-down mass spectrometry

Jean-François Greisch^{1,2}, Sem Tamara^{1,2}, Richard A. Scheltema^{1,2}, Howard W.R. Maxwell³, Robert D. Fagerlund³, Peter C. Fineran³, Stephan Tetter⁴, Donald Hilvert⁴ and Albert J.R. Heck^{1,2}

¹ *Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute of Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 Utrecht, The Netherlands*

² *Netherlands Proteomics Center, Padualaan 8, 3584 Utrecht, The Netherlands*

³ *Department of Microbiology and Immunology, University of Otago, PO Box 56, 9054 Dunedin, New Zealand*

⁴ *Laboratory of Organic Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 1-5/10, 8093 Zürich, Switzerland*

SUPPLEMENTARY INFORMATION

Table S1:

Name	Sequence, 5' – 3'	Notes
PF209	TCGTCTTCACCTCGAGAAATC	Sequencing pQE-80LoriT
PF210	GTCATTACTGGATCTATCAACAGG	Sequencing pQE-80LoriT
PF245	GTGGCACGACCAGCTC	Sequencing <i>cas8f</i>
PF246	ACCTCACTACTGCATCGCC	Sequencing <i>cas5</i>
PF401	AGGTCTGCAGTTAGAACCCACGGAACGGTG	R <i>cas6f</i> , PstI site
PF748	CTGAAAGAGGAGCTGGCATG	Sequencing <i>cas8f</i>
PF758	GCGAATCCGGTGAAGACTATTCGCTGCG	Sequencing <i>cas6f</i>
PF771	CTAAAAAGGACAACATCATGGCAAAG	Sequencing <i>cas5</i>
PF861	TAATACGACTCACTATAGGG	Sequencing pRSF-1b
PF1455	TTTTCAATTGAGGAGAAATTA ^{ACTATGTGGAGCCACCCGCAGT} TCGAAAAAGGCGCGAGAAATGGACTACCCG	F for StrepII-tag <i>cas8f</i> , MfeI site
PF2237	ATCCATGGGTTCACTGCCGTACAGGCAGCTTAGAAAGCAGAG ACTATCGATACGGTCTGG	F CRISPR1-Spacer1, NcoI site
PF2238	TCAAAGCTTTCTAAGCTGCCTGTACGGCAGTGAACGCATCCGT CCAGACCGTATCGATAG	F CRISPR1-Spacer1, HindIII site

*Restriction site in underlined

Table S2: Acquisition parameters used.

	Resolution 4375 & 8750 at <i>m/z</i> 200		Resolution 140000 at <i>m/z</i> 200	
	Microscans	Averaging	Microscans	Averaging
B-PE	1*	1000	1*	0
Csy	40*	0	1*	0
Wt-AaLS	1*	1000	1*	0

* This corresponds to the Xcalibur parameter, up to 1000 spectra where averaged using Thermo Qcal Browser or Thermo Freestyle softwares.

Table S3: Quantities relevant to UV photoabsorption by AaLS capsids.

Photon energy (eV)	6.43	Radius of wt-AaLS (Å)	75
Pulse energy inside HCD cell (μJ)	50	# peptide bonds	9180
Beam area inside HCD cell (mm ²)	π	σ_{193nm} of a peptide bond (Å ²)	0.26

Native Top-Down MS of B-phycoerythrin (B-PE)

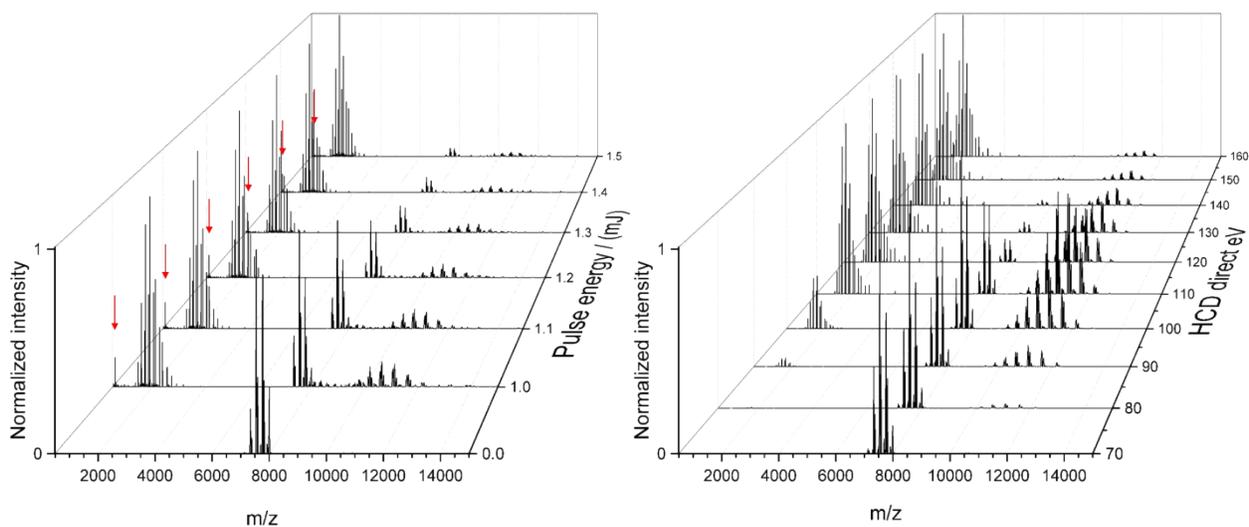


Fig. S1: Native top-down mass spectrometry of B-PE ($[37^+-33^+]$, $6 \cdot 10^{-10}$ mbar readout, 4375 resolution at m/z 200) using left) UVPD and right) HCD at increasing energies. The intact phycoerythrobilin chromophore fragment is exclusively observed upon UVPD (red arrows).

Table S4: Masses of the main product ions resulting from UVPD of B-PE (see Figure 2).

Ion	Mass (Da)	Expected mass (Da)
$(\alpha\beta)_6\gamma$ (I)	262377 ± 18	
$(\alpha\beta)_6\gamma$ (II)	263808 ± 22	
α (2 bilins)	18977.4 ± 0.7	
β (3 bilins)	20327.8 ± 1.1	
(I)- β	241692 ± 37	242050
(I)- α or (II)- β	243060 ± 15	243399 or 243481
(II)- α	244397 ± 128	244830

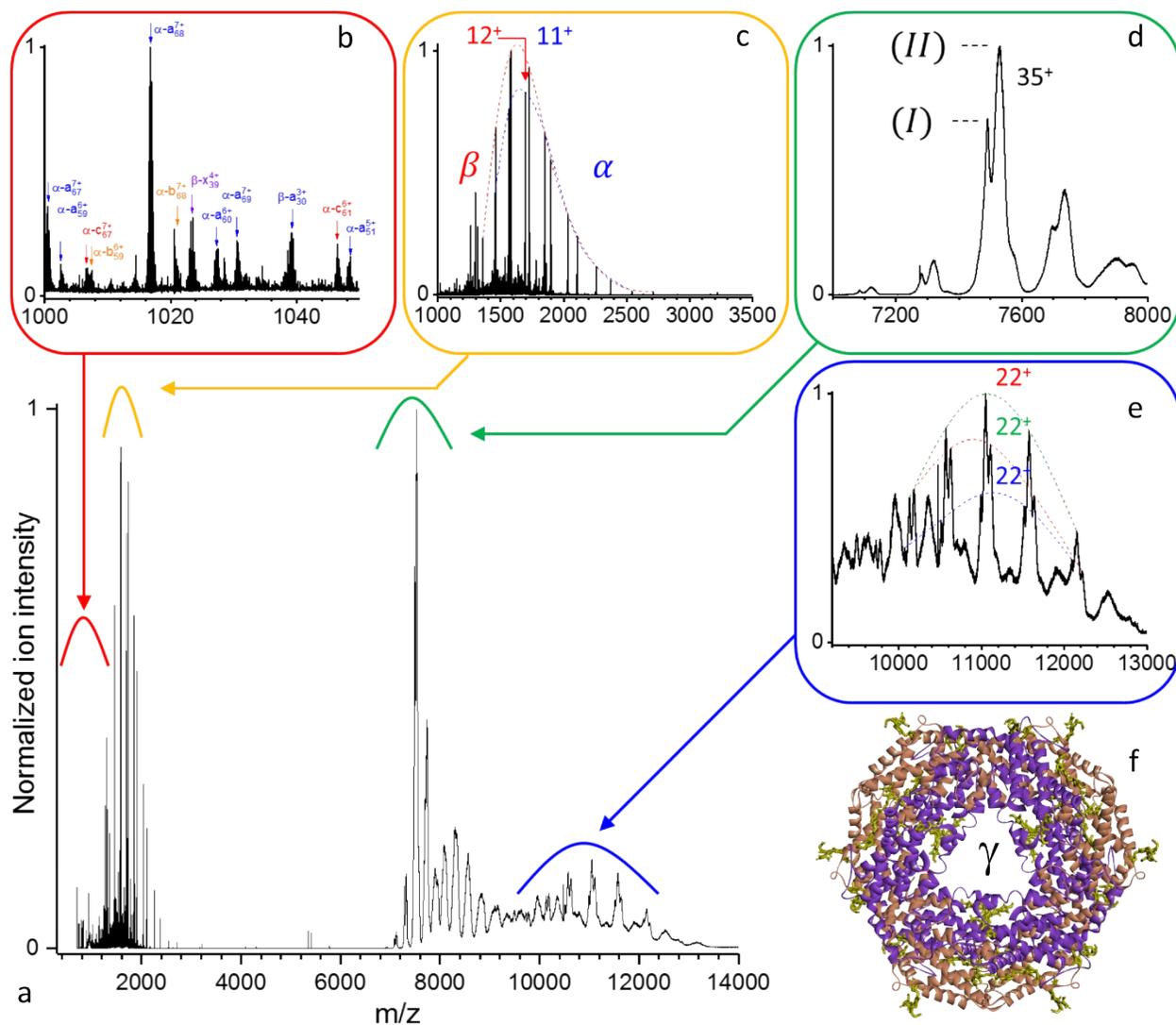


Fig. S2: (a) Native top-down mass spectrometry of mass selected B-PE 35^+ charge state using UVPD ($[35^+]$, 1.3 mJ/pulse, $2.4 \cdot 10^{-10}$ mbar N_2 readout, and resolution 140000 at m/z 200). The 35^+ charge state was selected using a 150 m/z window. (d) Precursor ion dissociated into (c) intact subunits, (e) larger fragments and (b) free chromophores. (f) Structural model of the B-phycoerythrin $(\alpha\beta)_6$ barrel with the more spurious (structurally non-resolved) γ unit at the center and the phycoerythrobilin chromophores highlighted in mustard green.

A detailed analysis of a higher resolution mass spectrum (140000 at m/z 200) yielding optimal detection for all m/z ranges, Fig. S2, provides evidence of the coexistence of at least two populations of proteoforms (Fig S2d, res. 8750 at m/z 200) with approximate masses (I) 262114 ± 18 Da and (II) 263479 ± 9 Da (as inferred from low intensity neighboring ions). These result from the sequence variability of the γ -unit threading B-PE $(\alpha\beta)_6$ barrel as well as losses of bilins upon prolonged storage in an ammonium acetate buffer. Dissociation of the complex yields intact subunits in the low mass range (Fig. S2c) with masses of 18977.4 ± 1.1 Da and 20327.8 ± 1.6 Da, respectively corresponding to α and β subunits each with two PEB bound. On the high m/z range (Fig. S2e), three concomitant charge distribution are observed with masses of 241775 ± 27 Da, 243076 ± 4 Da, and 244340 ± 11 Da. These respectively correspond to (I) minus β

(241787 Da), (I) minus α (243136 Da) or (II) minus β (243152 Da), and (II) minus (α) (244501 Da) respectively. PEB loss upon UVPD, mass 586.28 Da, can be seen in Fig S2b.

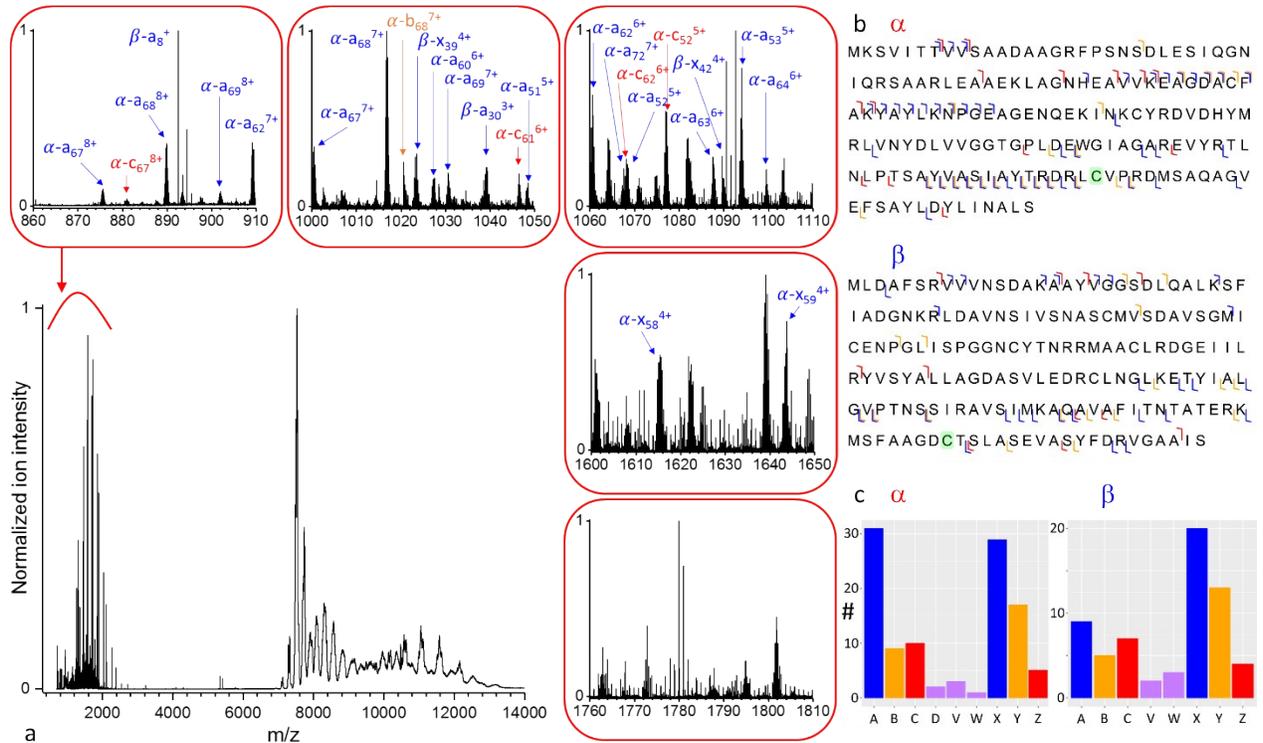


Fig. S.3: (a) Native top-down UVPD of B-PE optimized for sequence coverage ($[35^+]$, 1.3 mJ/pulse, $2.4 \cdot 10^{-10}$ mbar N_2 readout, and resolution 140000 at m/z 200). The 35^+ charge state of the precursor ions has been isolated using a 150 m/z window. Insets: illustration of fragments density for selected mass ranges. (b) Sequence coverages for the α - and β -subunits (5 ppm accuracy, 15% FDR). (c) Number of ions per ion type.

Native Top-Down MS of a CRISPR crRNA Csy complex

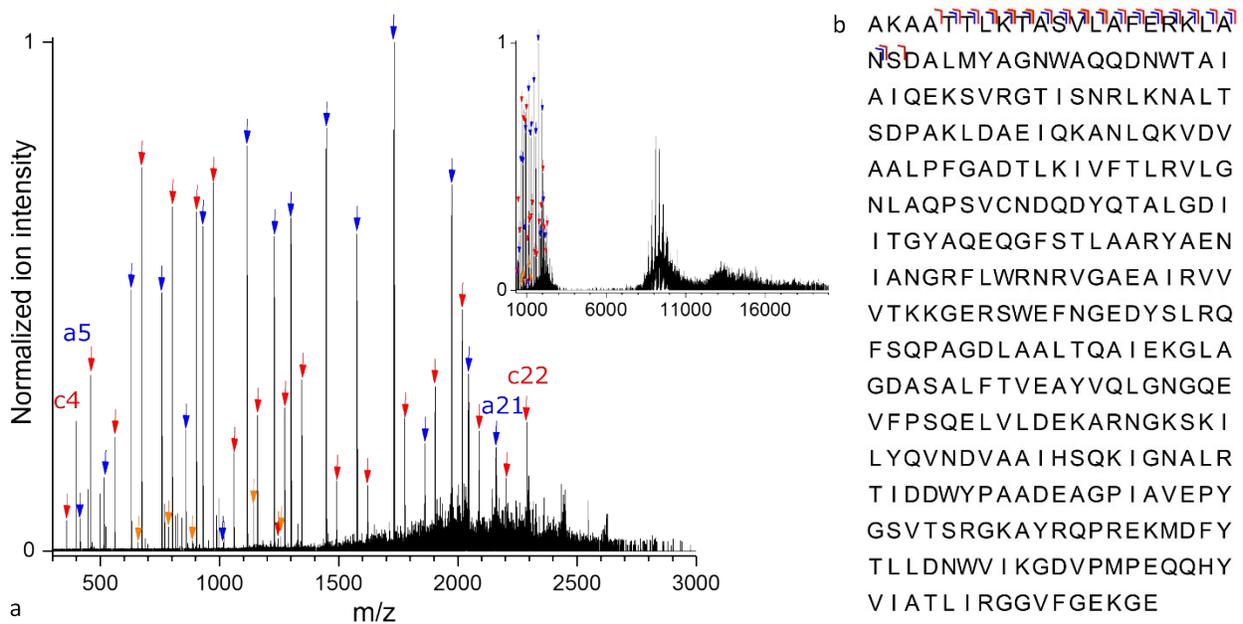


Fig. S.4: Csy UVPD spectrum optimized for sequence coverage ($[43^+-36^+]$, 2.5 mJ/pulse, $3.2 \cdot 10^{-10}$ mbar readout, and resolution 140000 at m/z 200). (a) The $[43^+-36^+]$ charge state distribution was selected using a 2000 m/z window and fragmented at a 2.5 mJ/pulse energy. The intact subunits have nearly fully vanished and peptide fragment ions dominate this m/z region. (b) Cas7 annotated sequence (see text). The a ions, in blue, start with a5 and end with a21. The 5 b ions, in orange, are b7, b8, b9, b12 and b13. The c ions, in red, start with c4 and end with c22 (a weak doubly charged is also observed).

Native Top-Down MS of the wt-AaLS Cargo-Protein

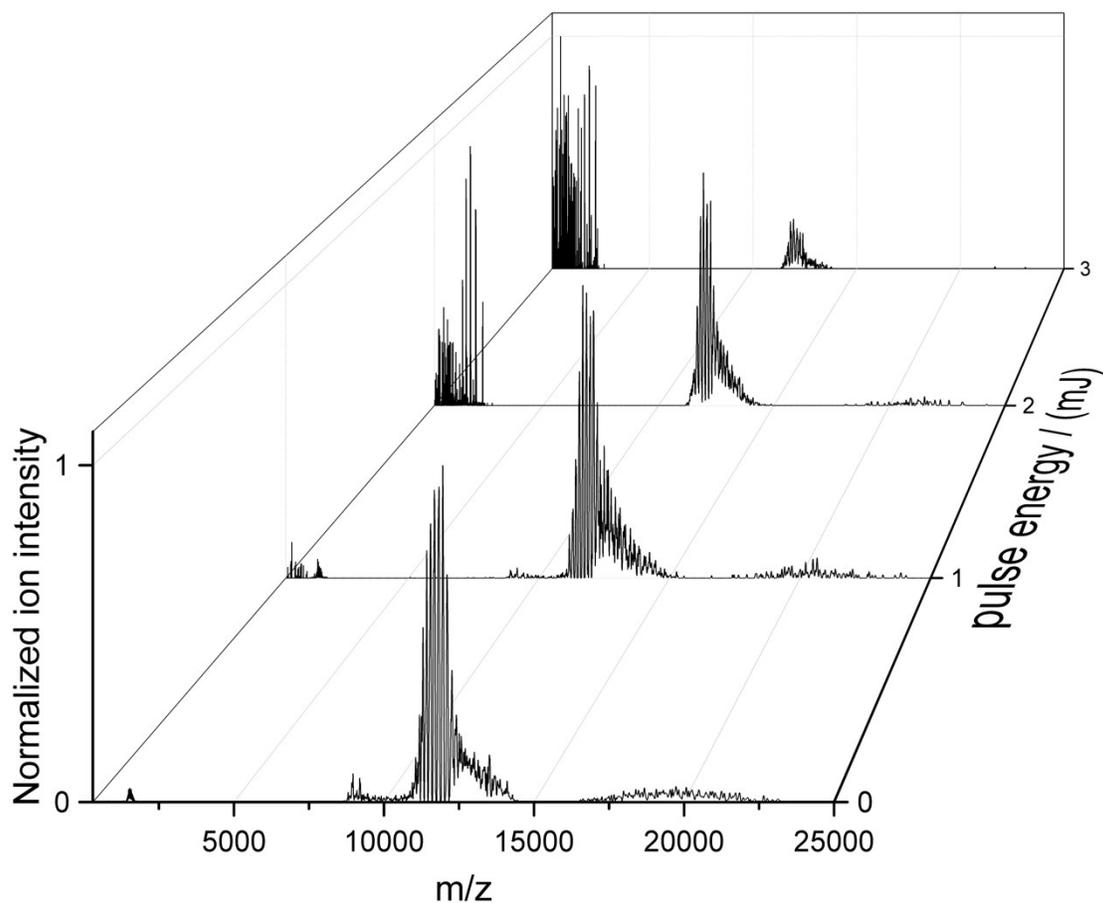


Fig. S5: Native top-down mass spectrometry of wt-AaLS virus-like nanocontainers ($[95^+-87^+]$, $5.5 \cdot 10^{-10}$ mbar readout, 4375 resolution at m/z 200) using UVPD at increasing laser pulse energies. Incomplete desolvation is responsible for the shoulder to the right of the 0 mJ/pulse charge distribution.

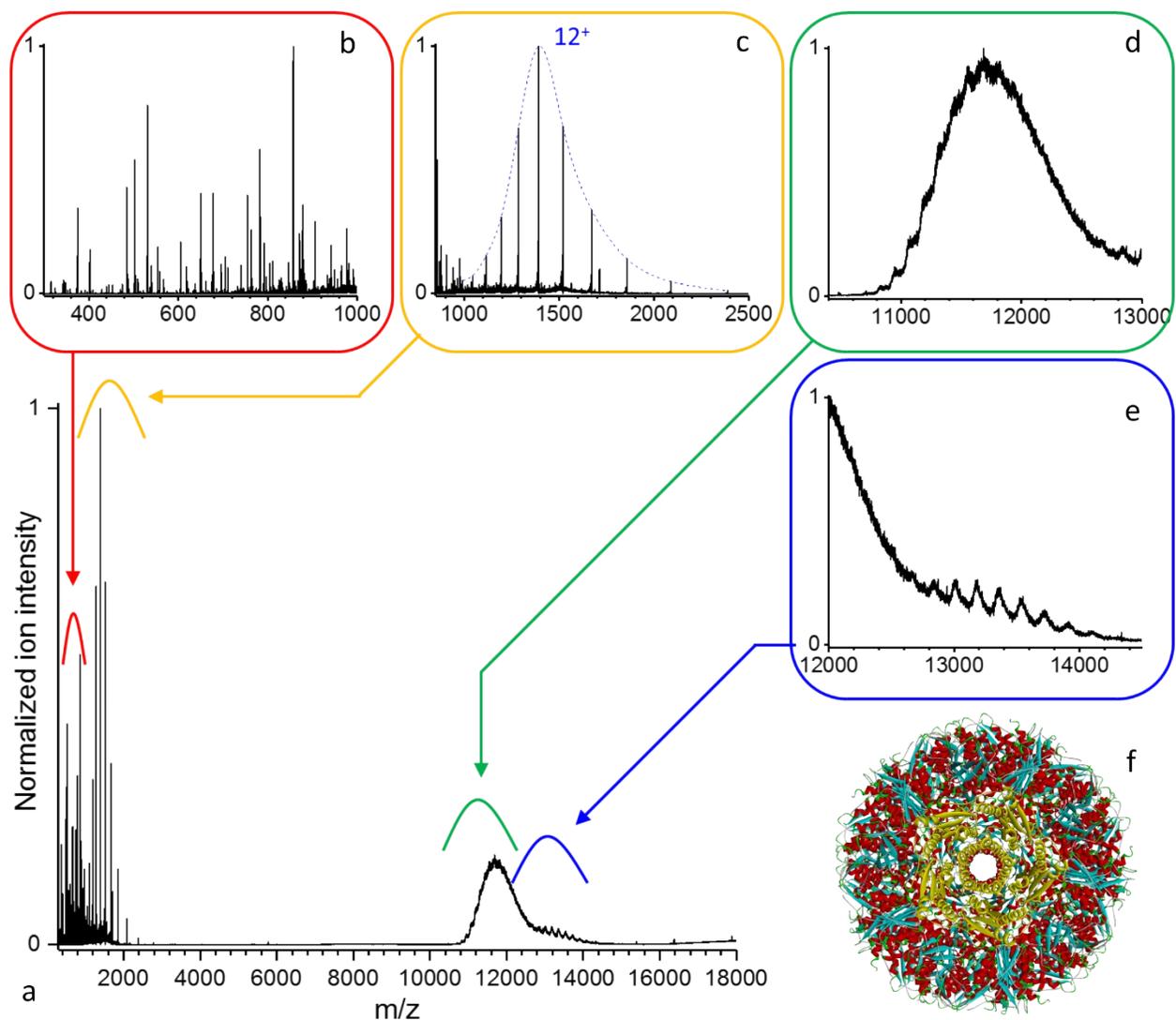


Fig. S6: (a) Native top-down mass spectrometry of mass selected wt-AaLS using UVPD ($[95^+-87^+]$, 0.7 mJ/pulse, $3.7 \cdot 10^{-10}$ mbar N_2 readout, and resolution 140000 at m/z 200). (d) Precursor ions are dissociated into (c) intact subunits, (e) larger fragments and (b) some low mass peptide fragments. (f) Structural model of wt-AaLS (PDB ID: 5mpp) with a pore formed by a pentamer of subunits highlighted in yellow.

Analysis of Fig. S6 (0.7 mJ/pulse, $3.7 \cdot 10^{-10}$ mbar N_2 readout), yields for the precursor (Fig. 6d) a low mass charge distribution (Fig. 6c) corresponds to the intact *Aquifex aeolicus* lumazine synthase with average mass of 16705.3 ± 0.7 Da in agreement with the average expected value of 16706.0 (NIST defined elemental av. Mass) as well as a concomitant high mass distribution (Fig. 6e).

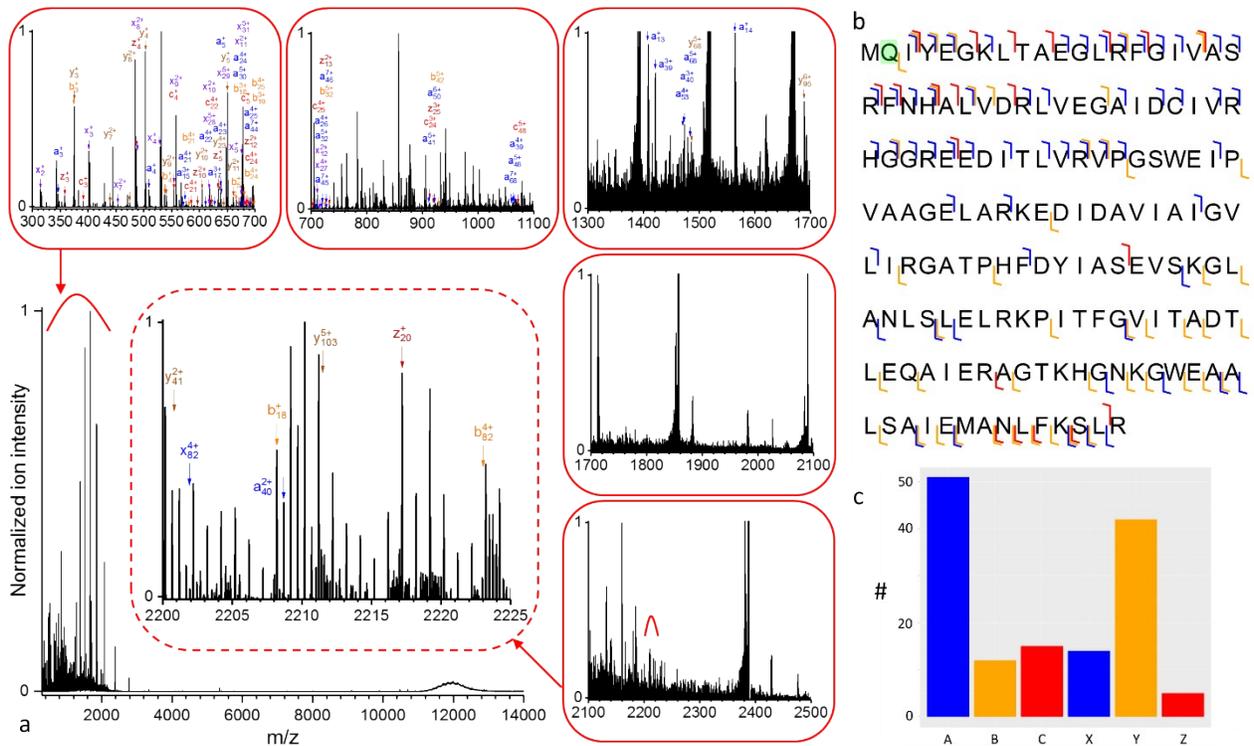


Fig. S7: (a) wt-AaLS UVPD spectrum optimized for sequence analysis ($[95^+-87^+]$, $3.9 \cdot 10^{-10}$ mbar, 1.6 mJ/pulse, and resolution 140k at m/z 200). wt-AaLS charge state distribution was isolated using a 4000 m/z window. Insets: illustration of fragments density for selected mass ranges. (b) 61% Sequence coverage of the subunit displaying Q2 deamination (4 ppm accuracy, 6% FDR). The coverage can be increased via manual assignment as illustrated in (a)-central inset. (c) Number of fragments per ion type. Comparison with Prosight Lite analysis is provided in Fig. S8.



Fig. S8: wt-AaLS 75% Sequence coverage of the subunit displaying Q2 deamination (standard 10 ppm accuracy, PCS 844, p-value $5 \cdot 10^{-73}$).