Enhancing Top-Down Analysis using Chromophore-Assisted Infrared Multi-Photon Dissociation from (Phospho)Peptides to Protein Assemblies

Jean-François Greisch[1,2], Saar A.M. van der Laarse[1,2], 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, 3584 Utrecht, The Netherlands*

[2] *Netherlands Proteomics Center, 3584 Utrecht, The Netherlands*

SUPPORTING INFORMATION

Compounds studied

Table S1. List of studied compounds.

Name	**Source**	**Sequence**	**%D,E**	**1st isotope [Da]**
angiotensin I (10 a.a.)	Merck KGaA A9650	DRVYIHPFHL	10	1295.677
[Glu1]-fibrinopeptide B human (14 a.a.)	Merck KGaA F3261	EGVNDNEEGFFSAR	29	1569.670
bradykinin (9 a.a.)	Merck KGaA B3259	RPPGFSPFR	0	1059.561
apo-colicin E9 DNAse (134 a.a.)	See Experimental	MESKRNKPGKATGKGKPVGDKWL DDAGKDSGAPIPDRIADKLRDKEFK SFDDFRKAVWEEVSKDPELSKNLN PSNKSSVSKGYSPFTPKNQQVGGR KVYELHHDKPISQGGEVYDMDNIR VTTPKRHIDIHRGK	16	15078.752
protein Aurora Borealis (Bora, 156 a.a.)	See Experimental	GAASMMGDVKESKMQITPETPGR IPVLNPFESPSDYSNLHEQTLASPSV FKSTKLPTPGKFRWSIDQLAVINPV EIDPEDIHRQALYLSHSRIDKDVEDK RQKAIEEFFTKDVIVPSPWTDHEGK QLSQCHSSKCTNINSDSPVGKKLTI HSEKSD	15	17459.755

"Pure" IRMPD of the monoclonal antibody Trastuzumab

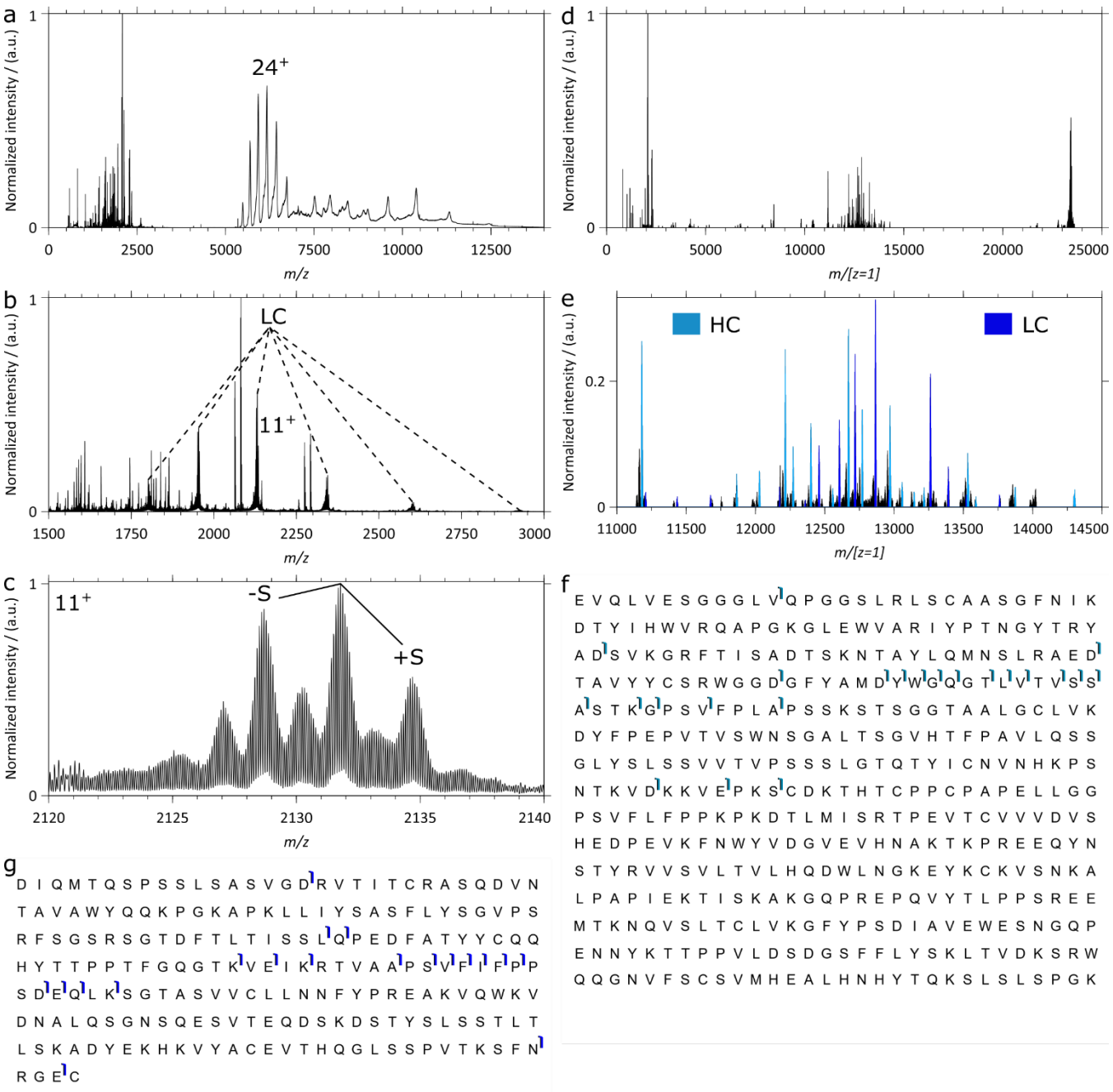

Fig. S1. IRMPD of the monoclonal intact antibody Trastuzumab (Roche, Penzberg, Germany). a) IRPMD mass spectrum of precursor ions spanning the 5500 to 6250 m/z range. b) Low m/z fragments among which intact light chain (LC) ions resulting from the cleavage of the disulfide bond between the LC and the heavy chain (HC). c) Pattern characteristic of the cleavage of a disulfide bond (in this case the bond between the LC and HC). d) Charge deconvoluted mass spectrum of the fragment ions, with the intact LC fragments to the right. e) b-ions corresponding to HC and LC fragments. Assignment was performed using a 3ppm tolerance (the 3 y-ions assigned under a 3ppm tolerance are omitted). f) b-ions assigned for the HC. g) b-ions assigned to the LC.

Phosphopeptides

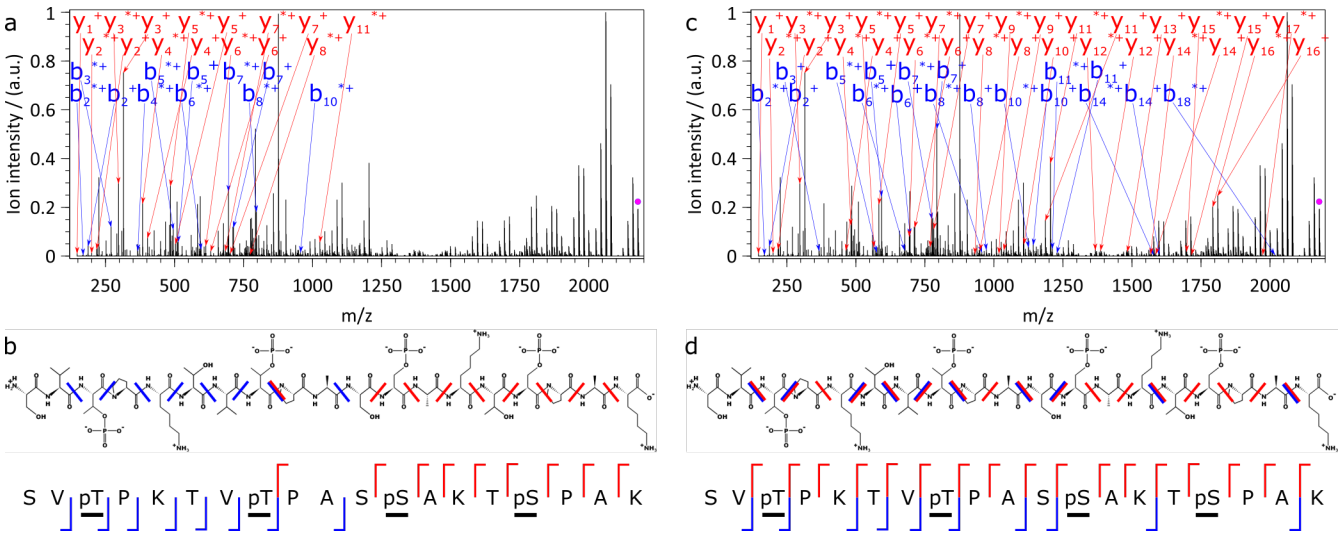

Fig. S2. IRMPD mass spectra of the quadruply phosphorylated SVpTPKTVpTPASpSAKTpSPAK peptide (see Fig. 2). a) Fragment ion assignment corresponding to the complete loss of HPO_3 (identical to the non-phosphorylated peptide). b) Fragment ion assignment corresponding to the phosphorylated peptide (assuming all phospho groups intact (all HPO_3 are retained). For an explanation of the fragments in the 1500 to 2200 range see Fig. S3.

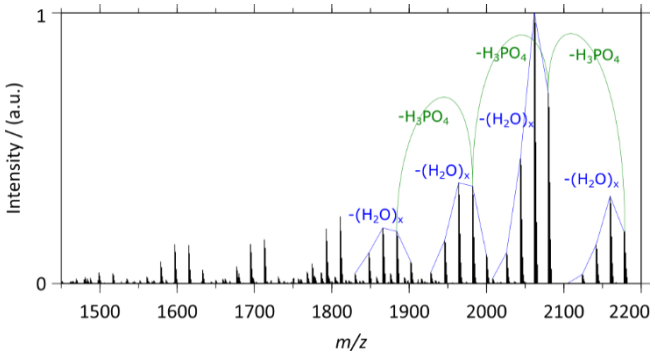

Fig. S3. IRMPD mass spectra of the quadruply phosphorylated SVpTPKTVpTPASpSAKTpSPAK peptide (see Fig. 2). a) Fragment ion assignment corresponding to the unphosphorylated peptide (assuming complete HPO_3 loss as displayed in Fig. 2). b) Fragment ion assignment corresponding to the phosphorylated peptide (assuming all phospho groups intact.

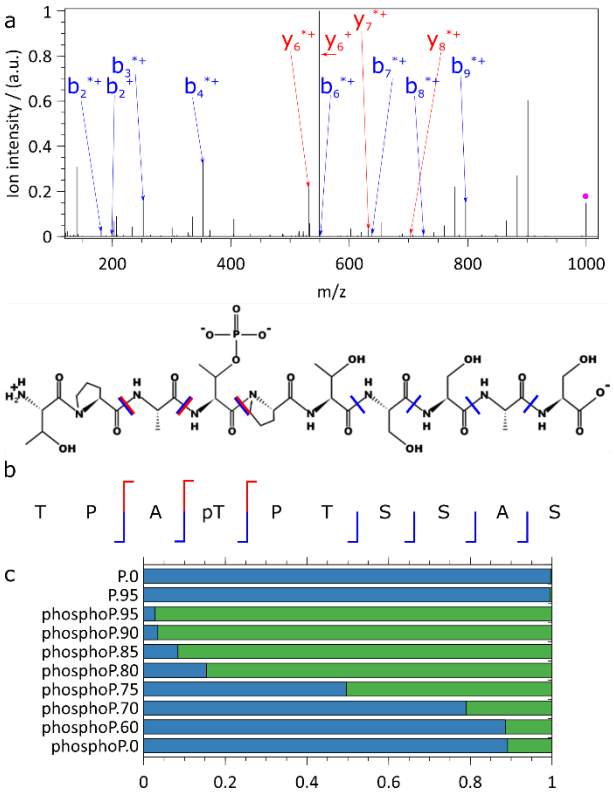

Fig. S4. IRMPD of the TPApTPTSSAS phosphopeptide: a) annotated charge-deconvoluted – all ions are 1^+ – mass spectrum, fragments labelled with * have lost H_2O, precursor highlighted by pink dot, b) annotated peptide sequence with y^+ ions in red and b^+ ions in blue, c) survival yield (blue) as function of the laser power (0-95 % range) for the non-phosphorylated (P.) and phosphorylated (phosphoP.) peptides. The complementary fragment fraction is displayed in green. 100 scans, 1 laser pulse of 80 ms per scan, $5.28 \cdot 10^{-11}$ mbar N_2 UHV readout, resolution of 140000 @ *m/z* 200.

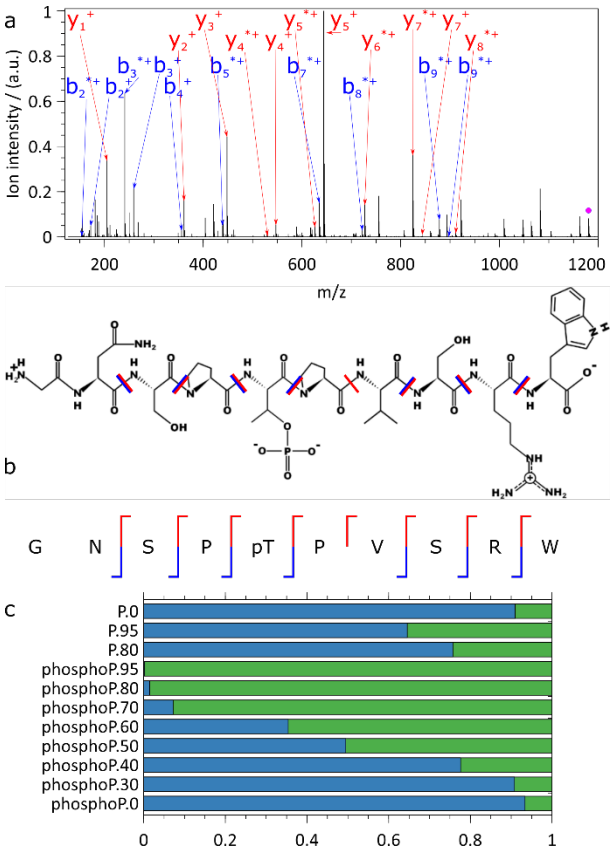

Fig. S5. IRMPD of the GNSPpTPVSRW phosphopeptide: a) annotated charge-deconvoluted – all ions are 1^+ – mass spectrum, fragments labelled with * have lost H_2O, precursor highlighted by pink dot, b) annotated peptide sequence with y^+ ions in red and b^+ ions in blue, c) survival yield (blue) as function of the laser power (0-95 % range) for the non-phosphorylated (P.) and phosphorylated (phosphoP.) peptides. The complementary fragment fraction is displayed in green. 100 scans, 1 laser pulse of 80 ms per scan, <5.00 · 10^{-11} mbar N_2 UHV readout, resolution of 140000 @ *m/z* 200.

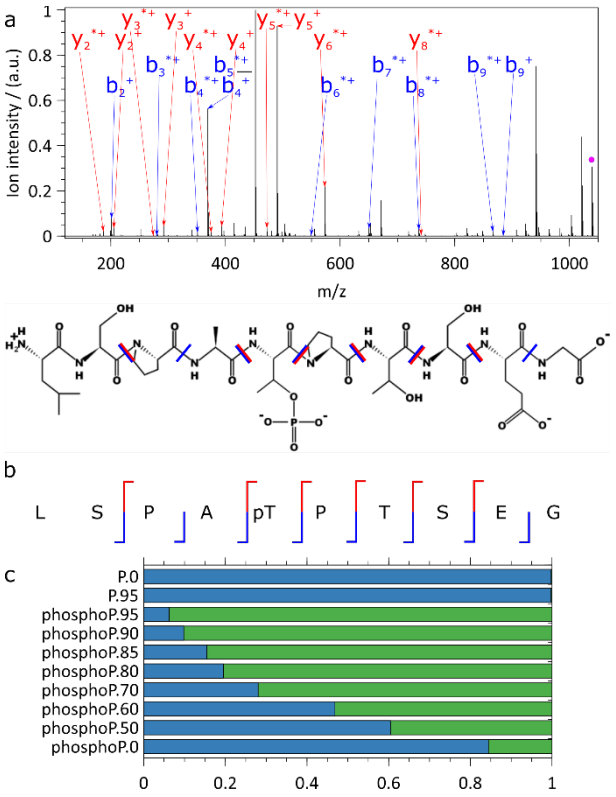

Fig. S6. IRMPD of the LSPApTPTSEG phosphopeptide: a) annotated charge-deconvoluted – all ions are 1^+ – mass spectrum, fragments labelled with * have lost H_2O, precursor highlighted by pink dot, b) annotated peptide sequence with y^+ ions in red and b^+ ions in blue, c) survival yield (blue) as function of the laser power (0-95 % range) for the non-phosphorylated (P.) and phosphorylated (phosphoP.) peptides. The complementary fragment fraction is displayed in green. 100 scans, 1 laser pulse of 80 ms per scan, $5.52 \cdot 10^{-11}$ mbar N_2 UHV readout, resolution of 140000 @ *m/z* 200.

Peptides complexing IP6

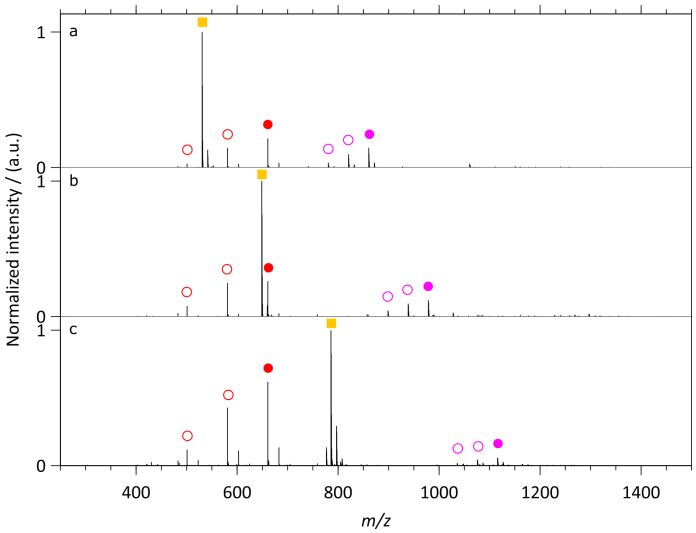

Fig. S7. Mass spectra (MS1) of a) bradykinin + IP6, b) Angiotensin I + IP6, c) [Glu1]-fibrinopeptide B + IP6: orange square = bare peptide, full red circle = IP6, empty red circles = IP6 – n × HPO_3, full pink = peptide + IP6, empty pink circle = peptide + IP6 – n × HPO_3. The IP6 – n × HPO_3 species are present in the original solution. Residual amounts of sodium can be found to the right of all IP6 (containing species). In c, the slightly less abundant IP6 adduct for [Glu1]-fibrinopeptide B is a direct consequence of the addition taking place for solutions diluted 30 times compared to a) and b).

In Fig. S5, we first report IP6-based IRMPD enhancement for bradykinin. Doubly charged ions of bare bradykinin (*m/z* 530.788) yield barely detectable fragments (relative intensities compared to precursor below 0.2%) at maximal laser power. 10.6 μm photo-activation of the bradykinin-IP6 complex (*m/z* 860.718, 1.68 J/pulse), on the other hand, produces extensive fragmentation. Besides the precursor ion (*m/z* 1720.428), IRMPD yields protonated bradykinin (*m/z* 1060.566), b-ions b_2^+, b_4^+ to b_8^+, as well as the y-ion series y_1^+ to y_4^+ and y_6^+ to y_8^+. In other words, IRMPD achieves full sequence coverage of bradykinin non-covalently bound to IP6 with mass separation enabling effective rejection of IP6-fragment retaining precursors.

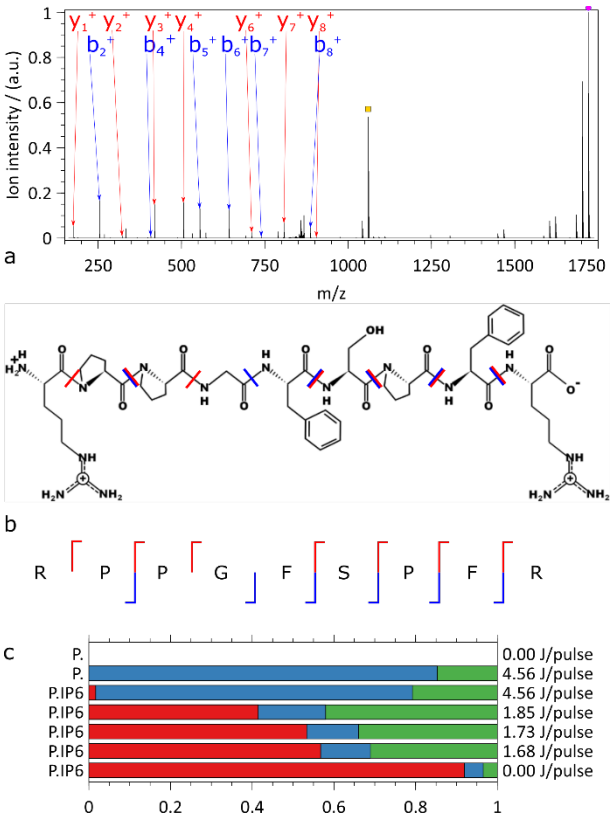

Fig. S8. (a) Charge-deconvoluted IRMPD mass spectrum of doubly charged bradykinin (exp. *m/z* 860.718) precursor with an IP6 adduct (charge-deconv. *m/z* 1720.428, pink). The bare precursor, *m/z* 1060.566, is marked in orange. The peaks immediately to the left of the *m/z* 1720.428 peak correspond to precursor ions retaining partially photo-fragmented IP6 adducts. The *m/z* gap between the fragments with IP6 and the bare precursor allows for easy rejection of ions carrying IP6 fragments (b) Annotated bradykinin structure with y^+ ions in red and b^+ ions in blue. (c) Relative abundances of precursor with IP6 (red), precursor without IP6 (blue), and sequence informative fragments (green) upon IRMPD of bradykinin with and without IP6 adduct at different pulse energies. It is interesting to note that high-energy excitation (4.56 J/pulse) primarily leads to IP6 ejection as a neutral while low energy excitation enables efficient intermolecular energy redistribution in the non-covalent complex leading to backbone fragmentation. ~1000 scans, 1 laser pulse of 80 ms per scan, $7.18 \cdot 10^{-11}$ mbar N_2 UHV readout, 140000 @ *m/z* 200 resolution.

Proteins complexing IP6

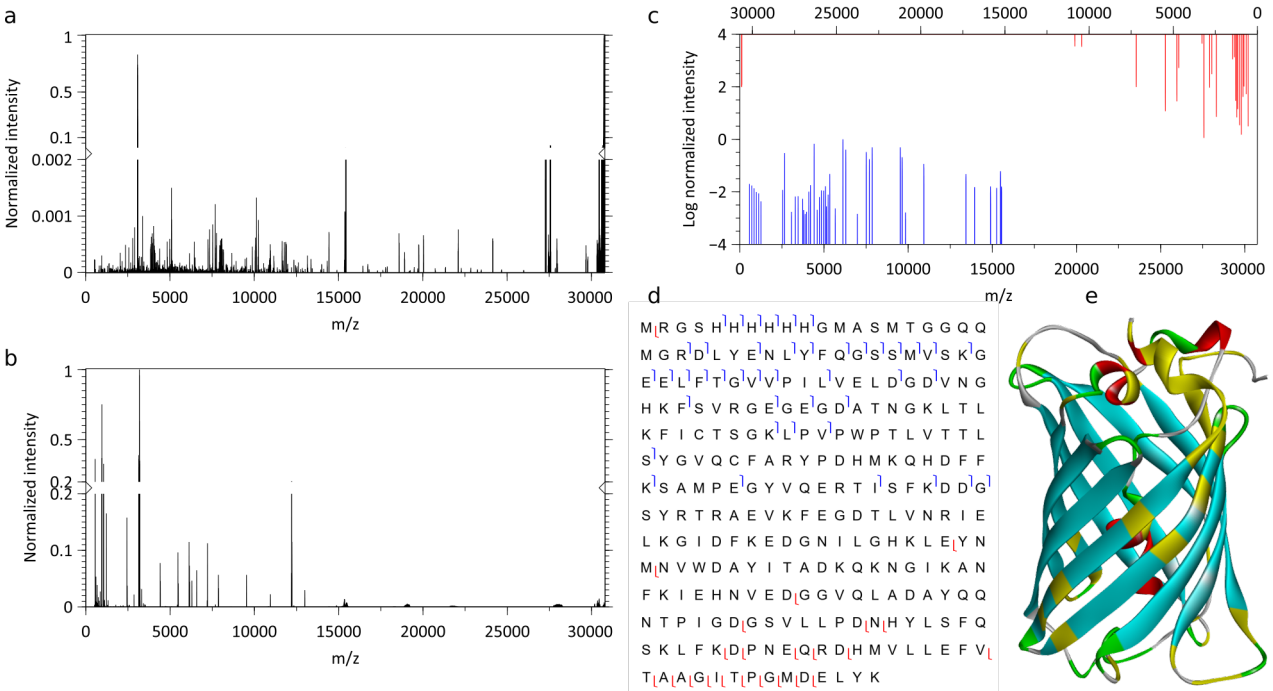

Fig. S9. IRMPD of native mAmetrine retaining its pBAD purification tag. (a) Charge-deconvoluted IRMPD mass spectrum of bare mAmetrine [10+], (b) annotated charge-deconvoluted MS spectrum of the mAmetrine protein with IP6 adduct [10+], note the different scales of the y-axis in (a) and (b). For the IRMPD of the mAmetrine protein with IP6 adduct are depicted in (c) the observed (b, y) fragment ions with y^+ ions in red and b^+ ions in blue, separately normalized. (d) The sequence coverage, and (e) homology model (without the pBAD tag) produced from the template PDB ID: 1FSJ with the cleavages in yellow. 1000 scans, one pulse of 4.56 J over 80 ms per scan, $6.93 \cdot 10^{-11}$ mbar N_2 UHV readout, 140000 @ m/z 200 resolution.

Phosphoprotein Bora

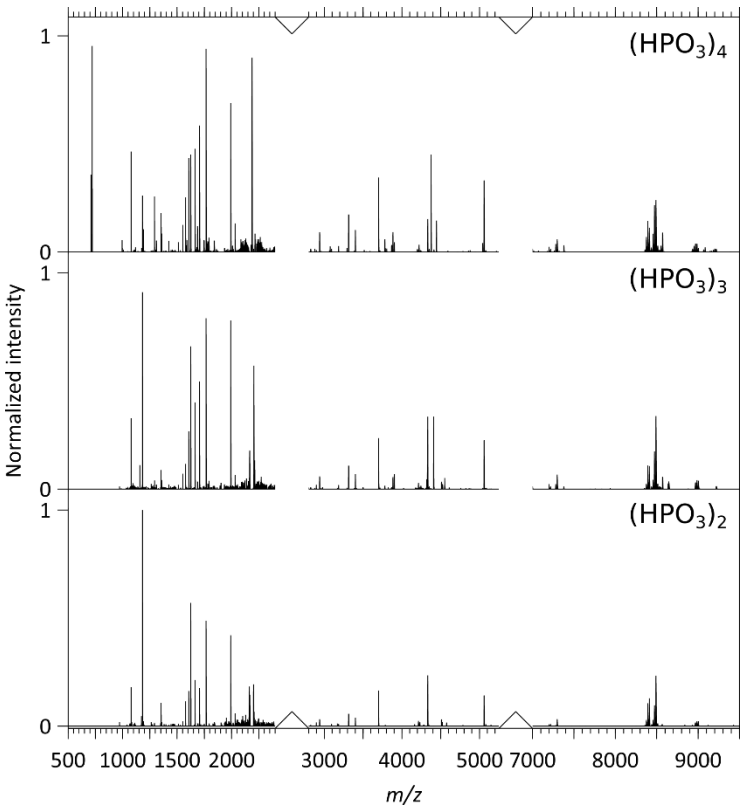

Fig. S10. Charge deconvoluted IRMPD mass spectra of phosphorylated Bora highlighting differences between the carried numbers of phospho-groups.

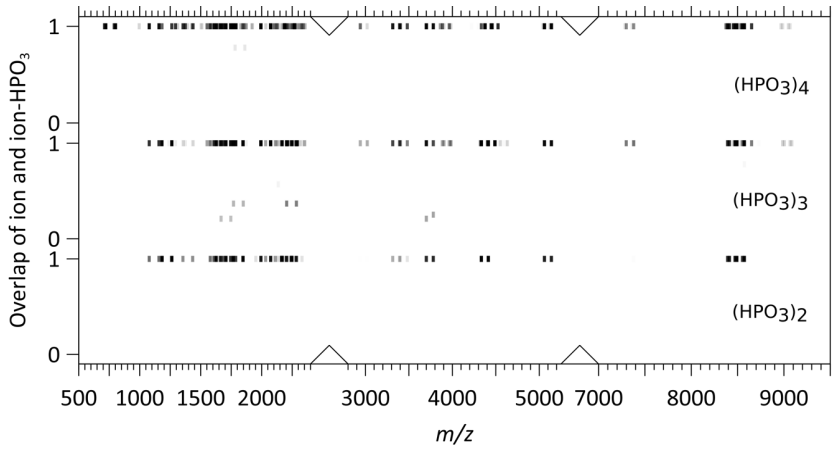

Fig. S11. IRMPD of the phosphorylation products of Bora (Fig. S10): identification of fragments carrying phospho-group. All three species, $(HPO_3)_4$, $(HPO_3)_3$, $(HPO_3)_2$, coexist in solution and likely corresponds to different, coexisting subsets of phosphorylated sites. IRMPD contributing to partial HPO_3 loss, IRMPD-produced fragments (of a precursor defined by a given number of phospho-groups) likely carry a number of phosphate groups ranging from 0 to 2, 3, or 4 depending on the precursor. As illustrated by Fig. S10, assigning all the fragments under such conditions is a complex and time-consuming task. We provide an assessment of the fragment ions carrying phosphate by looking specifically for a HPO_3 loss (mass difference of 79.966 Da). Detection involves computing the overlap between the isotopic distributions of a given spectrum and of the same spectrum shifted by the exact mass of a HPO_3 (79.966 Da). To account for mass calibration errors, the initially centroided isotopic peaks of the charge deconvoluted spectra are convoluted with a Gaussian corresponding to a 5 ppm standard deviation. The overlaps computed for each isotope distribution of the spectrum are then clustered according to their value. A perfect match between the peaks of isotopic distributions separated by 79.966 m/z corresponds to an overlap of 1 (while no match = 0) while the product of the ion pair intensities is displayed using a logarithmic gray scale covering 3 orders of magnitude (black = normalized squared intensity of 1, transparent = normalized squared intensity of 0.001).

Large protein complexes

Table S2. List of compounds detected upon IRMPD of SCRI104 CRISPR-Cas Csy complex in Fig. 6. Inset naming within Fig. 6, main CS corresponds to the charge state used to identify the compound in the figure, CS range spans the most abundant charge state detected.

Inset	Main CS	Name	Mass (Da)	CS range
a	38^+	CRISPR-Cas Csy complex (pr.)	346962±30	[42^+-35^+]
	21^+	Csy3	36778.2±1.2	[24^+-13^+]
	17^+	Cas6f	20458.57±0.85	[20^+-13^+]
	18^+	subunit fragment	20008.84±0.70	[18^+-13^+]
	40^+	pr.–Cas6f (high m/z CS distribution)	326476±24	[40^+-36^+]
	30^+	pr.–fragment	333580±38	[33^+-25^+]
	23^+	pr.–Cas6f (low m/z CS distribution)	326476±24	[26^+-19^+]
	19^+	pr.–Csy3	310152.9±8.5	[23^+-16^+]
b	38^+	CRISPR-Cas Csy complex (pr.)	346907±22	[41^+-34^+]
	21^+	Csy3 dimer	73553.7±2.2	[23^+-13^+]
	20^+	Csy1	51551.6±1.1	[23^+-19^+]
	15^+	Csy3	36776.4±0.8	[20^+-6^+]
	10^+	Cas6f	20456.9±0.8	[12^+-8^+]

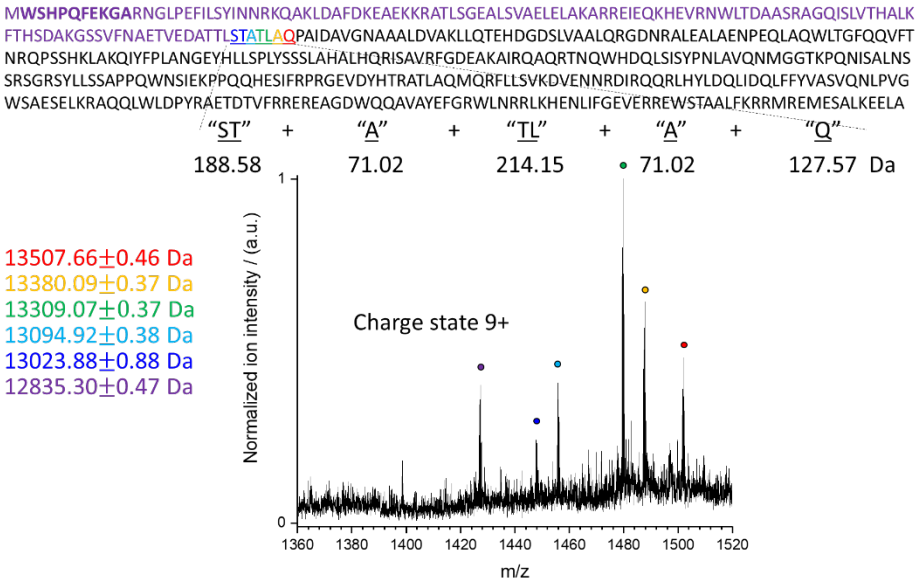

Fig. S12. IRMPD backbone fragmentation of the *P. atrosepticum* SCRI104 CRISPR-Cas Csy complex. An isotopically resolved continuous 9^+ ion ladder corresponding to backbone-cleavages of the Csy1 subunit can be found in the [1400-1520] *m/z* range.

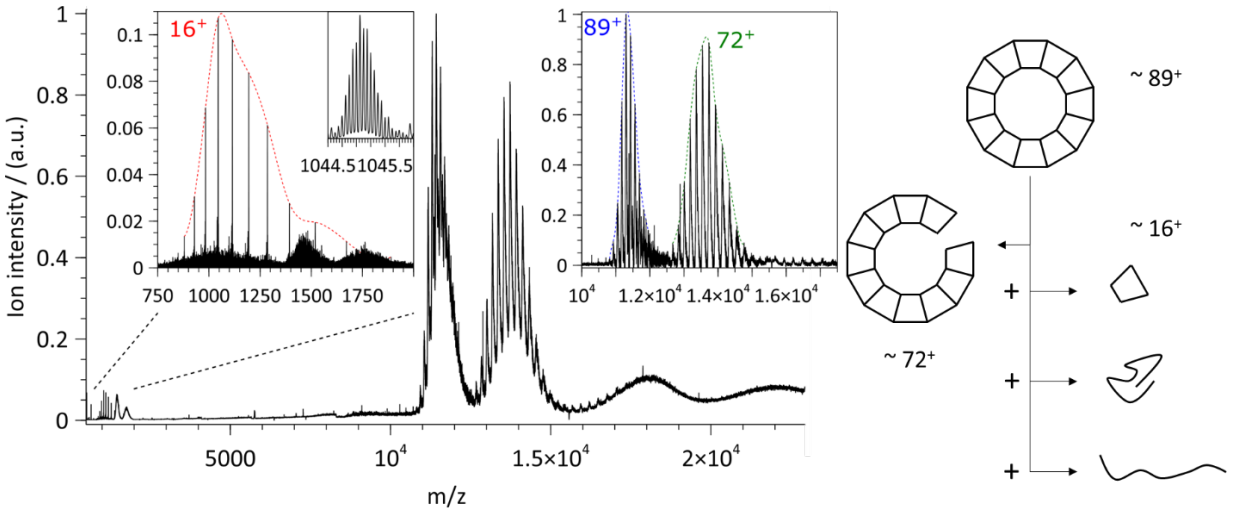

Fig. S13. IRMPD of wt-AaLS. The units are assembled into 12 subsets of 5, this gives ~7 to 8 charges per subset. Each unit is in contact with the units of another subset. Upon ejection of a subunit, it appears to take all the charges of the subset it belongs to as well as those of the neighboring one.