Characterization of the T4gp32-ssDNA complex by native, cross-linking, and ultraviolet photodissociation mass spectrometry

Molly S. Blevins[1], Jada N. Walker[1], Jeffrey M. Schaub[2], Ilya J. Finkelstein[2], Jennifer S. Brodbelt[1*]

[1]Department of Chemistry, [2]Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, United States

[*]Corresponding author: jbrodbelt@cm.utexas.edu

Supporting Information

Table of Contents

Figure S1. SDS-PAGE gel of ~33.5 kDa gp32 protein, and ~28.5 kDa gp32-ΔCTD protein.

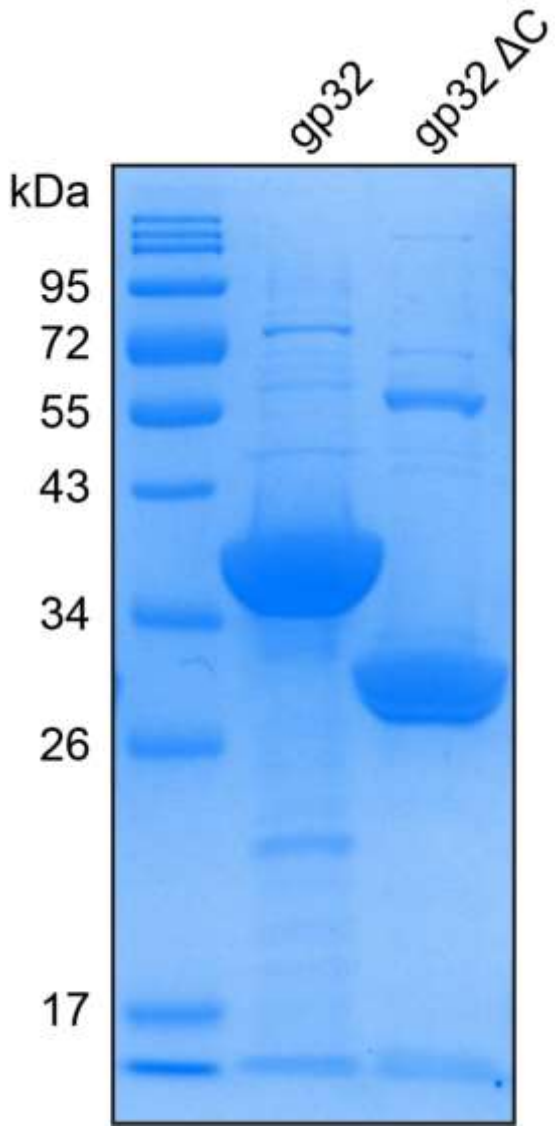


Figure S2. Optimization of UVPD conditions: Distributions of fragment ion sizes upon UVPD (1 pulse at 1 mJ, 2 mJ, or 3 mJ) of gp32 precursor (*m/z* 3061, 11+ charge state, containing 1 Zn) for gp32 apo ions binned by 5000 Da increments (for example, 5 kDa mass bin indicates fragments with masses between 0-5 kDa). Fragment ions do not contain Zn. 1p, 3mJ UVPD conditions generate the highest number of matched ions overall, as well as the highest number of matched ions in five of the seven mass bins, thus indicating the potential to obtain a large array of gp32 ions from solely the 1 pulse, 3 mJ UVPD condition, while simultaneously decreasing the rate of false positives by minimizing the amount of UVPD data that is combined.

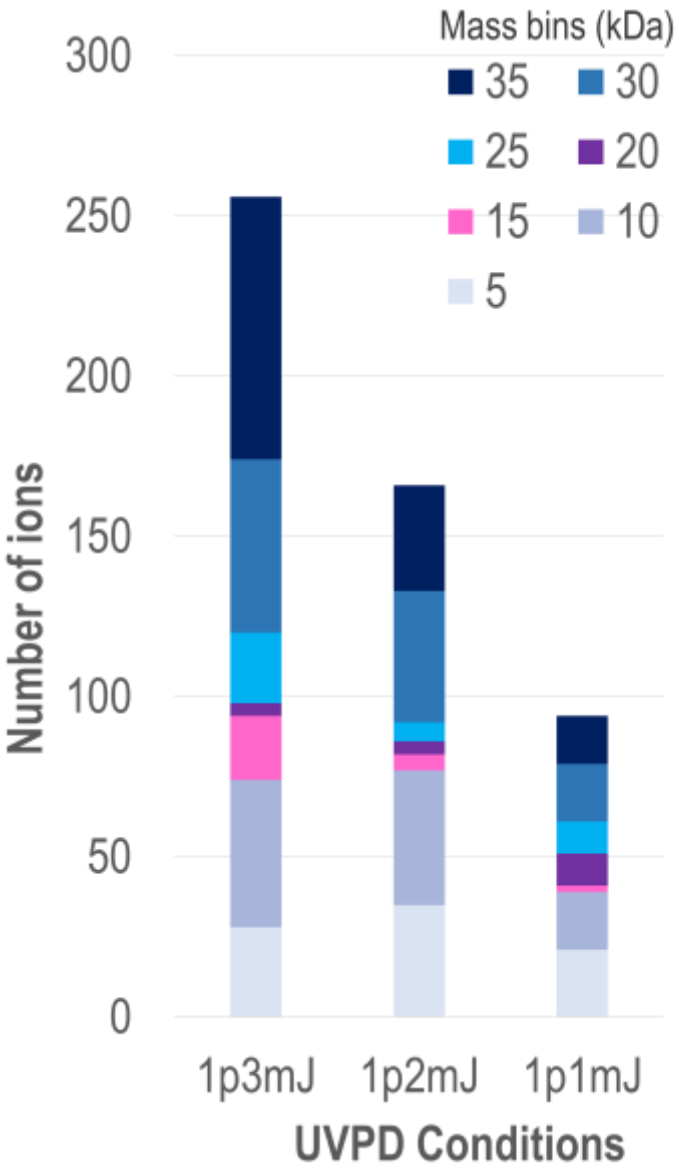


Figure S3. Representative UVPD mass spectra of gp32 (11+, *m/z* 3061) acquired using **(a)** 1p, 2mJ and **(b)** 1p, 1mJ, and corresponding deconvoluted spectra: **(c)** 1p, 2mJ and **(d)** 1p, 1mJ. Spectra collected at 1p, 3mJ are shown in Figure S6.

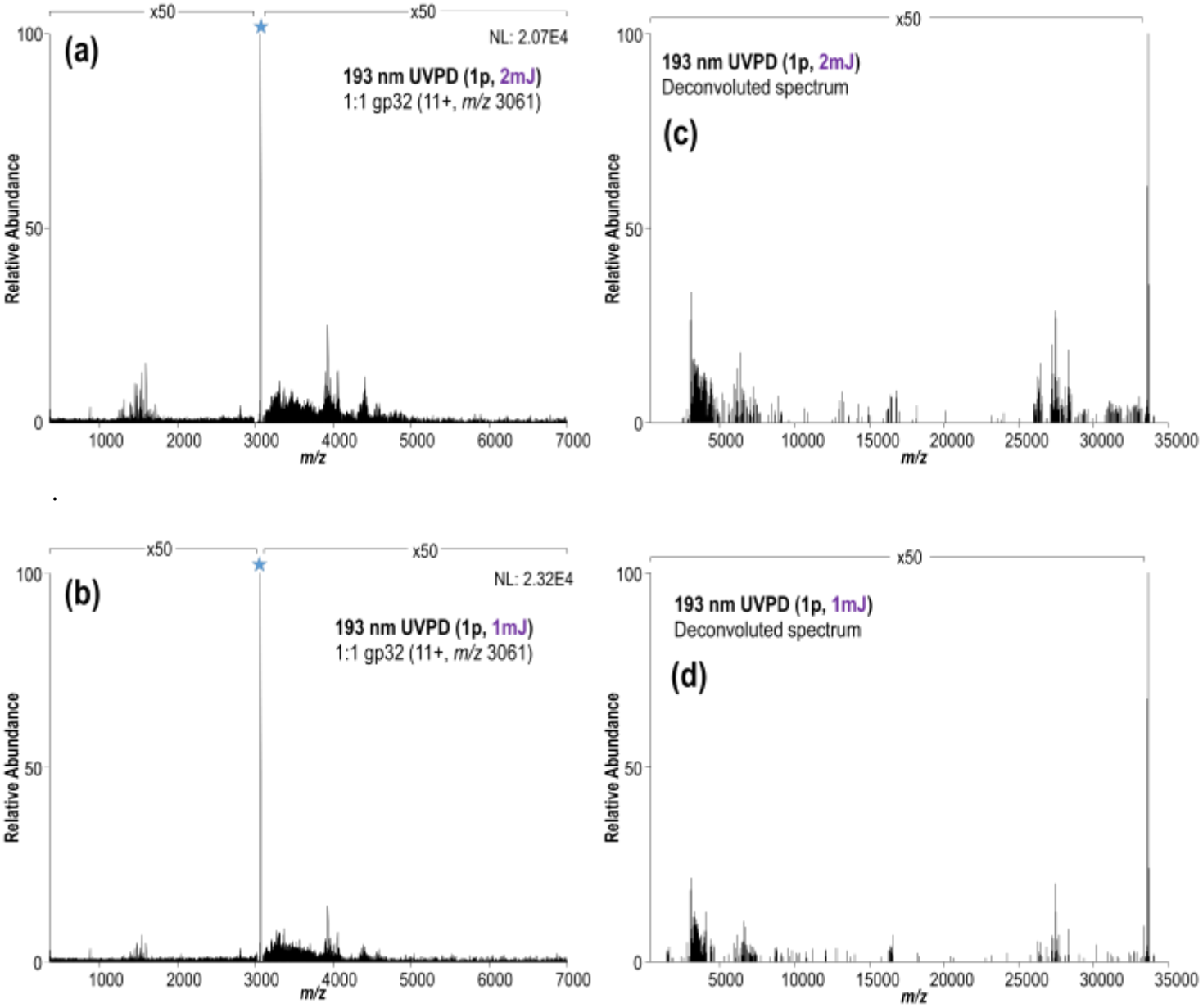


Figure S4. Native MS1 spectra of **(a)** dT12 and **(b)** dT20 showing the 3- and 5- charge states as most abundant, respectively.

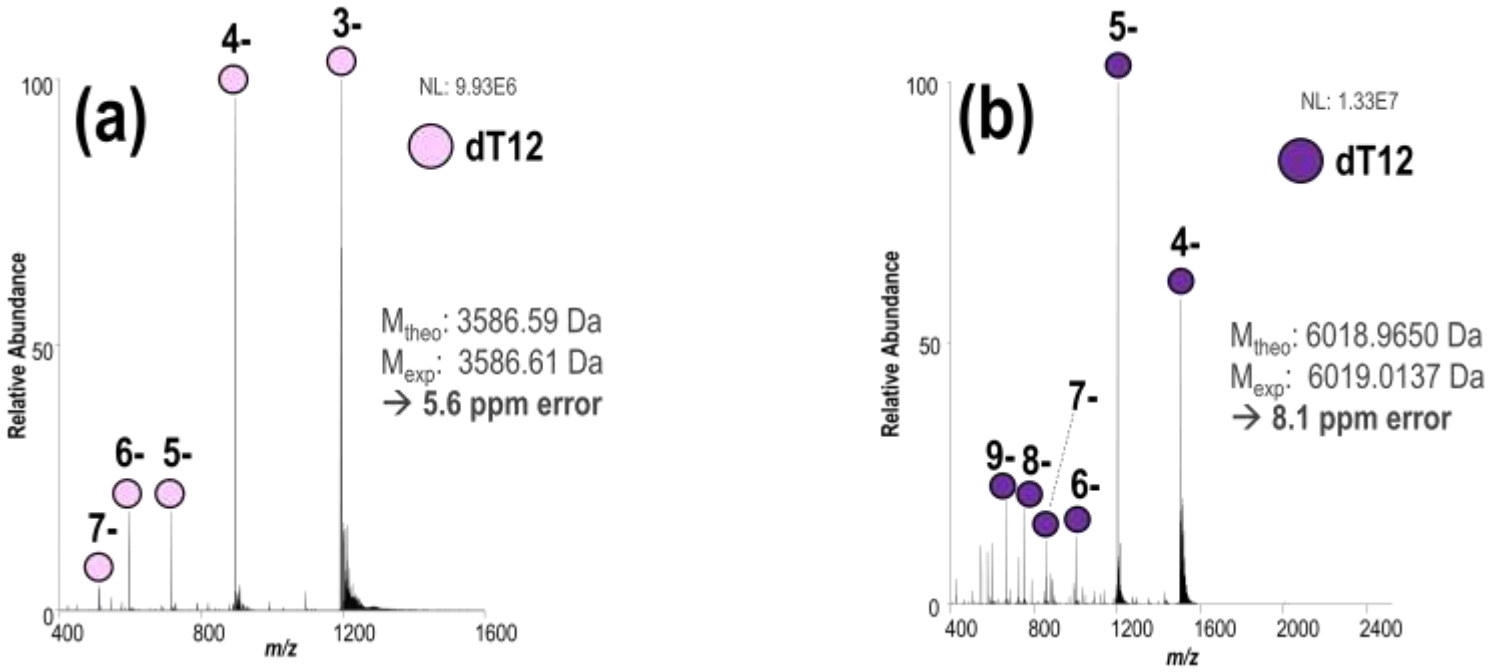


Figure S5. Deconvoluted low-resolution native MS1 spectrum of gp32 showing both monomer and dimer ion peaks.

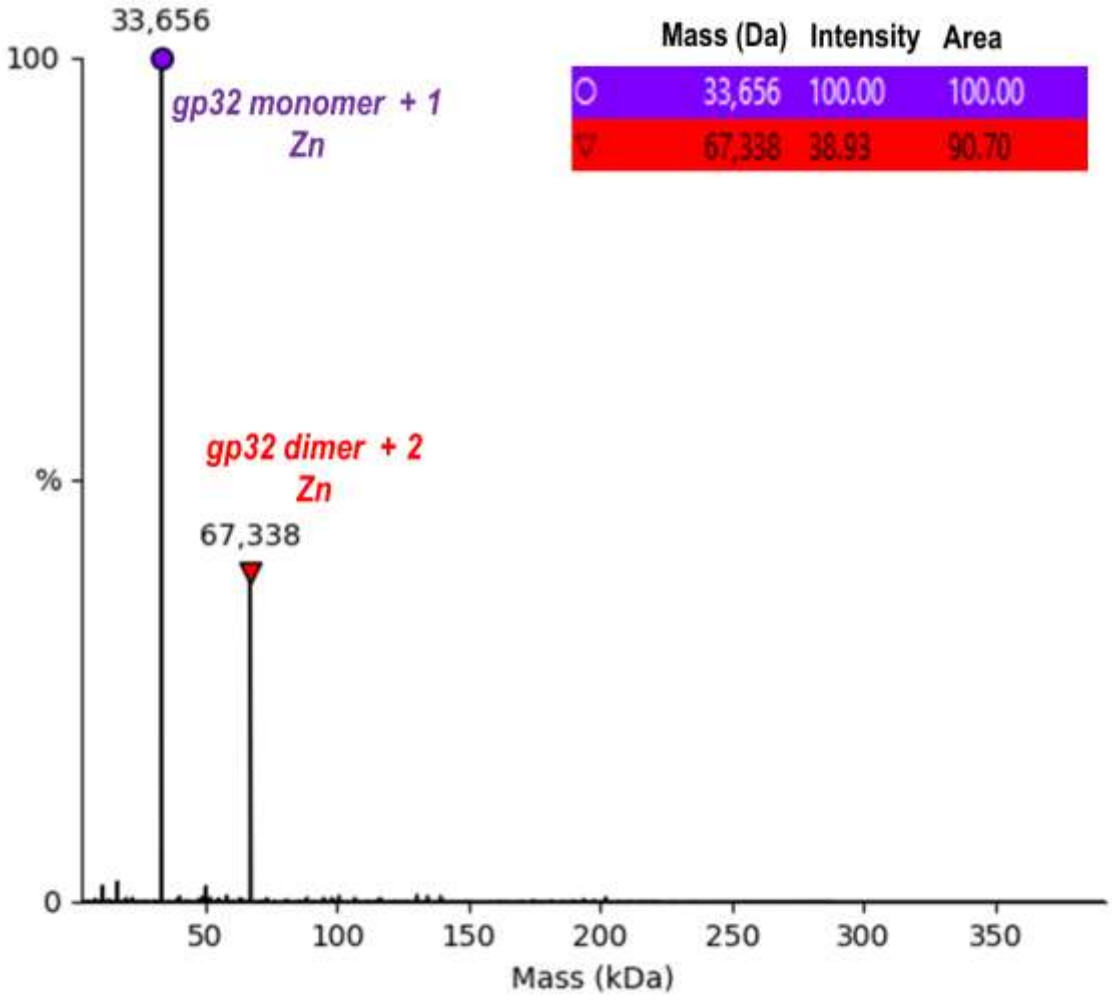
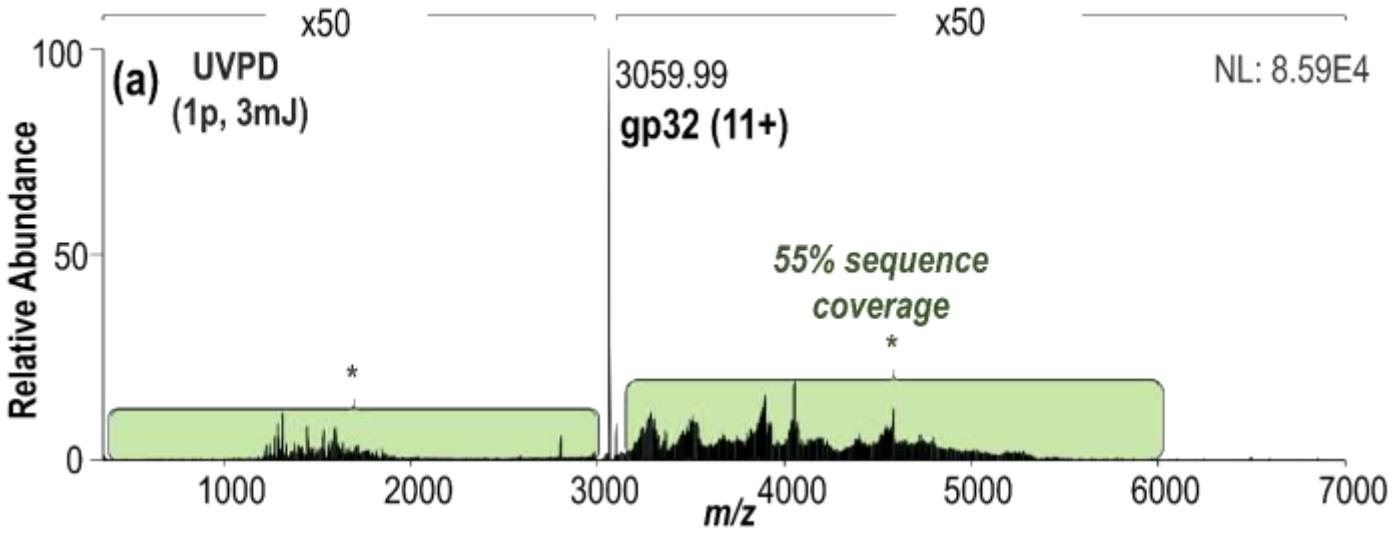


Figure S6. (a) Expanded view of 193 nm UVPD mass spectrum (1p, 3mJ) of native gp32 (11+, with one Zn), (b) deconvoluted UVPD mass spectrum with sequence coverage map from triplicate UVPD data (including apo ions (no Zn^{2+}) and Zn^{2+}-containing holo ions).



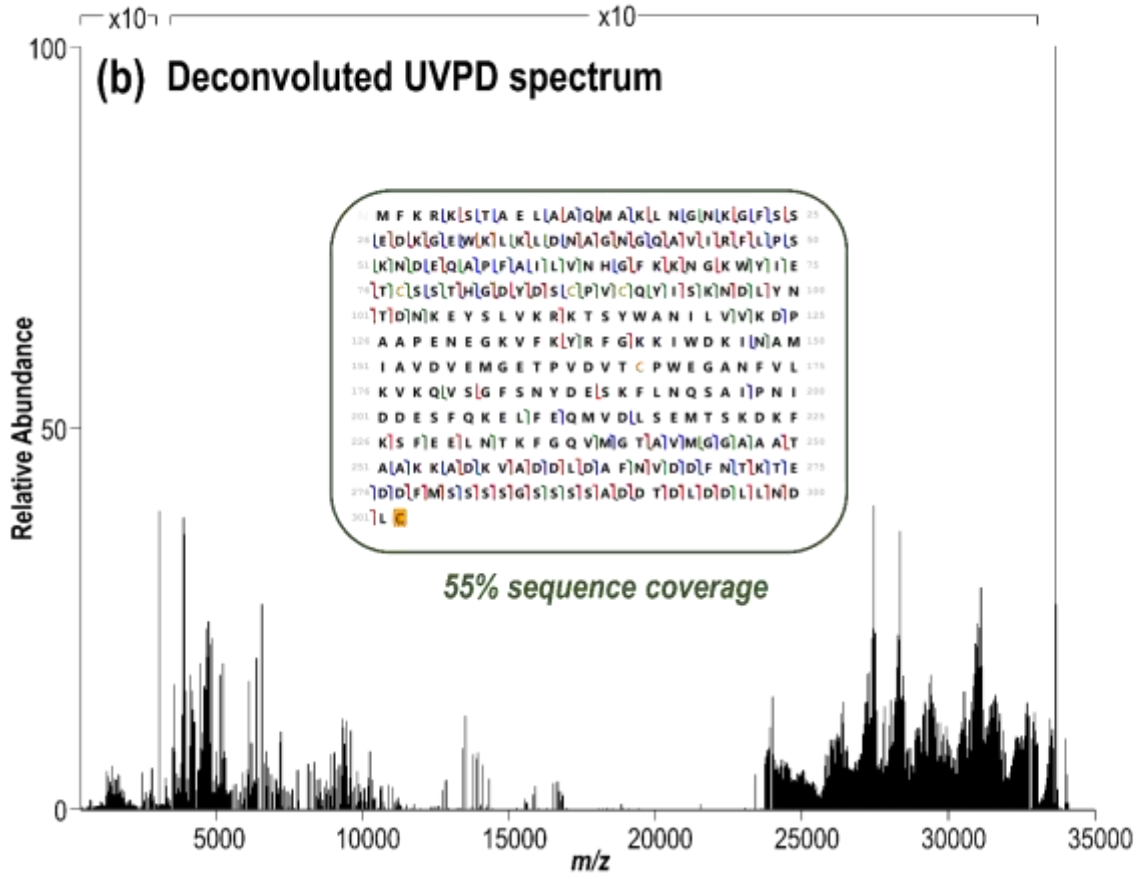


Figure S7. (a) HCD mass spectrum (CE = 130) of native gp32 monomer (11+, containing one Zn), **(b)** deconvoluted HCD spectrum with sequence coverage map.

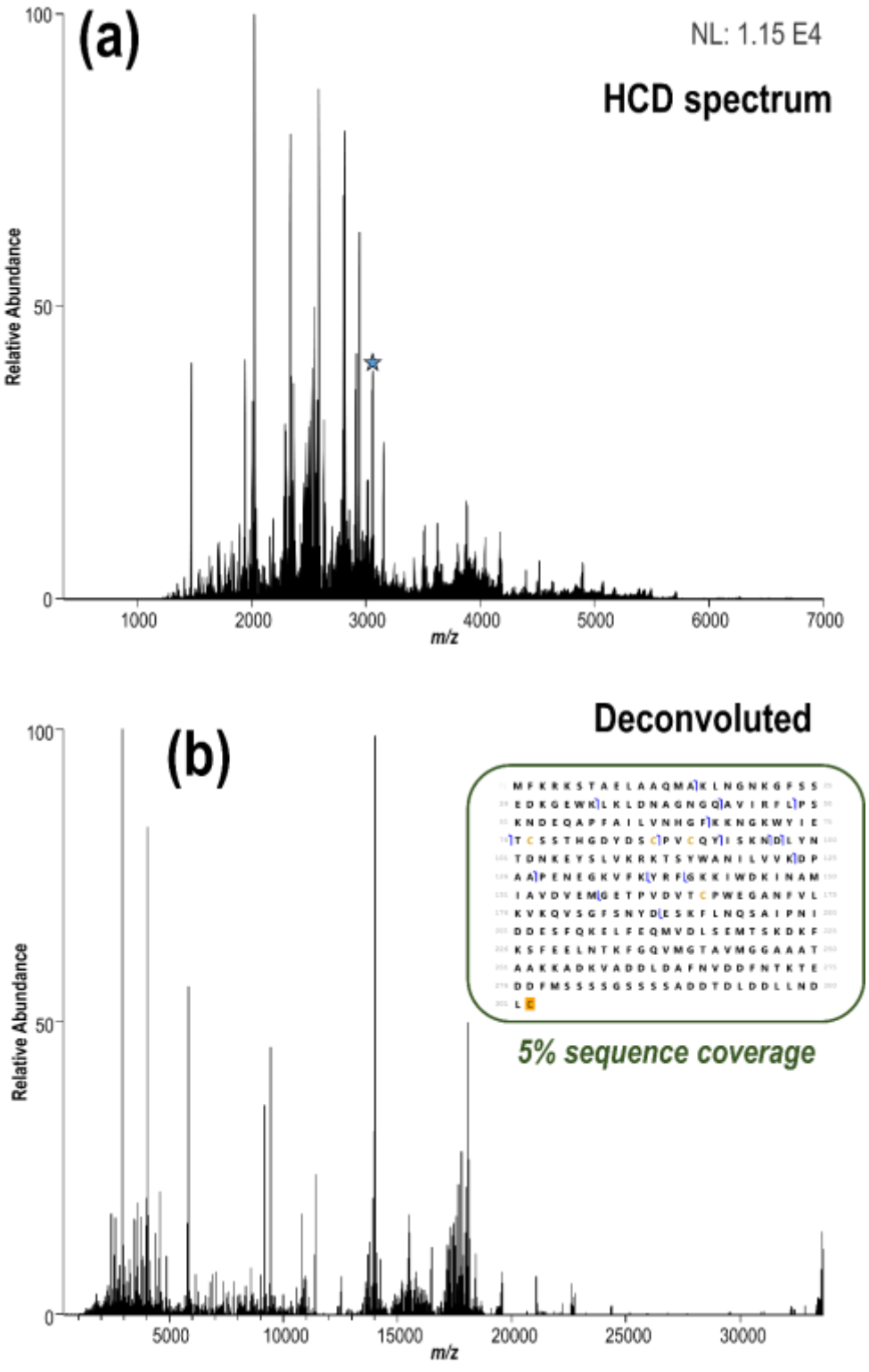


Figure S8. Comparison of sequence coverage of gp32 monomer for the **(a)** 10+ and **(b)** 11+ precursor for a single UVPD experiment using 1 pulse, 3 mJ

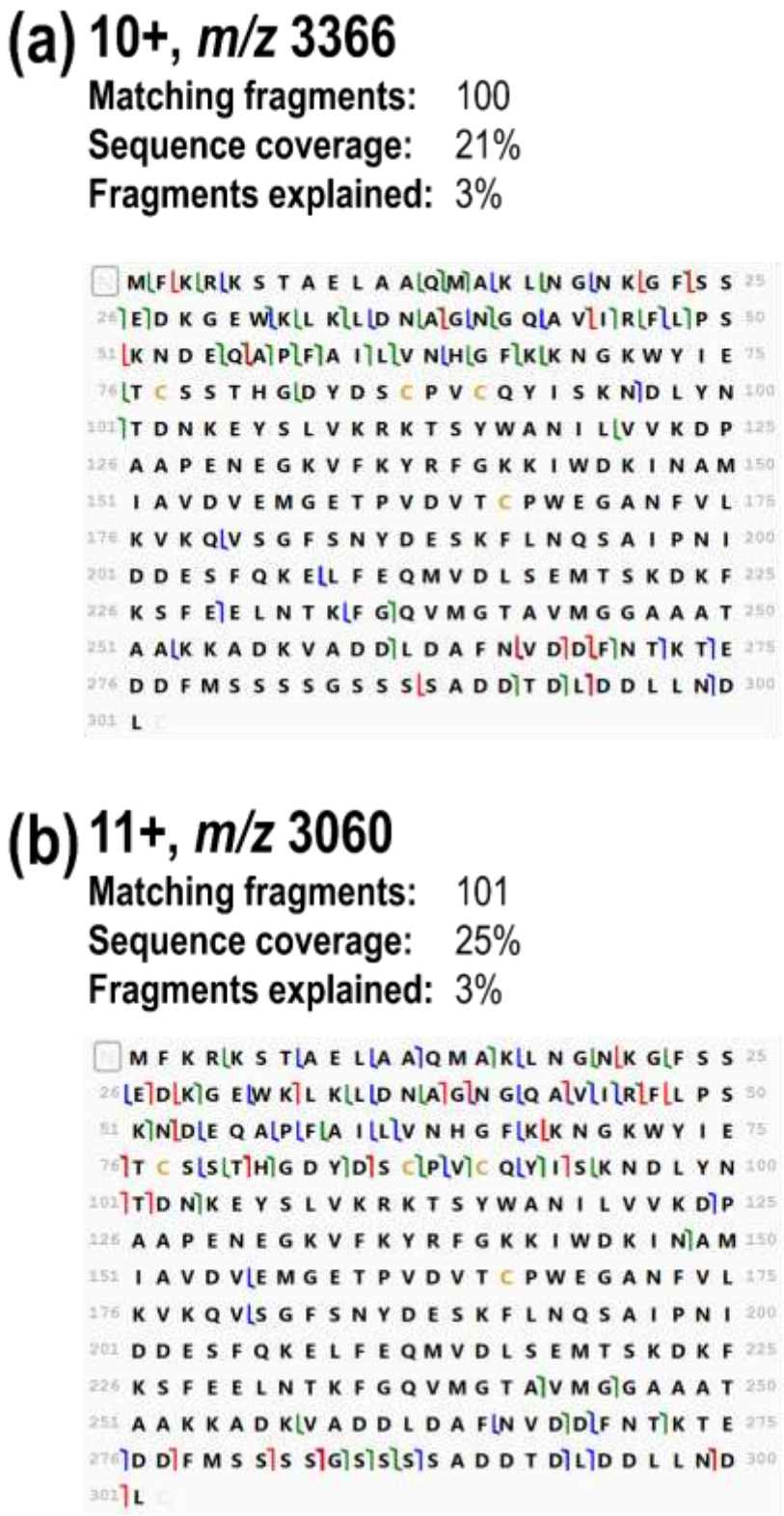
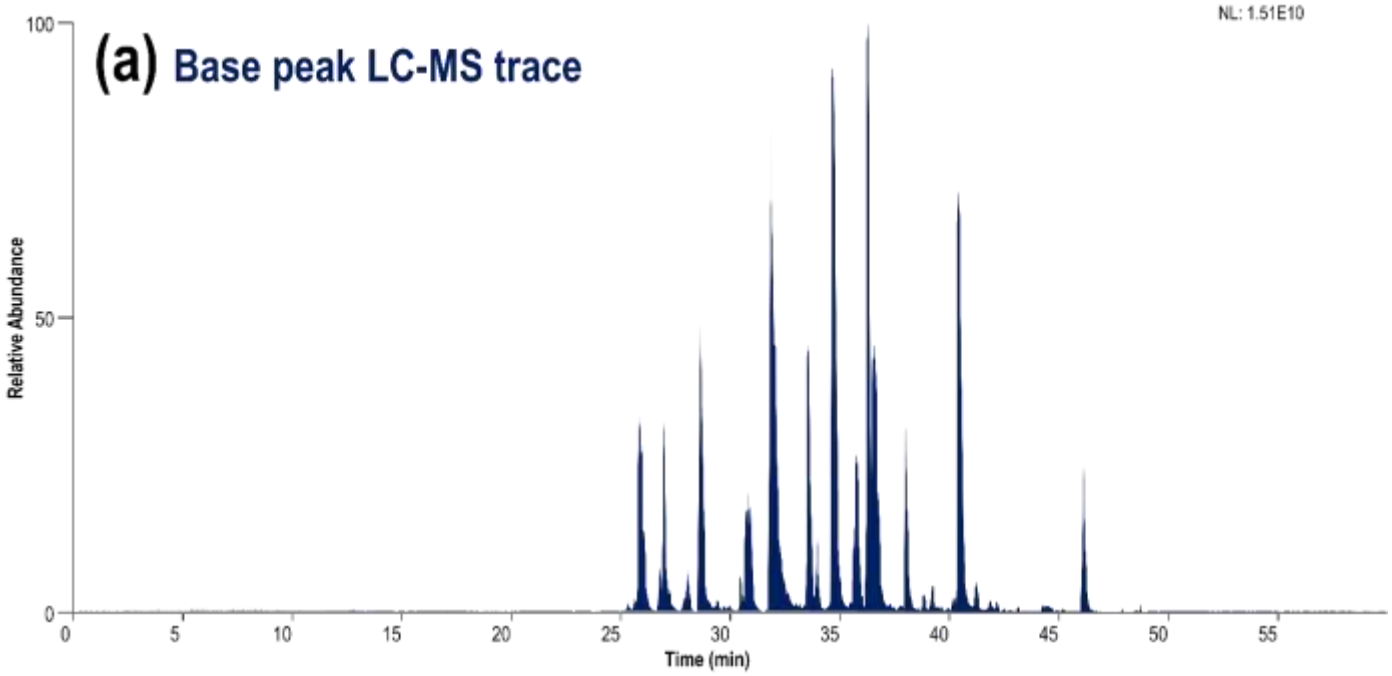


Figure S9. (a) Base peak C18 nanoLC-MS trace of bottom-up analysis of tryptic digest of gp32, **(b)** sequence coverage output from Byonic. Bottom-up sequence coverage both confirms the expected gp32 sequence and rules out a PTM (i.e., ~80 Da phosphorylation), thus localizing the 79.94 Da modification to the C-terminus of the protein.



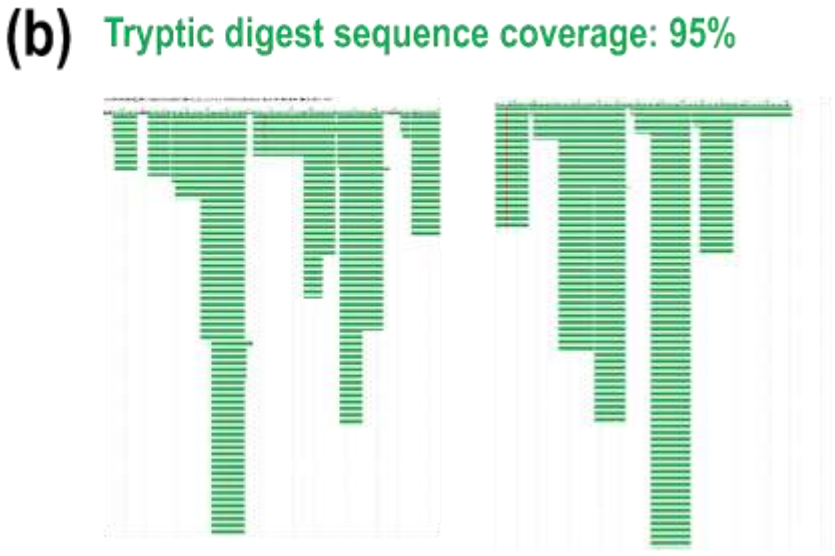


Figure S10. (**a)** MS1 spectrum of denatured gp32 with corresponding **(b)** UVPD (25+) and **(c)** HCD (25+) mass spectra and sequence coverage maps, **(d)** MS1 spectrum of denatured gp32 + TCEP with corresponding (e) UVPD (25+) and (f) HCD (25+) mass spectra and sequence coverage maps. Spectra highlight the presence of a Cys87-90 disulfide bond in the denatured gp32 which is not present in native or TCEP-reduced gp32.

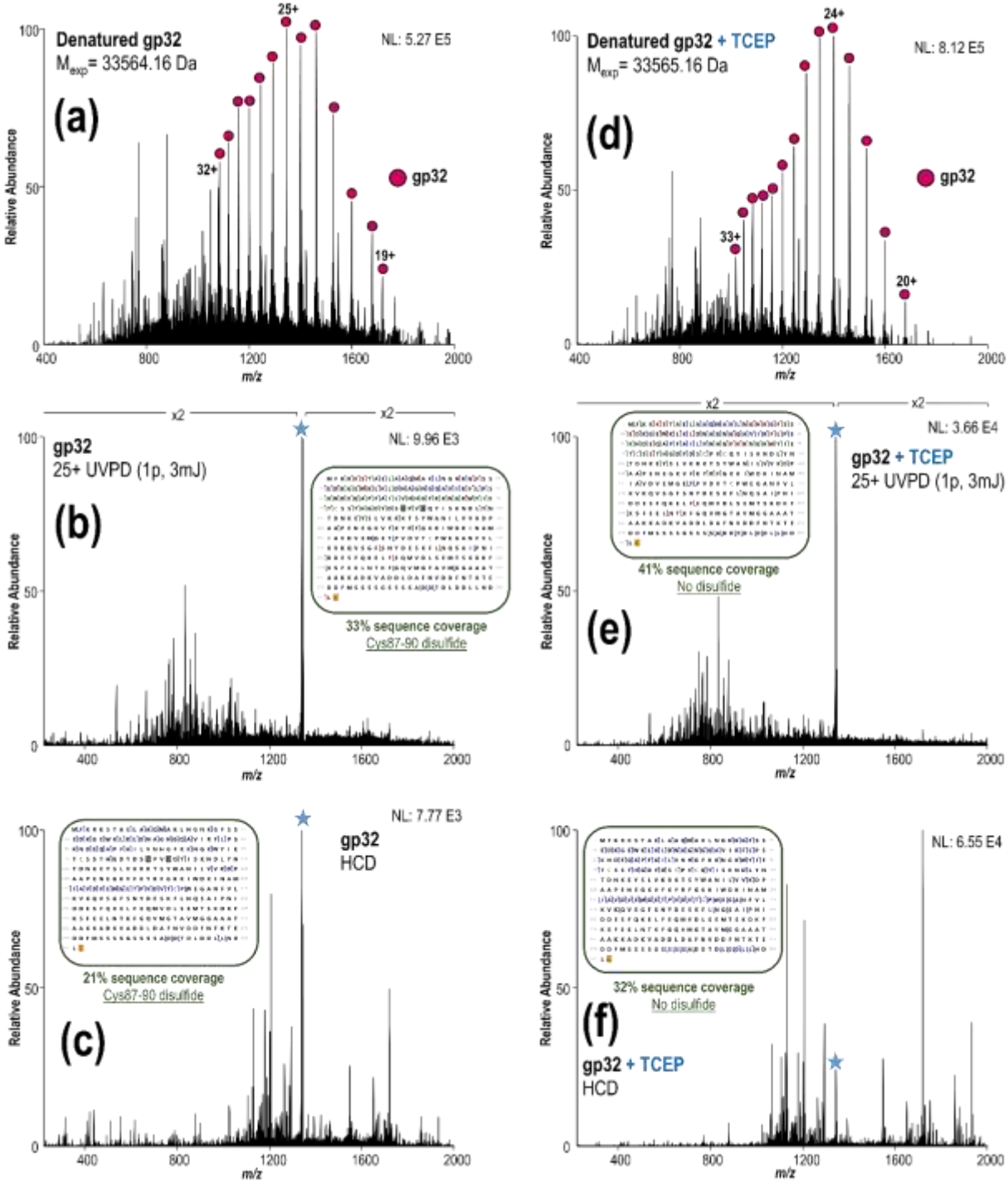
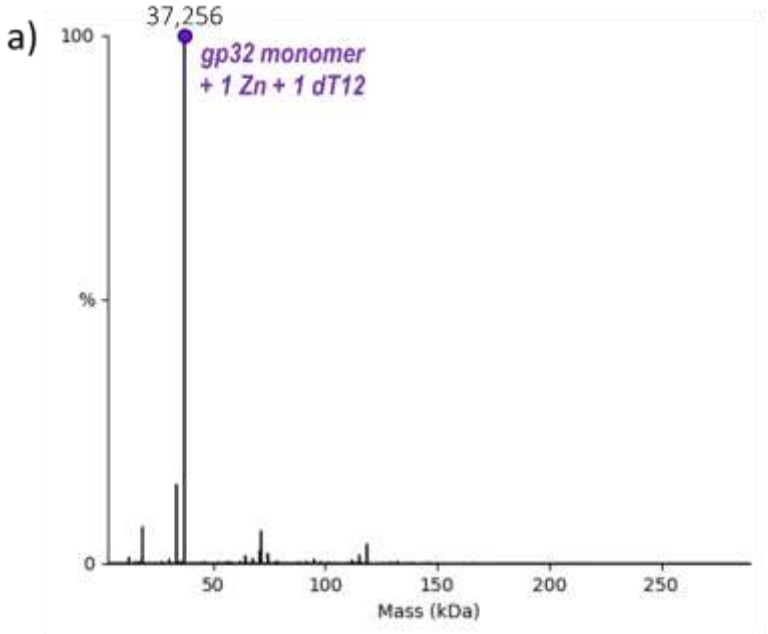


Figure S11. Deconvoluted low-resolution native MS1 spectrum of: **(a)** gp32 + dT12 , and **(b)** gp32 + dT20 solutions.



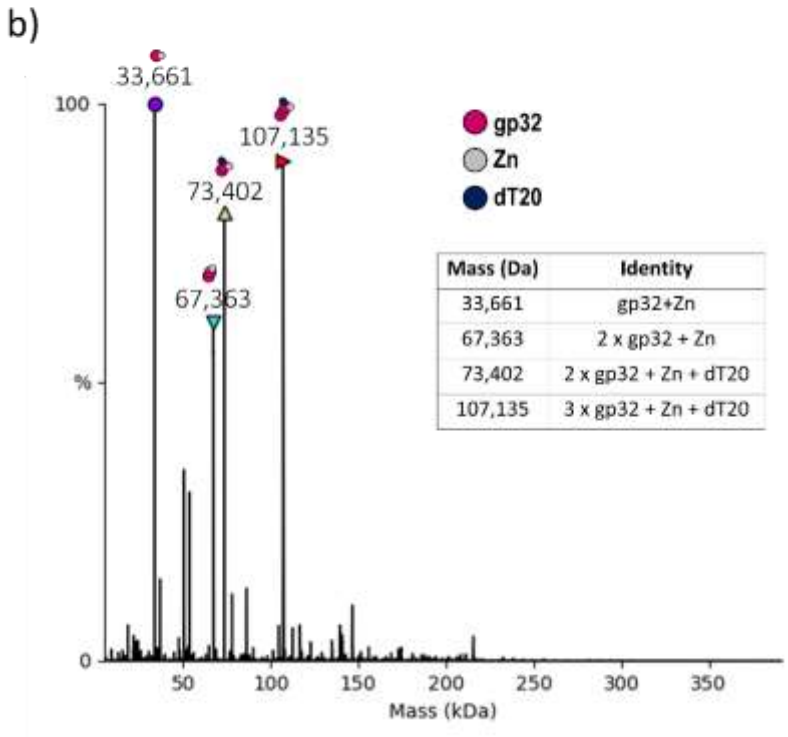
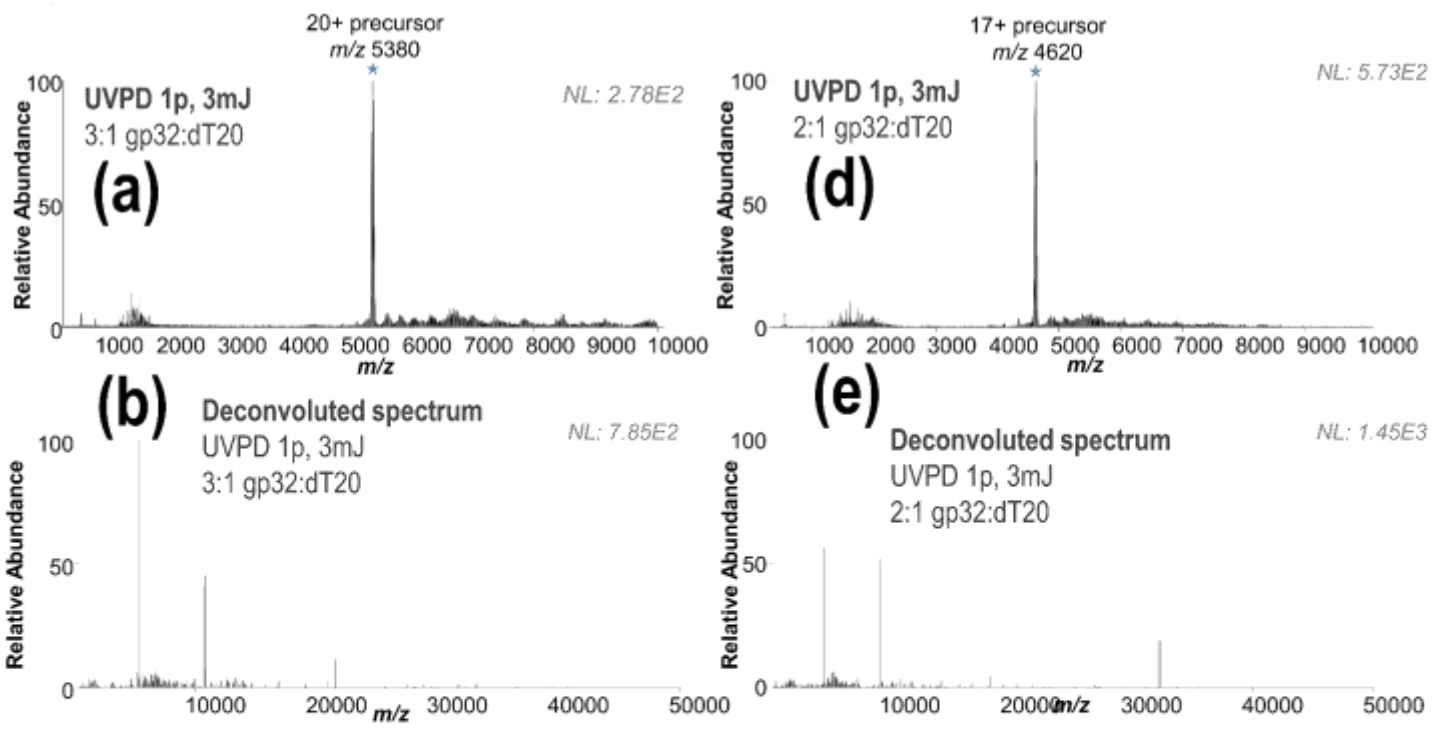
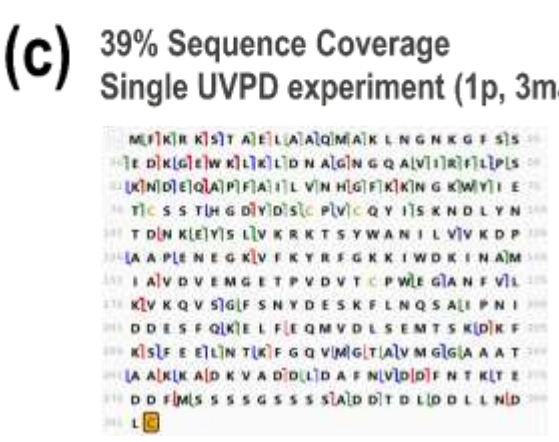


Figure S12. 193 nm UVPD mass spectra of **(a)** 3:1 and **(d)** 2:1 gp32●dT20 complexes with corresponding **(b,e)** deconvoluted mass spectra and **(c,f)** sequence maps. Sequence maps illustrate the backbone cleavages that result in identifiable fragment ions, include apo gp32 fragment ions as well as fragment ions that contain Zn, dT20, or Zn+dT20.





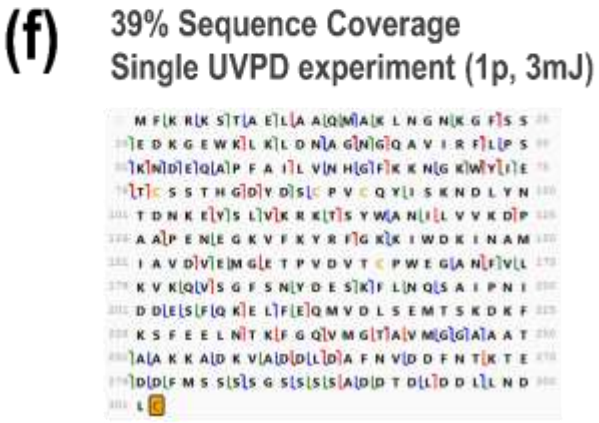


Figure S13. (a-c) Triplicate 193 nm UVPD mass spectra (1p, 3mJ) of native 1:1 gp32•dT12 complex (11+, *m/z* 3368) with **(d-f)** corresponding deconvoluted spectra.

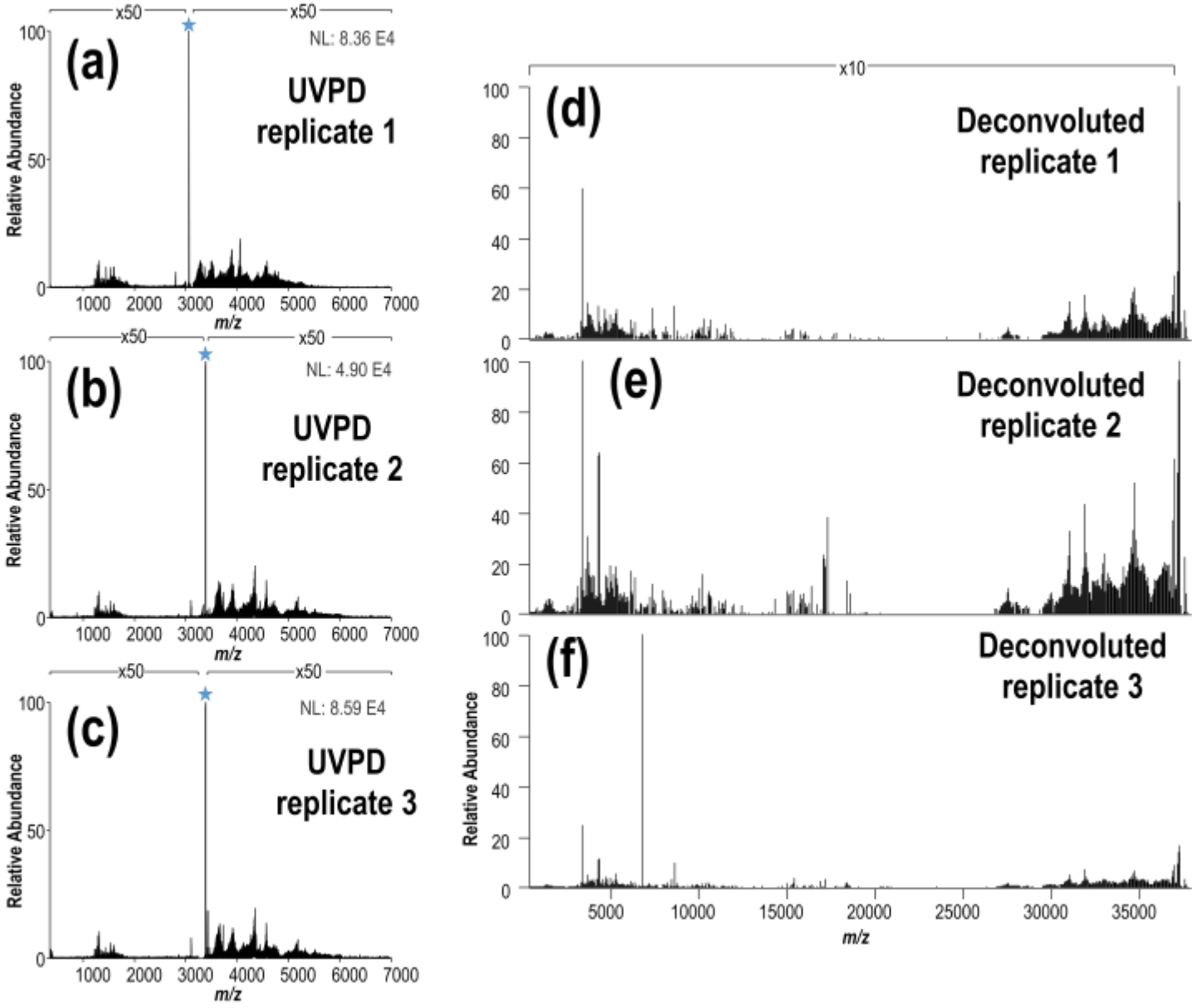
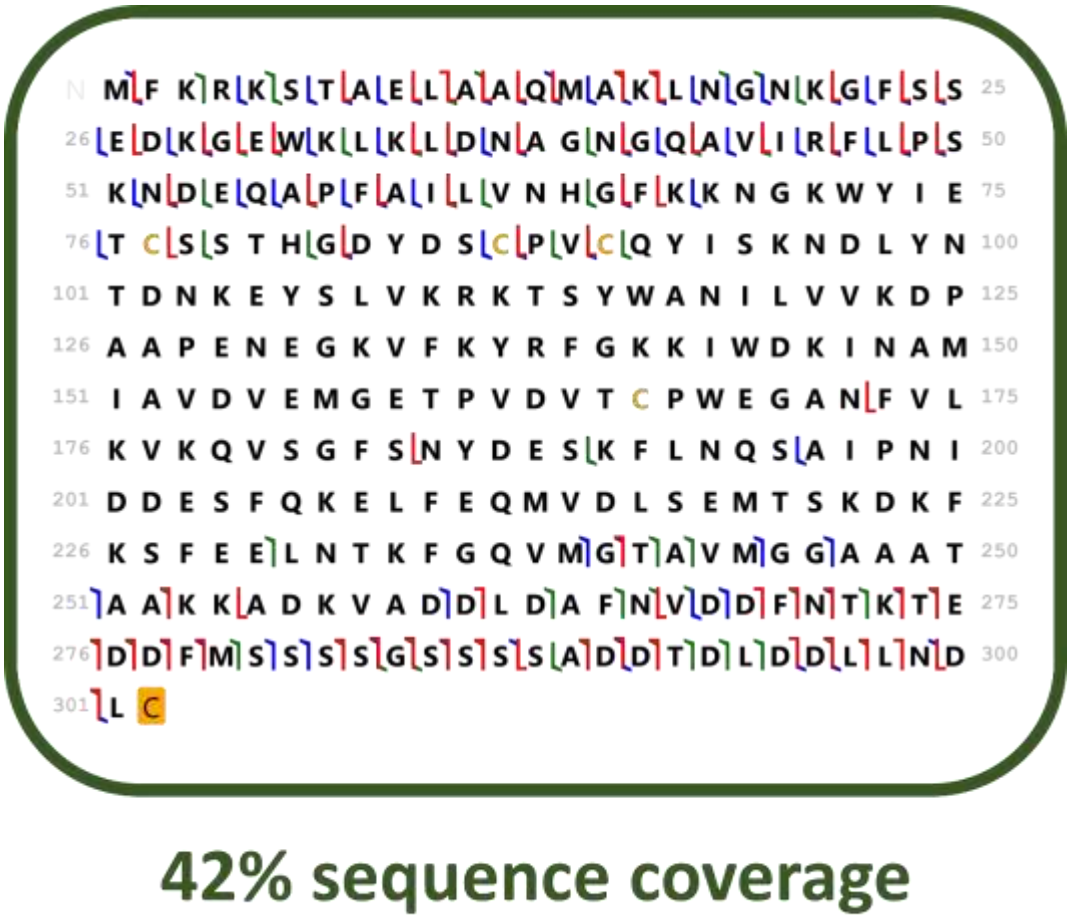


Figure S14. UVPD sequence coverage map for all DNA-containing holo ions (1 pulse, 3mJ) filtered by holo ions identified two or more times within the dataset. Data processed using ProSight Lite.



42% sequence coverage

Figure S15. (a) Sequence of gp32 with the backbone cleavage sites leading to N-terminal ((*a, a+1, b,c*) blue), C-terminal ((*x, x+1, y, y-1, z*) red) or bidirectional (green) dT12-containing holo fragment ions generated upon UVPD (combined data from 1 pulse at 3 mJ) of the 1:1 gp32•dT12 complex *(m/z* 3386, 11+ charge state, containing 1 Zn*).* Greyed out sequence areas correspond to regions of the sequence that are un-characterized in the crystal structure but contain many confirmatory backbone cleavages that lead to assignable sequence ions. The C-terminal cysteine is shaded in gold to denote the disulfur monoxide modification. Arrows indicate selected short dT12-bound holo ions that were confirmed via isotope fitting including **(b)** A5 and **(c)** Y-6. Experimental data is shown in the top panels while theoretical isotope patterns (generated with FreeStyle) are shown in the bottom panels.

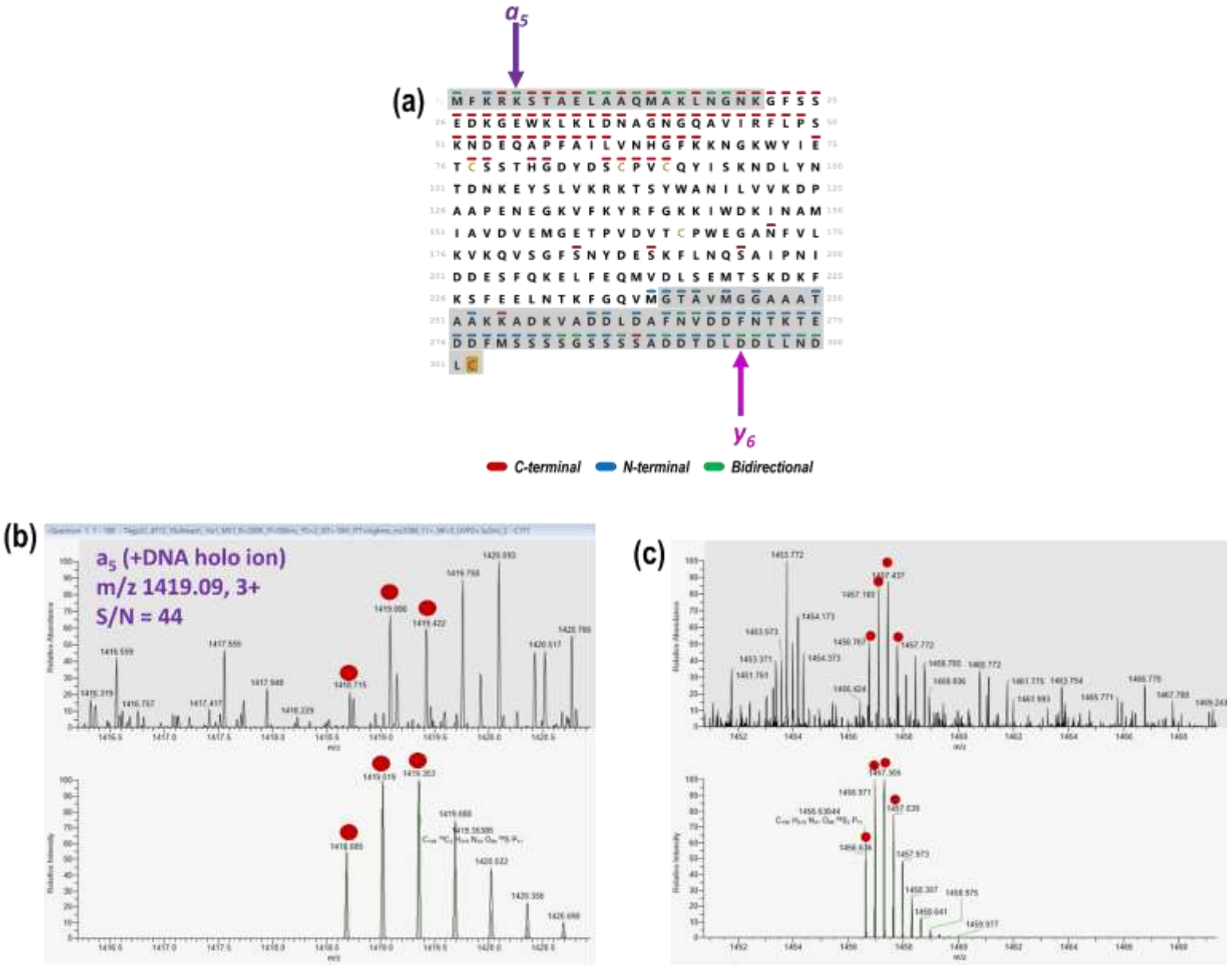


Figure S16. Sequence motif map of gp32 from PDB file 1GPC.

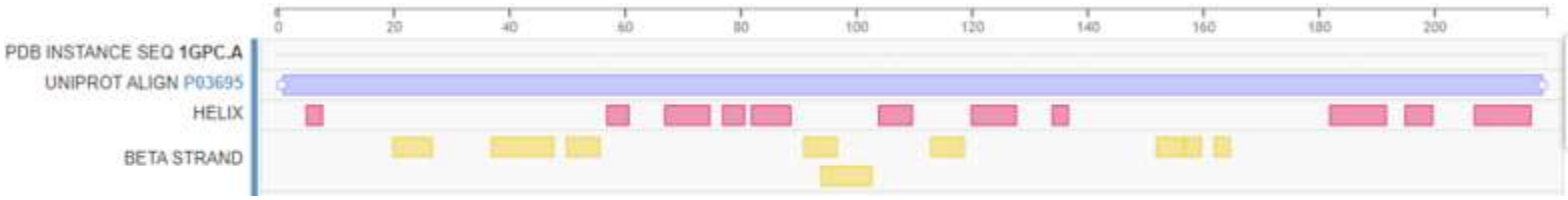


Figure S17. Deconvoluted low-resolution native MS1 spectrum of native cross-linked gp32 homodimer (50 mM AmAc solution, post 30kDa SEC cleanup) with an average mass of 69,755 Da (compared to an average mass of 67,338 Da for non-crosslinked dimer) resulting in a mass shift of 2,417 Da, corresponding to a maximum of 17 crosslinks or a maximum of 15 dead-end modifications (7-8 per monomer subunit) based on a dead-end mass shift of +158 Da and a crosslink mass shift of +138 Da. 4 identified crosslinks indicates 12 dead-end modifications (6 per monomer) along with 4 crosslinks.

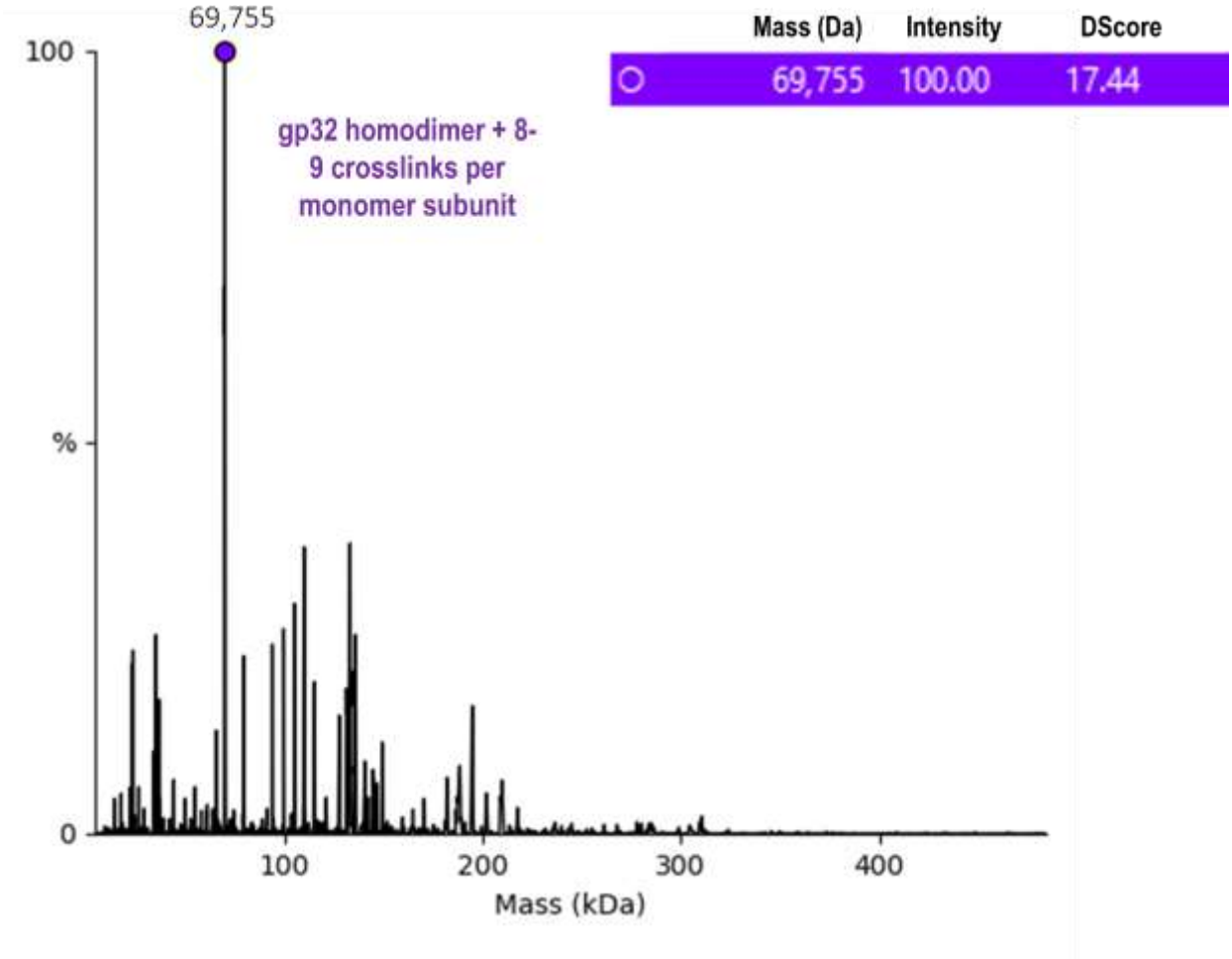


Figure S18. (a) Assignment of peaks using UniDec for low-resolution native (50 mM AmAc, post 30kDa SEC cleanup) MS1 spectrum of cross-linked gp32 + dT12 and **(b)** deconvoluted spectrum.

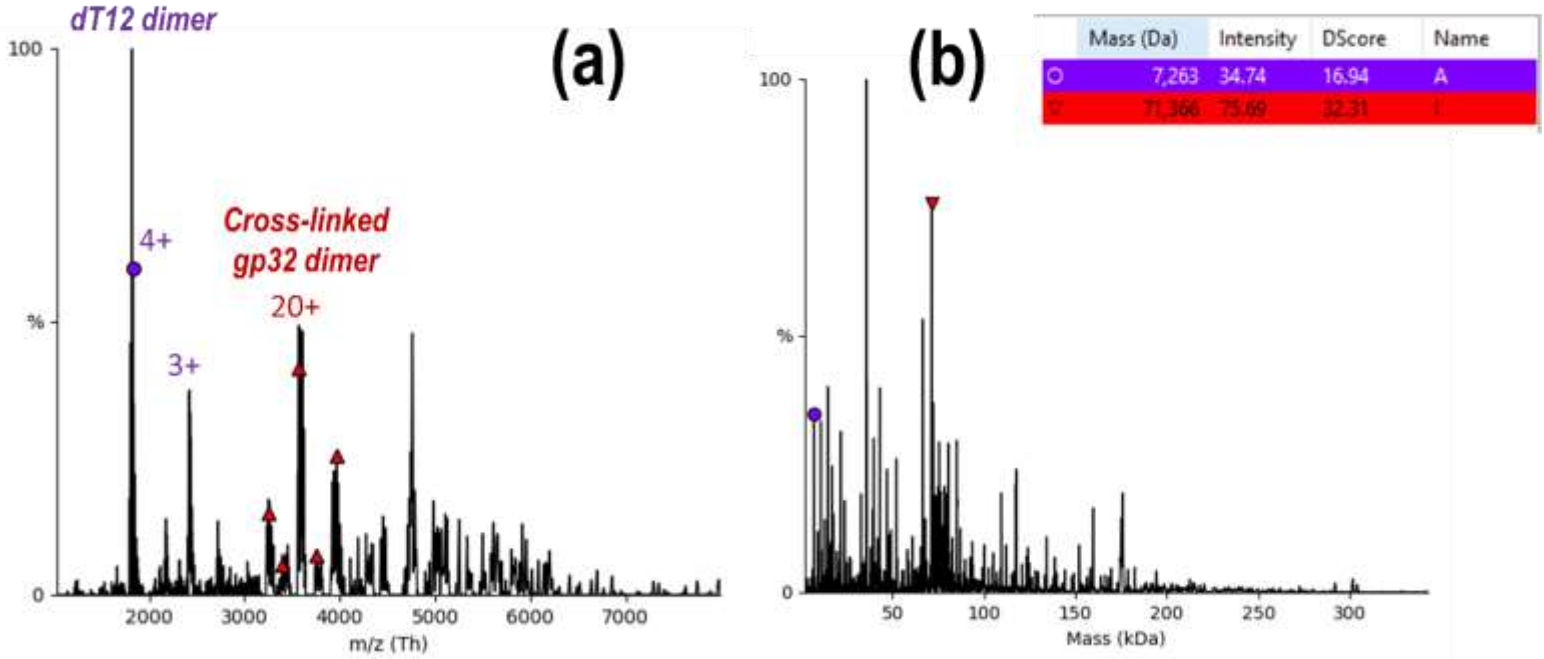


Figure S19. (a) Low-resolution native MS1 spectrum of solution containing dT20 and cross-linked T4gp32 dimer after 30 kDa SEC cleanup and **(b)** deconvoluted spectrum, with no evidence of binding between dT20 and the cross-linked gp32 homodimer.

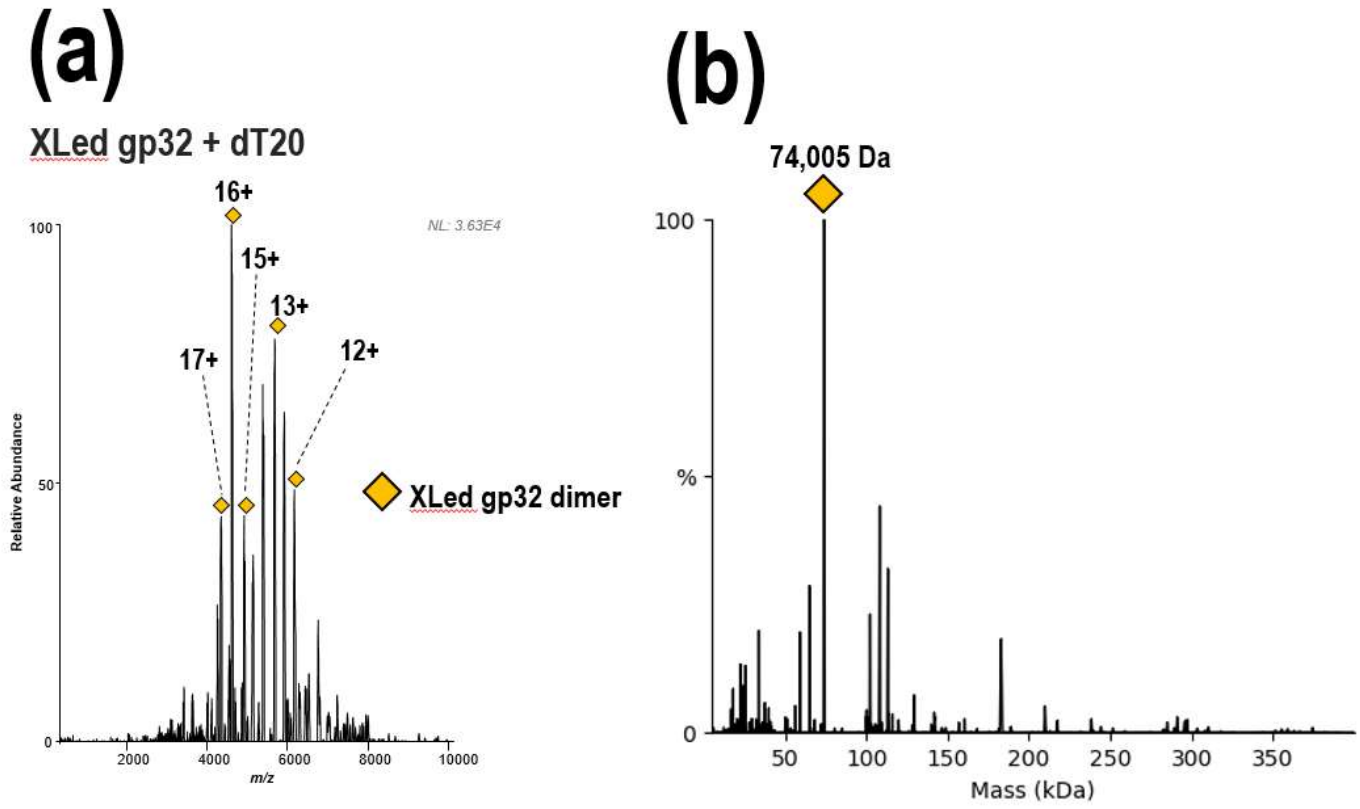


Figure S20. (a) Low-resolution MS1 spectrum of solution containing cross-linked T4gp32 dimer and dT12 acquired using denaturing conditions prior to 30 kDa SEC cleanup (acquired on UHMR mass spectrometer), **(b)** deconvoluted mass spectrum.

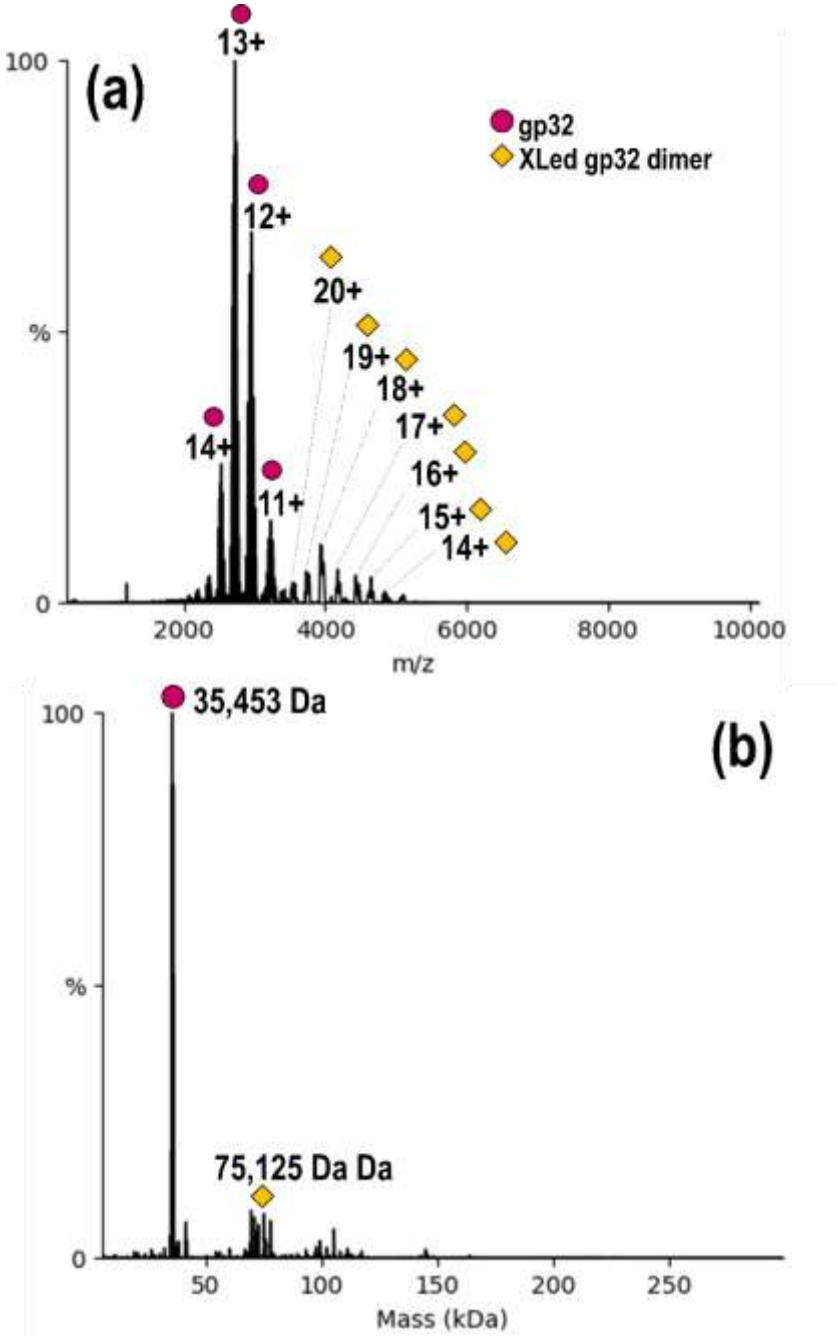


Figure S21. Low-resolution MS1 spectrum of T4gp32 acquired using denaturing conditions (acquired on UHMR mass spectrometer).

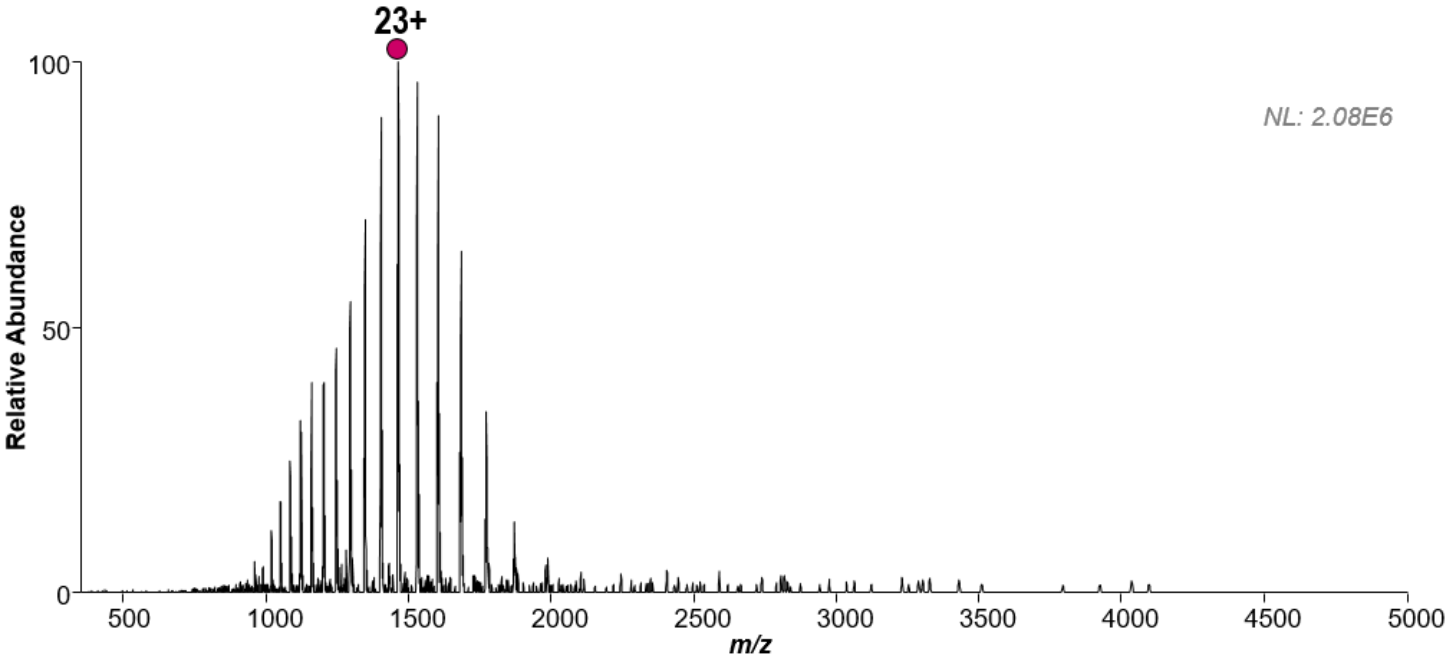


Figure S22. Base peak LC-MS chromatograms from tryptic digests of **(a)** BS3-crosslinked and **(b)** DMTMM-crosslinked gp32 homodimer for bottom-up XL-LC-MS experiments.

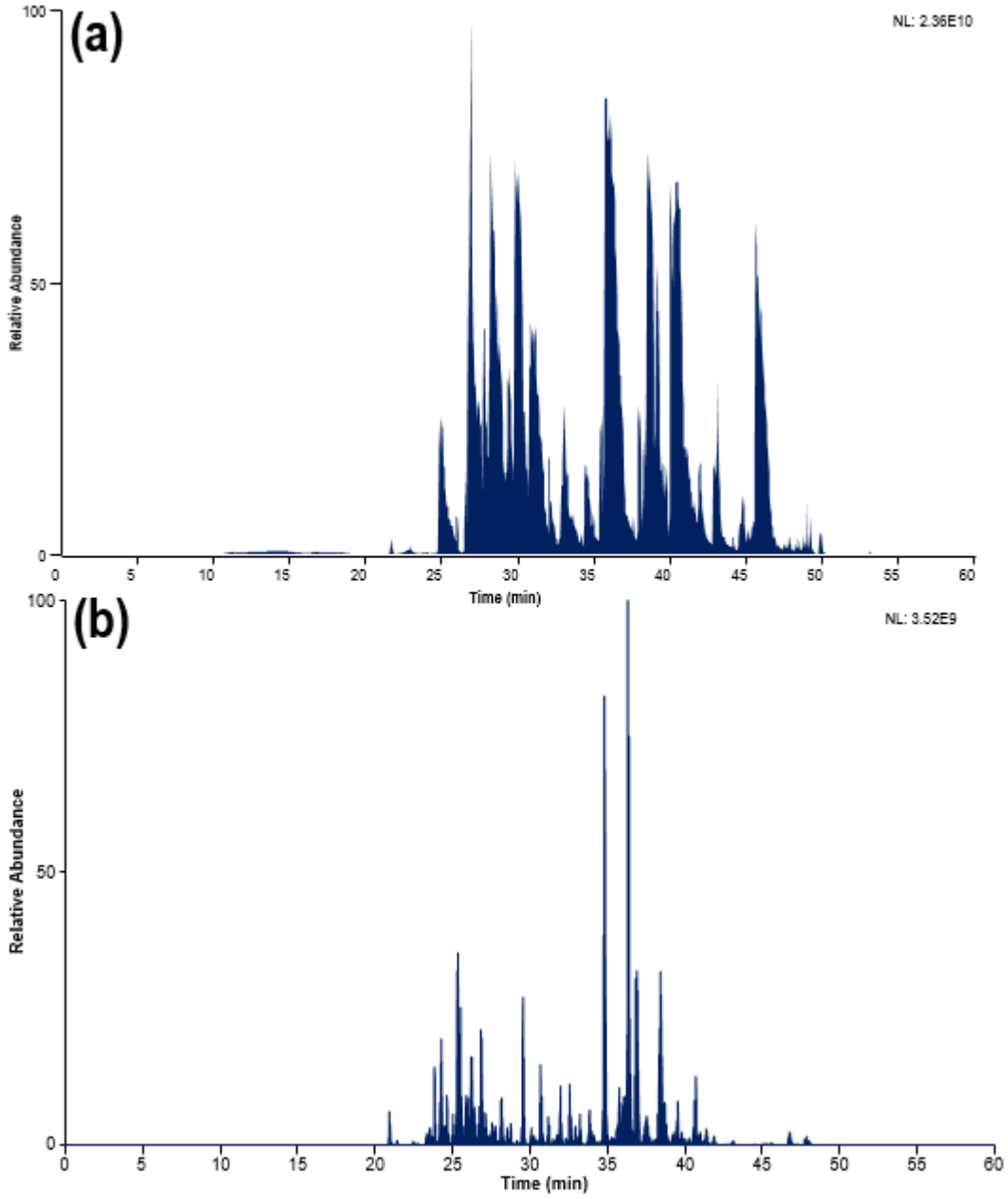
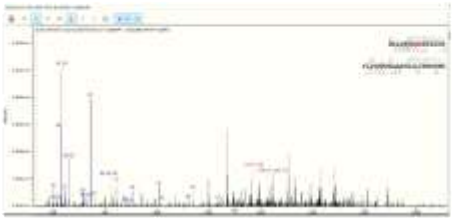


Figure S23. **(a)** Table of Byonic results for bottom-up gp32 XL-LC-MS/MS data with **(b-g)** annotated MS/MS spectra. "Sequence" and "Xlink Partner Peptide Sequences" columns indicate the two sequences of the two peptides that were identified as cross-linked, while the "Xlink Base Peptide Residue" and "Xlink Partner Peptide Residue" columns indicate the specific positions of the cross-linked lysine residues for that respective cross-linked peptide-peptide identification. Pep2D, Score, and ppm error metrics provide high confidence in the cross-linked peptide identifications. Model 1 and Model 2 columns indicate whether the crosslink was identified in either of the models (or both).

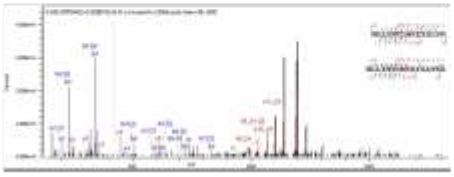
(a)

Cross linker	Peptide Xlink	Sequence	Xlink Partner Peptide Sequence	Xlink Base Residue	Xlink Partner Residue	Model 1	Distance (A)	Model 2	Distance (A)	z	Obs. *m/z*	Calc. *m/z*	ppm error	Score	PEP 2D
BS3	I	K.NDLYNTDN**K**EYSLVK.R	R.FLPS**K**NDEQAPFAILVNHGFKK.N	K104	K51	Y	24.3	Y	29.5	5	891.4594	891.6569	3.58	313.4	1.20E-07
	II	K.NDLYNTDN**K**EYSLVK.R	K.NDLYNTDN**K**EYSLVKR.K	K104	K104	N	N/A	Y	28.3	4	981.9806	981.9839	-3.33	386.5	5.90E-07
	III	K.V**K**QVSGFSNYDESKFLNQSAIPNIDDESFQK.E	K.V**K**QVSGFSNYDESKFLNQSAIPNIDDESFQK.E	K178	K178	Y	21.1	N	N/A	4	1966.448	1966.451	-1.52	447	5.30E-08
	IV	K.VKQVSGFSNYDES**K**FLNQSAIPNIDDESFQK.E	K.NDEQAPFAILVNHGF**K**KNGK.W	K190	K67	Y	22.1	Y	17.1	6	1669.481	1669.651	-1.34	346.6	1.90E-07
	V	K.VKQVSGFSNYDES**K**FLNQSAIPNIDDESFQK.E	K.NG**K**.W	K190	K71	Y	26.3	Y	21.3	5	1572.958	1572.958	0.09	307	1.20E-05
DMTMM	VI	K.NDLYNTDNK.E	R.FG**K**K.I	D102	K141	N	N/A	Y	14.7	3	624.9723	624.9704	2.93	342.7	9.70E-08

(b) XLink (I): K104/K51

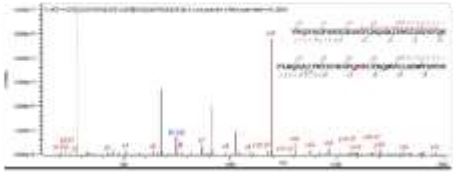


(c) XLink (II): K104/K104

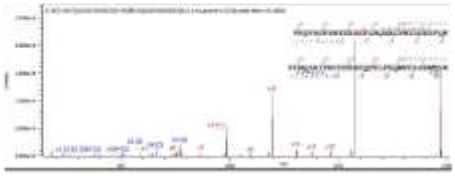


(d) XLink (III): K178/K178

(e) XLink (IV): K190/K67



(f) XLink (V): K190/K71



(g) XLink (VI): D102/K141

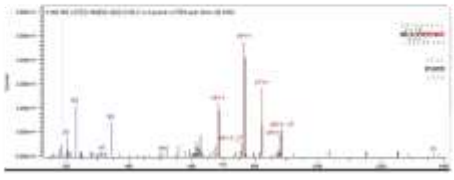


Figure S24. Native MS1 spectra of 10 µM gp32-ΔCTD (R=6,250) **(a)** without and **(b)** with addition of 10 µM dT12.