Investigation of Sequence Clipping and Structural Heterogeneity of an HIV broadly neutralizing antibody by a Comprehensive LC-MS Analysis

Vera B. Ivleva, Nicole A. Schneck, Deepika Gollapudi, Frank Arnold, Jonathan W. Cooper, and Q. Paula Lei

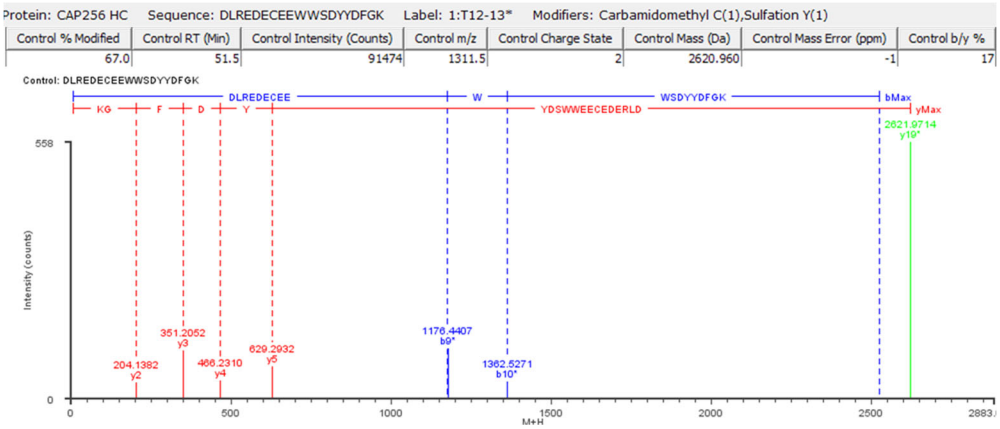
Vaccine Production Program, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Gaithersburg, Maryland, USA

SUPPLEMENTARY INFORMATION

Peptide Mapping LC-MS/MS data processing criteria and analysis

Identification of the CAP256 peptides was confirmed by the characteristic ion fragments generated by high energy MS^E-fragmentation. The data were processed automatically using BiopharmaLynx to confirm primary structure, and to characterize PTMs. For the MS-peaks which failed to get assigned by the automatic data processing, a *de novo* manual assignment was performed using the characteristic ion fragment of the MS^E spectra. To ensure the legitimate component assignment, a set of verification criteria was applied to filter the automatically processed results. The mass accuracy limit for the precursor ion was set to 5 ppm (heavily glycosylated peptides exhibit mass error up to 10 ppm), and 15 ppm for the fragment ions. In addition, all Cys residues were carbamidomethylated and treated as a static modification. The criteria for peptide identification included: 1) the retention time had to match the peptide length and hydrophobicity; 2) the glycosylated peaks eluted as a relatively tight peak cluster; 3) deamidated peptides, converted to aspartic and *iso*-aspartic acid analogs, displayed an expected chromatographic pattern, and their isotopic pattern was consistent with deamidation. For most peptides, a minimum 20% of the characteristic b/y-ion fragments was used to filter out invalid search hits, whereas no such filter was applied for the chymotryptic peptides (with less charge density), and relied on overall spectral quality. All MS^E spectra were visually scanned to make sure the MS^E profile looked reasonable, so that chemical noise peaks were not coincidentally taken into consideration. High energy spectra of the glycosylated peptides had to contain specific glycan fragment ions characteristic of the glycosidic bond cleavage, as opposed to the minimal 20% of b/y-ions for the non-glycosylated peptides. The combination of all these filtering criteria ensured an unambiguous sequence verification, free of false-positive assignments.

Supplementary Figure S1. Sulfation identification via automatic data processing with MS^E confirmation



Supplementary Figure S2. Automatic data processing with *pyro*QLPCAK included in the search as a standalone peptide. All the components were generated by BiophamaLynx data processing and sorted by intensity, including unknown peaks, without *pyro*QLPCAK included (top table); *pyro*QLPCAK was included in the search as a separate peptide (bottom panel), allowing Gln to be treated as a *N*-terminal amino acid.

Label	Peptide	Modifiers	Control % ...	Control RT (...	▼1 Control Intensit...	Control m/z	Control Charge ...	Control Mass (Da)	Control Mass Error...	Control b/y %
1:T30*	VVSVLTVLHQDW...	Deamidation N(1)	10.7	62.74	244709	603.7	3	1807.98	-0	60
1:T26c4*	TPEVTCVVVDVS...	Carbamidomethyl C(1)	99.0	42.25	242843	843.4	2	1684.76	0	89
1:T2*	VTISCSGNTSNI...	Carbamidomethyl C(1),Glycosylatio...	56.9	44.83	218825	1571.0	3	4710.01	-4	4
1:T9	FTISR		98.2	22.92	208193	623.4	1	622.35	3	75
1:T16*	TVAPTECS	Carbamidomethyl C(1)	99.2	15.32	187433	864.4	1	863.37	-0	71
1:T41-42n16	PENNYKTTPPVL...		100.0	64.59	184986	873.8	3	2618.30	17	5
1:T19n12*	SGALTSGVHTFP...	Carbamidomethyl C(1)	90.0	64.96	182495	1056.5	5	5277.66	-1	34
1:T45*	WQQGNVFSCSV...	Carbamidomethyl C(1)	100.0	44.39	172532	689.8	4	2755.29	1	55
				25.23	164869	699.3	1	698.34		
1:T11c14	NTLYLQMN		95.7	39.75	159003	996.5	1	995.48	1	86
1:T19c12	DYFPEPVTVSWN...		100.0	78.94	145640	1774.2	3	5319.61	-1	20

HC fragment

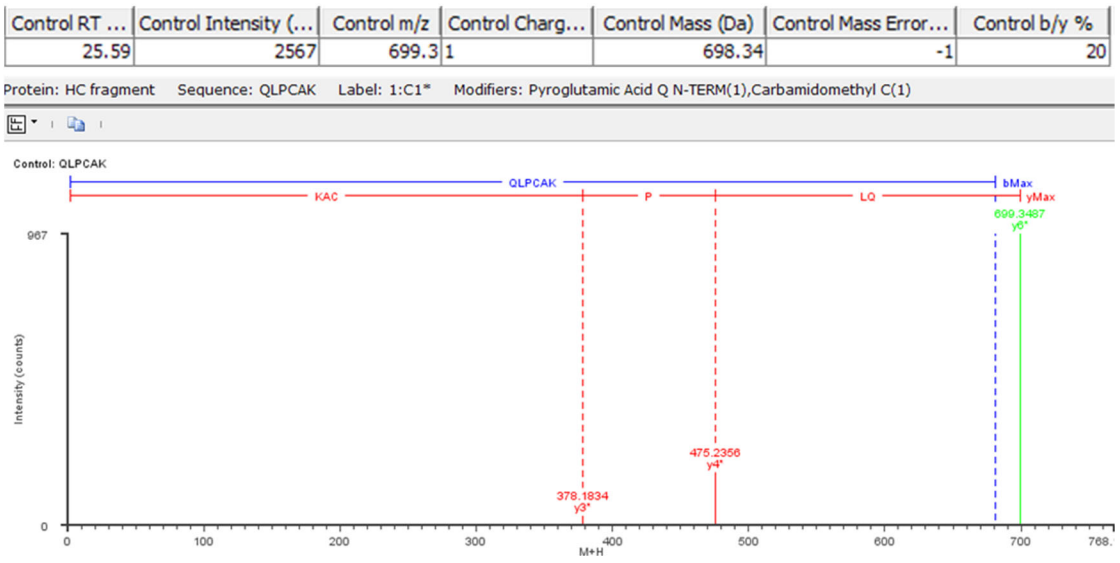
Control Coverage (%): 100.0 Combined Coverage (%): 100.0 Analyte Coverage (%): 0.0

Control Unique Coverage (%): 100.0 Common Coverage (%): 0.0 Analyte Unique Coverage (%): 0.0

1:1 to 6 QLPCAK

	Protein	Chain	Peptide	▲1 Label	Mass (Da)	Start	End	Modifiers
1	HC fragment	1	QLPCAK	1:T1*	698.3421	1	6	Pyroglutamic Acid Q N-TERM(1),Carbamidomethyl C(1)
2	HC fragment	1	QLPCAK	1:T1*x2	1396.6842	1	6	Pyroglutamic Acid Q N-TERM(1),Carbamidomethyl C(1)
3	HC fragment	1	QLPC	1:T1c2*	499.2101	1	4	Pyroglutamic Acid Q N-TERM(1),Carbamidomethyl C(1)
4	HC fragment	1	PCAK	1:T1n2*	289.1096	3	6	-Lysine C-TERM(1)

Supplementary Figure S3. Pyroglutamination of QLPCAK found in the chymotryptic digest and confirmed by MS^E fragmentation



Supplementary Figure S4. Verification of the pyroglutaminated modification at the Q118 site via peptide mapping analysis of the non-reduced CAP256: the peptide dimer with clipping identified and *N*-terminal Q-pyroglutamination confirmed.

Peptide	Modifiers	Mass (Da)	Mass Error (ppm)	b/y %
DLREDECEEWWSDYYDFGK=QLPCAK	Pyroglutamic Acid Q N-TERM(1)	3123.28	-4	53