

Application of Molecular Mass Spectrometry for The Structural Characterization of a DNA-Protein Cross-Links

Jiawei Gong, Morwena J. Solivio, Edward J. Merino, Julio A. Landero-Figueroa*

Agilent Technologies/University of Cincinnati Metallomics Center of the Americas, Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221-0172, USA

*Corresponding author, email: Landerjo@ucmail.uc.edu

Received November 28th, 2017; Accepted March 26th, 2018.

DOI for the main article: <http://dx.doi.org/10.29356/jmcs.v62i2.405>

Supplementary Calculations

Supplementary Calculations

Tryptic digest of RNase A will result to 45 peptide fragments as predicted by the UCSF MS Digest website: <http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msdigest>

The peptides are listed in Column 2 in the following table and their mono protonated mass to charge ratios are listed in Column 3. Lysine containing peptides are predicted to form adducts with guanine in DNA. Addition of one dG molecule will result in $\Delta m/z$ of 281.07. Because there are multiple lysine residues in the RNase A sequence, each of these residues can potentially form adducts with dG resulting to peptide fragment cross-linking with multiple equivalents of dG.

For example, if the molecular weight of one peptide is M, it is possible to observe $[M+H+ndG]^+$, $[M+2H+ndG]^{++}$, and $[M+3H+ndG]^{+++}$ (n could be 1 or 2 or 3). Let's take the first peptide **ETGSSKYPNTCAYKTTQANK** with no carboaminomethyl modification as an example. The mono protonated m/z of this peptide ($[M+H]^+$) is 2090.98. Since there are three lysines in this peptide, one or two or three dG can be crosslinked with the peptide. If there is only one dG crosslinked with the peptide, the calculation will be as follows:

$$[M+H]^+ = 2090.98$$

$$[M+H+dG]^+ = 2090.98 + 281.07 = 2372.05$$

$$[M+H+2dG]^+ = 2090.98 + 281.07 \times 2 = 2653.12$$

In addition, each of the M-dG, M-2dG, M-3dG can also be ionized with two or three protons (the mass of one proton is 1.007825). In this case, the calculation will be:

$$[M+2H+dG]^{++} = (2090.98 + 1.007825 + 281.07) / 2 = 1186.53$$

$$[M+2H+2dG]^{++} = (2090.98 + 1.007825 + 281.07 \times 2) / 2 = 1327.07$$

$$[M+2H+3dG]^{++} = (2090.98 + 1.007825 + 281.07 \times 3) / 2 = 1467.61$$

$$[M+3H+dG]^{+++} = (2090.98 + 1.007825 + 281.07) / 3 = 791.36$$

$$[M+3H+2dG]^{+++} = (2090.98 + 1.007825 + 281.07 \times 2) / 3 = 885.05$$

$$[M+3H+3dG]^{+++} = (2090.98 + 1.007825 + 281.07 \times 3) / 3 = 978.75$$

In the above calculation all possible combinations of m/z values of dG-peptide cross-links within **ETGSSKYPNTCAYKTTQANK** have been calculated. Similar tables were accomplished for the remaining 44 expected peptide. The data is shown in Table S-1 and was used to identify the site of cross-linking.

Table S1. Results of Crosslink Calculations

No	Peptide	m/z (mono)	Carbamid omethyl	K	[M+2H] ⁺	[M+3H] ⁺⁺	[M+H+d G] ⁺	[M+2H+d G] ⁺⁺	[M+3H+d G] ⁺⁺⁺	[M+H+2 dG] ⁺	[M+2H+ 2dG] ⁺⁺	[M+3H +2dG] ⁺⁺	[M+H+3 dG] ⁺	[M+2H+ 3dG] ⁺⁺	[M+3H+ 3dG] ⁺⁺⁺
1	ETGSSKYPNCAYKTTQANK	2090.98	0	3	1045.99	697.67	2372.05	1186.53	791.36	2653.12	1327.06	885.05	2934.19	1467.60	978.74
2	ETGSSKYPNCAYKTTQANK	2148.00	1	3	1074.51	716.67	2429.07	1215.04	810.36	2710.14	1355.58	904.05	2991.21	1496.11	997.74
3	CKPVNTFVHESLADVQAVCSQK	2403.18	0	2	1202.09	801.73	2684.25	1342.63	895.42	2965.32	1483.16	989.11			
4	CKPVNTFVHESLADVQAVCSQK	2460.20	1	2	1230.60	820.74	2741.27	1371.14	914.43	3022.34	1511.67	1008.12			
5	CKPVNTFVHESLADVQAVCSQK	2517.22	2	2	1259.12	839.75	2798.29	1399.65	933.44	3079.36	1540.19	1027.13			
6	KETAAAKFER	1150.62	0	2	575.81	384.21	1431.69	716.35	477.90	1712.76	856.88	571.59			
7	ETGSSKYPNCAYK	1447.65	1	2	724.33	483.22	1728.72	864.87	576.91	2009.79	1005.40	670.60			
8	YPNCAYKTTQANK	1501.71	1	2	751.36	501.24	1782.78	891.89	594.93	2063.85	1032.43	688.62			
9	KETAAAK	718.41	0	2	359.71	240.14	999.48	500.24	333.83	1280.55	640.78	427.52			
10	DRCKPVNTFVHESLADVQAVCSQK	2731.33	1	2	1366.17	911.11	3012.40	1506.70	1004.80	3293.47	1647.24	1098.49			
11	DRCKPVNTFVHESLADVQAVCSQK	2788.35	2	2	1394.68	930.12	3069.42	1535.21	1023.81	3350.49	1675.75	1117.50			
12	CKPVNTFVHESLADVQAVCSQKNVAC K	2918.43	0	2	1459.72	973.48	3199.50	1600.25	1067.17	3480.57	1740.79	1160.86			
13	CKPVNTFVHESLADVQAVCSQKNVAC K	2975.45	1	2	1488.23	992.49	3256.52	1628.77	1086.18	3537.59	1769.30	1179.87			
14	NVACK	534.27	0	1	267.64	178.76	815.34	408.17	272.45						
15	NVACK	591.29	1	1	296.15	197.77	872.36	436.68	291.46						
16	ETAAAK	590.31	0	1	295.66	197.44	871.38	436.20	291.13						
17	ETGSSK	608.29	0	1	304.65	203.43	889.36	445.18	297.12						
18	TTQANK	662.35	0	1	331.68	221.45	943.42	472.21	315.14						
19	SRNLTK	718.42	0	1	359.71	240.15	999.49	500.25	333.84						
20	NLTKDR	746.42	0	1	373.71	249.48	1027.49	514.25	343.17						
21	YPNCAYK	858.38	0	1	429.69	286.80	1139.45	570.23	380.49						
22	YPNCAYK	915.40	1	1	458.21	305.81	1196.47	598.74	399.50						

23	SRNLTKDR	989.55	0	1	495.28	330.52	1270.62	635.81	424.21						
24	ETAAAKFER	1022.53	0	1	511.77	341.51	1303.60	652.30	435.20						
26	QHMDSSSTAASSNYCNQMMK	2307.91	0	1	1154.46	769.98	2588.98	1294.99	863.67						
27	QHMDSSSTAASSNYCNQMMK	2364.93	1	1	1182.97	788.98	2646.00	1323.50	882.67						
28	QHMDSSSTAASSNYCNQMMKSR	2551.04	0	1	1276.03	851.02	2832.11	1416.56	944.71						
29	QHMDSSSTAASSNYCNQMMKSR	2608.06	1	1	1304.54	870.03	2889.13	1445.07	963.72						
30	NVACKNGQTNCYQSYSTMSITDCR	2687.13	0	1	1344.07	896.38	2968.20	1484.60	990.07						
31	FERQHMDSSSTAASSNYCNQMMK	2740.12	0	1	1370.56	914.05	3021.19	1511.10	1007.74						
32	FERQHMDSSSTAASSNYCNQMMK	2797.14	1	1	1399.08	933.05	3078.21	1539.61	1026.74						
33	NVACKNGQTNCYQSYSTMSITDCR	2744.15	1	1	1372.58	915.39	3025.22	1513.12	1009.08						
34	NVACKNGQTNCYQSYSTMSITDCR	2858.20	3	1	1429.60	953.40	3139.27	1570.14	1047.09						
35	NVACKNGQTNCYQSYSTMSITDCR	2801.17	2	1	1401.09	934.40	3082.24	1541.63	1028.09						
36	NGQTNCYQSYSTMSITDCRETGSSK	2818.17	1	1	1409.59	940.06	3099.24	1550.12	1033.75						
37	NGQTNCYQSYSTMSITDCRETGSSK	2761.15	0	1	1381.08	921.06	3042.22	1521.61	1014.75						
38	NGQTNCYQSYSTMSITDCRETGSSK	2875.19	2	1	1438.10	959.07	3156.26	1578.64	1052.76						
39	TTQANKHIIIVACEGPNYPVPHFDASV	2810.39	0	1	1405.70	937.47	3091.46	1546.24	1031.16						
40	TTQANKHIIIVACEGPNYPVPHFDASV	2867.41	1	1	1434.21	956.48	3148.48	1574.75	1050.17						
41	FERQHMDSSSTAASSNYCNQMMKSR	2983.26	0	1	1492.13	995.09	3264.33	1632.67	1088.78						
42	NGQTNCYQSYSTMSITDCR	2171.88	0	0	1086.44	724.63									
43	NGQTNCYQSYSTMSITDCR	2228.90	1	0	1114.95	743.64									
44	NGQTNCYQSYSTMSITDCR	2285.92	2	0	1143.46	762.65									
45	HHIVACEGPNYPVPHFDASV	2224.09	1	0	1112.55	742.03									