INCREASED CONFIDENCE FOR THE IDENTIFICATION OF N-LINKED GLYCOPEPTIDES USING AN OPTIMISED COLLISION ENERGY WORKFLOW

Yi Ju Chen[1], Lee A. Gethings[2], Alex Kuo[1], Yu Ju Chen[1]
[1]Academia Sinica, School of Chemistry, Taipei, Taiwan; [2]Waters Corporation, Wilmslow, United Kingdom.

INTRODUCTION

Glycosylation is responsible for the regulation of cell-cell interactions, recognition and diseases. The development of viable glycomarkers has been limited due to the technical challenges raised from sample preparation and analytical methodologies. Mass spectrometry (MS) is seen as an important tool for elucidating glycan structure but can prove challenging from a data acquisition and processing perspective. For example, N-glycosylated peptides often provide intense Y and B fragment ions corresponding to the carbohydrate moiety when fragmented by means of collision induced dissociation (CID). This is counteracted with inefficient fragmentation of the peptide backbone. Therefore having the ability to customise the collision energy (CE) applied to glycopeptides would be advantageous. Here we describe an LC-MS method which provides efficient glycan and peptide backbone fragmentation within a single acquisition using optimised collision energies.

METHODS

Sample preparation

α-fetoprotein (AFP) and HeLa membrane samples were reduced, alkylated, dephosphorylated and digested using trypsin. Resulting peptides were subsequently enriched using HILIC chromatography. The collected eluent was dried and prepared for LC-MS by resuspending with water:acetonitrile (97:3)/0.1% formic acid.

LC-MS setup

Glycopeptides were chromatographically separated over a 60 min gradient from 5 to 40% acetonitrile (0.1% formic acid) at 300 nL/min using an ACQUITY M-class, configured with HSS 1.8 µm C18 reversed phase 75 µm x 20 cm nanoscale LC column.

MS data were acquired using a Synapt G2-S*i* operating in both data dependent (DDA) and data independent (DIA) modes of acquisition (Figure 1).

Bioinformatics

Data were processed using either Mascot Distiller (DDA) or ProteinLynx Global Server (DIA) for peak picking prior to glycopeptide identification using Byonic and MAGIC (Figure 2).

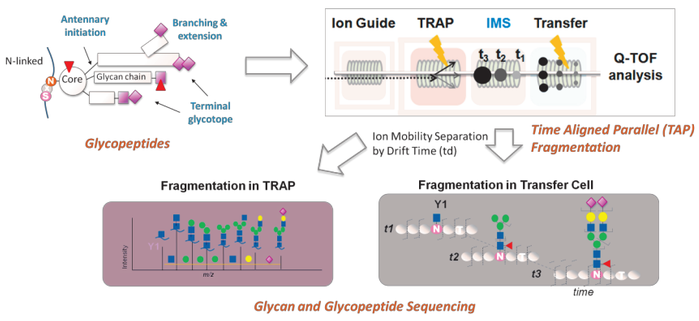


Figure 1. Glycopeptide data acquisition using a Synapt G2-Si Q-IMS QToF mass spectrometer.

RESULTS

Comparison of various CE for N-linked glycopeptides purified from AFP, showed a modest CE of 40eV provided optimal fragmentation from the peptide backbone (b and y ions), whilst a CE of 50eV (trap) yielded rich spectral information for intact glycopeptides. Comparison of transfer based fragmentation, showed that a CE of 50-60 eV yielded good peptide and glycan fragmentation (Figure 3). Confident glycopeptide identifications resulting from Byonic and MAGIC are provided in Figure 4.

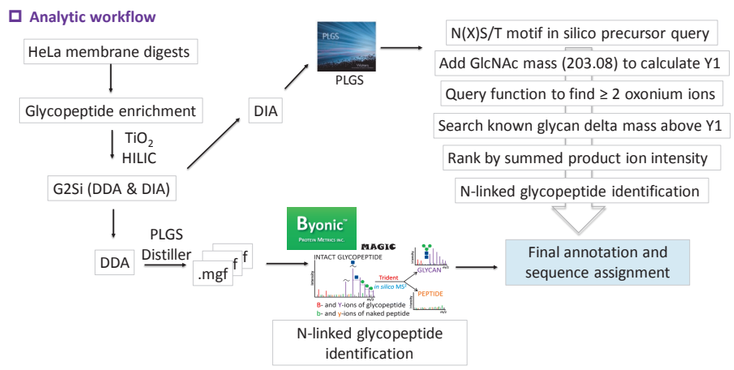


Figure 2. Experimental and data analysis workflow. Data processing uses either PLGS (DIA) or Mascot Distiller (DDA) for peak picking prior. Spectral outputs are further interrogated using Byonic and Magic software for glycopeptide identification and scoring.

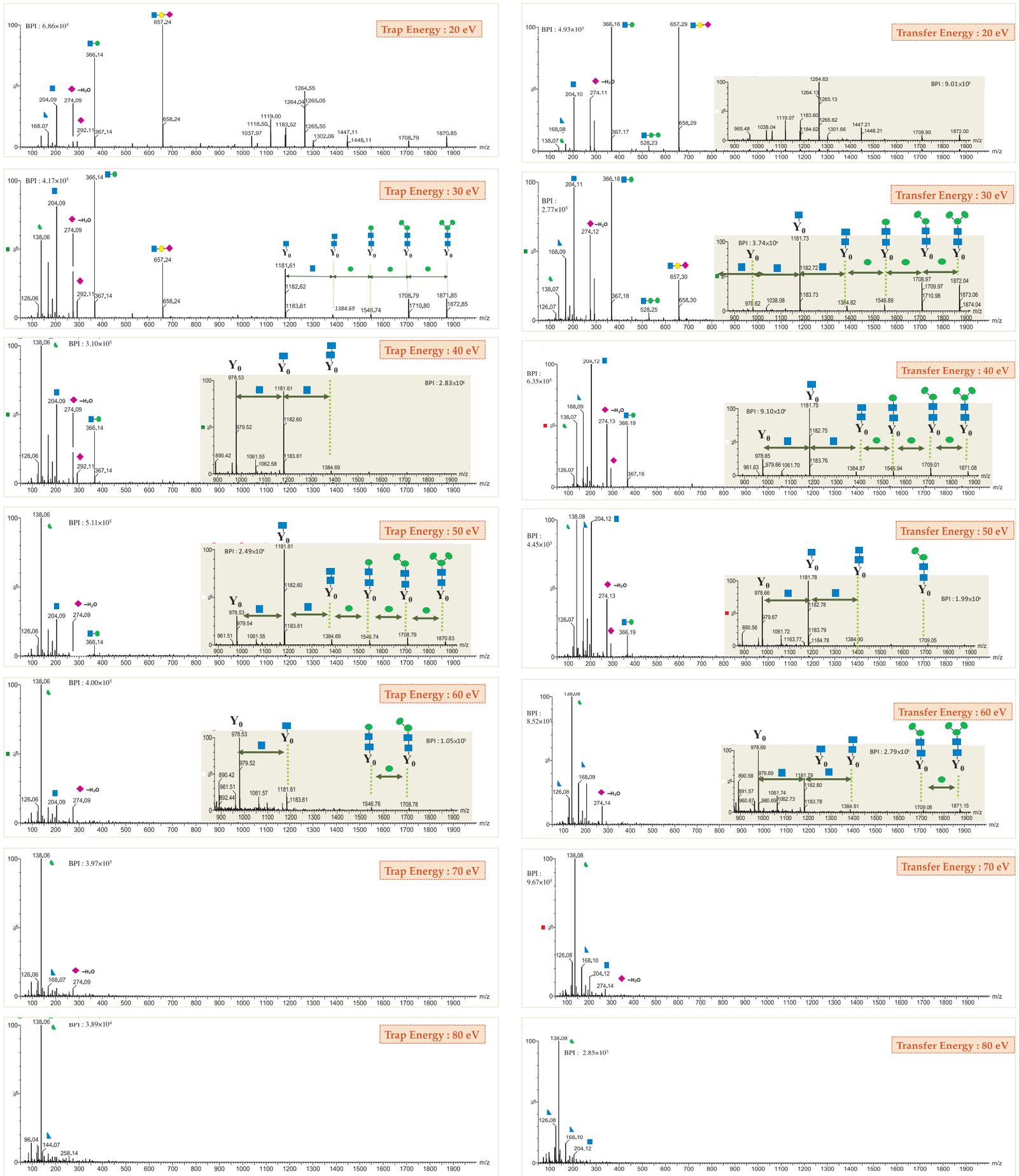
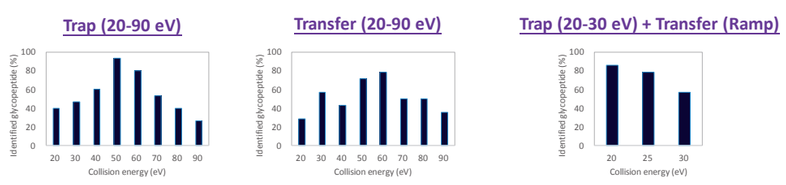


Figure 3. Glycopeptide fragmentation of α-fetoprotein (AFP) at various collision energies (20-80 eV). Comparative fragmentation profiles are shown for the trap and transfer regions of the Synapt G2-Si.

Peptide	Glycan	mz(z)	Energy in Trap (eV) 20	30	40	50	60	70	80	90
K.FTKVN*FTEI[Q]K.L	HexNAc(4)Hex(5)	993.452 (4+)				■	■			
	HexNAc(4)Hex(5)Fuc(2)	1090.822 (3+)				■	■	■		
	HexNAc(4)Hex(6)Fuc(2)	1144.814 (3+)				■	■			
	HexNAc(7)Hex(8)	1019.415 (4+)				■	■			■
K.FTKVN*FTEIQK.L	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1138.826 (3+)	■	■	■	■	■	■	■	
	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	1187.517 (3+)	■	■	■	■	■	■	■	■
	HexNAc(6)Hex(7)	1236.538 (3+)	■	■	■	■	■	■	■	■
	HexNAc(7)Hex(8)Fuc(1)NeuAc(3)	1019.415 (5+)				■	■			
K.LSQKFTKVN*FTEI[Q]K.L	HexNAc(6)Hex(4)Fuc(2)	993.452 (4+)				■	■			
	HexNAc(6)Hex(7)NeuAc(2)	1187.517 (4+)				■	■			
K.LSQKFTKVN*FTEIQK.L	HexNAc(6)Hex(7)Fuc(4)	1187.517 (4+)			■			■	■	
K.VN*FTEI[Q]K.L	HexNAc(4)Hex(5)	868.041 (3+)		■		■	■			
	HexNAc(4)Hex(6)Fuc(2)	1019.415 (3+)	■	■	■	■	■			
K.VN*FTEIQK.L	HexNAc(4)Hex(5)Fuc(2)	965.082 (3+)	■	■	■	■	■	■	■	
	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	1062.108 (3+)	■	■	■	■	■	■	■	■

Peptide	Glycan	m/z(z)	Energy in Transfer (eV) 20	30	40	50	60	70	80	90
K.FTKVN*FTEI[Q]K.L	HexNAc(4)Hex(5)	993.45(4+)				■	■			
	HexNAc(4)Hex(5)Fuc(2)	1090.82(3+)			■	■	■	■		
	HexNAc(4)Hex(6)Fuc(2)	1144.81(3+)		■		■	■			
	HexNAc(4)Hex(8)Fuc(1)NeuAc(1)	1138.83(3+)		■	■	■	■	■		
K.FTKVN*FTEIQK.L	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	1187.52(3+)	■	■	■	■	■	■	■	■
	HexNAc(6)Hex(7)	1236.54(3+)		■	■	■	■	■	■	■
	HexNAc(7)Hex(8)	1019.42(4+)					■			
K.LS[Q]KFTKVN*FTEIQK.L	HexNAc(6)Hex(4)Fuc(2)	993.45(4+)				■	■			
	HexNAc(6)Hex(7)NeuAc(2)	1187.52(4+)				■	■			
K.LSQKFTKVN*FTEIQK.L	HexNAc(6)Hex(7)Fuc(4)	1187.52(4+)		■				■	■	
K.VN*FTEIQK.L	HexNAc(4)Hex(5)Fuc(2)	965.08(3+)	■	■	■	■	■	■		■
	HexNAc(4)Hex(6)Fuc(2)	1019.42(3+)	■	■		■	■			
K.VN*FTEI[Q]K.L	HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	1062.11(3+)	■	■	■	■	■	■	■	■
	HexNAc(8)Hex(5)	1138.83(3+)							■	■



protein Entry	peptide sequence	Site	Glycan	mz	z	precursor intensity
ASPH	SSGNSSSSGSGSGSTSAGSSSPGAR	11	Hex1HexNAc5NeuAc2+Hex3HexNAc2	951.97	5	2737
CD166	NATGDYK	305	HexNAc1+Hex3HexNAc2	932.87	2	1699
ESYT2	LEWLTLMPNASNLDK	510	HexNAc1NeuAc1+Hex3HexNAc2	963.14	3	1769
G3BP2	GDMEQNDSDNRR	316	DeoxyHex1Hex2HexNAc2+Hex3HexNAc2	806.31	4	780
GNAS1	NCLYGNNMSGQR	4	DeoxyHex1Hex1HexNAc1+Hex3HexNAc2	945.70	3	1348
GSK3A	VIGNGSFGVVYQAR	126	HexNAc1NeuAc1+Hex3HexNAc2	1275.51	3	1328
IF4A3	EANFTVSSMHGDMPQK	300	Hex1HexNAc1NeuAc1+Hex3HexNAc2	1110.11	3	792
IQGA1	LIFQMPQNKSTK	995	DeoxyHex1HexNAc2NeuAc1+Hex3HexNAc2	1058.13	3	4048
ITA1	VQRNITVR	747	HexNAc1+Hex3HexNAc2	1041.00	2	2647
ITA3	NITVTGAPR	264	HexNAc1+Hex3HexNAc2	713.01	3	1229
ITB1	KNKNVTNR	91	HexNAc1+Hex3HexNAc2	690.65	3	10786
ITB1	NGVNGTGENGR	402	Hex1+Hex3HexNAc2	1064.94	2	1138
ITB4	LVNITIIK	979	DeoxyHex1Hex1HexNAc1+Hex3HexNAc2	580.04	4	1075
ITB4	RLVNITIIK	979	Hex1HexNAc1+Hex3HexNAc2	582.79	4	1816
K1C9	QVLDNLTMEK	265	DeoxyHex1Hex2HexNAc2+Hex3HexNAc2	1488.13	2	1002
LAMP1	NGNGTACIMANFSAAFSVNYDTK	34	HexNAc1+Hex3HexNAc2	1189.49	3	807
LAMP1	LLNINPNK	256	Hex1+Hex3HexNAc2	990.46	2	3576
LAP2A	MAAHTMGNATVGR	644	DeoxyHex1Hex2+Hex3HexNAc2	670.78	4	1284
MARK2	TTPTPSTNSVLSTSTNRSR	472	DeoxyHex1Hex1HexNAc1+Hex3HexNAc2	905.74	3	3727
MRP1	GVNLSGGQK	766	DeoxyHex1Hex2NeuAc2+Hex3HexNAc2	936.04	3	820

Figure 4. Identification of AFP glycopeptides. Peptides with associated glycans are tabulated (top/bottom left) indicating the collision energy required in order to observe sufficient diagnostic ions for confident identification. The percentage of identified glycopeptides based on collision energy (top right) show trap and transfer fragmentation to optimize at 50-60 eV. A large percentage of glycopeptides identifications are also shown to result from a combination of trap and transfer collision energies, whereby the transfer ramps 19-45 eV. Example identifications generated for the HeLa membrane digest using the same methodology is also demonstrated (bottom right).

CONCLUSIONS

- **Optimised and sequential CE (i.e. using both trap and transfer) is shown to be important for glycopeptide characterization, ensuring comprehensive profiling of both the peptide backbone and glycan moiety.**
- **A combined DDA and DIA approach has shown to provide good decomposition of glycan structures and peptide sequencing from intact glycopeptides, thereby providing a more comprehensive sequencing.**
- **193 intact glycopeptides corresponding to 112 glycoproteins from HeLa membrane digests were identified, including EGFR, CD63, LAMP1/2 and integrin family proteins.**

TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS