APPLYING LC-MS WORKFLOWS TOWARDS DEVELOPING A WELL-CHARACTERIZED RECOMBINANT SUBUNIT VACCINE CANDIDATE

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Hongwei Xie[1]; Catalin Doneanu[1]; Weibin Chen[1]; St John Skilton[1]; Joseph Rininger[2]; Martin Gilar[1]; Jeff Mazzeo[1]. 1: Waters Corporation, 34 Maple Street, Milford, MA 01757. 2: Protein Sciences, Meriden, CT

CHARACTERIZATION BY PEPTIDE MAPPING WITH RP LC-MS^E

- **Label-free quantification of rHA antigens and process-related impurity proteins with ProteinLynx Global SERVER™ (PLGS) Software for automated data processing[1] (Table 1);**
- **Confirmation of sequences of rHA/B proteins with BiopharmaLynx™ for automated data processing[2] (Figure 1);**
- **Confirmation of site-specific glycosylation and glycoforms on each rHA antigen (Diagram 1, Figures 2-4 and 7, Table 2);**
- **Identification of degradations on each rHA antigen (Fig. 5).**

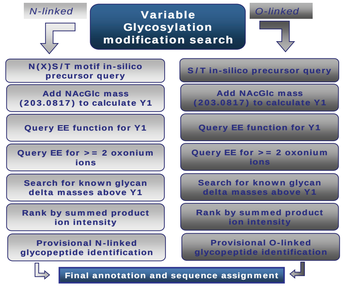

Diagram 1: Algorithm for the proposal of glycosylation in PLGS 2.4. The algorithm uses accurate masses to propose glycan masses matching identified peptides to within a reconstructed spectrum.

LC-MS^E METHODOLOGY

Sample: 4-hour tryptic digest of a research vaccine formulated with purified recombinant hemaglutinins rH1, rH3 and rB.

LC Conditions: Waters ACQUITY UPLC® System. 120-min gradient (0 to 40% ACN in 0.1% FA) on a 2.1 x 150 mm BEH C18 1.7 µm column at a flow rate of 0.2 mL/min and a column temperature of 60 °C.

MS Conditions: Waters SYNAPT® HDMS™ System. MS^E data acquired in ESI positive ion mode, collision cell energy alternating between low energy (5 V, for peptide precursor MS data) and elevated energy (rampping from 20 to 40 V, for fragments MS^E data).

Data processing: Protein identification and absolute quantification by searching against a combined sequence database with PLGS 2.4. Sequence verification, mining of site-specific PTMs (N-linked glycosylation and degradations) by BiopharmaLynx 1.2. Strict tryptic cleavage rule was applied.

Protein	Organism	M.W.	PLGS Score	Unique Peptides	Normalised Conc (%)*	RSD (%)**
H1-H1N1A/Brisbane/59/2007	rHA	61229	2166	45	64.8	14.5
H3-H3N2A/Brisbane/16/2007	rHA	61814	675	29	35.8	6.2
B-B/Florida/4/2006	rHA	61554	2606	44	100	9.9
P17501-Major envelope glycoprotein	A. California	58566	1410	16	6.7	12.1
P32651-Structural glycoprotein gp41	A. California	45381	204	7	2.7	16.9
P41678-Capsid protein p24	A. California	22110	223	3	2.3	3.9
Proteins Identified by Homology:						
Q05825-ATP synthase subunit beta	*Drosophila*	54108	555	8	2.5	14.2
Q14097-ATP synthase subunit beta (fragment)	*Drosophila*	34850	463	7	2.2	15.6
B1S02-GM10439	*Drosophila*	36921	161	3	1.9	4.0
Q5KMQ8-Expressed protein (proposed uncharacterized protein)	Yeast	30298	94	4	6.3	4.1

* Normalized to HA protein B and averaged from 4 repeated injections; Calculated using sum of the intensities of 3 most abundant identified unique peptides (ref 1)

** RSD - relative standard deviation

Table 1: Proteins identified from the tryptic digest of the recombinant vaccine candidate by PLGS 2.4. Protein concentrations were calculated using the intensity sum of the 3 most abundant unique peptides of each protein, and normalized to the most abundant protein (rB).

Figure 1: Sequence coverage of rB determined from the LC-MS^E data (98.8%) and displayed by BiopharmaLynx 1.2. Similarly high sequence coverage for rH1 and rH3 was obtained.

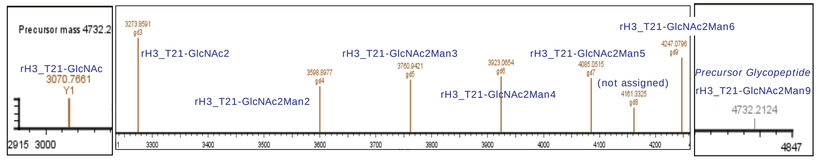

Figure 2: Glycan-related fragment ion spectrum of a glycosylated tryptic peptide rH3_T21 proposed by PLGS2.4. The 'Y1' (3071 Da) and precursor ions (4732 Da) are at different relative abundances and scaled separately. 8 of the 11 theoretical possibilities are automatically proposed as 'gd' [glycan difference], providing an almost complete fragment ladder with no pre-knowledge of the glycosylations, nor of the protein sequence.

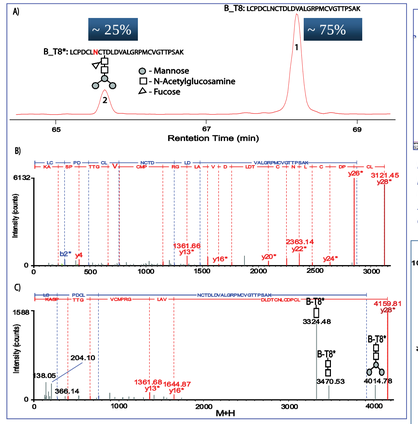

Figure 3: Evidence for assignment of glycosylations on a tryptic peptide of Hemaglutinin B from a baculovirus expression vector system[3]. A) XIC chromatograms of tryptic peptide rB_T8 using accurate masses. B) MS^E spectrum of unmodified rB_T8. C) MS^E spectrum of glycosylated rB_T8 by BiopharmaLynx 1.2 with proposed glycan assignments based on accurate masses.

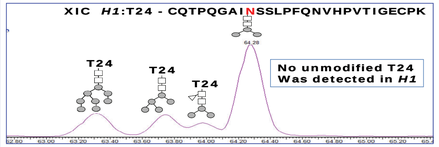

Figure 4: Separation (XIC) of multiple N-linked Glycoforms of tryptic peptide rH1_T24 (originally identified by BiopharmaLynx 1.2). The most abundant glycoform was confirmed by MS^E spectrum. No MS^E spectra were available for the other 3 low-level forms, but corresponding masses were evident in the HILIC LC-MRM traces.

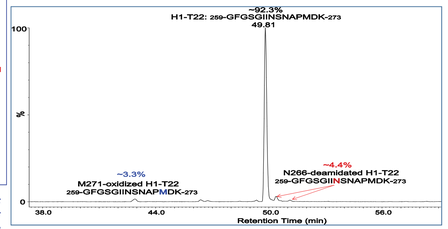

Figure 5: Separation and quantification of unmodified and modified tryptic peptide rH1_T22 (shown as XIC chromatogram, but originally identified and tabulated by BiopharmaLynx 1.2.)

INTRODUCTION

- **Recombinant vaccines have emerged as an alternative to egg-based production systems.**
- **The need for annual re-evaluation of influenza vaccines for efficacy calls for appropriate methods for characterization of hemaglutinin (HA) antigens and final vaccine products.**
- **Better characterization may help speed up vaccine development and reduce safety concerns.**
- **This study describes advanced LC-MS methods for characterization (by LC-MS^E) and monitoring (by LC-MRM) of a trivalent influenza vaccine candidate formulated from purified recombinant HA proteins (rH1, rH3, and rB corresponding to influenza A subtypes H1N1 and H3N2 and influenza B viruses, respectively).**

ANALYTICAL WORKFLOW

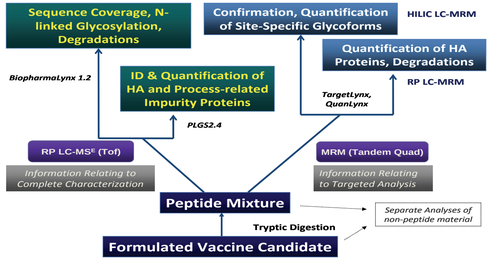

Diagram 2: An illustration of the different analyses possible from the same sample preparation starting with the formulated vaccine. Paths towards complete characterization are possible on tof-based instruments, and targeted analyses are performed on tandem quadrupole instruments

CONCLUSIONS

- **Peptide mapping by RP LC-MS^E was used to:**
  - **Verify rHA primary sequences with high sequence coverage,**
  - **Discover site-specific glycosylation for each rHA antigen;**
  - **Examine potential degradations on each rHA antigen;**
  - **Quantify rHA antigens;**
  - **Identify and quantify impurity proteins.**
- **LC-MRM was used to confirm certain critical attributes of vaccines:**
  - **Reverse phase (RP) LC-MRM was used to quantify rHA antigens**
  - **HILIC MRM was used to confirm site-specific glycoforms that were poorly separated by RP, and low abundance glycoforms without supporting MS^E spectra.**
- **LC-MS^E and LC-MRM are complementary analytical tools and contribute to assays appropriate for a well-characterized vaccine.**

RAPID QUANTIFICATION AND MONITORING BY LC-MRM

- **Quantification of rHA antigens and ratios of the antigens by RP LC-MRM (Figure 6);**
- **Confirmation of site-specific glycoforms on each rHA (Figure 7 and Table 2).**

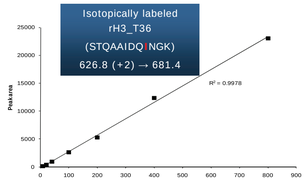

Figure 6: Calibration curve for quantification of rH3. Two additional MRM transitions were used to confirm and quantify the peptide (both native and isotopically labeled forms). The concentration of rH3 was calculated to be 7.9 ±0.8 pmole/ml. Similar methods were used to quantify rH1 and rB. The concentration ratio obtained between the antigens is consistent with the result quantified by RP LC-MS^E.

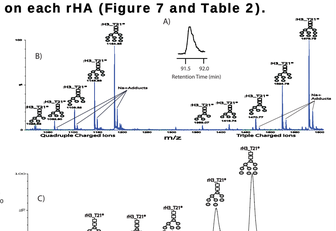

Figure 7: HILIC LC-MRM separation and confirmation of rH3_T21 glycoforms (C). The glycoforms were detected by RP LC-MS^E and BiopharmaLynx 1.2 (B), but have no separation (A) and except for the most abundant glycoform (rH3_T21-GlcNAc2Man9), no MS^E spectra were available for the confirmation of less-abundant glycoforms.

Protein	Peptide	Residue Numbers	Peptide Sequence containing -NXS/T- motif)	Peptide MW	Modification Site	Proposed Glycoforms*	Motif
rH1	T1	1-22	(-EDTICIGYHANNSTDVDTVLEKN)	2408.12	N11	3,3F,5,5F,7,8	NNSTD
	T2	23-40	(KJNVTVTHSVNLLENSINNGK)	1961.99	N23	3F,6,8,9	(KJNVTV
	T4	46-71	(NGQAIPLQLGNCSVAGWILGNPECELLISRE)	2894.50	N54	6,3,3F,5,8	GNCSV
	T5	74-102	(KQESWSYIVEKPNPENGTCYPGHFADYEELREQ)	3429.52	N87	2,3,5,6,7,8,9	ESWSY
	T8	130-145	(KJESWPNHTVTGVSASCSHNGESSFYRN)	2825.21	N125	1,1F,2,2F,3,3F	PNHTV
	T10	154-162	(KJNGLYPNLSKJS)	1004.53	N159	6,3,3F,5,6,7,8,9	PNLSK
	T24	278-304	(KJCCJTPQGAINSSLPFQNVHPVTIGECPK)	2884.39	N288	6,3F,5,8	INSSL
	T47	465-487	(KJNDSTYDYPKYS)	956.42	N480	3,3F,5,7,8,9	PQNDSTY
	T52	504-545	(KJLESMGVYQLAFJSTVASSLVLLVSLGAISFWMCSNGSLQCRI)	4509.28		nd	SNGSL
rH3	T2	3-27	(KSLPGNDNSTATLCLGHHAVPNGTIVKTJ)	2528.28	N8-N22	2F-2F	DNSTA/PNGTI
	T3	28-62	(KJITNDQEVTNATELVQSSSTGGICDSPHQILDGENCTLIDALLG DPQCDGFQNKK)	5889.71	N38-N61	3,3,3-3F	TNATE/ENCTL
	T8	110-141	(RJDLVASSGTLEFNNESFNWTGVTQNGTSSACJRGR)	3376.56	N122/N126/N133	6-6, 7-7, 7-8	NNESF/NWTG/QNGTS
	T10	143-159	(RJSNNVHYWIGL)	957.45	N144	6,3,3F,5,7,8,9	SNNVH
	T13	162-171	(KJYIQALSVYTMPNNEKJ)	1489.72	N165	3,3F,5,6,9	LSVYTM
	T21	230-253	(RJSSTVWTIVKPGDILLINSTGNLIAPRGG)	2866.43	N246	5,6,7,8,9	INSTG
	T27	277-289	(KJCNSECITPNGSIPNDKPFQNVNRJI)	2546.16	N285	3F,6,7,8,9	PNGSI
	T32	401-410	(KJNVCTYHHNVYRIJY)	1239.53	N405	3F,5,6F,8	(KJNVCT
rB	T3	18-38	(KJTATQCAYNVTQVFLTTPTVKS)	2127.14	N25	3,3F,5,6,7,8,9	VNVTQ
	T8	53-80	(RJLCPDCLNCTDLDVALGRPMCVGTTPSAK(A)	2892.33	N59	6,3F	LNCTD
	T16	137-149	(RJLGTSGCPNATSKJS)	1221.57	N145	3,3F,5,6,7,8,9	PNATS
	T19	187-197	(KJNATNFLTVVPJCTEGEIQFYWGFHSIDNQ*)	3477.68	N187	3,7,8,9	(KJNATN
	T30	299-304	(KJYGGLNK(S)	656.34	N303	3,3F,5,6,7,8,9	LNK(SK)
	T34	330-335	(RJLANGTK(Y)	602.34	N332	6,3,5,6,7,8,9	ANGTK
	T51	490-497	(RJCNQTCLDR(J)	951.39		nd	CNQTC
	T52	498-560	(RJEAAGTHNAGEFSLPTFDSLNITAAKLNDDGELDNHTILLVYSTAA SSLAVTLMLAIIIVYMVSR(D)	6699.38		nd	LNITA/DNHTI
	T63	561-569	(RJDNVSCSCJL(-)	952.40		nd	DNVSC

KEY: *-D - peptide without glycan identified; nd - no confident glycoforms detected; X - peptide with glycan containing -GlcNAc-GlcNAc-(Man)x; 3F - peptide with glycan containing -GlcNAc(Fuc)-GlcNAc-(Man)x; X-X or 3F-X - peptide with two glycans on two different N sites.

Table 2: Site-specific glycoforms detected by LC-MS^E and BiopharmaLynx 1.2, and verified by HILIC LC-MRM (except for those glycopeptides with multiple modification sites) for the 3 rHA antigens.

LC-MRM METHODOLOGY

ACQUITY UPLC Conditions: For quantification of rHA proteins, a 30-min gradient (0 to 35% ACN in 0.1% FA) on a 2.1 x 150 mm BEH C18 1.7 µm column at a flow rate of 0.3 mL/min and a column temperature of 35 °C ; For monitoring and verification of site-specific glycoforms, a 60-min gradient (90 to 70% B in 10 min, then to 50% B in 50 min) on a 2.1 x 150 mm BEH glycan 1.7 µm column at a flow rate of 0.2 mL/min and a column temperature of 60 °C . B—10 mM NH4-formate in 90/10 CAN/water and A—10 mM NH4-formate in water. **MS Conditions**: Waters® Xevo TQ MS. MRM acquisition was implemented in an ESI+ mode, Capillary voltage 3.5 V, Cone voltage 35 V, Dwell time 5 ms, collision energy was optimized for each transition.

Data process: QuanLynx for chromatographic peak integration and obtained peak areas were exported to MS Excel for subsequent analysis.

REFERENCES

1.Xie HW, Gilar M, Gebler JC. "Characterization of protein impurities and site-specific modifications using peptide mapping with liquid chromatography and data independent acquisition mass spectrometry", *Anal. Chem.*, **2009**, *81*:5699-5708

2.Xie HW, Chakraborty A, Ahn J, Yu YQ, Dakshinamoorthy DP, Gilar M, Chen WB, Skilton SJ, Mazzeo JR. "Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies", *mAbs*, **2010**, *2*:379-394

3.Tomiya NT, Narang S, Lee YC, Betenbaugh MJ. "Comparing N-glycan processing in mammalian cell lines and engineered lepidopteran insect cell lines", *Glycoconjugate J.* **2004**; *21*:343-360

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