MASS SPECTROMETRIC PROFILING OF HOST CELL PROTEINS IN BIOPHARMACEUTICALS

Catalin E. Doneanu[1], Weibin Chen[1], Alex Xenopoulos[2], Holly Prentice[2], StJohn Skilton[1] and Jeff Mazzeo[1]
[1] Waters Corporation, Milford, MA 01757; [2] Millipore Corporation, Bedford, MA, USA

OVERVIEW

A generic, efficient UPLC/MS^E assay for identification and quantification of HCPs over 4 orders of magnitude in concentration is presented.

INTRODUCTION

- Residual host cell proteins (HCPs) from recombinant production are a major component of biopharmaceutical process-related impurities. HCPs can elicit an unpredictable immune response in patients.
- The composition of HCPs is extremely heterogeneous and changes with production and purification procedures.
- Current analytical methods for measuring HCPs, (typically ELISA, gels, blots), are expensive, subjective, time-consuming to develop, and require prior knowledge about the identities of the contaminant proteins.

WORKFLOW

- Spike a mixture of 5 proteins in the concentration range of 10-1,000 ppm (ng/mg) in the protein biopharmaceutical
- Enzymatic digestion of the protein sample to peptides
- 2D-LC/MS^E with IDENTITY^E to discover HCPs
- Use the top 3 peptides from each protein for absolute label-free quantitation based on the known concentration of the spiked proteins
- Use VERIFY^E for mining the MS^E dataset to extract appropriate MRMs for monitoring HCPs by Xevo TQ MS
- Analyze the sample digests by tandem quadrupole MS for targeted, high-throughput quantitation (e.g. using isotopically labeled peptides)

METHODS

Sample Preparation

A chimeric anti-phosphotyrosine IgG1 mAb (PTG1) was expressed in two CHO cell lines (CHO-S and DG44) and purified by Protein A chromatography using two different protocols. Null cell cultures, (expressing no PTG1 protein), were grown under identical conditions and Protein A purified. Five protein standards were spiked in PTG1 and the resulting protein mixture was denatured with 0.1% RapiGest (15 min at 60 ^oC), reduced with 10 mM dithiothreitol (DTT) for 30 min at 60^oC, alkylated with 20 mM iodoacetamide (IAM) for 30 min (at RT) and enzymatically digested with Promega trypsin (1:20 w/w ratio) at 37°C overnight. Following digestion, the protein amounts loaded on-column on the 2D-LC system were: 4,000 fmoles LA (bovine alpha-lactoglobulin), 800 fmoles PHO (rabbit glycogen phosphorylase b), 320 fmoles ADH (yeast alcohol dehydrogenase), 80 fmoles BSA (bovine serum albumin), and 16 fmoles ENL (yeast enolase).

LC Conditions

A nanoACQUITY™ UPLC® system with 2D technology was used for peptide separations. A reversed-phase/reversed-phase (RP/RP) method was developed that uses the pH of the mobile phases to change the selectivity of a peptide separation in two separate dimensions [1,2].

- **First Dimension (1D) pH=10**: 1.0 mm x 50 mm XBridge C_{18} column (5 µm particles); flow 10 µL/min. Mobile phase 20 mM ammonium formate in water (Solvent A) and ACN (Solvent B).
- **Online dilution (**1:10) of the eluent from 1D before analyte trapping onto the 2D column.
- **Trap column**: A 5-µm Symmetry C_{18} trap (2.7 µl volume) was used to trap peptides between the two LC dimensions.
- **Second Dimension (2D) pH=2.4**: 0.3 mm x 150 mm analytical column BEH C_{18} 1.7 µm, kept at 65 ^oC and operated at 12 µL/min.

Fractions were eluted in ten steps (1: 10.8% Eluent B, 2: 12.4%, 3: 14.0%, 4: 15.4%, 5: 16.7%, 6: 18.6%, 7: 20.4%, 8: 25.0%, 9: 30.0%, 10: 50.0% B. Each step was mixed in a 1:10 ratio with 0.1% TFA in water (pH=2.1) before trapping. Low pH separations in the second chromatographic dimension used a 30 min gradient from 7 to 35% acetonitrile (0.1% FA-formic acid). See Figure 1 for the diagram of the 2D-LC system.

MS conditions

A) Discovery and Label-free Quantitation: Data independent, alternate scanning LC/MS^E experiments were performed with a SYNAPT HDMS™ mass spectrometer.

- Acquisition time was 0.5 sec
- m/z range: 50-1990
- Fixed CE at 5 eV for MS; CE ramp 15 - 35 eV for high-energy fragmentation
- Database search—ProteinLynx Global Server (PLGS) 2.4 with Identity^E Informatics [3].

B) Targeted Quantitation: MRM analyses were performed on a Waters XEVO™ TQ MS tandem quadrupole mass spectrometer coupled with the ACQUITY UPLC system.

- 2.1 x 150 mm BEH130 C_{18} column (1.7 µm particles) at 35 ^oC.
- Flow rate of 300 µL/min; 3-40% B over 10 min (B: acetonitrile with 0.1% FA; A: water with 0.1% FA)
- Other parameters: Potential 3.5 kV, CV 37 V, source temp 90 ^oC, MS1/MS2 isolation window 0.75 Da (FWHM), adjustable collision energy depending of precursors/fragment (15-30 eV) and 10 ms dwell time.

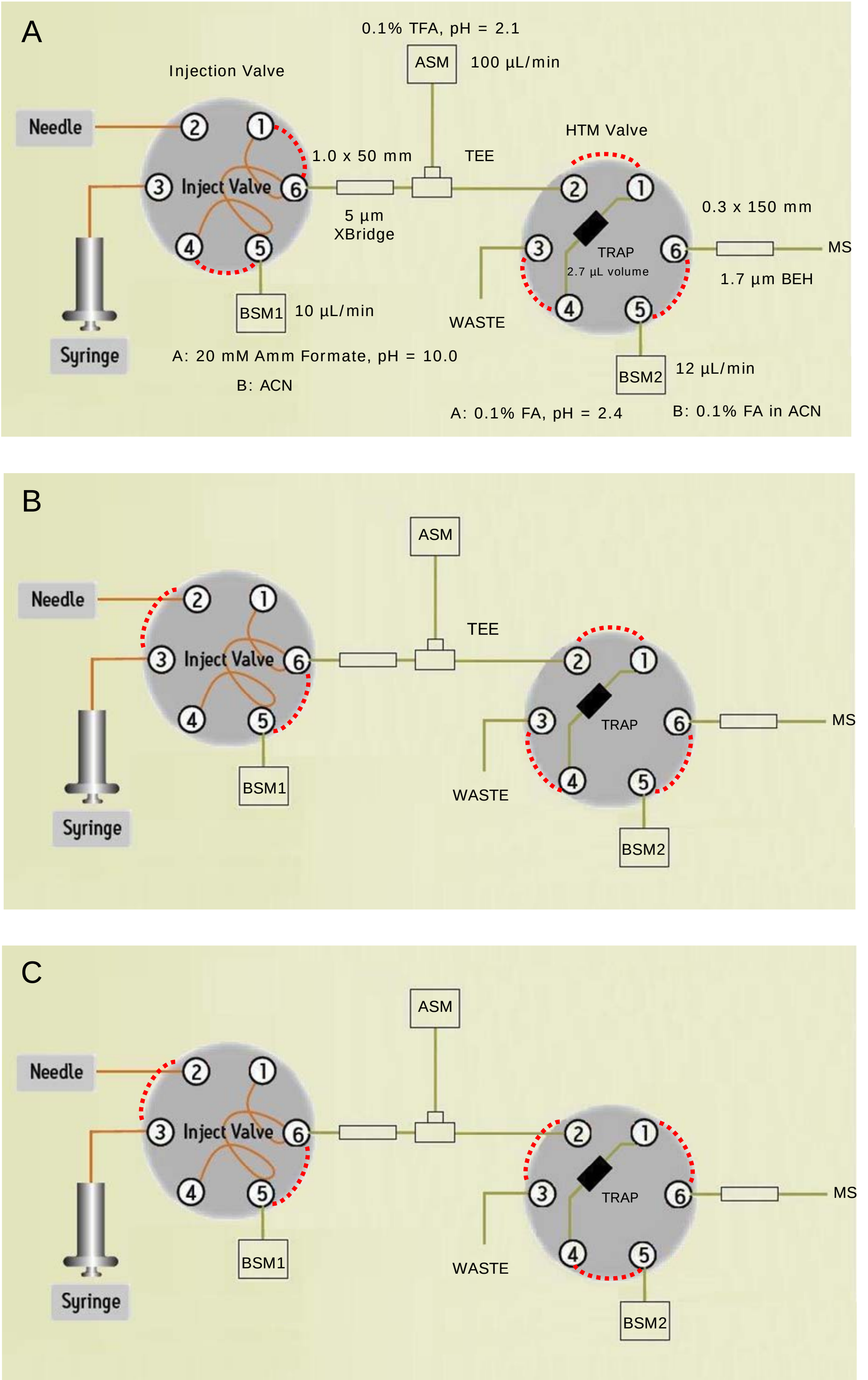

Figure 1: Fluidic configuration for 2-dimensional chromatography with on-line dilution: (A) Sample loading; (B) Step elution of peptides from 1D (in 10 steps) followed by trapping; (C) Peptide separation in the second dimension.

Figure 2. SYNAPT HDMS™ mass spectrometer coupled to the 2D nanoACQUITY UPLC® system.

RESULTS

Chromatographic Performance of the 2D-LC System

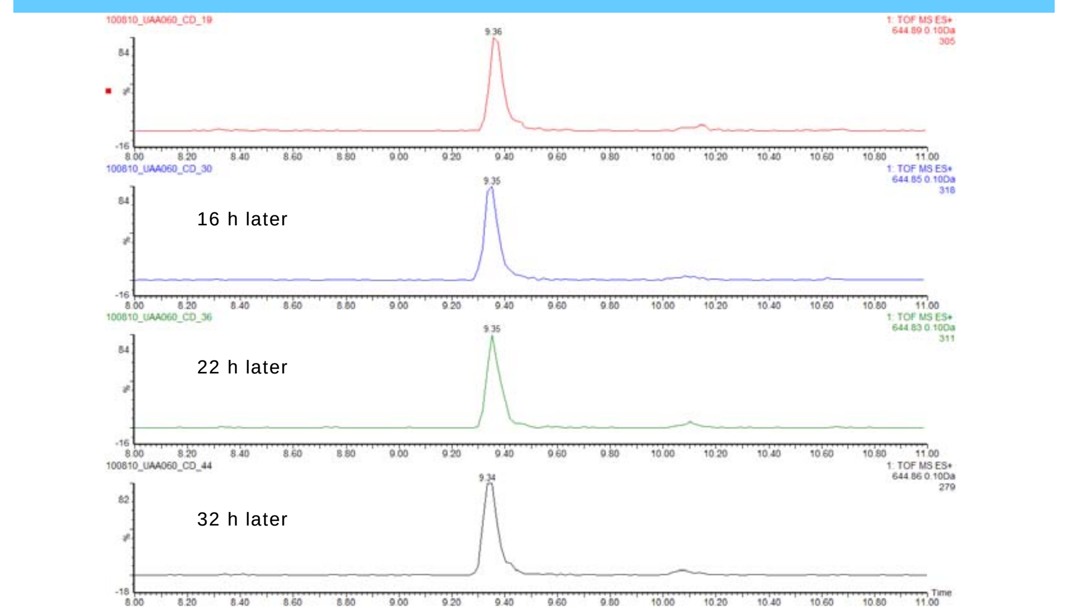

Figure 3. Reproducibility of 2D chromatography: extracted mass chromatograms of the ENL T43 peptide (VNQIGTLSESIK, $[M+2H]^{2+}$= 644.86) from four consecutive 2D-LC experiments. All second dimension chromatography runs were performed at 12 µL/min using a 30 min gradient (3-35% ACN, 0.1% FA). T43 peptide eluted only in Fraction 3/5 (16.7% B) in the 1st dimension chromatography. The amount of ENL digest loaded on column in each 2D experiment was 20 fmoles.

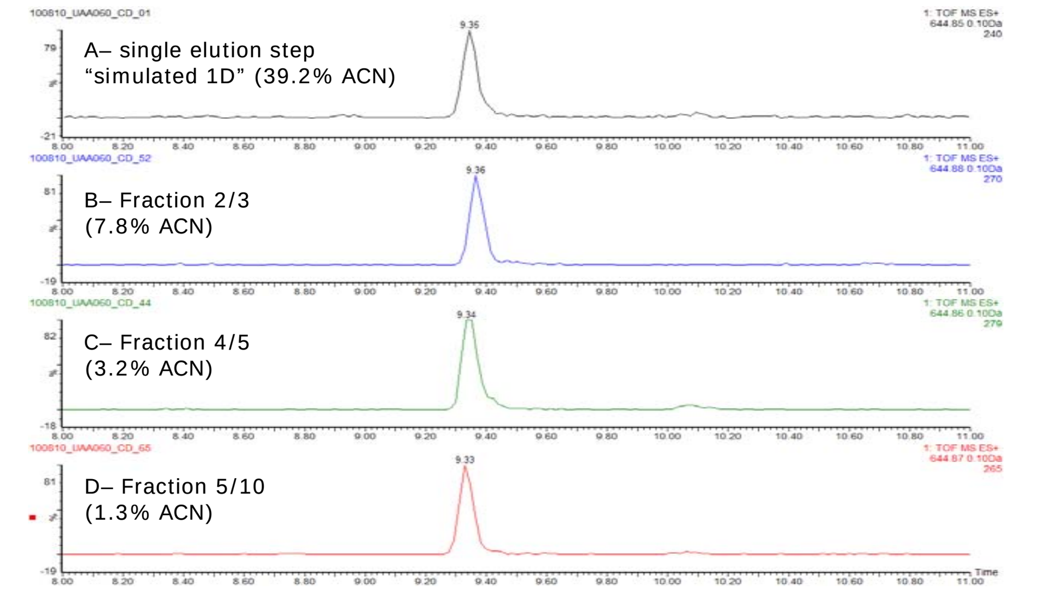

Figure 4. Chromatographic performance (e.g. RT reproducibility and peak width) is maintained during 1D fractionation: mass chromatograms of ENL T43 peptide obtained under four fractionation conditions:
(A) "simulated" 1D run using a single elution step (from 10.8 to 50% ACN)
(B) fraction 2 out of 3 (from 10.8 to 18.6% ACN)
(C) fraction 4 out of 5 (from 15.4 to 18.6% ACN)
(D) fraction 5 out of 10 (from 15.4 to 16.7% ACN)

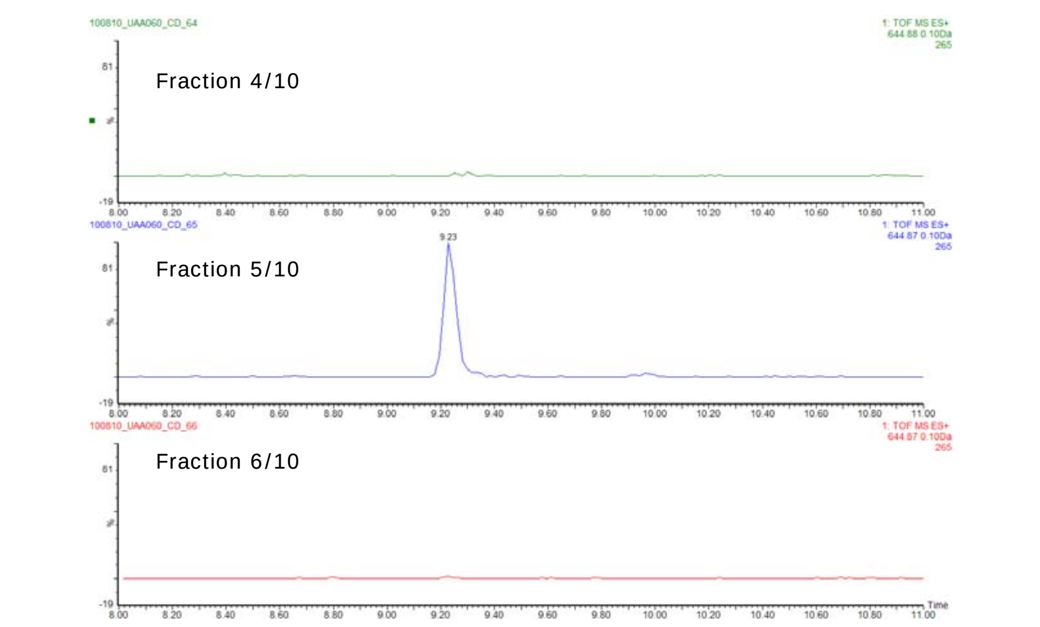

Figure 5. Performance of high-pH fractionation: mass chromatogram of ENL T43 peptide recorded for 3 successive fractions during a 10-step 2D fractionation experiment: fraction 4/10 corresponds to a step elution from 14.0 to 15.4% ACN, fraction 5/10 was eluted from 15.4 to 16.7% ACN and fraction 6/10 was eluted from 16.7 to 18.6% ACN.

HCP Identification Method

- **LC-MS^E Data**: a combined MS^E dataset collected from 10-step 2D RP/RP separations performed in triplicate.
- **Protein Identification Informatics**: PLGS 2.4 featuring Identity^E Bioinformatics.
- **Database**: protein database was compiled from 12,943 mouse Swiss Prot entries, 654 Golden hamster proteins, the sequences of LA, ADH, PHO, BSA, ENL, porcine trypsin, protein A (*S aureus*) and the heavy and light chain sequence of PTG1. The final randomized database contained 27,212 entries.

Prot no	Protein Description	CHO-S Cells				DG-44 cells					
		I	II	III	IV	A1	B1	A2	B2	C	D
1	Nascent polypeptide associated complex subunits Mus musculus		50		1768		5180		6756		
2	Nucleolin Mesocricetus auratus				261	351	3580		2342		
3	Heterogeneous nuclear ribonucleoprotein isoforms Mus musculus				1227		2847		2115		
4	Elongation factor isoforms Mus musculus	142	665		2049	1371	1597	380	2472		64
5	Procollagen C endopeptidase enhancer 1 Mus musculus						1493		2400		
6	Laminin subunits Mus musculus						1261	160	448		
7	Actin isoforms Mesocricetus auratus	287			129	1760	1035	1134	1180		115
8	Clusterin Mus musculus	537	232			1965	1009	1263	954		25
9	Nidogen 1 Mus musculus					484	826	474	601		208
10	Glyceraldehyde 3 phosphate dehydrogenase Mes auratus	153	59		75	2142	790	846	888		92
11	Splicing factor proline and glutamine rich Mus musculus		198				780				
12	**Glycogen phosphorylase b rabbit - PHO (800 fmoles)**	**778**	**778**	**778**	**778**	**778**	**778**	**778**	**778**	**778**	**778**
13	Plasminogen activator inhibitor 1 RNA binding protein Mus musculus				226		547		580		
14	Lipoprotein lipase Mesocricetus auratus					884	547	1460	970		22
15	78 kDa glucose regulated protein Mesocricetus auratus				210	628	546	251	244		56
16	Peptidyl prolyl cis trans isomerase B Mus musculus				53		406		261		
17	T complex protein 1 subunits Mus musculus					499	349	125			
18	Serine protease HTRA1 Mus musculus					488	338	707	682		
19	High mobility group protein isoforms Mus musculus		74		78		329		165		
20	40S ribosomal protein S3 Mus musculus					595	328	442	389		
21	**Alpha-lactalbumin Allergen Bos - LA (4000 fmoles)**	**200**	**130**	**502**	**281**	**479**	**319**	**150**	**224**	**447**	**358**
22	Lysosomal alpha glucosidase Mus musculus					708	314	499			
23	Hepatoma derived growth factor Mus musculus		73		227		304		285		
24	Nuclease sensitive element binding protein 1 Mus musculus				58		245		429		
25	Pyruvate kinase isozyme M2 Mus musculus					612	232	302	207		39
26	Activated RNA polymerase II transcriptional coactivator p15				24		191		191		
27	Heat shock protein HSP 90 beta Mus musculus					210	182	95			26
28	Nucleophosmin Mus musculus				123		171		139		
29	Insulin like growth factor binding protein 4 Mus musculus		64		50		153		217		9
30	Complement C1q tumor necrosis factor related protein 4				68	93	123	67	237		
31	Eukaryotic translation initiation factor isoforms Mus musculus		21		38		116	33	347		
32	**Alcohol dehydrogenase Yeast - ADH (320 fmoles)**	**101**	**96**	**111**	**96**	**109**	**98**	**90**	**118**	**108**	**97**
33	Interleukin enhancer binding factor 2 Mus musculus						96		103		
34	Guanine nucleotide binding protein subunits Mus musculus					158	92	65			
35	Vimentin Mesocricetus auratus						88				69
36	Endothelial differentiation related factor 1 Mus musculus				50		71				
37	**Serum albumin precursor Allergen Bos - BSA (80 fmoles)**	**61**	**57**	**67**	**56**	**47**	**68**	**57**	**69**	**55**	**28**
38	Cofilin Mus musculus					73	63	19	31		
39	Peroxiredoxin 1 Mus musculus					60	58	39			24
40	Histone isoforms Mus musculus		320		599		58		171		28
41	Peroxiredoxin 6 Mus musculus					95	48				8
42	Peptidyl prolyl cis trans isomerase A Mus musculus					92	44				39
43	Nonhistone chromosomal protein HMG 17 Mus musculus		95		94		36				
44	60S ribosomal protein L23a Mus musculus				17		35				
45	DNA binding protein A Mus musculus				16		28		161		
46	60S ribosomal protein L30 Mus musculus		30				23				
47	Ubiquitin associated protein 2 Mus musculus				563				109		
48	Tubulin isoforms Mus musculus	128	75			271		156			
49	Thrombospondin 1 Mus musculus										19
50	Polyadenylate binding protein 1 Mus musculus				373						
51	Heat shock protein HSP 90 alpha Mus musculus					153		89			
52	Heat shock cognate 71 kDa protein Mus musculus					390		152			168
53	**Enolase yeast - ENL (16 fmoles)**			**6**	**10**					**8**	**8**
54	Endoplasmin Mesocricetus auratus					294					

Table I. HCP concentrations (ng/mL) measured in PTG1 mAb expressed in CHO (samples I-IV) and DG-44 cells (samples A-D). Two protein A purification protocols were compared across ten samples: I vs II - CHO cells grown under regular conditions; III vs IV - CHO cells non-expressing the PTG1 target; A1/2 vs B1/2 biological replicates for DG-44 cells grown under regular conditions; C vs D - non-expressing DG-44 cells. High abundance proteins are highlighted in RED (>1,000 ng/mL), medium abundance are either in YELLOW (500-1000 ng/mL) or GREEN (100-500 ng/mL) and low-abundance HCPs are GREY (< 100 ng/mL).

Sample ID	Cell line	PTG1 expressed?	No of HCPs Identified	Common HCPs	HCP concentration (ng/mL)	
					Mass Spec	ELISA
I	CHO	Y	5	4	4,870	7,013
II	CHO	Y	14	4	9,752	12,228
III	CHO	N	0	0	0	30
IV	CHO	N	31	0	9,093	4,138
A1	DG-44	Y	34	21	61,138	3,684
B1	DG-44	Y	50	21	109,437	8,042
A2	DG-44	Y	22	15	35,031	4,282
B2	DG-44	Y	32	15	105,123	8,041
C	DG-44	N	0	0	0	6
D	DG-44	N	28	0	1,454	1,318

Table II. Comparison of total HCP concentrations (ng/mL) measured by the mass spectrometry and ELISA (using the Cygnus Technologies CHO HCP mammalian kit).

The MS^E data was automatically mined using VERIFY^E software for the best proteotypic peptides from 20 HCPs. The MRM assay, implemented on a triple quadrupole instrument, used a 15-min gradient for monitoring 58 MRMs from 29 peptides (2 transitions per peptide), belonging to twenty HCPs.

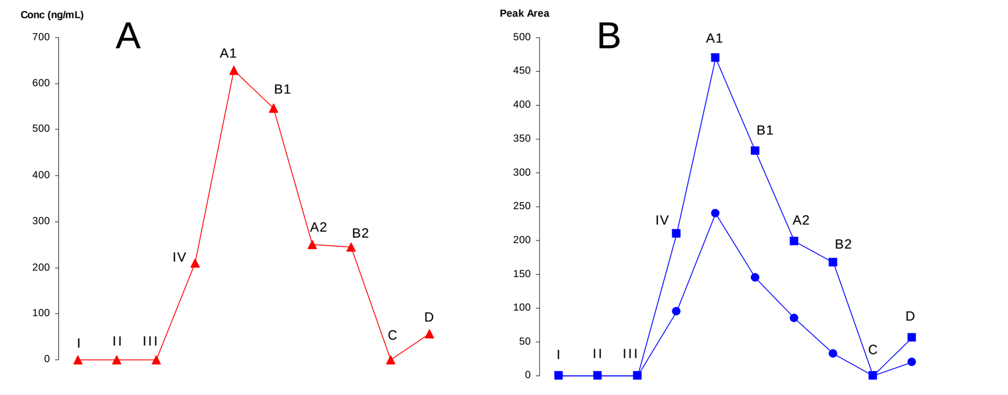

Figure 6. Correlation between the MS^E and MRM data: (A) concentration of 78 kDa glucose regulated protein measured across 10 different PTG1 preparations (samples I-IV and A-D) measured using the MS^E data; (B) peak areas produced by two MRM transitions of peptide VEIIANDQGNR from the same protein.

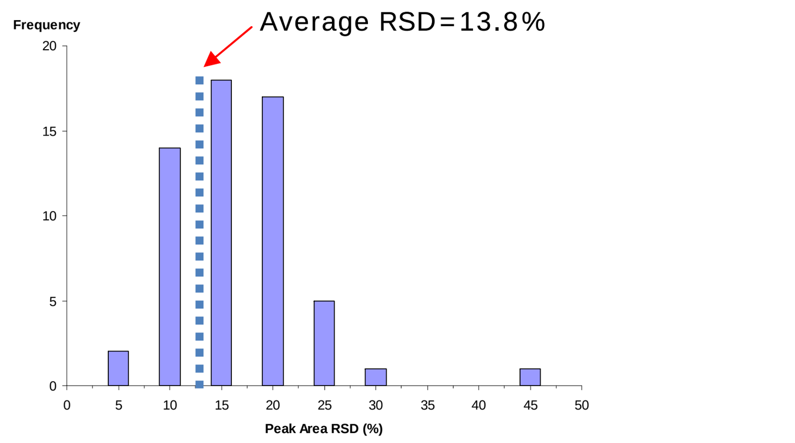

Figure 7. Histogram plot showing the peak area RSDs for 58 MRM transitions measured in 5 replicates for sample B1. The average RSD was 13.8%.

CONCLUSIONS

- The LC/MS^E-based assay allows the identification and quantification of low-abundance HCP contaminants in biopharmaceuticals over four orders of magnitude in concentration.
- Protein A purification of mAbs using different purification protocols produces different HCP patterns. Cell lines used for mAb purification can affect the HCPs identity and concentration significantly.
- ELISA assay seems to underestimate total HCP concentration.

References

1. Gilar M, Olivova P, Daly AE, Gebler JC J. *Sep. Sci*, **2005**, 1694.
2. Stapels MD, Fadgen K Current Trends in Mass Spectrometry, *Spectroscopy* supplement, March **2009**.
3. Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ *Proteomics*, **2009**, 1696.

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