Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Meeting Increased Pressures to Detect and Quantify Host Cell Proteins in Biotherapeutics by LC/MS

Sean M McCarthy, PhD

Introduction – Importance of HCPs

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

- Residual host cell proteins (HCPs) from recombinant production are classified as process-related impurities.
  - HCPs can elicit an unpredictable immune response in patients.
  - Need to be monitored as part of regulatory guidelines
- European regulations in effect since 1997:
  - '6.2 Validation of the purification procedure - …. The ability of the purification process to remove other specific contaminants such as host-cell proteins … should also be demonstrated'

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

- Current analytical methods (typically ELISA, gels, blots) are:
  - expensive,
  - subjective,
  - time-consuming to develop,
  - Require prior knowledge about the contaminant proteins
- Composition of HCPs is extremely heterogeneous and changes with production and purification procedures.

Business Impacts of HCP Issues

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

- During the development of Omnitrope, an immunogenicity issue emerged with an early version of the product.
- Up to 60% of patients enrolled in two clinical studies developed anti-hGH antibodies, and 100% developed anti-HCP antibodies.
- The cause was excess host cell protein levels, which was resolved by the manufacturer with purification process changes.

What is the cost of halting clinical trials or losing drug efficacy over time?

Today's approach to HCP's: 12 months to develop a product-specific quantitative immunoassay

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Immunoassays are sensitive but inflexible

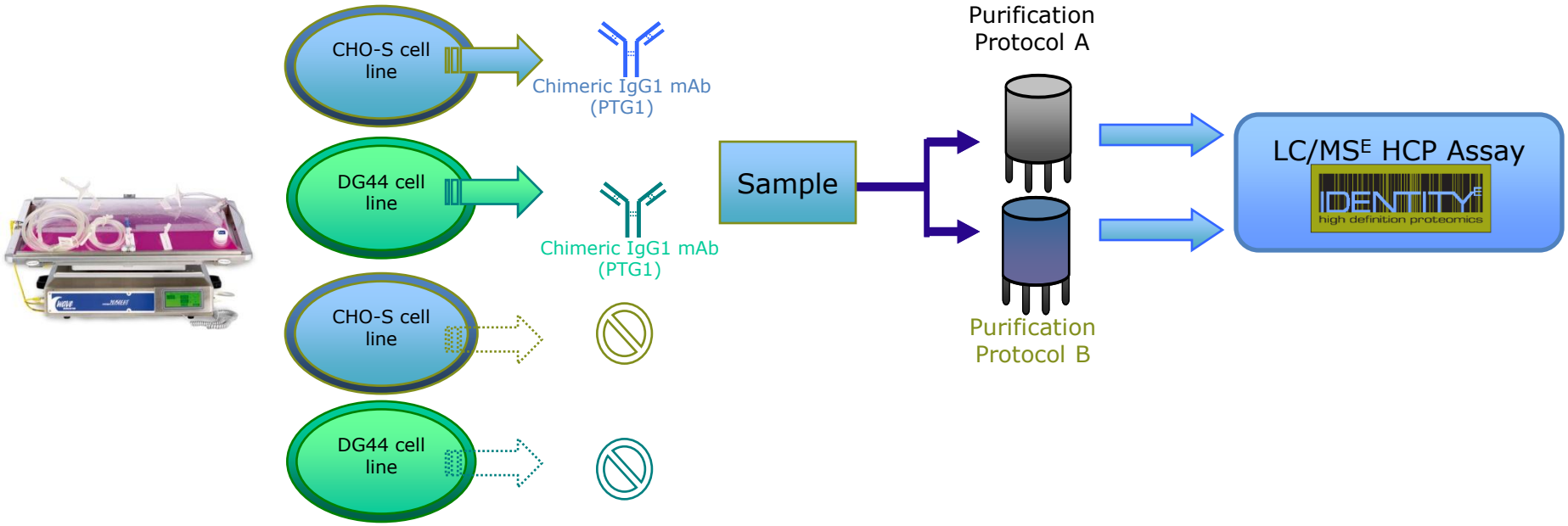
Figure 2. Development steps of a quantitative host cell protein assay

Antigen	Immunization	Qualitative Test of Sera	Quantitative Assay
about 2 months	3–9 months	1 month	1–2 months
Mock-run	Control of preimmune sera	Purification of antibodies (Protein A, affinity chromatography)	
Selection of antigen	Immunization (rabbit, goat etc.)	1-D and 2-D electrophoresis and Western blot	
Characterization of antigen	Control of sera (e.g., for cross reaction against product protein)		

Biopharm International, Volume 13, Number 6, pp. 38-45, May 2000

- **Process Changes may require new assay development.**
- **Hurdle to Biosimilar approval**
- **Regulators not all that happy with this current technology.**

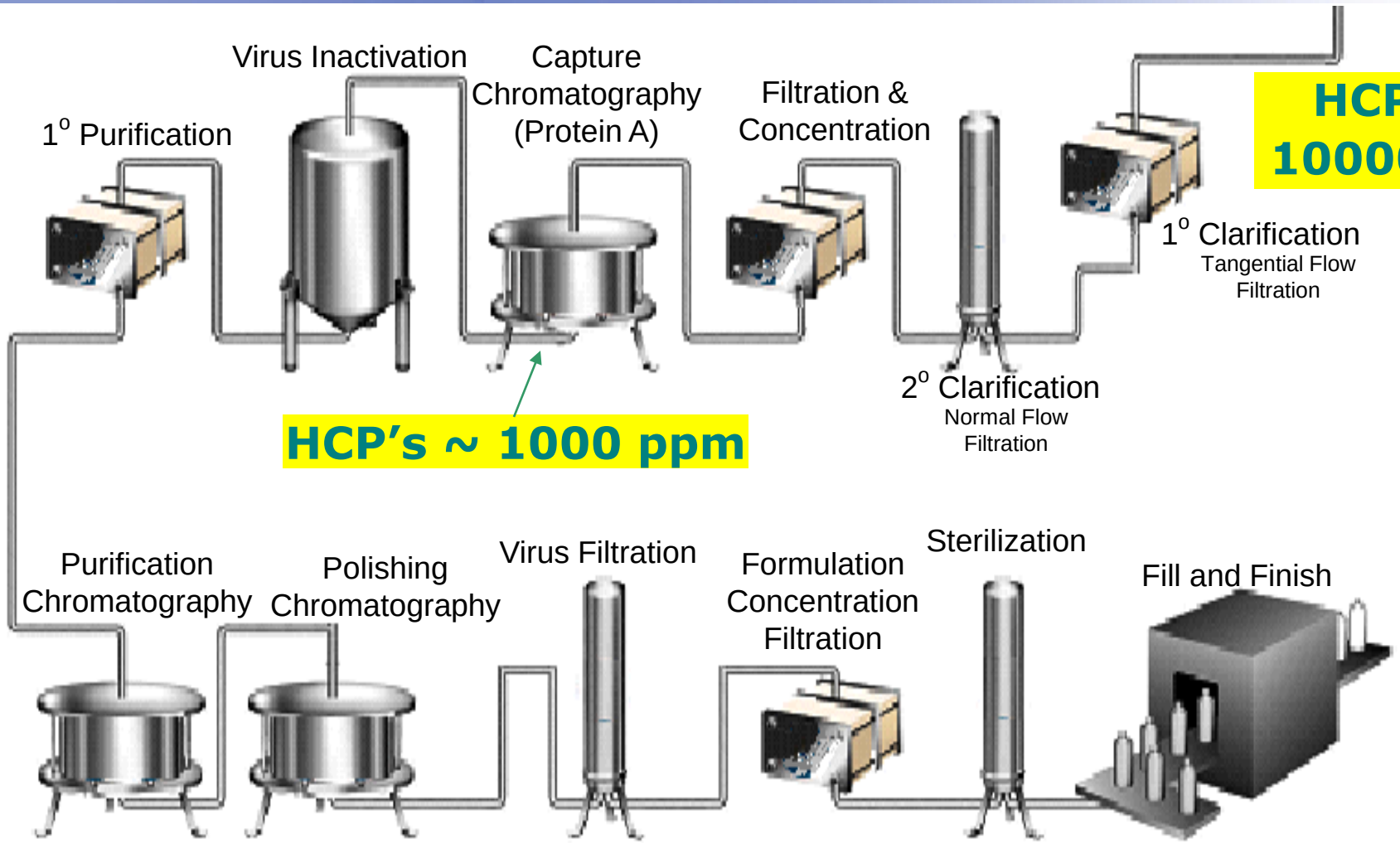
Antibody Expression



- Chimeric anti-phosphotyrosine IgG1 mAb (PTG1) expressed in 2 cell lines (CHO-S and DG44)
- Purified by Protein A chromatography (two different protocols).
- Two cell lines (DG44 and CHO-S) containing no PTG1 vector were also grown under the same conditions and purified on a Protein A column.
- Five proteins were spiked in PTG1
  - 4,000 fmoles LA (bovine beta-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).
- Samples analysed by LC/ MS^E

"Catalog" Host Cell Proteins (HCP's) early in downstream processing

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

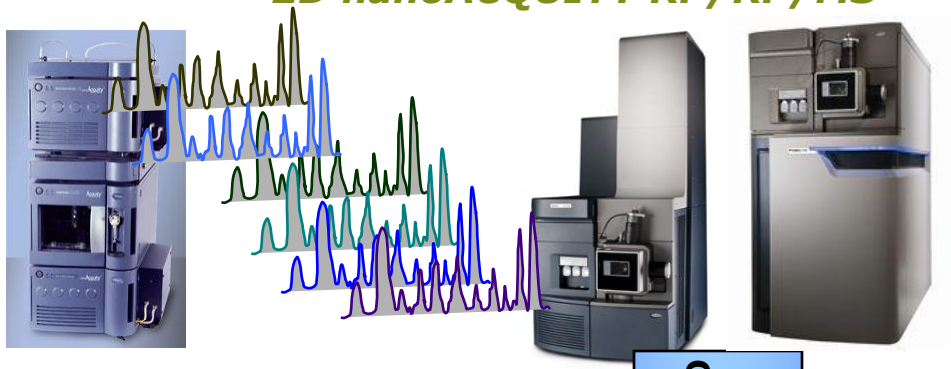


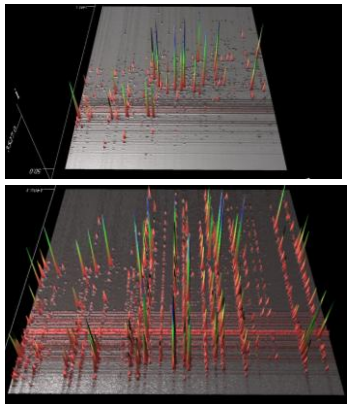
Cataloging HCP's in a biotherapeutic protein sample

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Tryptic Digestion

2D nanoACQUITY RP/RP/MS^E



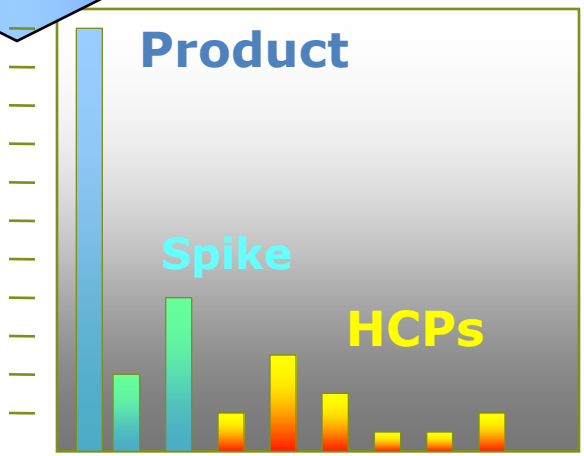


Identify HCPs
Quantify by Hi3

"Catalog" of HCP Process-impurity proteins

Routine Sensitivity ~ 50 ppm

ID no	Protein ID and species	Concentration (ppm or ng/m DG44 cells	CHO-S ce
1	78 kDa glucose regulated protein (hamster)	778	235
2	Actin isoforms (hamster)	3237	4275
3	Alcohol dehydrogenase (yeast) - 560 ppm (ADH)	90	188
4	Clusterin (mouse)	3306	5584
5	Cofilin 1 (mouse)	109	189
6	Elongation factor 1 and 2 (hamster)	1684	1247
7	**Enolase (yeast) - 13 ppm (ENL)**	**23**	**16**
8	Glyceraldehyde 3 phosphate dehydrogenase (hamster)	4594	6359
9	Glycogen phosphorylase (rabbit) - 530 ppm (PHO)	480	607
10	Heat shock 70 kDa protein (mouse)	9907	6532
11	Heat shock cognate 71 kDa protein (mouse)	628	322
12	Heat shock protein HSP 90 alpha(mouse)	**266**	**874**
13	alpha lactalbumin (bovine) - 1,200 ppm (LA)	654	725
14	Lipoprotein lipase (hamster)	1024	2358
15	Lysosomal alpha glucosidase (mouse)	2153	3385
16	Nucleolin (hamster)	1431	2891
17	Pyruvate kinase isozyme (mouse)	**683**	**450**
18	Serine protease HTRA1 (mouse)	741	952
19	**Serum albumin precursor (bovine) - 100 ppm (BSA)**	**100**	**100**
20	T complex protein subunits (mouse)	**8120**	**10563**
21	Tubulin alpha and beta chain (mouse)	**2464**	**7158**
	Percent composition for 21 common proteins	4.25	5.71



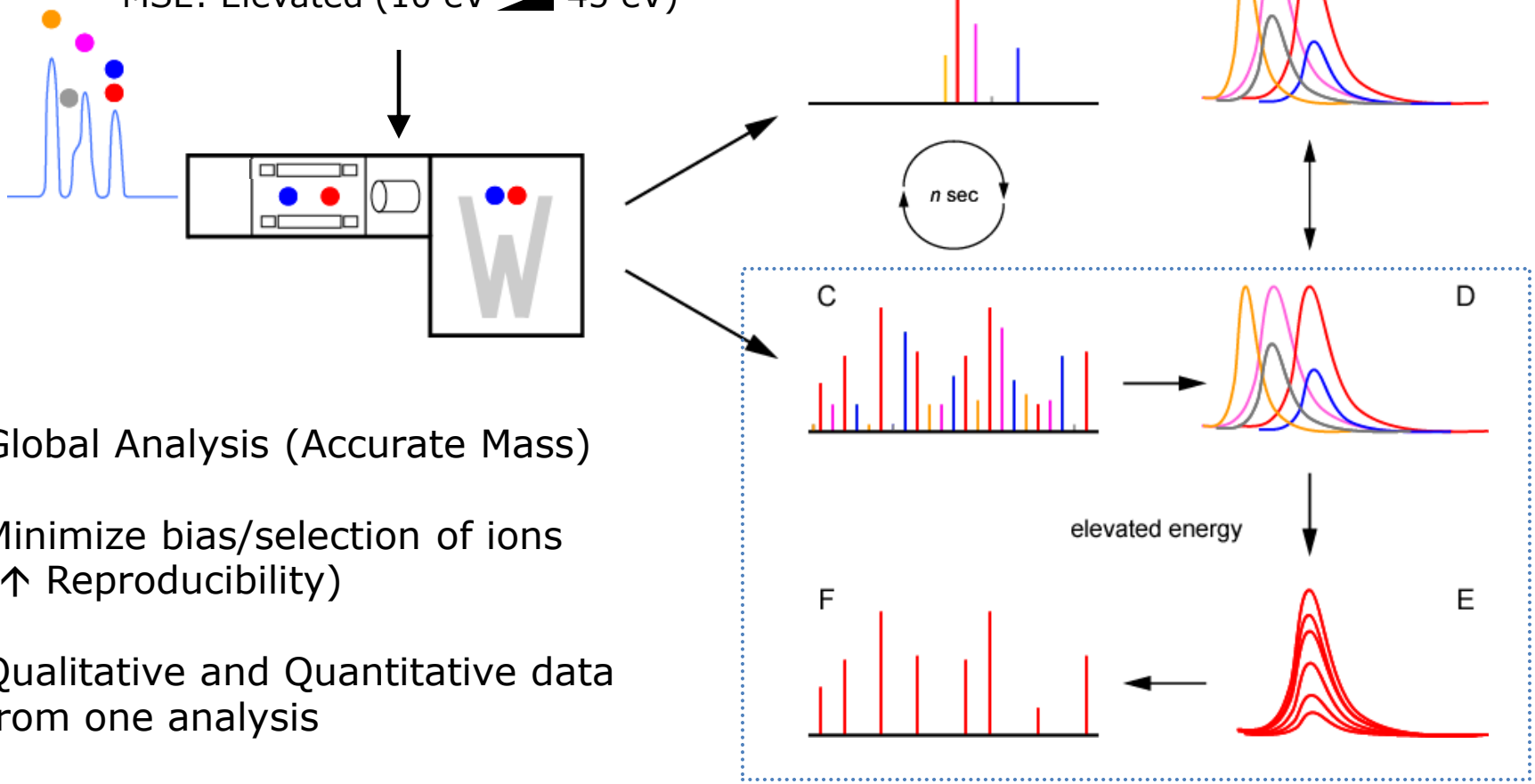
MS^E: An Alternate Scanning Methodology for Acquiring Peptide and Fragmentation Data

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Collision Cell Energy Alternates:

MS: Low (5 eV)

MSE: Elevated (10 eV ◢ 45 eV)

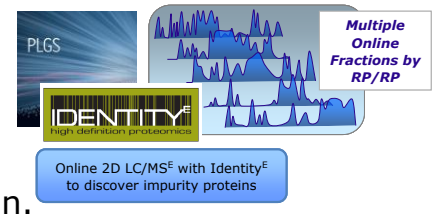


- Global Analysis (Accurate Mass)
- Minimize bias/selection of ions (↑ Reproducibility)
- Qualitative and Quantitative data from one analysis

Silva J.C. et al (2005) Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs Analytical Chemistry 77(7): 2187-2200.

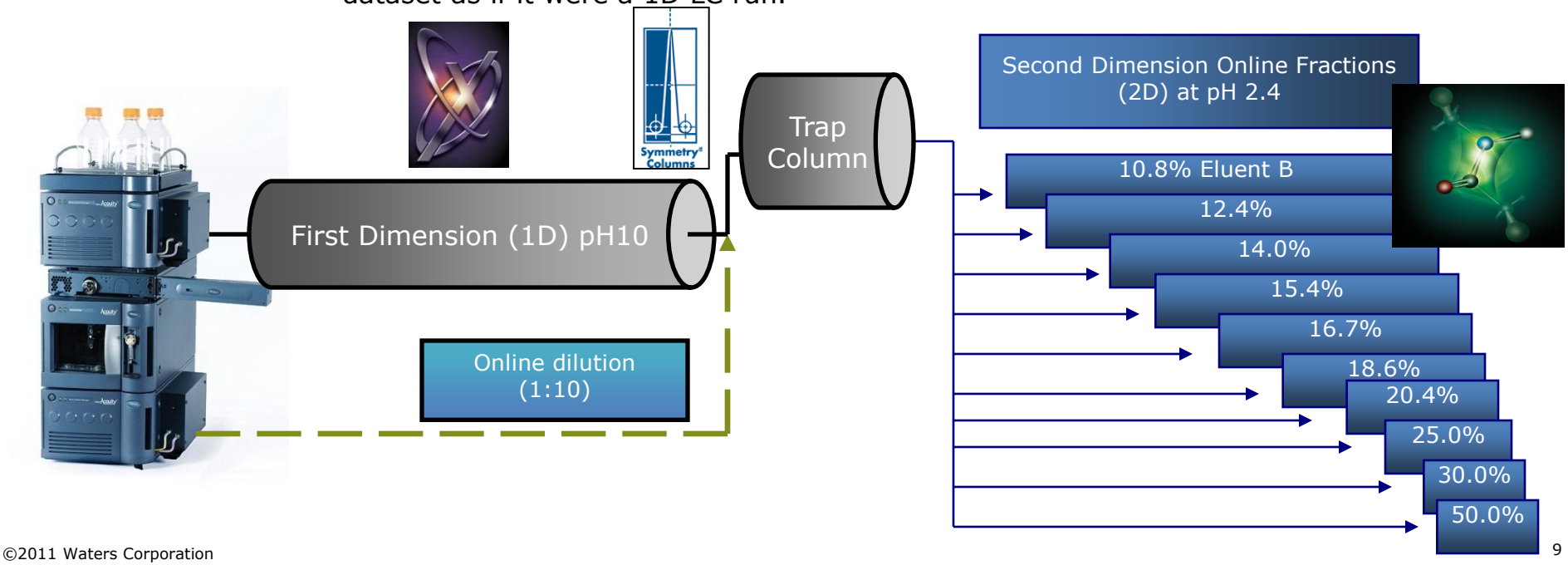
LC Conditions for HCP discovery

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™



nanoACQUITY™ UPLC® system with 2D technology

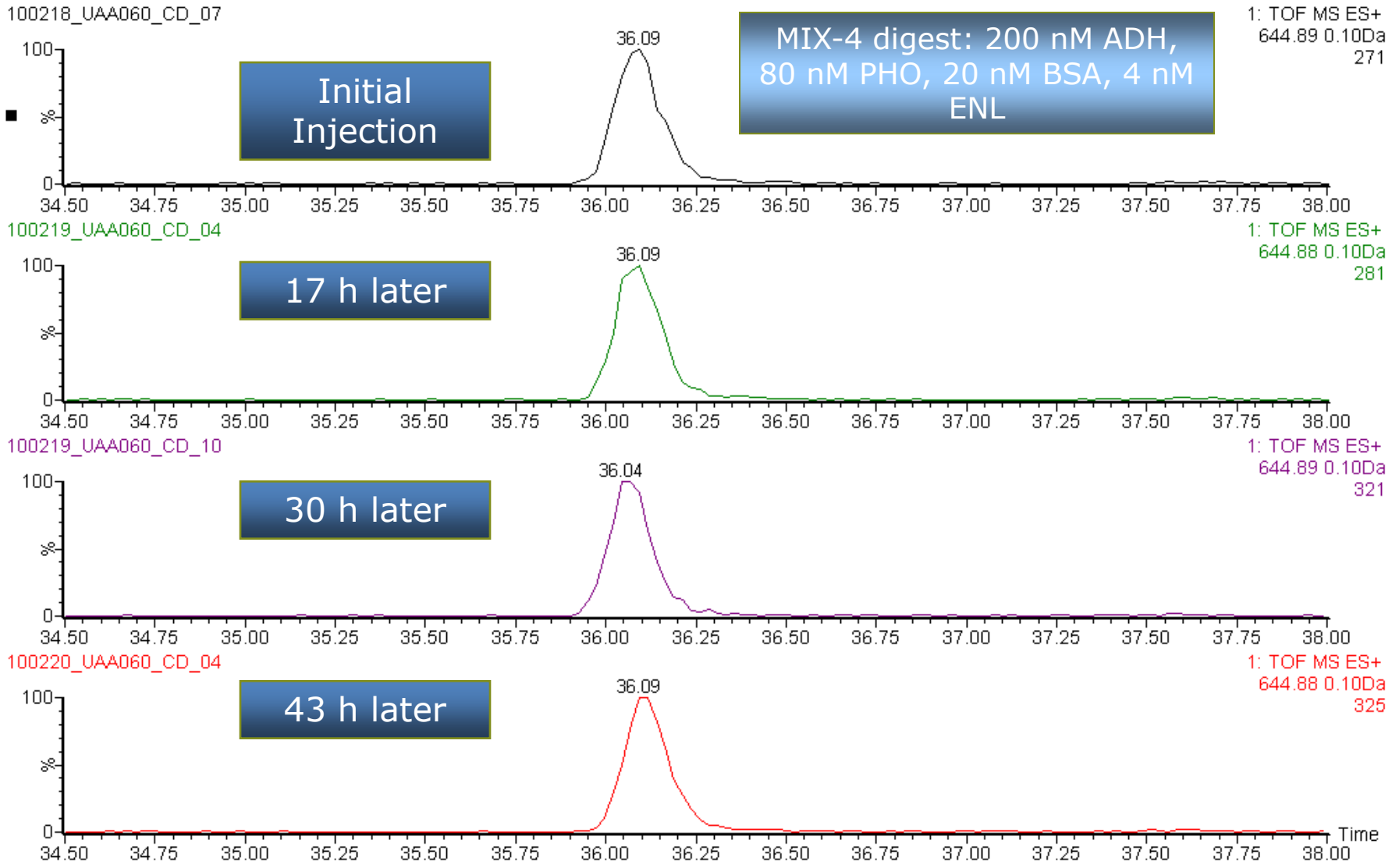
- First Dimension (1D) pH=10: High Loading Capacity, high pH resistance (XBridge)
- Online dilution (1:10) of the eluent from 1D before analyte trapping onto the 2D column.
- Trap column: 5-µm Symmetry for high carbon loading – facilitates peptide retention
- Second Dimension (2D) pH=2.4: analytical chromatography of each fraction in turn with longer gradient to maximize resolution for each fraction.
- Digital Fraction 'Merge': digitally group peptides that fall into different fractions by treating the dataset as if it were a 1D LC run.



Reproducibility of 2D Chromatography:

T43 ENL (VNQIGTLSESIK), 24 fmoles on column, Fraction 3/5
90 min gradients

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™



Absolute Protein Quantitation (Hi3) by LC-MS^E

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Research

Absolute Quantification of Proteins by LCMS^E

A VIRTUE OF PARALLEL MS ACQUISITION

Jeffrey C. Silva‡§, Marc V. Gorenstein‡, Guo-Zhong Li‡, Johannes P. C. Vissers¶, and Scott J. Geromanos‡

Relative quantification methods have dominated the quantitative proteomics field. There is a need, however, to conduct absolute quantification studies to accurately model and understand the complex molecular biology that results in proteome variability among biological samples. A new method of absolute quantification of proteins is described. This method is based on the discovery of an unexpected relationship between MS signal response and protein concentration: the average MS signal response for the three most intense tryptic peptides per mole of protein is constant within a coefficient of variation of less than ±10%. Given an internal standard, this relationship is used to calculate a universal signal response factor. The universal signal response factor (counts/mol) was shown to be the same for all proteins tested in this study. A controlled set of six exogenous proteins of varying concentrations was studied in the absence and presence of human serum. The absolute quantity of the standard proteins was determined with a relative error of less than ±15%. The average MS signal responses of the three most intense peptides from each protein were plotted against their calculated protein concentrations, and this plot resulted in a linear relationship with an R^2 value of 0.9639. The analyses were applied to determine the absolute concentration of 11 common serum proteins, and these concentrations were then compared with known values available in the literature. Additionally within an unfractionated *Escherichia coli* lysate, a subset of identified proteins known to exist as functional complexes was studied. The calculated absolute quantities were used to accurately determine their stoichiometry. *Molecular & Cellular Proteomics 5:144–156, 2006.*

The study of proteins is crucial in understanding and combating disease through identification of proteins, discovering disease biomarkers, studying protein involvement in specific metabolic pathways, and identifying protein targets in drug discovery (1, 2). An important technique that is used in these studies to quantify and identify peptides and/or proteins present in simple and complex mixtures is ESI-LCMS.

To date a majority of the quantitative proteomic analyses have been performed using stable isotope labeling strategies such as ICAT (3), iTRAQ™ (4), SILAC (stable isotope labeling by amino acids in cell culture) (5), and ^{18}O labeling (6, 7). These methodologies require complex, time-consuming sample preparation and can be relatively expensive.

Recently there have been numerous reports applying label-free methods to monitor the relative abundance of protein between different conditions (8–11). Relative quantification provides information regarding specific protein abundance changes between two conditions caused by an induced perturbation (environment-induced, drug-induced, and disease-induced). These studies require comparison of identical proteolytic peptides in each of the two experiments to accurately determine relative ratios of the particular protein(s) of interest. Relative abundance values for each peptide to a given protein can then be obtained to quantitatively characterize the differential expression of proteins between different sample states. Many of these methods are based on determining the ratios of the peak area of identical peptides between different conditions. One critical factor limiting the quantitative reproducibility of these methods includes the ability to efficiently cluster the detected peptides. This in turn relies on the accuracy of the mass measurement and the chromatographic reproducibility. Although relative quantification monitors changes in protein abundance between two conditions, it does not determine the absolute quantity of these proteins.

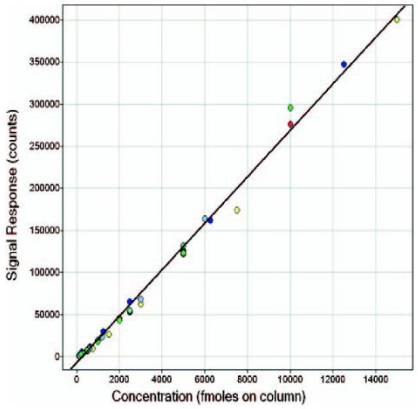
The ability to determine the absolute concentration of a protein (or proteins) present within a complex protein mixture is valuable for the understanding of the underlying molecular biology guiding the response to an applied perturbation. Cellular responses are often controlled through direct and indirect interactions of proteins present in the cell. These coordinated interactions allow the cell to communicate a response across many cellular compartments. The cell can thereby execute an efficient and expeditious recruitment and production of critical proteins needed for adaptation. A method for determining the absolute quantity of proteins in a complex sample would enable determination of the stoichiometry of proteins within a sample and would facilitate understanding of the complicated biological network of cooperative protein interactions that guide cellular responses.

To date a technique capable of determining the absolute concentration of proteins in complex mixtures from a simple LCMS analysis without using specific internal standards for

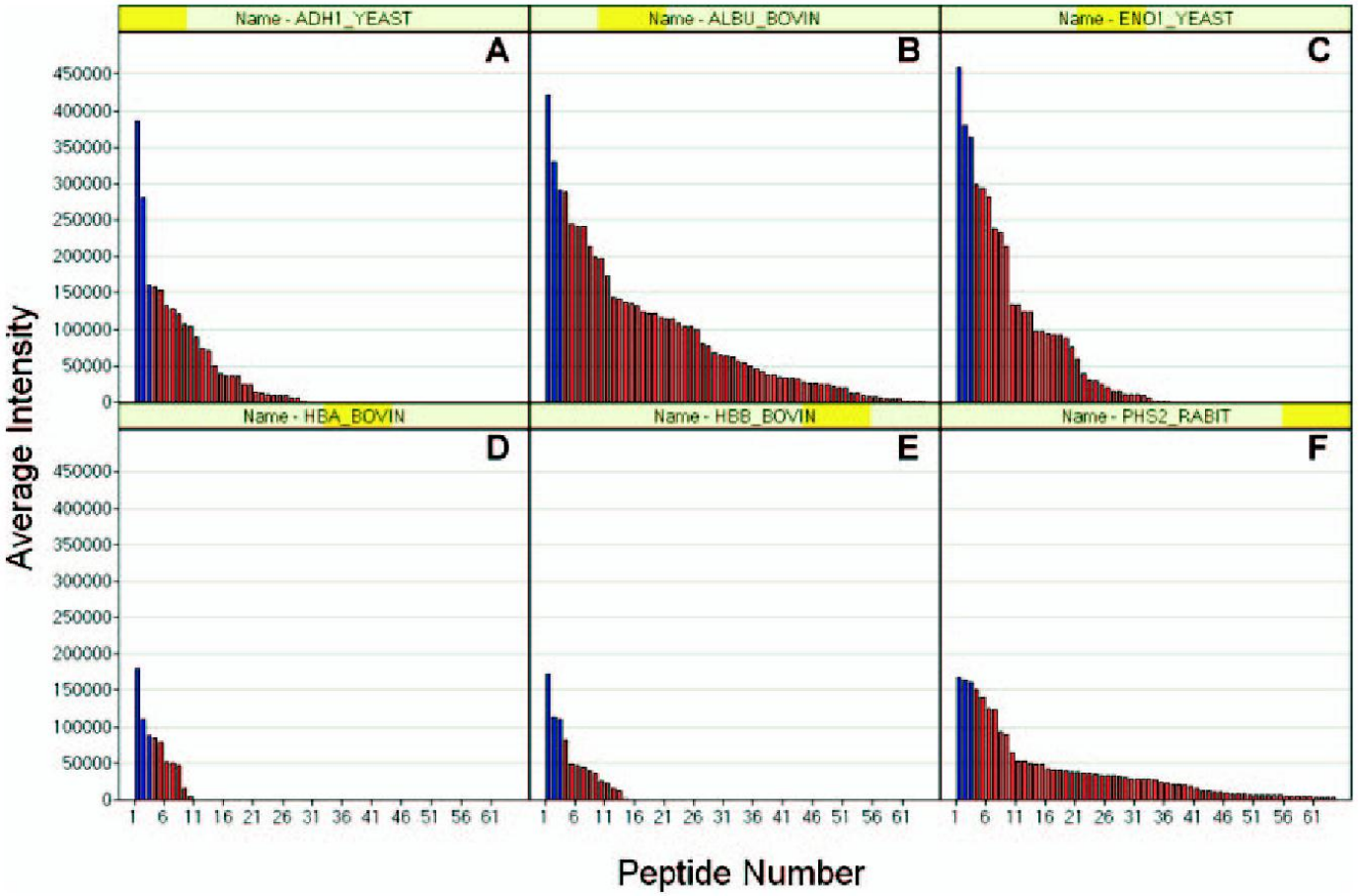
From the ‡Waters Corporation, Milford, Massachusetts 01757-3696 and ¶Waters Corporation, Transistordreef 18, 1322 CE Almere, The Netherlands

Received, July 26, 2005, and in revised form, September 23, 2005
Published, MCP Papers in Press, October 11, 2005, DOI 10.1074/mcp.M500290-MCP200

144 *Molecular & Cellular Proteomics 5.1*



Basis: Intensity of Top 3 Peptides of each Protein



Molecular & Cellular Proteomics 5:144–156, 2006.

Example: Comparing HCPs in a therapeutic mAb isolated by two Protein A purification methods

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Post-Protein A Purification Step	Method A	Method B
No. proteins identified	49	73
HCP Levels	1.0x	2.7x
Total HCP by ELISA Assay (ppm)	1.0x	3.5x

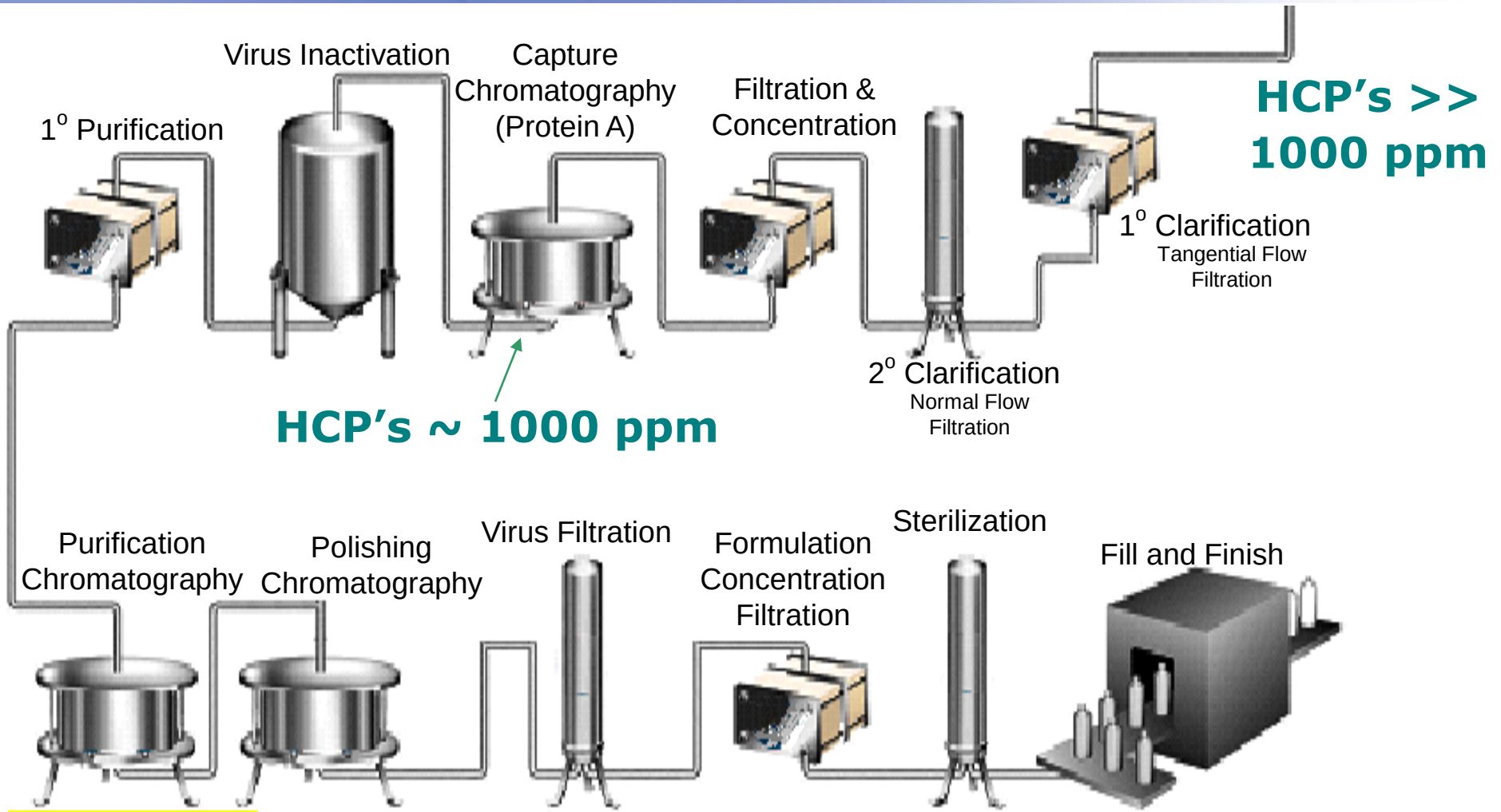
Protein ID and species	Protocol A (ppm or ng/mg)	Protocol B (ppm or ng/mg)
40S ribosomal protein S3 (mouse)	663	561
	216	83
	580	1265
	2416	1844
	560	690
	2467	1899
	82	87
	51	219
	1256	3204
	89	451
	3428	2386
	530	652
	7393	**325**
	468	448
	199	754
	1,200	1480
	764	731
Lysosomal alpha glucosidase (mouse)	1607	541
Nucleolin (hamster)	**1068**	**14191**
Peroxiredoxin 1 (mouse)	122	229
Procollagen C endopeptidase enhancer (mouse)	**89**	**2995**
Pyruvate kinase isozyme (mouse)	510	1624
S methyl 5 thioadenosine phosphorylase (mouse)	34	70
Serine protease HTRA1 (mouse)	553	705
Serum albumin precursor (bovine) - 80 fmoles IS	100	123
T complex protein subunits (mouse)	**6060**	**777**
Transcription factor HES 5 (mouse)	104	293
Tubulin alpha and beta chain (mouse)	1839	10072

Determine:

- ✓ **Total HCP Level**
- ✓ **HCP Complexity**
- ✓ **Which Proteins?**
- ✓ **At what Levels?**

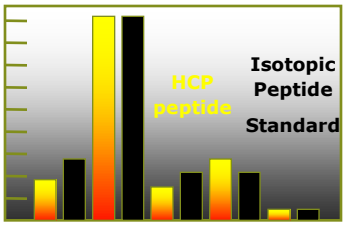
Monitor Host Cell Proteins (HCP's) later in downstream processing using MRM

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™



MRM Monitoring of HCP's

- Verify facilitates development of MRM methods for multiple HCP's via "signature" peptides
- Absolute quantification can be performed by using stable isotopically labeled peptide standards
- **Assay extremely simple to update/ change**



Targeted MRM Quantification of HCP Signature Peptides

ID no	Protein ID and species	Concentration (ppm or ng/mg) DG44 cells	CHO-S cells
1	78 kDa glucose regulated protein (hamster)	778	2354
2	Actin isoforms (hamster)	3237	4275
3	Alcohol dehydrogenase (yeast) - 560 ppm (ADH)	90	188
4	Clusterin (mouse)	3306	5584
5	Cofilin 1 (mouse)	109	189
6	Elongation factor 1 and 2 (hamster)	1684	1247
7	Enolase (yeast) - 13 ppm (ENL)	23	16
8	Glyceraldehyde 3 phosphate dehydrogenase (hamster)	4594	6359
9	Glycogen phosphorylase (rabbit) - 530 ppm (PHO)	480	607
10	Heat shock 70 kDa protein (mouse)	9907	6532
11	Heat shock cognate 71 kDa protein (mouse)	628	322
12	Heat shock protein HSP 90 alpha(mouse)	266	874
13	alpha lactalbumin (bovine) - 1,200 ppm (LA)	654	725
14	Lipoprotein lipase (hamster)	1024	2358
15	Lysosomal alpha glucosidase (mouse)	2153	3385
16	Nucleolin (hamster)	1431	2891
17	Pyruvate kinase isozyme (mouse)	683	450
18	Serine protease HTRA1 (mouse)	741	952
19	Serum albumin precursor (bovine) - 100 ppm (BSA)	100	100
20	T complex protein subunits (mouse)	8120	10563
		2464	7158
		4.25	5.71

IDENTITY^E high definition proteomics

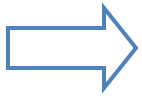
VERIFY^E high definition proteomics

Development of an HCP MRM Assay

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Discovery Phase (QTof)

Identify HCPs using LCMS^E



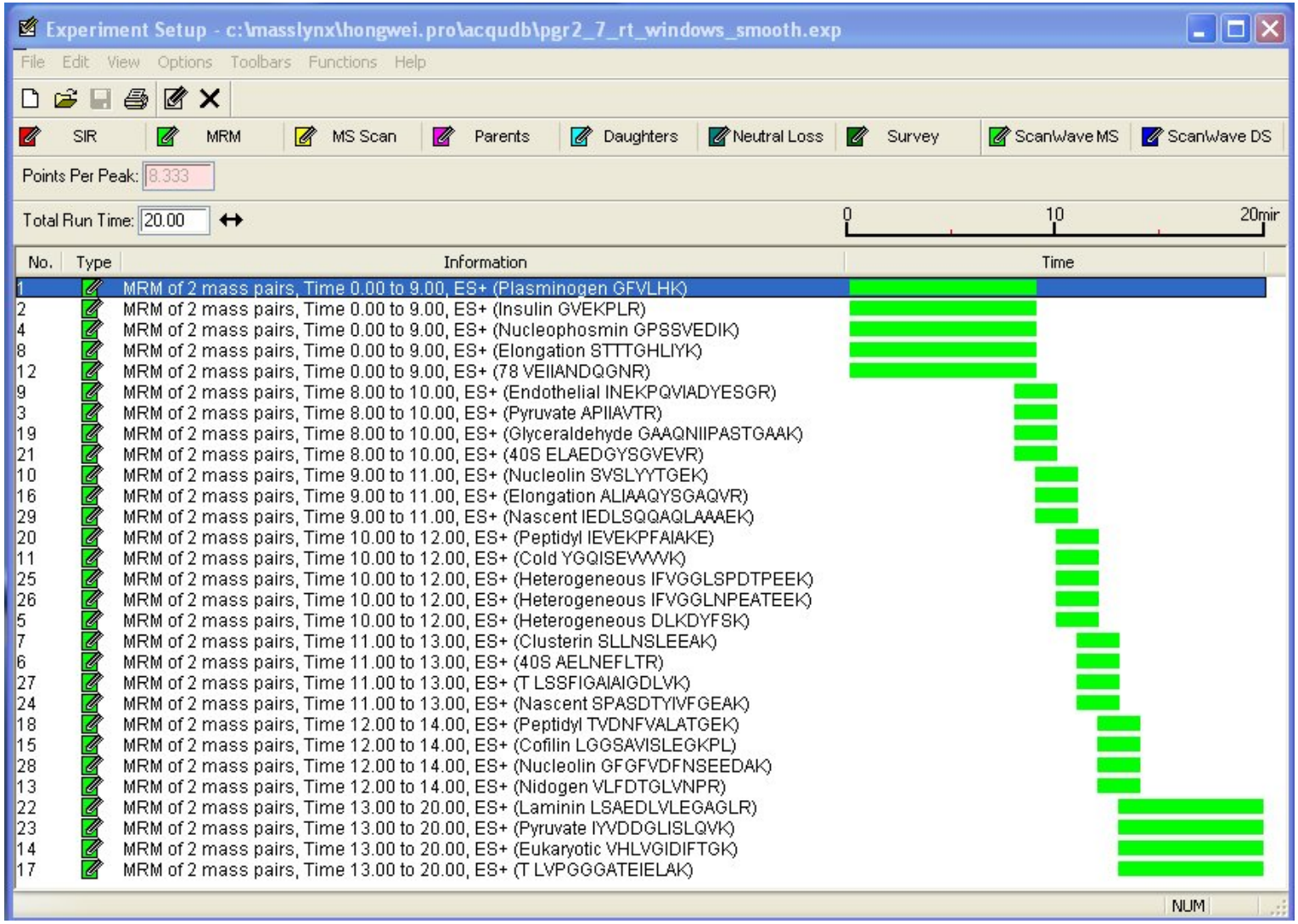
HCP Protein ID
Peptides (RT, Intensity)
MS^E Fragmentation

Monitoring Phase (QQQ)

- Select 3-5 peptides per protein (best ionizing, good fragmentation)
  - Produce recombinant HCP?
- Produce synthetic peptides
- Identify best 2-3 MRM transitions per peptide (Infusion Experiment)
  - Optimize Cone Voltage
  - Optimize Collision Energy
- LC/MS method development
  - Spike peptides into biotherapeutic digest to identify product interferences
  - (Optional) Stable isotope labeled peptide for absolute quantitation
  - Investigate the linearity of the MRM assay

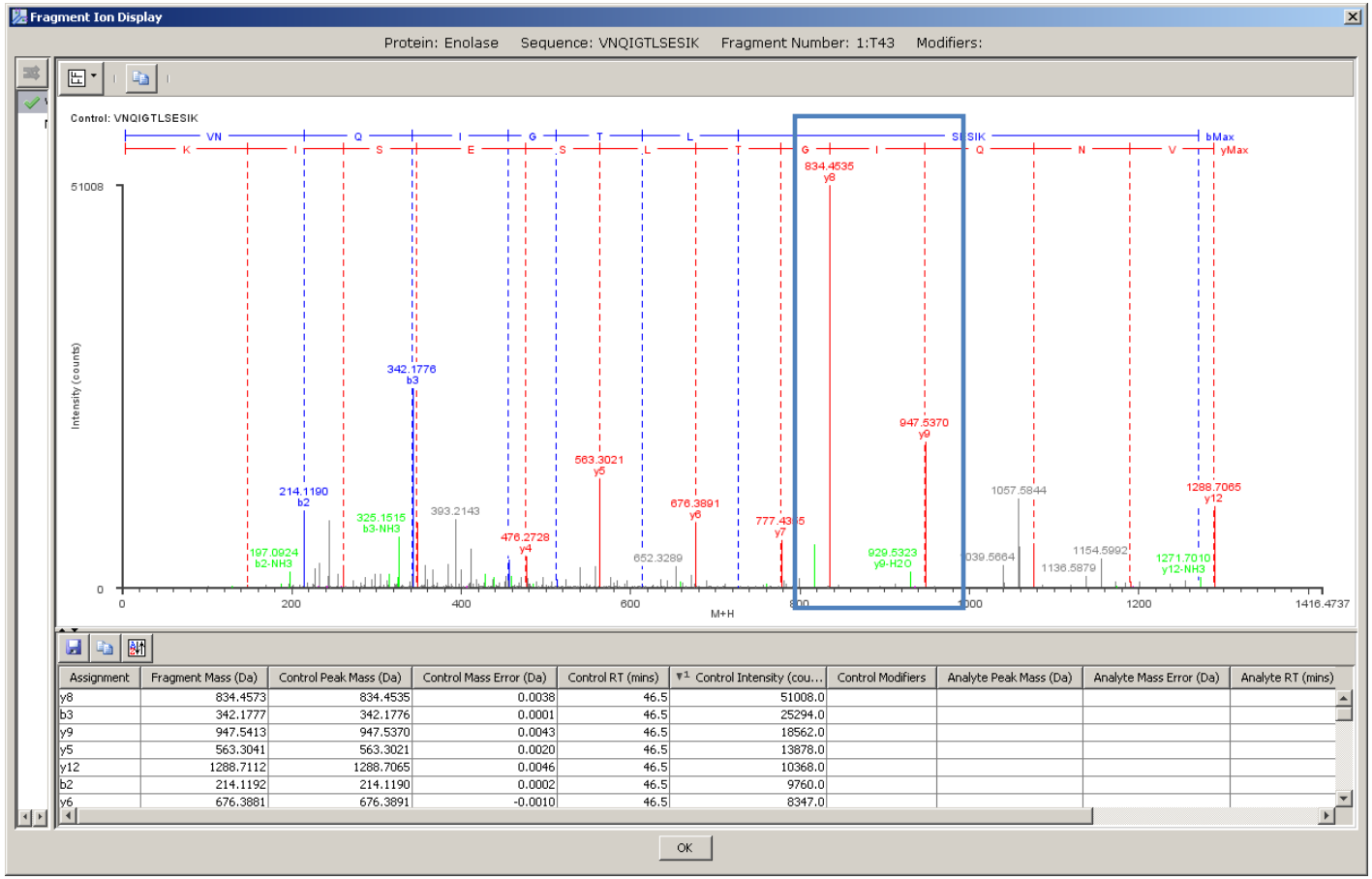
QQQ MRM method to monitor Top 20 HCP Proteins (29 Peptides, 58 Transitions) from method 1

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

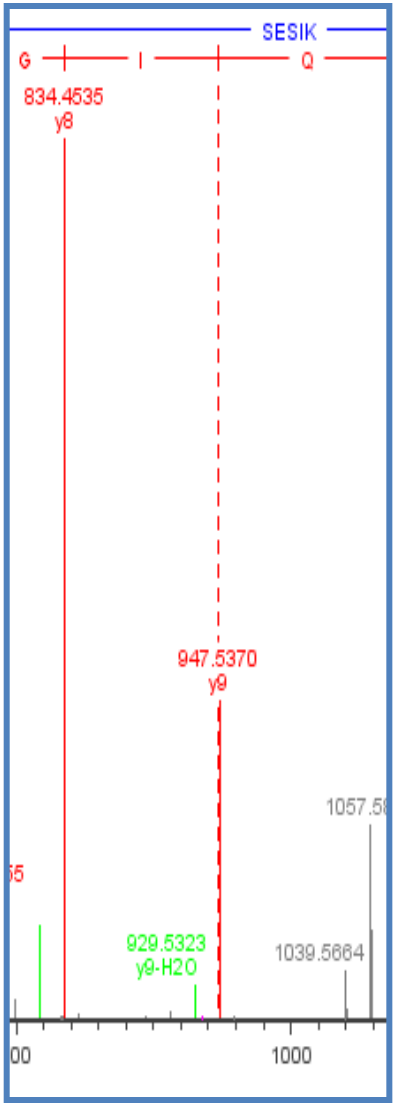


HCP T44 Peptide MS^E fragmentation

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

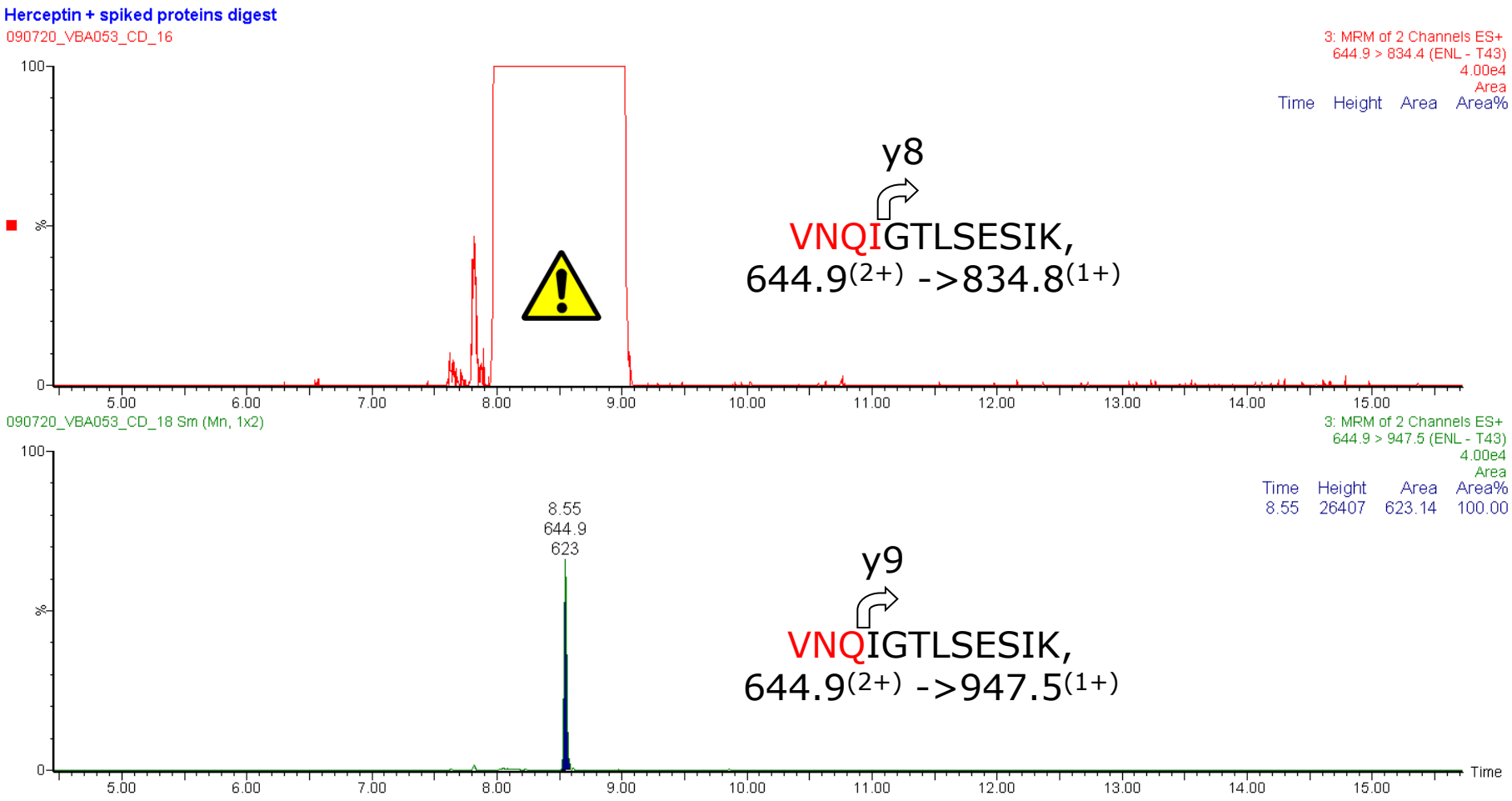


Assignment	Fragment Mass (Da)	Control Peak Mass (Da)	Control Mass Error (Da)	Control RT (mins)	Control Intensity (cou...	Control Modifiers	Analyte Peak Mass (Da)	Analyte Mass Error (Da)	Analyte RT (mins)
y8	834.4573	834.4535	0.0038	46.5	51008.0				
b3	342.1777	342.1776	0.0001	46.5	25294.0				
y9	947.5413	947.5370	0.0043	46.5	18562.0				
y5	563.3041	563.3021	0.0020	46.5	13878.0				
y12	1288.7112	1288.7065	0.0046	46.5	10368.0				
b2	214.1192	214.1190	0.0002	46.5	9760.0				
y6	676.3881	676.3891	-0.0010	46.5	8347.0				

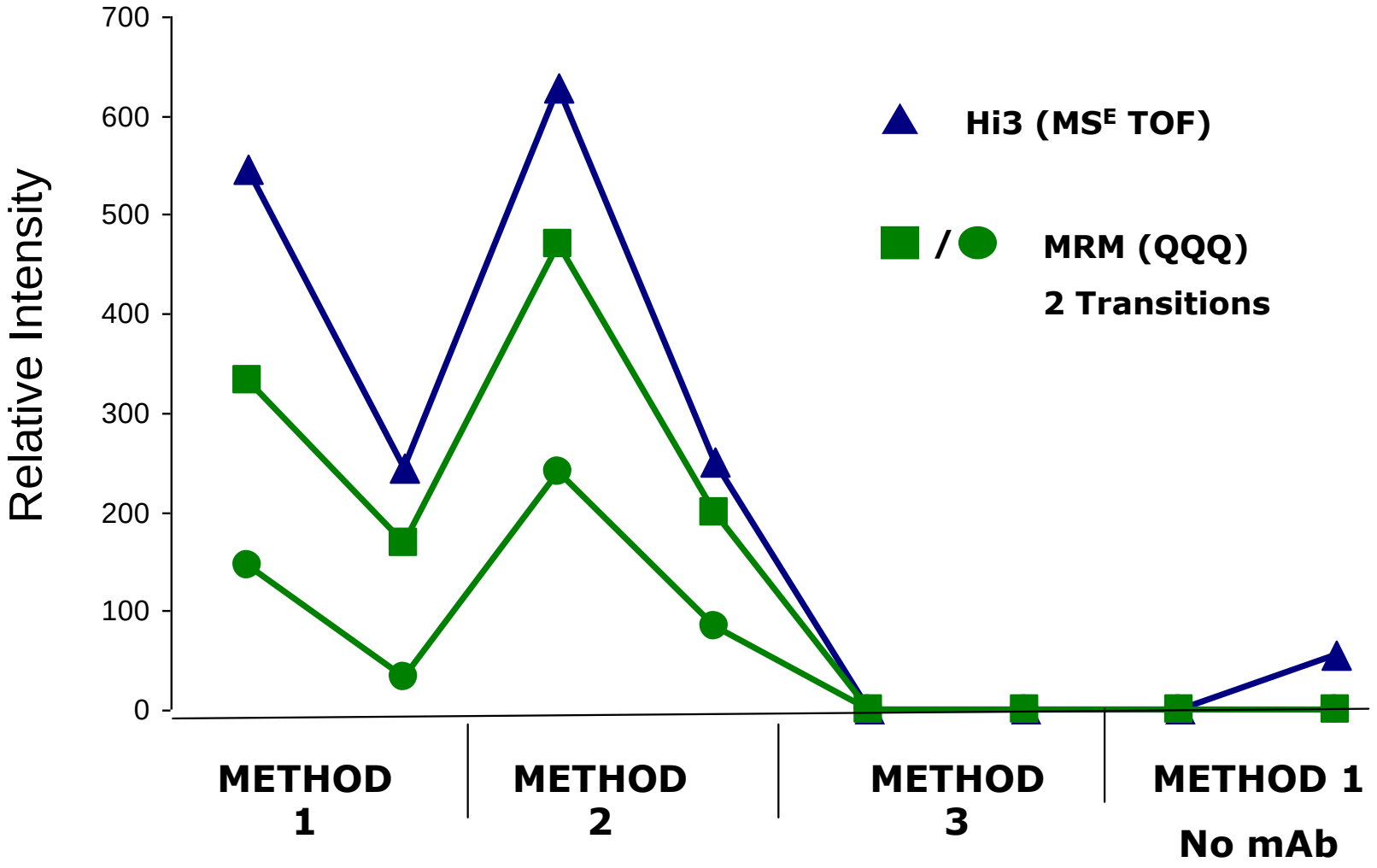


MRM Interference for a HCP Peptide MRM Transition

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™



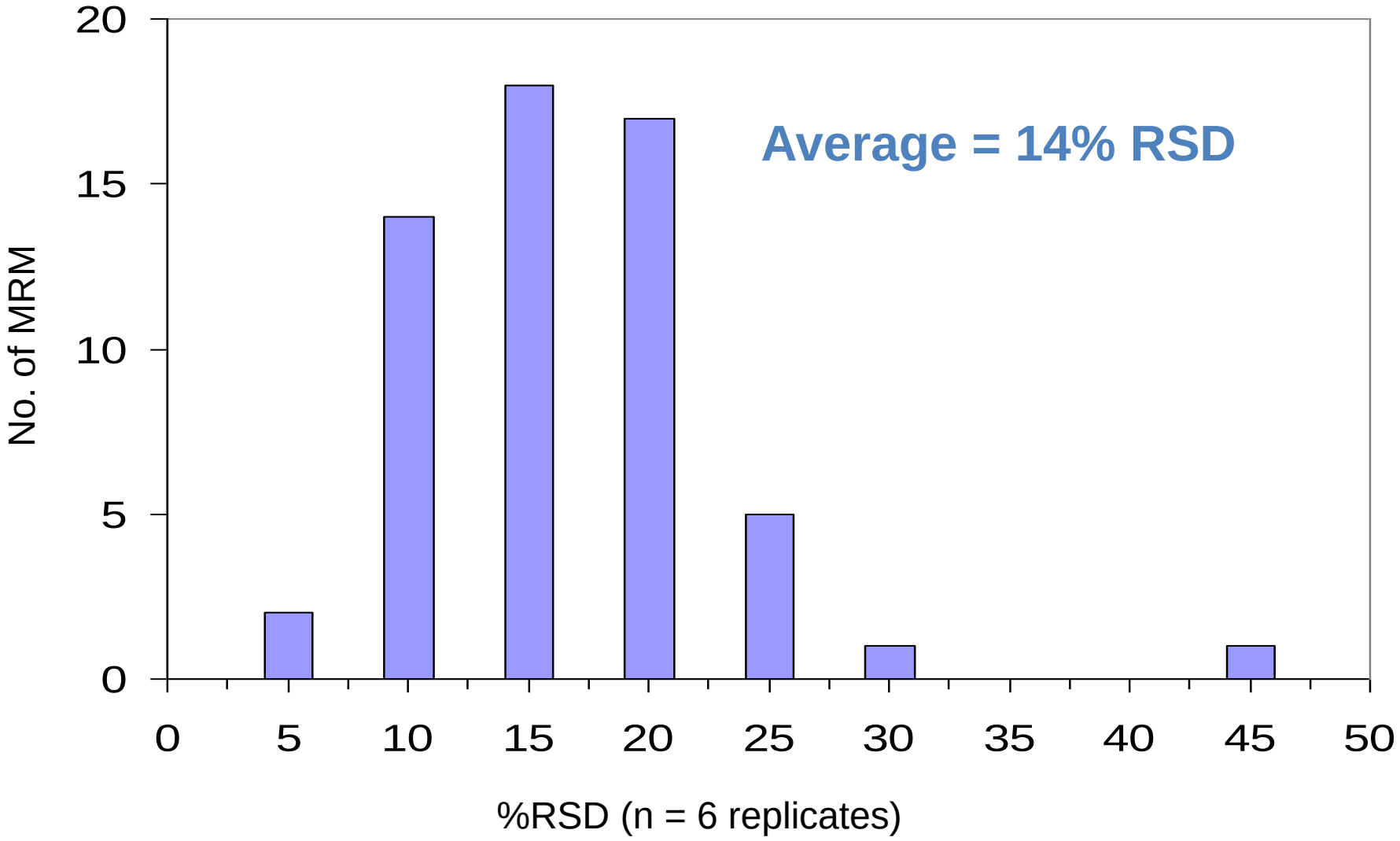
Hi3 vs. MRM quantitation data for a 58 kD HCP demonstrates common results



Analytical Reproducibility for 29 HCP Proteins (58 MRM Transitions)

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Method 1 Sample



More Flexibility for HCP Analysis

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

- Using the same technologies you use for protein characterization you can also identify HCP's in your product.
  - 2DLC/MS^E and the most accurate bioinformatics platform are all required to produce sensitive high quality results
  - Catalog (Identify and quantitate) and compare HCP's in early stages of processing.
- Easily developed MRM assays can monitor low ppm level HCP's in later processing steps and in the purified drug.
  - These MRM assays are readily modified to facilitate faster process development/improvement cycles.
  - Such MRM assays are commonly used today for regulated analyses, and could be readily validated for future QC needs.

Waters

THE SCIENCE OF WHAT'S POSSIBLE.™