Combination of Isobaric Tagging Reagents and Cysteinyl Peptide Enrichment for In-Depth Quantification

Protein Expression Analysis using the TripleTOF® 5600 System and iTRAQ® Reagents

Vojtech Tambor[1], Christie Hunter[2], Sean L. Seymour[2], Marian Kacerovsky[3], Jiri Stulik[1] and Juraj Lenco[1]

[1]Institute of Molecular Pathology, Faculty of Military Health Sciences, University of Defence, Hradec Kralove, Czech Republic, [2]SCIEX, USA, [3]Department of Obstetrics and Gynecology, University Hospital Hradec Kralove, Czech Republic

The high complexity, wide concentration range and dynamic nature of proteins in human samples raises many challenges in biomarker identification and quantification. Using mass spectrometry, the detection and quantification of the lower abundance proteins is usually prohibited by inherently low signal levels, ion suppression and instrument sampling issues. The high MS/MS sensitivity and resolution of the TripleTOF 5600 system enables high acquisition rates (up to 50 MS/MS per second) for in-depth analysis of complex proteomes.

The dynamic range challenge of biological samples has also led to the development of a range of enrichment techniques. In 1999, an innovative strategy for global quantification of proteins was introduced, the ICAT® Reagents[1]. This reagent specifically labeled cysteine residues with an isotope-coded affinity tag, for subsequent enrichment of cysteine labeled peptides, significantly reducing the complexity of the proteomic sample. In this work, the enrichment benefits of cysteine capture are combined with the high multiplexing properties of isobaric iTRAQ reagents, to create a powerful new quantitative workflow[2]. The workflow is based on the standard iTRAQ reagent protocol, enhanced by a subsequent cysteinyl-peptide enrichment (CPE)[3]. In contrast to the ICAT reagent approach, the separation of enrichment and reagent labeling enables quantitative analysis of both cysteinyl and non-cysteinyl peptide fractions.

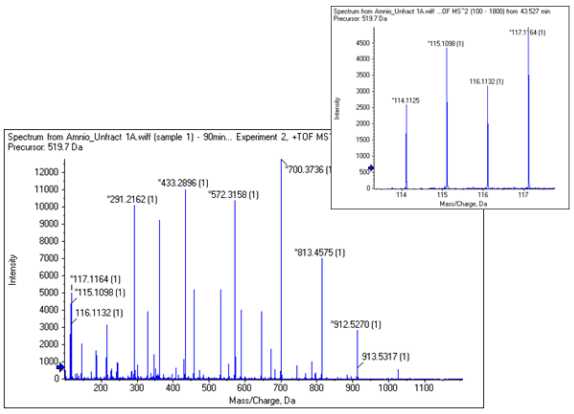

Figure 1. High Quality MS/MS Acquisition. High quality MS/MS is acquired using high sensitivity mode at a rate of 15 spectra per 1.8 sec. Resolution of >15,000 resolution is typically obtained on the fragment ions and reporter ions ensuring confident identification and quantification.

Here, the combination of the high-resolution TripleTOF 5600 system with this powerful sample preparation strategy significantly increased the number of identified and quantified proteins in preparations of human amniotic fluid.

Complete Solution for Protein Expression Analysis

- TripleTOF 5600 System
  - High speed acquisition of MS/MS for greatest depth of coverage, maintaining fragment ion resolution of >15,000
  - High resolution and mass accuracy, especially at low mass provides for accurate quantification of reporter ion signals
  - High sensitivity CID fragmentation and TOF detection provides excellent reporter ion intensities for quantitation
- SCIEX nanoLC Ultra™ and cHiPLC® system provides robust, high quality separations for greater sample coverage
- ProteinPilot™ Software
  - Powerful identification and quantification algorithms with enhanced scalability for large protein expression datasets.
  - ProteinPilot Descriptive Statistics Template[6,7] for detailed analysis of quantification and identification metrics.
  - Protein Alignment Template[7] for comparing results across multiple samples
- A complete data acquisition and processing solution scalable meet the throughput demands of contemporary high throughput proteomics and systems biology laboratories.

Experimental Section

Sample Preparation: Amniotic fluid samples were supplied by the Department of Obstetrics and Gynecology, University Hospital, Hradec Kralove. Four amniotic fluid samples were analyzed; two samples in which intraamniotic infection (IAI) was confirmed (IAI positive, n=2), and two samples in which infection was ruled out, (IAI negative, n=2). Respective samples were pooled and then split into two equal parts to generate duplicates. Samples were then depleted from 14 highly abundant proteins using immunoaffinity chromatography (Mars Hu-14, Agilent). Flow-through proteins were reduced using TCEP, digested with trypsin and labeled with iTRAQ® reagents according to the SCIEX protocol. The two IAI-negative samples were labeled with 114 and 116, the two positive samples were labeled with 115 and 117. The CPE workflow is based on the protocol originally described by Liu *et al*.[3] and was further optimized here (Figure 2)[2].

Chromatography: Separation of unfractionated, cysteinyl and non-cysteinyl samples was performed on a SCIEX NanoLC Ultra™ 2D System, and the cHiPLC® system in trap elute mode. In each injection, (1 µg) was desalted on a 200 µm x 0.5 mm trap chip and then eluted onto a 75 µm x 150 mm column chip for MS analysis. Both the trap and column chips were filled with ChromXP™ C18-CL 3µm 120Å phase. Peptides were separated using a linear gradient formed by A (2% ACN, 0.1% FA) and B (98% ACN, 0.1% FA), from 12–32% of B over 90 minutes at a flow rate of 300 nL/min.

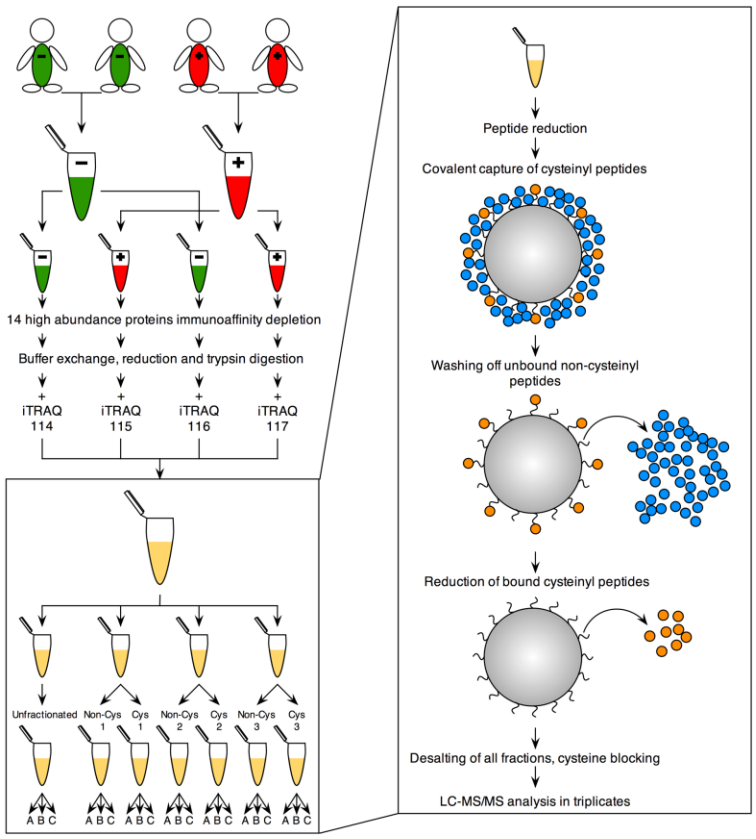

Figure 2. Cysteine Capture Workflow for Biomarker Discovery using iTRAQ Reagent in Amniotic Fluid Samples. Representative IAI positive and IAI negative samples were processed as described[2]. Resulting peptides were iTRAQ reagent labeled, combined, desalted, split into four aliquots and dried. Three of these aliquots were fractionated using CPE. The right part of the picture shows the enrichment of cysteinyl peptides in closer detail.

Mass Spectrometry: The MS analysis was performed on a SCIEX TripleTOF® 5600 system in Information Dependent Mode (IDA). MS spectra were acquired in high resolution mode (>30,000) using 250 ms accumulation time *per* spectrum. A maximum of 15 MS/MS *per* cycle were acquired with a 100 msec minimum accumulation time for each precursor and dynamic exclusion for 20 sec. Tandem mass spectra were recorded in high sensitivity mode (resolution >15,000) with rolling collision energy on and iTRAQ reagent collision energy adjustment on.

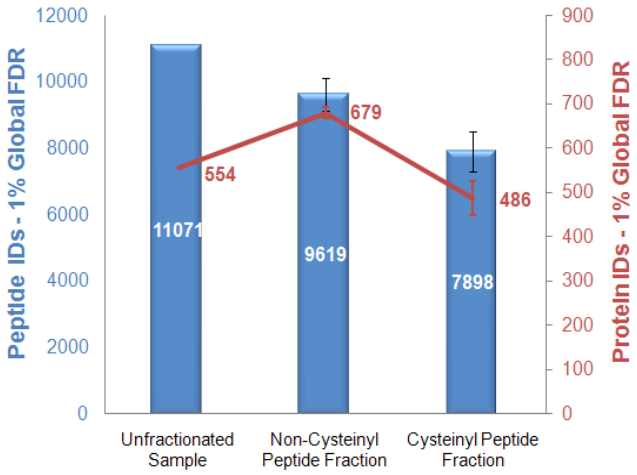

Figure 3. Comparison of Identified Proteins and Peptides in Sample Preparation Fractions. Three injections of unfractionated sample were searched and evaluated as one data set and were then compared with the average yield of identified proteins/peptides obtained from three injections of the three non-cysteinyl peptide fractions, and three injections of the three cysteinyl-peptide fractions (data are shown as mean ± SD).

Data Processing: Peptide identification and quantification was performed using ProteinPilot™ 4.0 software[4] searching the UniProtKB/Swiss-Prot database (canonical and isoforms included, June 22, 2010 + SCIEX 2007 contaminant DB). Only proteins at 1% global FDR and distinct peptides at 5% local FDR were used for further analysis[5]. The ProteinPilot Descriptive Statistics Template (beta v3.001p) was used for estimation of FDR associated with quantification results[6,7]. The Protein Alignment Template was used for alignment of multiple results at the protein level[8].

High Quality Identification at High Acquisition Rates

The high MS/MS sensitivity of the TripleTOF® 5600 system enables high acquisition rates for in-depth analysis of complex proteomes. The enhancement of the time-of-flight detector with a fast (40GHz) sampling rate means that there is no decrease in resolution as acquisition rates are increased. In this study, the high sensitivity MS/MS mode was used, delivering >15,000 (FWHM) resolution on all product ions, with excellent mass accuracy (Figure 1). While acquisition rates of up to 50 MS/MS in a second are commonly employed on the TripleTOF 5600 System for protein identification, iTRAQ® reagent quantification experiments are acquired at a slightly slower rate to ensure high quality reporter ion signals for the most reliable quantification results.

ProteinPilot™ Software was used for identification and quantification of the acquired LC-MS/MS data. The software automatically performs false discovery rate (FDR) analysis based on a forward and reversed protein database search approach. In this study, only proteins at 1% global FDR and distinct peptides at 5% local FDR were used for further analysis. The evaluation of three injections of the unfractionated sample yielded 544 proteins and 11,071 distinct peptides. In the evaluation of three injections of the non-cysteinyl peptides fraction and cysteinyl peptides fraction, an average of 679 proteins (9,619 peptides) and 486 proteins (7,898 peptides) were identified, respectively (Figure 3).

Protein Alignment Template for Intersection Analysis

Further understanding of the impact of this sample fractionation strategy can only be gained by analyzing the assigned protein and peptide populations in each fraction. Intersection analysis between the various sample preparation fractions was performed to assess the impact on unique protein identifications using the Protein Alignment Template.

The Venn diagrams shown in Figure 4 illustrate the benefits of the fractionation on uniquely identified proteins. One injection of each non-cysteinyl and cysteinyl fractions detected 38% more unique proteins than two injections of the unfractionated sample, showing that the CPE enrichment clearly led to increased proteome coverage. In the comparison of three injections of the unfractionated sample vs. three injections of either the cysteinyl or non-cysteinyl peptides fraction, the number of identified proteins increased by 20.9% and 40.2%. Even though fewer peptides (28.6 %) and proteins (10.6%) were identified in the cysteinyl peptide fraction, the analysis of this fraction still delivered a 20% increase of unique proteins assigned compared to the unfractionated sample alone. Ultimately, the combined evaluation of three injections of both cysteinyl and non-cysteinyl peptides fractions resulted in identification of 60.8% more proteins compared to three injections of the unfractionated sample (Figure 4). The alignment example of this comparison is shown in Figure 5.

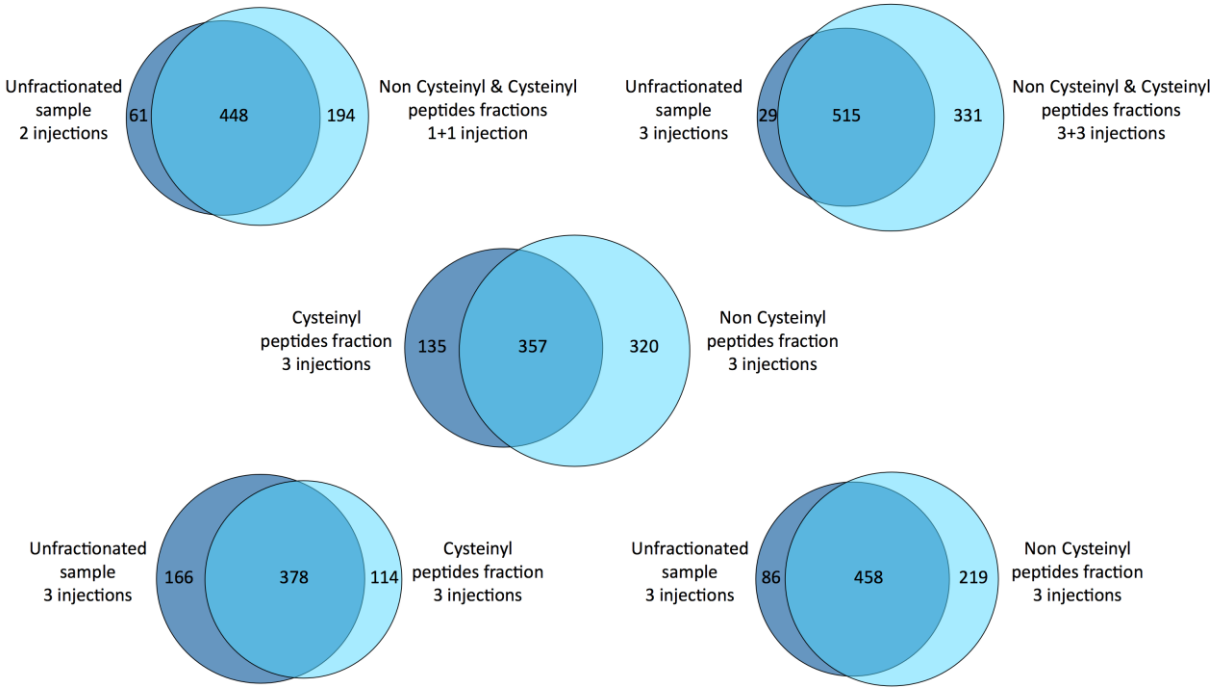

Figure 4. Protein Identification Overlap in Individual Analysis Combinations. The Venn diagrams illustrate yields of identified proteins in various combinations of analyzed sample preparation fractions. The numbers represent the median number of identification numbers achieved from the fractions indicated. The protein distribution and overlap across multiple fractions was assessed using the Protein Alignment Template[7].

Protein Alignment for Sequence Coverage Comparison

In addition to aligning the proteins between the samples, the sequence coverage for each protein can be compared on a sample by sample basis. This provides further color to usefulness of this fractionation strategy. Figure 5 shows the comparison between 3 replicates of the unfractionated sample and 3 replicates of the combined Cys and Non-Cys fractions. The comparison clearly shows that the fractionation strategy provides significantly higher sequence coverage *per* protein. Higher sequence coverage yields higher confidence quantitation, increased detection of post-translational modifications and more information about individual proteins of interest.

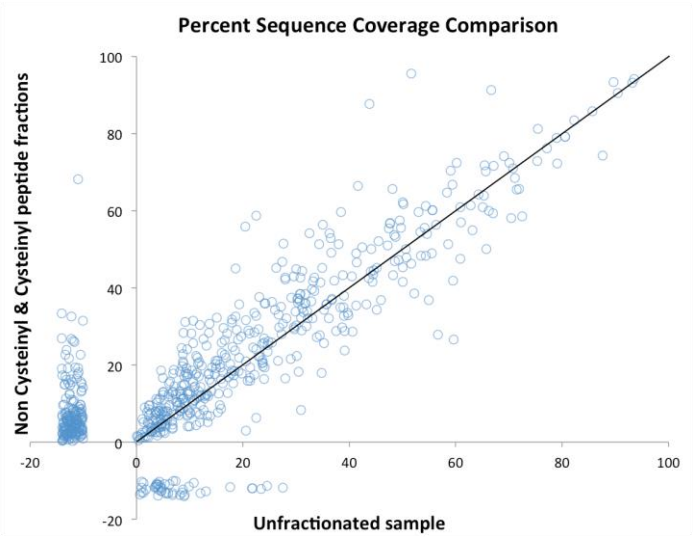

Figure 5. Comparison of Sequence Coverage Obtained Unfractionated and Cysteine and Non-Cysteine Fractionated Samples. Higher sequence coverage is observed for most proteins and more proteins are found in total when the Cysteinyl fractionation strategy is used and both fractions are measured.

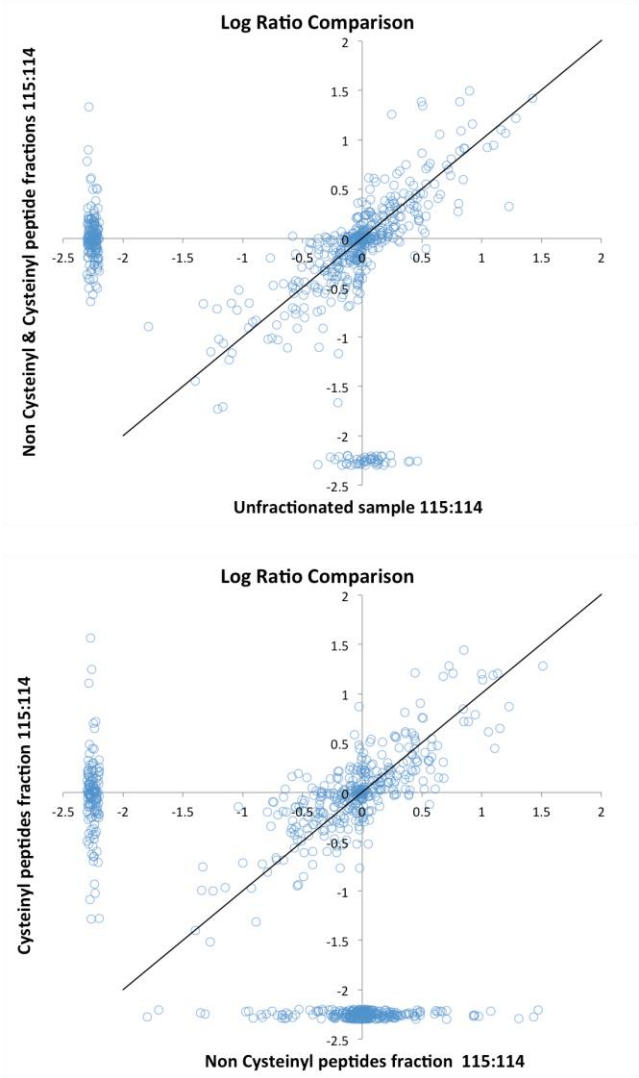

Figure 6. Two Sample Quantification Comparison. Log Ratio comparison shows quantification reproducibility across sample preparation. Each protein ratio is shown as a dot, with proteins unique to the respective fraction shown along each axis. Pearson correlation factor (*r*) for both comparison were 0.80 and 0.84, respectively. The comparison was performed using the Protein Alignment Template.

Protein Alignment for Quantification Comparison

In addition to the alignment and coverage tools, the quantification results for each protein across a set of samples can also be aligned and compared using the Protein Alignment Template. After aligning at the protein level, the protein ratios across all samples can now be compared. Any quantification strategy used and analyzed by ProteinPilot™ Software can be aligned and studied using the Protein Alignment Template.

In this study, the protein ratios from the iTRAQ® reagent data were aligned between the sample preparation strategies to confirm that no bias was introduced by the Cysteine capture fractionation strategy. First, the protein quantitation ratios between the unfractionated sample and the combined Cys and non-Cys fractions from one replicate were compared (Figure 6, top). A good correlation was observed with an *r* value of 0.80. In addition, the Cysteine peptide fraction and the Non-cysteinyl peptide fraction were aligned and compared and a similar strong correlation was seen (*r* = 0.84, Figure 6, bottom)

Protein Expression Analysis in Amniotic Fluid

In this study, two amniotic fluid samples with negative intra-amniotic infection (IAI) were labeled with 114 and 116, the two positive samples were labeled with 115 and 117, and the 114 sample was used as the denominator for ratio computation. As the differences between the negative infection samples were not of interest in this study, the 116 sample was set to be the decoy in the quantitative false discovery rate analysis. The quantitation FDR computation was done using the PDST[6] and is shown in Figure 7. Based on this analysis and using a cut-off of 5% global FDR of differential expression, there were 83 proteins of interest found to have altered expression in all three replicate analyses.

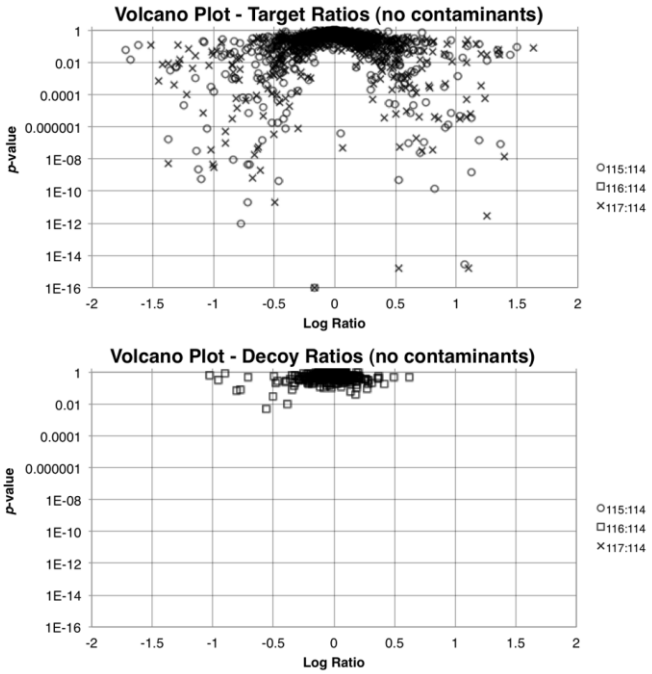

Figure 7. Volcano Plots and Quantitative FDR assessment. A volcano plot is a representation of the extremity of change vs. the certainty of change and therefore is a useful visualization tool. As the 116/114 ratio represents both IAI negative samples and minimal differences are expected, this was used as the decoy in the quantitative FDR assessment and ratios between these two samples is shown in the bottom pane. The upper pane shows the IAI positive samples / IAI negative sample and here, differential proteins can be detected. Based on these data, the p<0.05 threshold in turn corresponds to an FDR of differential expression of <5%

Conclusion

In this study, the usefulness of using iTRAQ® reagent quantitation in combination with a cysteinyl capture strategy for increasing the depth of coverage in quantitative proteomics has been demonstrated. Several key advantages have been shown by this technique:

- The ability to quantify four samples in one analysis (application of the latest version of iTRAQ reagent would enable analysis of eight samples)
- The quantification is performed in MS/MS mode and thus the complexity of MS spectra is not increased.
- Due to the fact that the quantification feature of the method is not linked to any cysteine-targeting tag, this approach enables flawless quantitative analysis of both cysteinyl and non-cysteinyl peptide fractions, providing more proteins and more protein sequence coverage for high confidence quantitation.

References

1. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) *Nat. Biotech.* **17(10)**, 994-999.
2. Tambor V, Hunter CL, Seymour SL, Kacerovsky M, Stulik J, Lenco J (2012) *J .Proteomics* **75(3)**, 857-67
3. Liu T, Qian WJ, Strittmatter EF, Camp DG 2nd, Anderson GA, Thrall BD, Smith RD (2004) *Anal Chem.* **76(18)**, 5345-5353.
4. Shilov IV *et al.* (2007), *Mol. Cell. Proteomics,* **6**, 1638-1655.
5. Tang W et al. (2008), *J. Prot. Res.* **7(9)**, 3661–3667.
6. The ProteinPilot Descriptive Statistics Template has now been incorporated into the ProteinPilot Report in ProteinPilot™ Software 5.0
7. ProteinPilot Report for ProteinPilot™ Software. SCIEX technical note RUO-MKT-02-1778-A.
8. Protein Alignment Template can be downloaded from the website – http://sciex.com/support/software-downloads

.

Publication number: RUO-MKT-02-2513-A

Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices