IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | | |
|---|---|---|---|
| In Re Patent of: | SAMPALIS, Fotini | Confirmation No.: | 1897 |
| Control No.: | 95/001,774 | Group Art Unit: | 3991 |
| Filed: | October 19, 2011 | Examiner: | CAMPELL, Bruce R. |

FOR: **INTER PARTES REEXAM OF U.S. PATENT 8,030,348: NATURAL MARINE SOURCE PHOSPHOLIPIDS COMPRISING POLYUNSATURATED FATTY ACIDS AND THEIR APPLICATIONS**

---

**Mail Stop Declaration**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

## DECLARATION OF FAUSTINUS YEBOAH, PH.D. UNDER 37 C.F.R. § 1.132

I, Faustinus Yeboah, declare as follows:

1. I am a Canadian citizen.

2. I am the Director of KABS Laboratories, Inc., a company that offers a broad range of product development services to the bio-pharmaceutical industry worldwide. The services offered include strategic planning, pre-clinical development, analytical testing, formulation development, manufacturing of prototypes and clinical supplies, distribution of clinical supplies to clinical sites, and chemistry, manufacturing and controls (CMC) aspects of regulatory affairs. I am also the Founder and Principal Consultant of PDMC Pharma Consulting.

3. I obtained my Ph.D., concentrating on protein and carbohydrate chemistry, and my M.Sc., concentrating on food chemistry, from McGill University in Montréal, Québec. I was a post-doctoral fellow at the Biotechnology Research Institute of the National Research Council in Canada.

4. I have authored twenty five papers, many of which concern mass spectrometry and extraction of biomolecules, and I am an inventor of one U.S. patent and six patent applications. I am a Faculty Lecturer at McGill University in the Department of Food Science, and I serve as a Journal Reviewer for the *Journal of Food Composition and Analysis*, the *Journal of Agriculture and Food Chemistry*, and the *Journal of Environmental Toxicology*. I also serve as a Grant Reviewer for the National Science and Engineering Research Council of Canada (NSERC). My Curriculum Vitae is enclosed as **Appendix A**.

5. I am considered an expert in the area of mass spectrometry and extraction of biomolecules.

6. I was engaged by counsel for Neptune Technologies and Bioressources, Inc. ("Neptune") of Québec, Canada to analyze the Corrected Request for Reexamination (US 95/001,774) filed by Aker Biomarine ASA ("Aker"). I am being compensated at my customary hourly rate for my time spent on developing, forming, and expressing the facts and opinions in this declaration. I have no personal interest in the ultimate outcome of the reexamination proceedings involving U.S. Patent 8,030,348 ("the '348 patent").

7. Specifically, I was asked to review the Declaration of Thomas Gundersen, which was filed in support of the Request for Reexamination filed by Aker and the Declaration of Earl L. White, Ph.D., which was submitted by Neptune in the prosecution of the '348 patent, and to opine on the validity of the results provided therein. Further, I was asked to express an opinion on quantitative aspects of the data discussed herein.

**Gundersen Presents Incomplete and Unreliable Data**

8. I have read and reviewed the Declaration of Thomas Gundersen submitted by Aker, and it is my opinion that it suffers from considerable technical deficiencies and errors which render its conclusions completely unreliable. I summarize these deficiencies and errors below.

*Gundersen Presents Clearly Erroneous Data*

9. My review of the Gundersen Declaration leads me to conclude that it contains incomplete and unreliable data. Gundersen erroneously presents data as distinct when it is merely an exact

copy of another chromatogram. In my opinion, there is clear error in the Gundersen Declaration and, whether reflective of a sloppy study marred by negligence or a fraudulent study submitted with deceitful intent, one cannot rely on this data at all.

10. Specifically, referring to Appendix B of the Gundersen Declaration, the chromatograms labeled "Sample P308-8" and "Sample P308-9" are identical (*see* Gundersen Declaration, Appendix B, pp. 18-19). Similarly, the chromatograms labeled "Sample P308-10," "Sample P308-11," and "Sample P308-12" are identical (*see* Gundersen Declaration, Appendix B, pp. 19-20).

11. Gundersen provided Table 1, which states that "Sample P308-8," "Sample P308-9," "Sample P308-10," "Sample P308-11," and "Sample P308-12" are distinctly different samples (*see* page 3 of the report appended to the Gundersen Declaration). Below, I reproduce, in part, Table 1 (emphasis added, note the "Marking of Sample" Column):

| Vitas ID | Fraction number | Temperature treatment (°C) | Time (min) | Marking of sample |
|---|---|---|---|---|
| P308-8 | IIa | 70 | 5 | **E.superba Fraction IIa 70 degr 5 min** |
| P308-9 | IIa | 125 | 15 | **E.superba Fraction IIa 125 degr 15 min** |
| P308-10 | IIb | - | - | **E.superba Fraction IIb not heated** |
| P308-11 | IIb | 70 | 85 | **E.superba Fraction IIb 70 degr 5 min** |
| P308-12 | IIb | 125 | 15 | **E.superba Fraction IIb 125 degr 15 min** |

12. I have also attached the incomplete and unreliable chromatograms as enumerated above as a series of figures to make this point clear (*see* Figures 1-5). By reviewing Figures 1 and 2 side by side (on pages 14-15 of this Declaration), it is apparent that the retention times, areas under the curve, and sample identification information are exactly the same. For convenience, I have magnified the bottom peak of each of the chromatograms labeled "Sample P308-8" and

"Sample P308-9" and displayed them in Figures 1 and 2. The bottom chromatograms both have areas under the curve of 124384 and retention times of 3.758 minutes. It is my opinion that it would be impossible for two distinct samples to provide the exact same data. Even if "Sample P308-8" and "Sample P308-9" were merely repeats of each other, which they are not according to Gundersen's Table 1 (above), the data would have *at least some deviation*. The unreliability of this data is further underscored by the fact that the same sample identification number appears on the chromatograms labeled "Sample P308-8" and "Sample P308-9" and displayed in Figures 1 and 2 (on pages 15-16 of this Declaration). Both chromatograms have the following sample identifier: "MSD1 826. EIC=825.7.826.7 (C:\PAD\AKER BIOMARIN\P308\AB 110929\AB 110929\AB 110929 2011-09-29 10-35-22\004-0401D) ES." By providing the exact same data twice yet referring to the data as originating from two different experimental samples, Gundersen himself firmly demonstrates that his data is not credible. This is summarized below:

| Vitas ID | Marking of sample | Retention Time | Area Under the Curve |
|---|---|---|---|
| P308-8 | "E.superba Fraction IIa 70 degr 5 min" | 3.758 | 124384 |
| P308-9 | "E.superba Fraction IIa 125 degr 15 min" | 3.758 | 124384 |

13. The chromatograms labeled "Sample P308-10," "Sample P308-11," and "Sample P308-12" also present identical data for allegedly distinct samples and therefore provide further incomplete and unreliable data. By reviewing Figures 3, 4, and 5 (on pages 16-18 of this Declaration) side by side, it is apparent that the retention times, areas under the curve, and sample identification information are exactly the same. For convenience, I have magnified the bottom peak of each of the chromatograms labeled "Sample P308-10," "Sample P308-11," and "Sample P308-12" and displayed them in Figures 3, 4, and 5. The bottom chromatograms of all three have areas under the curve of 58821.8 and retention times of 3.770 minutes. It is my opinion that it would be impossible for three distinct samples to provide the exact same data. Even if "Sample P308-10," "Sample P308-11," and "Sample P308-12" were merely repeats of each other, which they are not according to Gundersen's Table 1 (above), the data would have *at least some deviation*. The unreliability of this data is further underscored by the appearance of the same sample identification number on the chromatograms labeled "Sample P308-10," "Sample P308-11," and "Sample P308-12" and displayed in Figures 3-5. All three chromatograms carry the following sample identifier: "MSD1 826. EIC=825.7.826.7

(C:\PAD\AKER BIOMARIN\P308\AB 110929\AB 110929\AB 110929 2011-09-29 10-35-22006-D601 D) ES." As with the samples discussed in Paragraph 12 above, by providing the exact same data three times yet referring to it as three different experimental samples, Gundersen establishes that his data is not credible. This is summarized below:

| Vitas ID | Marking of sample | Retention Time | Area Under the Curve |
|---|---|---|---|
| P308-10 | "E.superba Fraction IIb not heated" | 3.770 | 58821.8 |
| P308-11 | "E.superba Fraction IIb 70 degr 5 min" | 3.770 | 58821.8 |
| P308-12 | "E.superba Fraction IIb 125 degr 15 min" | 3.770 | 58821.8 |

14. I note that I limited my review in this section to data presented in Appendix A of the Gundersen Declaration, as the data presented in Appendix B of the Gundersen Declaration is so poorly reproduced that I cannot discern most of the alleged peaks that are presented. For example, I cannot see any data on the chromatograms labeled "p308-4" to "p308-7" and can only make out faint images on the remainder of the figures presented in Appendix B of the Gundersen Declaration.

*Gundersen's Results are Highly Variable and Reflective of a Rushed Experiment.*

15. Besides the incomplete and unreliable data presented, there is a puzzling complete absence of data in the Gundersen Declaration for the ethyl acetate extract of *E. superba* krill lipid samples (*see* Gundersen Declaration ¶ 5). The fact that no data was generated for these experimental samples suggests that the experiment was not conducted carefully (*see* Gundersen Declaration, Appendix B, pp. 16-17 for chromatograms labeled "Sample P308-4," "Sample P308-5," and "Sample P308-6"). In fact, Gundersen himself states that the experiment "should have been repeated but there was not enough time for this" (*see* Gundersen Declaration ¶ 5; emphasis added). Further, on page 2 of Exhibit 2 of the Gundersen Declaration, Gundersen states that "the analysis of the samples took place between 28 September and 4 October 2011." Therefore, all within the span of just five business days, the data was acquired, the data was analyzed, a report was generated, and a declaration regarding analysis of the data was written. A

complicated experimental analysis such as this would require far more time if it were conducted properly. These facts suggest a rushed experiment and data analysis.[1]

*The Gundersen Data Suffers From the Memory Effect.*

16. It is not only conventional but also necessary, when conducting mass spectrometry experiments, to run solvent-only (a/k/a "blank" or "negative") controls at the beginning of an experiment and between runs which involve different experimental species. Gundersen does not provide these crucial controls.

17. These omitted controls are crucial because they correct for a technical difficulty in mass spectrometry known as the "memory effect" or "carry-over effect."[2] For instance, contaminant species from a preceding experimental run skews the subsequent profiles.

18. This deficiency of the Gundersen Declaration is *amplified* because Gundersen employs multiple reaction monitoring (MRM), a highly sensitive technique used to detect *very small quantities of signal*.[3] As the detection is very sensitive, any cross-contamination of species may completely blur actual data. Therefore, the failure of Gundersen to run solvent-only controls, in a very sensitive mass spectrometry method, casts significant doubt upon the data presented in the Declaration.

19. Gundersen's failure to control and correct for the memory effect, in my opinion, probably resulted in incorrect conclusions regarding the identification of species. One cannot, therefore, exclude with any confidence that signals from the positive or reference standards were not carried-over to sequential sample runs, which would result in false positive data. For instance, in the data presented in the chromatograms of "Sample P308-15" and "Sample P308-16" (same experimental conditions, *see* Gundersen Declaration, Appendix B, p. 22) there are considerable differences in the peak intensities observed where one would expect essentially similar

[1] It is respectfully noted that some of the samples did not arrive in the laboratory where the Gundersen Declaration experiments were undertaken until September 30, 2011. This would leave a maximum of three business days to generate data, analyze data, write a report, and write a declaration regarding analysis of these samples.

[2] *See, e.g.*, Hughes *et al.*, "Determination of Carryover and Contamination for Mass Spectrometry-Based Chromatographic Assays." *The AAPS Journal* 2007; 9 (3) Article 42, **Appendix B.**

[3] *See, e.g.*, Elliott *et al.*, Current Trends in Quantitative Proteomics. *J. Mass. Spectrom.*, 44 (12): 1637–1660 (2009), **Appendix C.**

intensities. However, in the former, the peak intensity is in the tens of thousands, while in the latter, the peak intensity is in the low hundreds (compare, for example, the bottom chromatograms of both). Given this highly diminished intensity, I cannot reasonably exclude the possibility that the higher signals bled into the subsequent run's results via the memory effect. A solvent-only control, which Gundersen failed to run, would have excluded the memory effect.

20. Further, another potential manifestation of the "memory effect" is seen in the highly variable retention times presented in the Gundersen Declaration for what is alleged to be the same compound. For instance, selecting a random pair of chromatograms, the bottom traces of "Sample P308-1" and "Sample P308-7," are purported to be the same chemical species, yet have disparate retention times of 4.181 and 3.747 minutes, respectively. This is a retention time deviation of close to 26 seconds. Therefore, in one run ("Sample P308-1") the same compound is about 10% slower than in another ("Sample P308-7"). In my opinion, this should not be. Because of Gundersen's failure to run the solvent control, I cannot exclude the memory effect. In fact, these data are also indicative of two different chemical species incorrectly labeled by Gundersen as the same chemical species.

*The Gundersen Data Suffers From Poor Resolution.*

21. Gundersen also presents experimental data peaks that are <u>not</u> the sharp and clearly defined peaks that are reflective of clean chromatography. On the contrary, the peaks show a poor resolution that suggests that multiple species may be present.

22. For instance, the experimental runs labeled "Sample P308-1," "Sample P308-2," "Sample P308-3," "Sample P308-14," and "Sample P308-16" are particularly broad and asymmetrical, which is indicative of a population of unknown species in a "peak" as opposed to a single species. In contrast, a single species is expected to give a sharp peak in the form of a tight Gaussian-shaped curve.

23. In summary, it is my opinion that the Gundersen Declaration suffers from a number of deficiencies, including incorrect and unreliable data, rushed execution, and experimental errors which render its conclusions completely unreliable and not credible. As an expert I cannot give the results presented in the Gundersen declaration any weight.

**The White Declaration of May 31, 2011 Presents Sound Raw Data and a Reasonable Overall Conclusion.**

24. I have also read the Declaration of Earl L. White, Ph.D. that was submitted in the prosecution of the '348 patent on May 31, 2011 and I have been asked to provide an independent evaluation of the correctness of its conclusions. Unlike the Gundersen Declaration, the White Declaration provides data that is experimentally sound and has a reliable conclusion.

25. I understand Dr. White's conclusion to be that it is his "opinion that the Beaudoin Oil Fractions received and tested by [him] do not contain PLs [phospholipids] which have attached to them DHA and DHA, EPA and EPA, DHA and EPA, or EPA and DHA, at the detection limits described [LC/MS and MS/MS techniques]." Therefore, Dr. White could not find any species that could be definitively identified as those disclosed and claimed in the '348 patent. This is a reasonable conclusion based on the experimental data.

26. Some commentary on the White Declaration of May 31, 2010 is helpful for understanding it.

27. First, it is my expert opinion that the *raw data* collected by Dr. White is correct and does not suffer from the many flaws seen in Gundersen's data.

28. Second, Dr. White's Declaration of May 31, 2010 does have a couple of minor *interpretative* errors that do not affect the validity of the raw data or the overall conclusion (*i.e.* that Dr. White could not, within the limits of detection, find any species that could be definitively identified as those disclosed and claimed in the '348 patent).

29. A first minor error lies in Figure 10 of the Declaration of Dr. White of May 31, 2010. I understand that this figure was included because the 2011 Beaudoin oil did not produce product ions such that Dr. White could not identify species that might have been those disclosed and claimed in the '348 patent. Dr. White included Figures 9 and 11 in support of this conclusion. However, it appears that because Dr. White did not detect characteristic product ions that are needed to definitively identify the fatty acids attached to the phospholipids at the molecular weights of 826 and 852 he included, for comparison, Figure 10, "from a previous Beaudoin oil," as a "representative MS/MS spectrum for a [phosphatidylcholine] with a molecular ion at *m/z*

826." *See* White Declaration of May 31, 2011, Figure 10 legend and ¶ 12. This figure shows what a species at the weight of a phospholipid of EPA/EPA would look like if it could be fragmented such that product ions could be detected. This explains why Dr. White separated this figure from the others in ¶ 12 of the White Declaration, which states: "MS/MS spectra for a representative Beaudoin Oil Fraction sample are shown in Figures 4-9, and 11. Figure 10 is a representative MS/MS spectrum for a [phosphatidylcholine] with a molecular ion at *m/z* 826.5" (emphasis added). This, to me, is a clear indication that Dr. White did not intend to present Figure 10 as part of the experimental set upon which he was opining in May, 2011. This is bolstered by a statement in the figure legend of Figure 10, in which it is stated that the analyzed sample is "from a previous Beaudoin oil." In fact, Dr. White also included the date of the experiment in the upper left corners of each of the spectra included in the figure set. Figure 10 of the White Declaration of May 31, 2011 has the label "09Dec07" indicating that it is from a December, 2009 experiment while all of the other figures (1-9, 11) have date labels of April or May, 2011, indicating that they are from the experiments commissioned for prosecution of the '348 patent.

30. I have reviewed an experimental report from the 2009 extraction (**Appendix D**) and conclude that, because of the experimental errors in reproducing Beaudoin, the 2009 samples of Figure 10 are not "Beaudoin oil." Thus, Dr. White's statement that Figure 10 is "from a previous Beaudoin oil" is erroneous. As described in the experimental report, the procedure removed the water from the oil and did not adequately heat. Therefore it would be expected that the oil of Figure 10 shows enough *m/z* 826 species to allow product ion detection (*See* **Appendix D**).

31. Specifically, I note that the 2009 extraction did not replicate Beaudoin because a series of changes to the Beaudoin protocol were inadvertently made (*see* **Appendix D**):

- In the acetone extraction ("step 1"), extensive evaporation of the solvent *and water* was undertaken, in contrast to what was described in Beaudoin. Specifically, the experimenter "[e]vaporate[d] the acetone;" "[s]eparat[ed] the fat from the water by decantation after addition of 1 volume of acetone;" "[r]ecover[ed the] oil;" and "[e]vaporated under vacuum." In contrast, Beaudoin's extraction includes *only* rotary

evaporation to remove acetone and leaves 10.0% moisture and volatiles (Beaudoin I, Table 13). Such water content causes hydrolysis of the phospholipid upon heating.[4]

- Further, the experimenter *failed to heat the acetone extract at 125°C for 15 minutes*, in contradistinction from the step taught by Beaudoin. Heat was not applied at all to Fraction I.

- The experimenter undertook an ethanol evaporation to yield Fraction II as in Beaudoin ("step 2") but *did not heat Fraction II at 125°C for 15 minutes*. Instead, Fractions I and II were merged to create a fraction not taught by Beaudoin: a fraction having much less water than the Beaudoin fractions. This non-Beaudoin oil was then treated by "[e]vaporat[ing] under vacuum" and "[h]eat[ing] at 125 °C."

As Beaudoin does not teach the experimental features enumerated above, the 2009 experiment did not correctly replicate the Beaudoin extraction and thus resulted in an oil that is not Beaudoin oil. I understand that it was oil from this incorrectly-executed krill oil extraction that was sent to Dr. White for mass spectrometric analysis. Dr. White characterized this non-Beaudoin oil in 2009 and included Figure 10 from this analysis in his Declaration of May 31, 2011.

32. For completeness, I also note that there is a slight mathematical error in Figure 10, as Dr. White presented it in his Declaration of May 31, 2011. While his point was to simply show how a detectable ion fragmentation looks, Dr. White mistakenly indentified the two product ion peaks at *m/z* 524 and 542 as C20:5 (EPA) and C18:0, respectively. *See* White Declaration of May 31, 2011, Figure 10. My review of this figure shows that these two species cannot result from the fragmentation of a peak at *m/z* 826. On the contrary, this fragmentation pattern is reflective of a species bearing two EPA molecules. The fragment at *m/z* 524 represents a neutral loss of the free acid of EPA from the parent phospholipid (*m/z* 826), and the fragment at *m/z* 542 represents a neutral loss of a ketene form of EPA from the parent phospholipid. However, as the 2009 experiment was an incorrect reproduction of Beaudoin, in that water was removed and heat was, for the most part, *not used*, this species *should contain* measurable levels of a phospholipid with

[4] *See, e.g.*, Herman and Groves "The Influence of Free Fatty Acid Formation on the pH of Phospholipid-Stabilized Triglyceride Emulsions" *Pharmaceutical Research* 10(5): 774-776 (1993), **Appendix E**.

EPA and EPA. This mistake, being in a figure that does not even reflect Beaudoin, has no bearing on Dr. White's conclusions.

33. Therefore, it is my opinion that Figure 10 is a failed attempt at providing a comparative example between the extract produced in Beaudoin and the composition of the '348 patent. This figure, in my opinion, has no bearing on the conclusions of Dr. White's study and therefore I view the May 31, 2011 White Declaration as experimentally sound despite it.

34. A second minor error lies in Tables 1 and 2 of Dr. White's Declaration of May 31, 2011; particularly in Dr. White's identifications of the potential fatty acids that could represent different *m/z* signals. It is apparent that Dr. White made his calculations without the knowledge that krill do not possess fatty acids shorter than C14.[5] As the *m/z* values of the composition of the '348 patent are unambiguous, this has no bearing on Dr. White's overall conclusion as it does not at all alter the soundness of the raw experimental data that he generated.

**Comments on Quantification of White and Gundersen Data**

35. I have reviewed both the Gundersen and White raw data and have been asked to express an opinion on quantitative aspects of both.

*The Beaudoin Oil Contains, If Any, a De Minimis Amount of the Phospholipids Carrying Two of EPA and DHA*

36. The conclusion of the White Declaration, that Dr. White could not find any species of phospholipids carrying two of EPA and DHA within the detection limits of the experiment conducted, is sound. The species detected at *m/z* values of 826 and 852 represent amounts in a range on the order of only 0.1 to 2.8% of the *phospholipids of the oil*.[6] I understand that phospholipids represent about 40% of the total lipids in krill oil[7] and therefore, the raw data of Tables 1 and 2 of the White Declaration shows that the amount of phospholipids carrying two of

[5] *See, e.g.* Winther *et al.*, Elucidation of Phosphatidylcholine Composition in Krill Oil Extracted from *Euphausia superba Lipids* 46 (1): 25-36 (2011), **Appendix F**.

[6] This is based on the raw data presented in Tables 1 and 2 of the White Declaration of May 31, 2011.

[7] *See,* Winther *et al.*, Elucidation of Phosphatidylcholine Composition in Krill Oil Extracted from *Euphausia superba Lipids* 46 (1): 25-36, page 29 (**Appendix F,** an Aker paper that stated: "the PtdCho content of the undiluted krill oil was determined to be 34 ± 5% (w/w)"); *see also* Beaudoin I, Table 14; '348 patent, Column 15, lines 32-35.

EPA and DHA *in the total Beaudoin oil* is only about 0.05 to 1.1%. This is a *de minimis* amount of phospholipids carrying two of EPA or DHA.

37. Unfortunately, due to the myriad experimental errors of the Gundersen Declaration, I cannot quantify its data. Given this, it is my opinion that even if there is any of the phospholipids carrying two of EPA or DHA, it is likely a very small amount. Specifically, it is noted that Gundersen needed to turn to LC-MRM detection to allegedly find the species. LC-MRM is a tandem mass spectrometric technique that is *several orders of magnitude more sensitive* than standard LC-MS or LC-MS$^{n}$.[8] LC-MRM is usually used when the detection and quantification of extremely low or trace levels of analyte is desired.

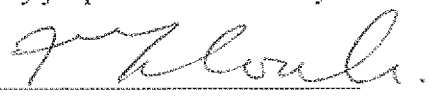
38. Having established that there is, at most, a *de minimis* amount of phospholipids carrying two of EPA or DHA in Beaudoin, I have also ascertained what percentage of *intact phospholipids, regardless of the identity of the fatty acid chains,* in the Beaudoin phospholipid sample is intact. Based on Tables 1 and 2 of the White Declaration of May 31, 2011, and the fact that a C14/C14 phospholipid is the lowest intact molecular weight species possible in krill, I calculate this value to be at about 35% of the phospholipids.[9] Therefore, only about 15% of the total Beaudoin oil, regardless of the identity of the fatty acid chains, is intact.

[remainder of page left intentionally blank]

[8] *See* Elliott *et al.*, "Current Trends in Quantitative Proteomics." *J. Mass. Spectrom.*, 44 (12): 1637–1660 (2009), **Appendix C**.

[9] I used Dr. White's raw data to do this calculation. In making my assignments, I utilized the fact that krill does not contain fatty acids longer than C14 and therefore, a phosphatidylcholine bearing two C14 moieties would be the lowest molecular weight intact phosphatidylcholine possible. Such a species would be expected to have a molecular weight of about 677 (this is the sum of two C14 chains, glycerol, and choline). Therefore, any *m/z* value of 677 or higher was classified as intact. The same would apply if phosphatidylethanolamine is considered (molecular weight cutoff would be about 635- the sum of two C14 chains, glycerol, and ethanolamine).

39. I further declare that all statements made herein of my own knowledge are true, and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of U.S. Patent 8,030,348.

By: 

Faustinus Yeboah, Ph.D.

Dated: March 16, 2012

93792 v1/BN

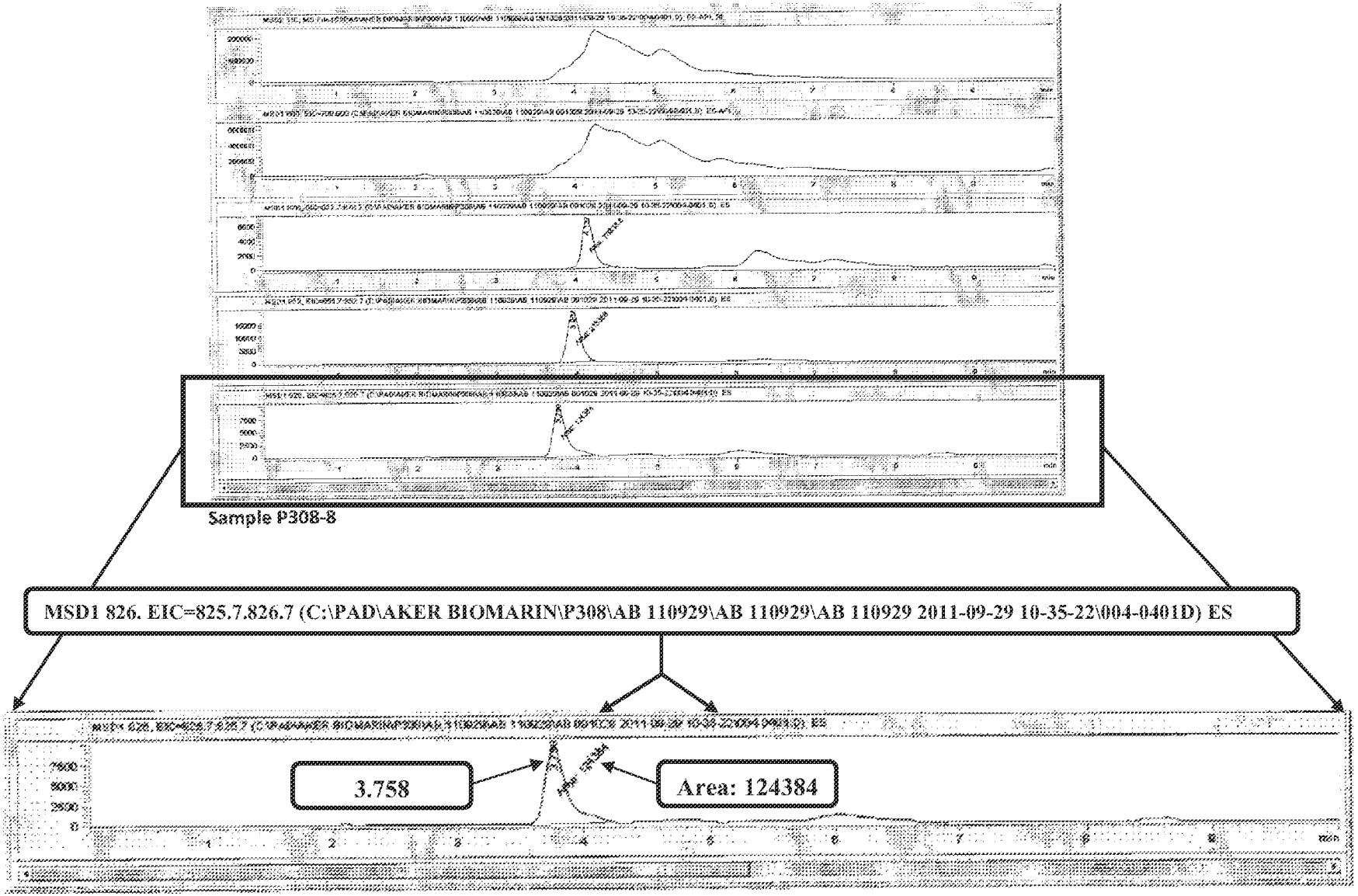



FIGURE 1

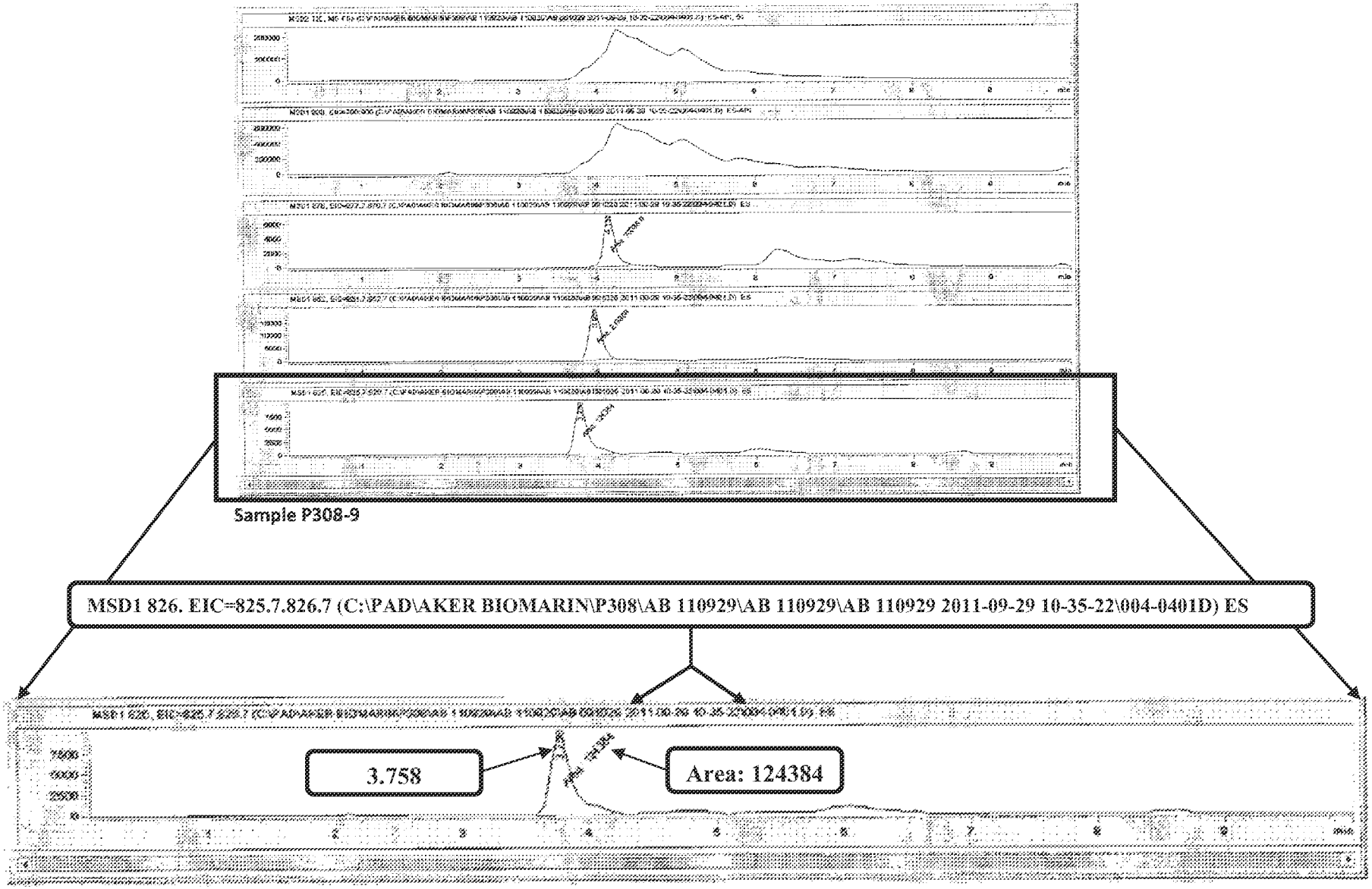

Sample P308-9
MSD1 826. EIC=825.7.826.7 (C:\PAD\AKER BIOMARIN\P308\AB 110929\AB 110929\AB 110929 2011-09-29 10-35-22\004-0401D) ES
3.758
Area: 124384

FIGURE 2

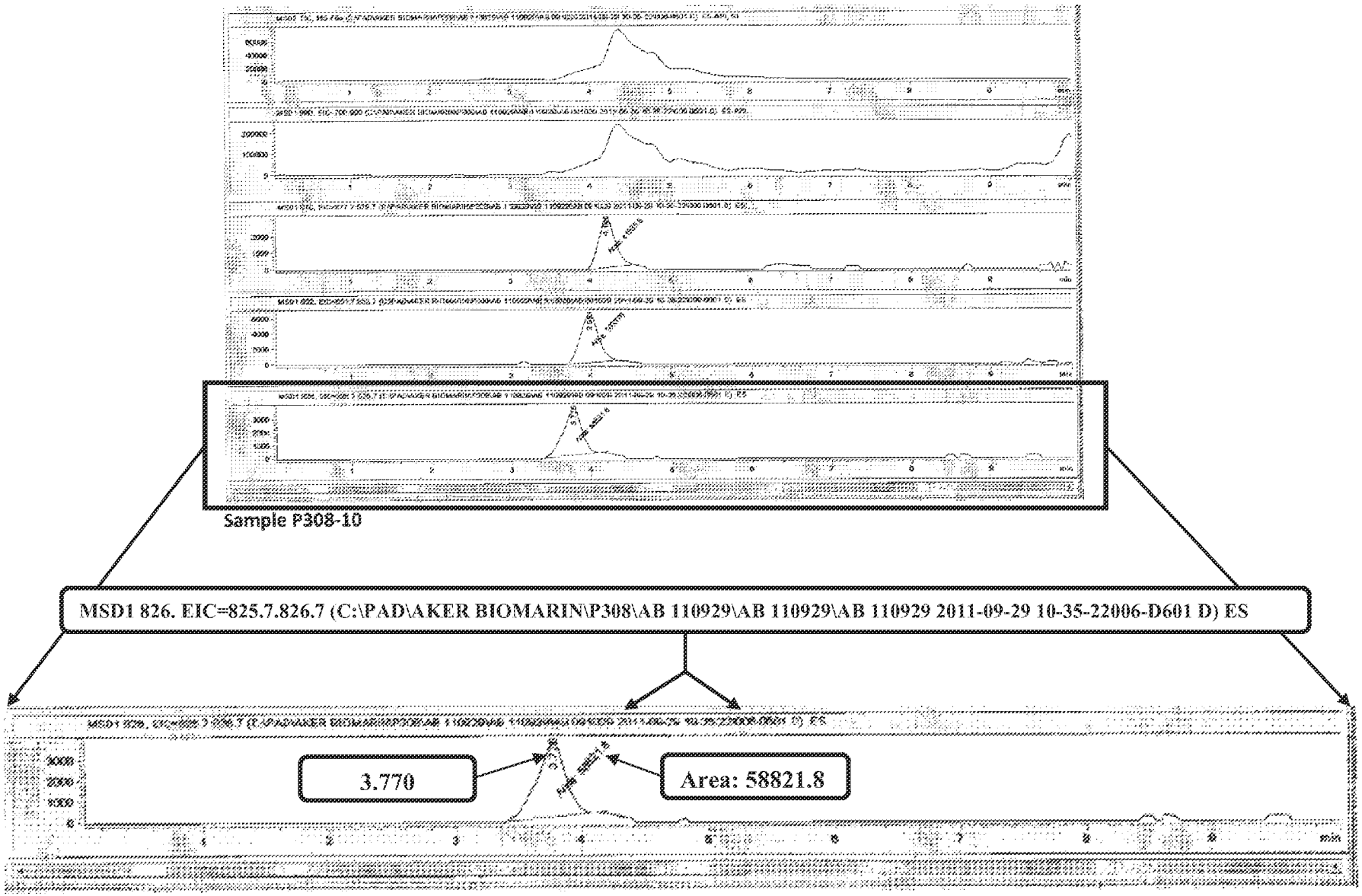



FIGURE 3

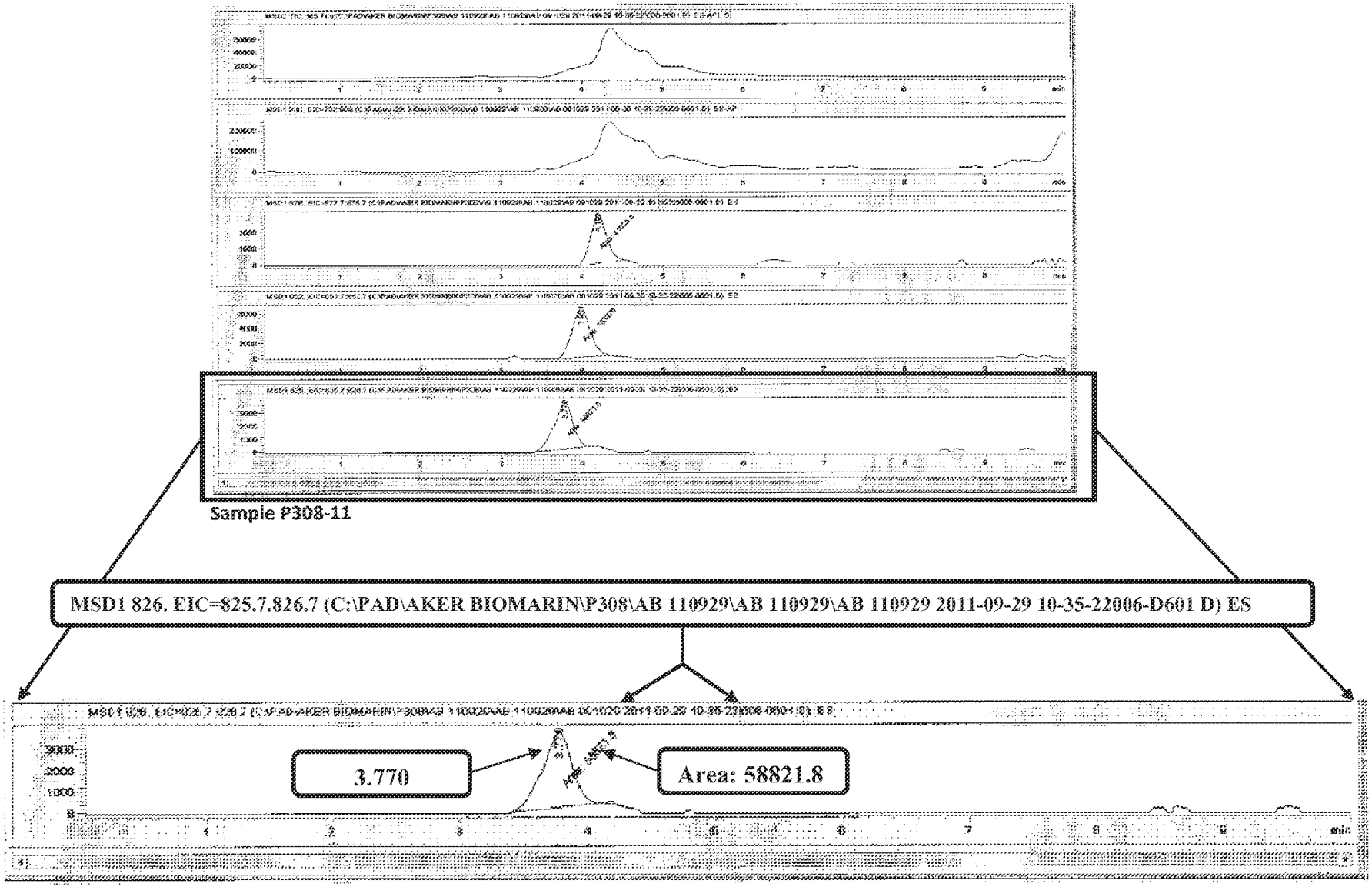

Sample P308-11
MSD1 826. EIC=825.7.826.7 (C:\PAD\AKER BIOMARIN\P308\AB 110929\AB 110929\AB 110929 2011-09-29 10-35-22006-D601 D) ES
3.770
Area: 58821.8

FIGURE 4

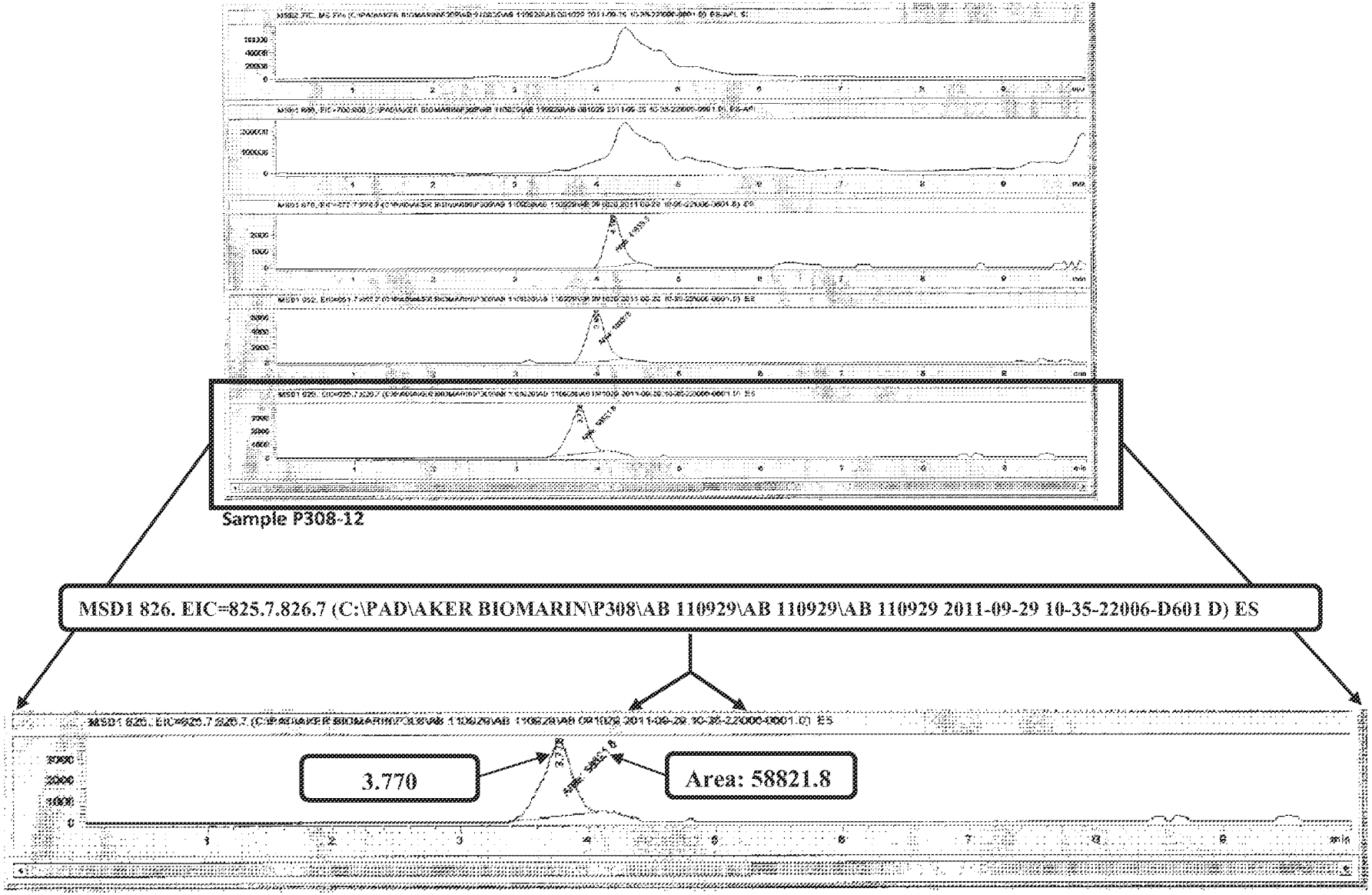

Sample P308-12
MSD1 826. EIC=825.7.826.7 (C:\PAD\AKER BIOMARIN\P308\AB 110929\AB 110929\AB 110929 2011-09-29 10-35-22006-D601 D) ES
3.770
Area: 58821.8

**FIGURE 5**

# Appendix A

## Curriculum Vitae of Dr. Faustinus Yeboah

# *CURRICULUM VITAE*

**Faustinus Kwabena Yeboah**
1154 Rue Des Hirondelles, Longueuil
J4G 4C6. Quebec.
Tel.: (450)-928-0162
E-mail.: yfaustinus@ymail.com

## SUMMARY:

I currently provide consultation in:

- Bio-process development and optimization (Recombinant proteins production and purification, Protein folding, and Natural products characterization),
- Bio-Process scale-up
- Bio-analytical method development and validation (Protein and Peptide drugs, small molecule drugs)
- Impurity identification and characterization
- Natural products drug discovery and characterization (Traditional African Medicinal Plants),
- Pharmaceutical Development (Chemistry Manufacturing and Control)
- Protein chemistry (Extraction, purification Structure and functional characterization)

## EDUCATION:

- **Ph.D. (Protein/Carbohydrate Chem).** McGill University Montreal, Canada (1999)
- **MSc. (Food Chem).** McGill University Montreal, Canada (1994)
- **BSc. (Hon's) Chem.** University of Science & Technology (UST), Ghana (1986)

## RESEARCH TRAINING:

- **Post Doctoral Fellow (Medicinal Chemistry)** Chemical Biology Group, BRI-NRC, Montreal (2000-02)
- **Research Fellow (Natural Product Chem.)** UST, Ghana (1989-91)

## EMPLOYMENT HISTORY:

**Director** (Bioprocess), KABS Pharmaceutical Inc. (Jul, 2011 to Present)
Leading the development of bioprocess R&D and cGMP manufacturing at KABS.

**Founder & Principal Consultant** PDMC Pharma Consulting, (Dec, 2009 to Present)

- Pharmaceutical Development,
  - Bio-Process Development and Validation, cGMP Manufacturing

  - Purification process scale-up (Recombinant proteins, Protein folding)
  - Bioanalytical Method Development and Validation (protein and peptide drugs)
  - Drug product Characterization and Release
  - Impurity Identification and Characterization
- Natural Product Drug Discovery and Development.

**Associate Director** (CMC, Drug Development) Thallion Pharmaceuticals Inc., Montreal (Nov, 2004 – Nov, 2009).

- Led the development and implementation of the manufacturing and control strategies for clinical candidates
- Directed API and drug product manufacturing and release, and the supply chain management of clinical candidates.
- Contributed extensively to the development and implementation of CMC regulatory strategies and submissions (IND, CTA), and served as CMC representative at meetings with regulatory agencies (FDA, Health Canada).
- Managed local and international collaborations relating to outsourcing of production and research activities, including process development, optimization and validation, and tech-transfer.
- Contributed extensively to the development of research tools for drug profiling and provided technical oversight for preclinical and clinical ADME studies.

**Major Achievements**:

- Developed a novel, scalable production and purification process from laboratory-scale (milligram scale) to pilot production-scale (kilogram scale) of cGMP material of API, and successfully oversaw the transfer of the process technologies to external and international partners.
- Developed a production department and developed and led the production of clinical materials.
- Reduced the cost of drug production by over 80% over a 6 month period through innovative improvements to the production process and production scale-up, while at the same time reducing the environmental impact of the production process.
- Developed drug product formulations and a validated process for the sterile production of injectables, as well as IV infusion delivery systems.
- Developed and implemented a quality systems approach to the cGMP production process that incorporated rigorous risk management, and controls, at all stages of production and supply of API and drug product.
- Developed and managed a network of internal, external and international resources to ensure the delivery of high quality, safe and consistent drug supply for phases-I and II clinical trials of two cancer drug candidates.
- Co-authored 4 patents and helped develop a portfolio of Intellectual Property (IP) around the company's clinical assets.

**Manager** (Process Development and Manufacturing) Thallion, Montreal (Apr, 2004 – Nov, 2004).

- Established a production department and lead the development of processes, protocols, and control systems for the production and profiling of drug candidates for clinical development.
- Develop and manage external collaborations and outsourcing of pilot scale cGMP production of clinical candidates.
- Developed quality systems to ensure quality assurance and control in process development, production and documentation.
- Managed the transfer process technologies, protocols and controls to CROs and CMOs

**Associate Research Officer** (Chemical Biology Group) Biotechnology Research Institute (BRI/NRC), Montreal (Aug, 2002 – Apr, 2004)

- Managed the Mass spectrometry facility (Q-TOF Micromass; Sciex API III) for the Research centre.
- Developed High-Throughput Screening methods and systems for natural products drug discovery
- Establish a natural product library from African medicinal plants to enable high throughput discovery of leads from natural products.
- Developed an integrated technology using a combined analytical LC-MSMS, chemotaxonomy and bio-informatics for the discovery of novel bioactives from medicinal plants and the identification of active principles in nutraceuticals, as a means of controlling the medicinal efficacy of "natural medicines".
- Developed High-Throughput protocols for the discovery of novel antiglycation agents, and methods to study glycation induced protein-cross-links that lead to the formation of advanced glycation-end-products and the development of diabetes related heath complications.
- Produced a patent on antiglycation agents and their application in preventing smoking and diabetes related health complications. Patent is currently being developed for Eye drops for the delivery of an anti-glycation agent that prevents and treats diabetes- and age-related ophthalmologic pathologies.
- Developed a successful spin-off business case for the application of LCMSMS/chemotaxonomy for the characterization Natural Health Products and nutraceutical products..

**Post Doctoral Research Fellow** Chemical Biology Group, BRI/NRC, Montreal. (Nov, 1999-Aug, 2002)

**Research Activities**:

**Protein Glycation Research:**

- Development of high throughput assay to screen drug libraries for anti-glycation agents as potential drugs for the treatment of age and diabetes related health complications.
- Study of the mechanism of inhibition of anti-glycation agents.
- Elucidation of protein-carbohydrate interactions in biological systems, that leads to the formation protein cross-links and protein fragmentation.
- Mechanistic based design and synthesis of anti-glycation agents.

**Natural Product Drug Discovery:**

- Developed a natural products library from tropical medicinal plants that is amenable to high throughput screening.
- Development of a natural product drug discovery platform that combine chemotaxonomy, mass spectrometry (dereplication) of medicinal plant extracts, and informatics (natural product databases).
- Development and optimization of a high-load displacement chromatographic procedure for obtaining pure minor bioactive components from medicinal plant extracts for the development of high-value natural product libraries.

**Mass Spectrometry:**

- Development of multi-stage mass spectrometric methods for the identification, characterisation, and the structure elucidation of natural products (alkaloids, terpenoids, and steroids).
- Protein mass spectrometry (Structure/function relationship, post-translational modification, proteomics).
- Development of Mass spectrometric methods for the analysis of drug metabolites

**Research Associate / Lab Manager:** Centre for the Nutrition and Environment of the Indigenous Peoples (CINE), Macdonald Campus of McGill. (1993-96).

- Oversaw the setup of a food and environmental research laboratory for the analysis of environmental pollutants in food systems.
- Managed the lab and provided analytical service to First Nations Peoples in Canada.
- Developed novel analytical methods for the analysis of pesticides, PCB, and polyaromatic hydrocarbons in food products used by the Native Peoples of Canada.
- Provided training to graduate students and laboratory staff on instrumental techniques, lab safety, and GLP standards.

## TEACHING EXPERIENCE

**Faculty Lecturer:** McGill University, Department of Food Science. (1999- Present).
**Chemistry Teacher:** Tema Secondary School, Ghana (1986-89).

**Awards:**

-Max Stern Recruitment Fellowship, (A McGill University major fellowship: $14,000 per year). 1996-1999
-National Science and Engineering Research Council of Canada, ($39000 per year). 1999-2002

**Charity Work**:

Founding Member of Ghana International Health Foundation (GIHF): A Registered Canadian Charitable Organization.
Involved in the raising of funds and providing medical equipment and supplies to Clinics and health service centres in Ghana (West Africa)

**Hobbies:** Chess, Basketball, Table Tennis, Soccer.

## PUBLICATIONS:

Refereed Publications

Jean Watchueng, Pierre Kamnaing, Jin-Ming Gao, Taira Kiyota, **Faustinus Yeboah**, Yasuo Konishia,. Efficient purification of paclitaxel from yews using high-performance displacement chromatography technique. J. Chromatogr. A 1218 (2011) 2929–2935.

Gao, H. Y., Yaylayan, V., Yeboah, F. Oxalic acid-induced modifications of post-glycation activity of lysozyme and its glycoforms. *J. Agric. Food Chem.*, 2010, 58, 6219,-6225.

Allaoua Achouri, Joyce Irene Boye , Denis Belanger, Tiphaine Chiron, Varoujan A. Yaylayan, **Faustinus K. Yeboah**. Functional and molecular properties of calcium precipitated soy glycinin and the effect of glycation with j-carrageenan. *Food Research International* 43 (2010) 1494–1504

S.-J. Cho, G. Roman, **F. Yeboah** and Y. Konishi. The Road to Advanced Glycation End Products: A Mechanistic Perspective. Current Medicinal Chemistry, 2007, 14, 103-112.

Y Konishi, T Kiyota, C Draghici, JM Gao, **F Yeboah**, Molecular Formula Analysis by an MS/MS/MS Technique To Expedite Dereplication of Natural Products. *Anal Chem.* 2007 Feb 1;79(3):1187-97

Charles P. Zuta, Benjamin K. Simpson and **Faustinus K. Yeboah**. Synthesis of acylglycerols from w-3 fatty acids and conjugated linoleic acid isomers. *Biotechnol. Appl. Biochem.* (**2006**) 43, (25–32).

Kebba Sabally; Salwa Karboune; **Faustinus K. Yebaoh**; Selim Kermasha, Enzymatic esterification of dihydrocaffeic acid with linoleyl alcohol in organic solvent media. *Biocatalysis and Biotransformation*, January 2005, vol. 23, no. 1, p. 37-44.

BM Lue, S Karboune, **FK Yeboah,** S Kermasha. Lipase-catalyzed esterification of cinnamic acid and oleyl alcohol in organic solvent media Journal of Chemical Technology & Biotechnology, *J Chem Technol Biotechnol* . **2005** - 80:462–468.

Karboune S, Safari M, Lue BM, **Yeboah FK**, Kermasha S. Lipase-catalyzed biosynthesis of cinnamoylated lipids in a selected organic solvent medium. *J Biotechnol.* 2005 Sep 29;119(3):281-90.

Sabally K, Karboune S, **Yeboah FK**, Kermasha S. Lipase-catalyzed esterification of selected phenolic acids with linolenyl alcohols in organic solvent media. *Appl Biochem Biotechnol.* **2005** Oct;127(1):17-27

A Achouri, JI Boye, VA Yaylayan, **FK Yeboah**, Functional Properties of Glycated Soy 11S Glycinin, *Journal of Food Science.* **2005**. Volume 70 Issue 4, Pages C269 - C274

**Yeboah FK.,** Alli, I., Yaylayan, VA., Yasuo K., Chowdhury SF., Purisima EO. Effect of Limited Solid State Glycation on the Conformation of Lysozyme by ESI-MSMS Peptide Mapping and Molecular Modelling. *Bioconjug Chem.* **2004**, 15(1), 27-34.

**Yeboah FK**, Alli I, Yaylayan VA, Yasuo K, Chowdhury SF, Purisima EO. Effect of limited solid-state glycation on the conformation of lysozyme by ESI-MSMS peptide mapping and molecular modeling. *Bioconjug Chem.* **2004** Jan-Feb;15(1):27-34

**Yeboah, FK**., and Yasuo, K. Mass Spectrometry of Biomolecules: Functional foods, Nutraceuticals and Natural Health Products. *Anal. Lett.* **2003**, 36(15), 3271-3307.

Skopp S, Oehme M, Chu FL, **Yeboah F**, Chan HM. Analysis of "total toxaphene" and selected single congeners in biota by ion trap HRGC-EI-MS/MS using congener-optimized parent ion dissociations. *Environ Sci Technol.* **2002**, 36(12): 2729-35.

**Yeboah FK.,** Yaylayan VA. Analysis of glycated proteins by mass spectrometric techniques: qualitative and quantitative aspects. *Nahrung/Food.* **2001**, 45 (3) 164-171.

Stefanowicz P., Prasain, JK; **Yeboah, FK.,** Konishi, Y Detection and Partial Structure Elucidation of Basic Taxoids from *Taxus wallichiana* by Electrospray Ionization Tandem Mass Spectrometry. *Anal. Chem.* **2001**, 73, 3583-3589.

**Yeboah, FK.,** Alli, I., Yaylayan, V.A., Yasuo K., Stefanowicz P. Monitoring glycation of Lysozyme by electrospray ionization mass spectrometry. *J. Agric Food Chem.* **2000**, 48, 2766-2774.

Chan, HM., **Yeboah, FK**. Total toxaphene and specific congeners in fish from the Yukon, Canada, *Chemosphere*, **2000**, 41(4): 507-515.

**Yeboah, FK.,** Alli, I., Yaylayan, **VA**. Reactivities of D-glucose and D-fructose during glycation of bovine serum albumin. J. *Agric. Food Chem.* **1999**, 47, 3164-3172.

**Yeboah, FK.,** Alli, I Simpson, BK., Yasuo, K., Gibbs, FB.. Tryptic fragments of Phaseolin from hydrolysates of protein isolates of *Phaseolus* Beans. *Food Chem.* **1999**, 67, 105-112.

Chan, M., Zhu, J., **Yeboah, FK**., Determination of toxaphene in biological samples using high resolution GC coupled with ion trap MS/MS. *Chemosphere.* **1998**, 36(9), 2139-2148

Kuhnlein, HV., **Yeboah, FK.,** Sedgemore, M., Sedgemore, S., Chan, HM. Nutritional qualities of ooligan grease: A traditional food fat of British Columbia First Nations. *J. Food Compos. Anal.* **1996**, 9: 18-31.

Belinsky, D., Kuhnlein, HV., **Yeboah, FK.,** Chan, HM. Composition of fish consumed by the James Bay Cree. *J. Food Compos. Anal.* **1996**, 9, 148-162.

Patents:

P Morris, **F Yeboah**, J McAlpine, M Ranger, E Roux, Crystalline Forms of a Farnesyl Dibenzodiazepinone. EPO Application EP1983997, Filed: 01/26/2007, Published: 10/29/2008. U.S. Application Pub No. US20090186878, Filed: 01/26/2007, Published: 07/23/2009.

**Faustinus Yeboah**, Mustapha Aouidate, James B. McAlpine, Ashraf Ibrahim, Arjun H. Banskota, Maxime Ranger, Gregory L. White. Phosphate Prodrugs of a Farnesyl Dibenzodiazepinone: Processes for their production and their use as Pharmaceuticals. U.S. Patent US7384935, Filed: 09/27/2006, Issued: 06/10/2008

**Faustinus Yeboah**, Mahmood Piraee, Process for producing farnesylated dibenzodiazepinone by fermentation. U.S. Patent App. 11/785,226, **2007**, US 2007/0254346 A1.

**Faustinus Yeboah,** Yasuo Konishi, Sung Ju Cho, Jittiwud Lertvorachon, Taira Kiyota, Popek Tomasz. Anti-Glycation Agents for Preventing Age-, Diabetes-, and Smoking Related Complications. U.S. Application Pub No. US20080139664, Filed: 02/11/2008, Published: 06/12/2008.

Lee Soo Young Stanford; Wong Tin Kam Kathy; Jardin Barbara; **Yeboah Faustinus**; Kamen Amine; Konishi Yasuo. Preparation, characterization and quantification of active fractions that promote cell growth and recombinant-protein expression. U.S. Application Pub No. US20050250177, Filed: 08/07/2003, Published: 11/10/2005

Conference Presentations:

My-Anh Wioland, Ashraf Ibrahim, Francis Beaudry, Faustinus Yeboah, Maxime Ranger, James McAlpine and Henriette Gourdea. Pharmacokinetic and ADME studies of the antitumor compound ECO-4601. Amer Assoc Cancer Res, Volume 47, 2006.

El-Ramahi ,R., Alli, I., Konishi, Y., Kermasha, S., and **Yeboah, F.** Characterization of a High Molecular Weight Flaxseed Protein. Institute of Food Technologists, Annual Conference, (IFT) **2003.**

Achouri A., Boye JI., Yaylayan VA., and. **Yeboah FK**, Functional Properties of 11S Glycinin Glycated Through the Maillard Reaction with Glucose. American Oil Chemist Society Annual Conference (**2002**).

Boye JI., Li, P. Lenay J., Yaylayan VA., **Yeboah FK**., and Achouri A. Soy Protein Isolate Glycation and Its Effect on Protein Functionality. American Oil Chemist Society Annual Conference (**2002**).

**Yeboah FK**., Stefanowicz P., Konishi Y. Analysis of Steroids by ESI-Tandem Mass Spectrometry: Based on the Fragmentation of the Steroid Scaffold ion. American Society of Mass Spectrometry 49$^{th}$ Conference, **(2001)**.

**Yeboah FK.,** Alli I., Tryptic Fragments of Phaseolin. CIFST Conference (Montreal) June. **(1997)**.

**Yeboah, FK.,** Alli, I., Simpson BK., Isolation and Characterisation of Proteins from Phaseolus Beans: Poster Presented at the CIFST Conference (Toronto) June **1993.**

Invited Lectures:

Grand Manan Conference (Lyceum Research Company, Business New Brunswick) Title: *Innovative Research – From Biomass to Bioactives.* 2003, October. Subject: **Value of Research in Natural Health Products Development.**

Lyceum Research Company. 2002, October. Subject: **Enhancing Natural Products Drug Discovery: Application of Modern Technologies.**

McGill University (Dept. of Food Science) Lecture Series**: Application of Mass spectrometry in food and environmental analysis** (2000 to present)

Journal Reviewer:

*Journal of Food Composition and Analysis.*
*Journal of Agriculture and Food Chemistry.*
*Journal of Environmental Toxicology.*

Grant reviewer:

National Science and Engineering Research Council of Canada (NSERC).
Fonds zur Förderung der Wissenschaftlichen Forschung (Austria).

# Appendix B

Hughes et al., "Determination of Carryover and Contamination for Mass Spectrometry-Based Chromatographic Assays." The AAPS Journal 2007; 9 (3) Article 42

*The AAPS Journal* 2007; 9 (3) Article 42 (http://www.aapsj.org).

*Themed Issue: Bioanalytical Method Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays*
*Guest Editors - Mario L. Rocci Jr., Vinod P. Shah, Mark J. Rose, Jeffrey M. Sailstad*

# Determination of Carryover and Contamination for Mass Spectrometry–Based Chromatographic Assays

*Received: August 21, 2007; Final Revision Received: September 27, 2007; Accepted: October 15, 2007; Published: November 2, 2007*

Nicola C. Hughes,[1] Ernest Y.K. Wong,[1] Juan Fan,[1] and Navgeet Bajaj[1]

[1]Biovail Contract Research, Toronto, ON M1L 4S4

## Abstract

The Third American Association of Pharmaceutical Scientists/Food and Drug Administration Bioanalytical Workshop, held in 2006, reviewed and evaluated current practices and proposed that carryover and contamination be assessed not only during the validation of an assay but also during the application of the method in a study. In this article, the potential risks of carryover and contamination in each stage of a bioanalytical method are discussed, to explain to the industry why this recommendation is being made.

**Keywords:** Carryover, contamination, extraction, chromatography, detection, bioanalysis, accuracy, precision, memory effect

## Introduction

Sample carryover is a major problem that can influence the accuracy and precision of high performance liquid chromatography (HPLC), liquid chromatography-mass spectrometry (LC-MS), and liquid chromatography-tandem mass spectrometry (LC-MS/MS) bioanalysis, with the consequences being more pronounced at lower concentrations.[1] The continuous increase in sensitivity of new-generation LC-MS/MS instruments, with detection limits in the low pg/mL range and the possibility of using wider calibration ranges (>$10^4$), has also drastically increased the risk of carryover during bioanalysis.[2] Reduction of carryover during assay development consumes time and resources and can lead to reduced productivity and delays in the drug discovery and development process.[3,4]

Carryover in general is serial in nature and is caused by residual analyte from a sample analyzed earlier in the run. It does not necessarily involve only the next sample in the sequence and can affect several samples in a sequence, if many samples above the calibration ranges are analyzed. Carryover can also be random, where carryover from late-eluting residues on chromatographic columns may affect chromatograms several samples later. Carryover from analyte residues can also occur via dislodgment from a sample's flow path through a chromatographic system and mass spectrometric detection system.

Contamination, conversely, tends to be more random, and precautions should be taken to avoid contamination during sample preparation techniques (extraction) using both manual and automated procedures. The potential for contamination and carryover is highly dependent on the calibration range selected for a given assay.

Carryover and contamination can affect both the accuracy and precision of a method and should be investigated and minimized or eliminated during method development, assessed during method validation, and monitored routinely in study samples analysis. It is critical that unexpected or random carryover and contamination not go unchecked. Unless this random carryover and contamination occurs in samples with known analyte concentrations, such as calibration standards, quality control samples, or placebo/predose samples, the contamination will go undetected and potentially erroneous results will be reported for individual samples, or an entire bioanalytical batch. When blanks or low-concentration samples follow, or are in close proximity to, high-concentration samples, there is a potential risk of contamination and carryover. This article will review the potential risks of carryover and contamination during 3 stages of a bioanalytical method (extraction, chromatography, and detection) and provide some important considerations that should be used to assess and prevent them.

## Carryover and Contamination: Sample Preparation (Extraction)

For a bioanalytical assay, the major sources of cross-contamination during sample preparation (extraction) are spills, aerosols, and drips during the liquid transfer steps.[5] Table 1 lists the steps required to perform 3 common bioanalytical sample preparation techniques for small molecules and the potential risk of carryover and cross-contamination.[5] For solid phase extraction there is a moderate risk of carryover during the sample aliquoting,

**Corresponding Author:** Nicola C. Hughes, Biovail Contract Research, Toronto, ON M1L 4S4. Tel: 416-752-3636; Fax: 416-752-7610; E-mail: Nicki.Hughes@biovail.com

**Table 1.** Degree of Risk of Carryover and Cross-Contamination During Sample Preparation*

| Preparation Steps | Carryover | Cross-Contamination | SPE | LLE | PPT |
|---|---|---|---|---|---|
| Aliquot sample, addition of internal standard and reagent, mixing | Medium | Medium | √ | √ | √ |
| SPE elution | Low | High | √ | √ | — |
| Vigorous mixing | High | High | — | √ | √ |
| Transfer of supernatant/extract | Medium | High | — | √ | √ |
| Evaporation of extract | Medium | High | √ | √ | √ |
| Dilute extraction and mixing | Medium | Medium | √ | — | √ |

*SPE indicates solid phase extraction; LLE, liquid-liquid extraction; PPT, protein precipitation. Check marks (√) indicate the at-risk steps involved in SPE, LLE, or PPT.[5]

evaporation, and reconstitution steps. However, the chances of cross-contamination are quite high during the elution and evaporation steps. The risk of cross-contamination is also very high during the vigorous mixing of organic solvents, supernatant transfer, and evaporation steps for liquid-liquid extraction (LLE) and protein precipitation (PPT).[5]

### *Manual Extractions*

Since the early 1990s there has been a shift toward the use of automated liquid handlers to carry out extractions.[4] Some bioanalytical laboratories, however, still carry out these extractions manually. Speed and throughput are compromised in extractions done manually, but problems due to carryover and contamination are generally less pronounced. It follows that it is easier to limit, or avoid, these mitigating effects when the sample preparation and extractions are done manually.

There are several ways to overcome these problems during manual extractions. For instance, to reduce or eliminate carryover, the glassware in which analyte stock solutions are prepared should not be reused for preparing other solutions, such as buffers, working internal standard solutions, and dilute analyte solutions (spiking solutions). Those flasks should be cleaned separately (not with the other glassware) to prevent carryover of analytes. Workbenches, pipettes, vacuum manifolds, evaporation needles, and other items should be cleaned with appropriate reagents before each extraction. Moreover, when performing extractions for HPLC assays, bioanalytical scientists need to be extra vigilant if they share equipment or glassware with others. The poorer selectivity of HPLC detection techniques means that if reagents/solvents or common glassware are contaminated with analytes, albeit from a different assay, they may be detectable by HPLC and influence the selectivity and accuracy of the assay. Cross-contamination between assays may also affect quantitation for MS-based assays, if the cross-contamination analyte co-elutes with the analyte of interest, potentially causing sequential or random ionization suppression/enhancement.

Pipetting using handheld devices should be done slowly to minimize foaming and aerosol formation. Pipettes with aerosol barrier tips are commercially available and may be used. The airflow through these tips reduces the flow of aerosols or liquid into the pipette barrel, which helps to prevent carryover and contamination.[6]

The selection of appropriately sized test tubes is imperative to avoid splashing during the vortexing steps of sample preparation. Contamination from extraction solutions can be avoided by using separate refillable bottles for extraction solvents. These bottles should be emptied and refilled daily. In some cases contamination or interference could arise from impurities in buffers/organic solvents, such as methanol and acetonitrile, and the use of high-purity reagents is recommended.

### *Automated Extractions*

When extractions are performed using automated liquid handlers, the potential of carryover and cross-contamination increases because the samples are clustered together in a 96- or 384-well format. This physical characteristic, with each sample being in close proximity, leads more readily to cross-contamination. Using fixed tips is less expensive than using disposable tips, but fixed tips are more likely to lead to carryover problems. This effect is more pronounced when the analyte is "sticky" and prone to adsorption to the surface of the tip. Appropriate methodology involving washing and rinsing solutions can be used for fixed tips to lower the risk of carryover considerably in most cases, but the requirement for extensive washes between steps will ultimately affect sample throughput.

Currently, there are several automated liquid handlers that can control the dispensing height, dispensing speed, position of tips, and adjustment of air gap to prevent dripping

and thereby limit contamination. Nevertheless, transfer of organic solvents is a potential source of contamination due to dripping. During PPT or LLE, the mixing step may generate aerosols or allow organic solvents to climb over the barriers between wells because of capillary action. For example, the capillary action in polypropylene microtiter plates is highest for heptane > ethyl acetate > 75% methanol or acetonitrile > water > 50% dimethyl sulfoxide. Capillary action thus reduces the usable volume of the wells, thereby affecting accuracy and precision.[7]

To avoid cross-contamination during the mixing steps in a PPT or an LLE, heat-sealing films can be used. Heat-sealing films are also available with pierceable sealing foil, which further limits contamination. Caution should be exercised while removing the films because of the potential for contamination from the droplets on the film. An additional step of centrifugation could be performed to remove the droplets.

Some automated liquid handlers can mix the sample using disposable tips, which helps eliminate the risk of contamination from sealing films. The bioanalytical scientist should consider these factors in designing the analytical method and determining when it is appropriate to use 96-well plates with larger volume, fixed tips or disposable tips, or square well or round well plates, and should also consider displacement of solution from tips when tips are used for sample mixing.[5]

The use of surrogate markers or contamination markers for LC-MS/MS is becoming very popular in tracking the cross-contamination when extraction is performed in a 96- or a 384-well format. A surrogate marker, often an analog of the analyte, is ionizable at the MS interface, extracted with the analyte, and eluted in the HPLC method but not co-eluted with the analyte or the internal standard. The method is developed for the analyte and the extraction recoveries, and chromatography is determined for the surrogate marker. To monitor the cross-contamination, high concentrations of markers are spiked in a checkerboard pattern as shown in Figure 1. The markers are added to a clean 96-well plate and evaporated (if required); then the spiked plate is used for sample preparation.[5] The presence of both markers in any well indicates cross-contamination has occurred. The response of an unspiked marker in the well is subsequently measured, with the result indicating the degree of contamination.[4] Routine application of this technique does add to the time and expenses required to develop a bioanalytical method, as extraction and chromatographic conditions for the surrogate as well as the analyte of interest will need to be developed. Despite this limitation, the most notable advantage of the application of this technique is that cross-contamination can be assessed for all samples individually. If significant cross-contamination is observed, only those affected samples, rather than the whole batch, would be failed (deactivated).

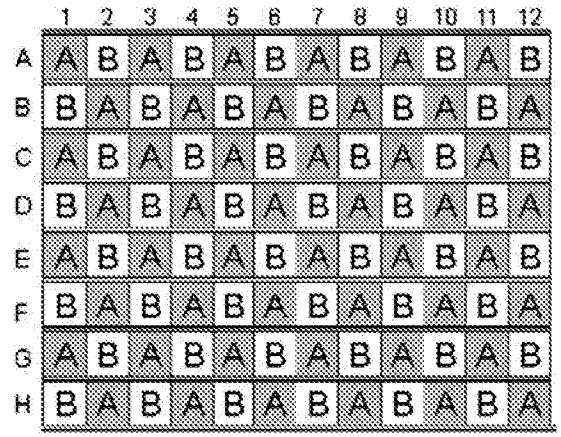



**Figure 1.** Use of 2 additional analytes (A and B) as surrogate markers[5]

In cases when extraction contamination and carryover are not observed in control samples but are suspected (eg, upon random sample repeat), additional investigational analysis may be required. This will allow the cause to be identified and appropriate and corrective action to be performed, to ensure the integrity of the results of other samples in the batch, and subsequent analysis of batches.

## Carryover and Contamination: Chromatography

Carryover and contamination from a chromatographic system can be caused by residues of a previously injected sample that are absorbed on, or trapped within, the autosampler. Carryover can also be caused by residues on columns that may randomly affect chromatograms several samples later. There are many publications that describe measures to deal with autosampler carryover, but only a few discuss column carryover. This section discusses autosampler carryover, the origins of carryover, and the means to overcome issues associated with column carryover.

### *Types and Features of HPLC Carryover*

The primary causes of HPLC carryover can be divided into 2 categories: autosampler carryover and column carryover. Autosampler carryover results from the residue of a previously injected sample absorbed on and/or trapped in the autosampler needle, injection port, transfer tube, sample loop, or injector valve. Typical autosampler carryover has a similar retention time to that of the analyte. This often introduces a positive bias (% relative error) and consequently has a major impact on the accuracy of quantitation, most significantly at lower analyte concentrations. Column carryover, however, can be caused by the residue of a previously injected sample on the column, both in its original form and occasionally in different forms of the analyte (eg, analyte:reagent adducts and analyte dimers)[8] that can

decompose in the ion source back to the original form of the analyte. Typical column carryover has uncertain analyte retention times and often generates random error that affects mainly the method precision.

### *Interaction Mechanisms and Solutions Used to Reduce HPLC Carryover*

Autosampler carryover is largely associated with the interaction of an analyte with the flow path components of the system; it has a close relationship with the chemical/physical characteristics of both the analyte and the analysis system. Analysis of extremely basic and hydrophobic compounds can be particularly problematic, because of their tendency to be present in a charged form and to adsorb to the sample path of an autosampler through ionic interaction with metallic surfaces and through hydrophobic interaction with plastic materials.[9] Great efforts have been made by scientists and engineers to reduce carryover in 2 ways: by removing it by rinsing, and by preventing it in the first place.[9-14] Rinsing can be effective, but selection of the most effective rinsing solution, optimized for time, is no trivial matter. Rinse solution chemistry can have a huge impact and should be carefully considered to best counteract carryover. "Like dissolves like" is the primary rule to follow. Generally speaking, acetonitrile or 90% acetonitrile is an acceptable choice for rinsing/removing analytes adsorbed by hydrophobic interaction (eg, lipophilic compounds). A more protic solvent, such as methanol or 90% methanol, is an alternative for more polar lipophilic compounds. Acidified acetonitrile, alkalized acetonitrile, or methanol/isopropanol/water solution is quite efficient and universally used to dissociate analyte adsorption caused by dipole-dipole and ionic interaction (hydrophilic compounds).

Matching the pH to the organic/water or buffer ratio of the rinsing solution can dramatically reduce carryover since the pH of the rinsing solution influences the analyte charge state. For example, a basic compound exists in a positively charged state under acidic and neutral conditions and is uncharged in alkaline conditions. An acidified organic/water or an alkalized organic needle/valve wash solution is useful in removing it, but selection of an acidified organic or alkalized organic/water solution will greatly compromise the rinsing effectiveness. This effect occurs because when charged (ionized), a basic compound easily dissolves in organic/water or acidified organic/water solutions. However, in an uncharged state, it has more affinity toward pure organic or alkalized organic needle/valve wash solutions.

The $pK_a$ of an analyte is a good indicator that should be considered when making pH adjustments to the needle/valve wash solutions. For an analyte that is hard to dissolve in common solvents (methanol, acetonitrile, or aqueous mixtures thereof), strong solvents such as tetrahydrofuran, dimethylsulfoxide, or a halohydrocarbon (eg, methylene chloride) can be used. Use of such strong solvents can, however, cause nonmetallic tubing to swell, which greatly reduces the rupture pressure of the tubing and should be avoided under ultra performance chromatography (UPLC) conditions. An ion pair reagent such as perchloric acid can be used as a rinsing solution, to reduce sample adsorption caused by ionic or coordination interactions, but the possible effect of the counterion should be considered in MS-based assays, as it may suppress ionization. Also, the introduction of any nonvolatile ion pair reagents into the MS system must be avoided.

Most modern autosamplers are equipped with 2 or more needle- and valve-wash lines, allowing multiple rinses to be performed. The first rinsing solution removes analyte residues and involves a weaker solution or mobile phase. The last rinsing solution has better compatibility with the detection system. If only 1 needle- or valve-wash for the autosampler is available, the options for selecting suitable rinsing solutions are more limited, and the compatibility of the rinsing solution with the mobile phase must be considered.

### *Autosampler Design*

Many improvements have been made in autosampler design, materials, and techniques to prevent or limit carryover. The first is the "push-to-fill" design, which is an automated version of a manual injection. In this design, a needle attached to a motor-driven syringe is moved to the sample vial, is filled, and then transfers the sample to the injection loop. The valve rotor is moved, and the sample is injected. Any sample residue left inside the needle, the syringe, or the connecting tubing can be flushed out with a wash solvent or rinsing solution. Another setup involves the "needle-in-loop" design, which combines the needle and loop as 1 component, so that both the needle and the loop are flushed with the mobile phase during the sample elution and no additional internal rinsing of the needle is required. As rinsing takes place during the chromatographic run, it is best to leave the loop in the inject position during the entire run for maximum flushing, especially during gradient elution chromatographic methods. An alternative design is the "load-ahead" autosampler, in which the loop is removed from the inject position before the run is complete. This may have the potential for less thorough flushing of the inside of the loop.

Carryover can also result from sample residue left on the outside of the sample needle. The vial septum is the first line of defense to remove any residue on the outside of the needle. A well-chosen septum will act as a "squeegee" and wipe the outside of the needle. Polymeric septa, such as silicone or polytetrafluoroethylene-faced silicone, work well

in this regard. In the "needle-in-loop" design, there is normally no valve wash but there is an external needle wash to avoid injection seal contamination. There are 2 common techniques for external needle wash. The first technique is "dip only," which is a static dip approach used to wash the external needle by dipping it into a vial of wash solvent. The second technique is the "active rinse," in which the needle is dipped into a wash station with wash solvent flowing on the outside of the needle. This approach is slower but more effective than a static dip technique, but the static dip is better than no rinsing at all.

Over the years, injection needle coatings have been developed to prevent carryover caused by basic or ionic compounds adsorbed to metallic needle surfaces (eg, stainless steel alloy) by ionic or coordination interaction. Three kinds of common needle coatings are commercially available: Teflon, polyetheretherketones (PEEK), and platinum. Teflon coating is mechanically weak (coating layers can peel off after ~300 injections). PEEK is a thin-layer coating (of a few dozen micrometers) that is technically complex, is chemically stable, and has utility across a broad pH range. Platinum coating is also a thin layer (of a few micrometers) and due to a special coating process is very durable and can last more than 20 000 injections.

Adsorption of lipophilic analytes, via hydrophobic interaction, with resinous materials on rotor seals can be a significant cause of carryover. Vespel is common material employed in rotor seals with excellent durability, but unfortunately it has a strong affinity for lipophilic molecules. Delrin is another common material that can be used with an alkaline mobile phase with little adsorption of hydrophobic compounds. PEEK seals are also available and can be used with the mobile phase across the entire pH range with little adsorption of lipophilic compounds.

### *Column Carryover*

Column carryover is very compound-dependent and is related mainly to analyte:reagent interaction. The so-called sticky analytes often have unique chemical and physical characteristics. Compounds having active positive carbon atoms in the molecule, or strong electron withdrawing groups (eg, fluoride ions), have a strong tendency to form adducts with common organic, acid, salt, and solvent ions. Compounds that contain dipolar ions or are rich in hydroxy groups can form low-molecular-weight polymers (typically dimers) at high concentrations. The different adduct or polymer forms of an analyte can then decompose in the ion source (by in source collision-induced dissociation) back to the original analyte form and cause random carryover- and contamination-like effects that can affect the quantitation of the assay. These 2 cases may be thought of as late-eluting interference effects but should also be considered as a special case of carryover due to analyte interaction with the mobile phase, extending analyte retention on the column. This type of carryover can be observed as a highly variable analyte response, particularly at low analyte concentrations. The potential for this type of analyte-adduct formation or polymerization should be taken into account during the method development process. Precautions should be taken to avoid adduct formation or analyte polymerization during extraction, chromatography, and detection. Gradient elution could be considered an option for removing such effects when adduction or polymerization cannot be minimized effectively. The extended interaction of basic compounds, caused by ionic interaction with active acidic sites on silicone-based stationary phases, is well known. Careful selection of column chemistry will provide many good options to overcome this kind of problem.

### *Assessment and Accepted Criterion for Autosampler Carryover*

Carryover can be assessed by injecting 1 or more blank samples after a high-concentration sample or standard.[1] The commonly accepted criterion for carryover is that the peak area of the analyte in a blank sample that follows a standard prepared at the upper limit of quantitation (ULOQ) must be less than 20% of the peak area of the lower limit of quantitation (LLOQ) sample. This criterion is closely correlated to the dynamic range of a bioanalytical assay. Considering that carryover is proportional to the concentration of analyte in the preceding sample, the higher the concentration of the preceding sample, the higher the peak area will be in the sample that follows. Therefore, the selection of the LLOQ of an assay is directly related to the ULOQ and any subsequent carryover. In addition, because the peak response from carryover in the blank sample is also directly related to the sensitivity of the detector, the absolute peak response may vary from day to day or from system to system for the same analyte. Therefore, autosampler carryover evaluation should be performed for each analytical run to ensure that it does not affect the accuracy of quantitation. An assessment of autosampler carryover may be challenging when the response of an analyte at the LOQ is close to the limit of detection, where it may be difficult to accurately differentiate carryover from background noise. In such cases, additional experiments may be required when considering the impact of any carryover on the integrity of the data.

## Carryover and Contamination: Mass Spectrometry Detection

### *Artifactual Contamination Caused by Cross-Talk*

"Cross-talk" is caused by the slow removal of ions from the collision cell.[15] This can become a problem if different

analytes of interest have the same monitored fragment ions. For example, cross-talk occurs when fragment ions from the first mass transition scan event of an analyte have not cleared the collision cell before a second mass transition scan event of another analyte takes place. The impact of this cross-talk leads to signal/response artifacts in the next mass transition, so it has an impact on the quantitation of the analytes of interest. Modern triple-quadrupole mass spectrometers have been redesigned so that collision cells evacuate the ions quickly before the next mass transition scan event takes place.[16] For the old mass spectrometers, this problem still remains, but it can be resolved by adding a "dummy ion transition" scan event between the 2 analytes of interest, and thereby allowing time for the collision cell to empty of the common fragment ion, which eliminates the "artifactual contamination" caused by cross-talk.

Intersprayer cross-talk[17,18] has also been reported using multiplexed electrospray technology. An evaluation of the cross-talk effect using this type of mass spectrometer platform should be considered in the development and application of methods that use this technique.

### *Memory Effect I: Column Carryover*

Memory effect I is observed as an elevated, downward-drifting baseline in a blank sample analyzed after a high-concentration sample.[19] This suggests that the analyte from the previous injection was still eluting off the column at the time when the blank injection was made. The elevated baseline is in fact the tail of the peak from the previous injection (Figure 2). This is common for analytes that exhibit strong interactions with silanol groups on the chromatographic column, and that have a very short run time, such that the analyte peak has had insufficient time to fully elute from the column. This raised baseline in the subsequent samples may affect the analyte if present at low concentrations—that is, the peak becomes hard to accurately differentiate from the background noise. This problem can be improved by selection of end-capped columns to minimize the residual silanol effects, careful selection of the mobile phase pH, and adjustment of the chromatographic run time.

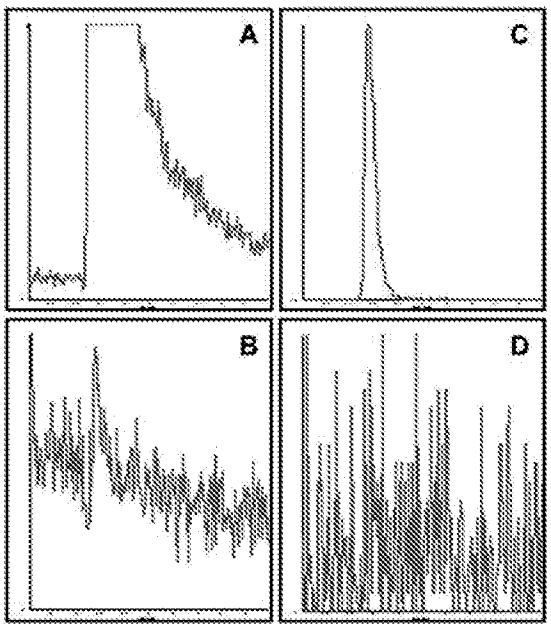



**Figure 2.** Memory effect due to chromatographic peak tailing: (a) ULOQ sample magnified to show peak tailing; (b) peak tailing from ULOQ (a) causes memory effect in blank sample that follows; (c) modification of chromatographic conditions for ULOQ; (d) no memory effect in blank that follows ULOQ (c). ULOQ indicates upper limit of quantitation.

### *Memory Effect II: Additives Such As Triethylamine*

If triethylamine (TEA) has been used in the mobile phase for 1 assay, any residual TEA that remains in the system may carry over and have a negative impact on the quantitation of an analyte of interest for subsequent analysis.[20] TEA strongly adsorbs on the surfaces of the mass spectrometer and can produce ion suppression of other analytes, particularly for those present in low concentrations with low detection limits. Hence an evaluation of the impact on the quantitation of these analytes is required and if necessary a thorough cleaning of the system may be indicated to remove or reduce the impact of carryover and contamination from residual TEA.

### *Chip-Based Technology*

In chip-based technology, electrospray ionization (ESI) is integrated into a chip format to form an array of ESI nozzles.[21] This technology is similar, in principle, to flow injection analysis in that each sample has its own unique spray (ESI nozzle) and no chromatography. This MS-based approach has the advantage of directly introducing each sample into the mass spectrometer without the mobile phase or any common sample flow path. The possibility of injection and chromatographic carryover is therefore completely eliminated, and extended calibration ranges can be used. The major disadvantage of this chip-based analysis format is that the analyte may co-elute with its metabolites or there may be endogenous matrix interferences, because of the absence of chromatographic separation. Ion suppression can be significant, and if it is not consistent from matrix to matrix, quantitation of the analyte can be affected. This type of approach is also not suitable for the differential quantitation of isomers (structural or enantiomers), because without chromatographic separation, the isomers cannot be differentiated by the mass spectrometer alone. This approach has significant limitations for quantitative application of bioanalytical methods to support human clinical trials. Conversely, the lack of carryover, and hence the time required to

minimize it, is particularly advantageous during drug development. Using this technique, high-throughput screening of a large number of samples over wide calibration ranges is achievable, with no risk of an impact from carryover and contamination.

### *Purity of Stable Isotopic-Labeled Internal Standards*

The purity of stable isotopic-labeled internal standards, which are commonly used in bioanalytical assays, is an important consideration. For example, if the $D_0$ of a deuterated internal standard is present in a significant amount, "apparent contamination" from the internal standard can affect the quantitation of an analyte, and in such cases the concentration of the internal standard used needs to be carefully selected relative to the LOQ of a given assay.

## CONCLUSION

It is clear that each stage of bioanalysis (extraction, chromatography, and detection) is susceptible to risk from carryover and contamination. These effects can be both serial and random. During method development and validation, these risks should be understood, and steps need be taken to ensure they are eliminated or minimized. While there is no standard acceptable magnitude of carryover and contamination for a passing bioanalytical run, it is most typically assessed in blanks analyzed after the highest calibration standard. During the routine application of bioanalytical methods in support of preclinical and clinical trials, this type of assessment must be performed for each batch of analysis. It is imperative to ensure that carryover and contamination do not affect the in-process accuracy and precision of the method and thereby guarantee the integrity of the results generated for all samples analyzed. When unexpected/unplanned occurrences of carryover and contamination do occur, the bioanalytical scientist must interpret the impact on the results and carry out the appropriate corrective action to eliminate further occurrences.

## REFERENCES

1. Viswanathan CT, Bansal S, Booth B, et al. Workshop/conference report—quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *AAPS J [serial online]*. 2007;9:E30-E42.

2. Zeng W, Musson DG, Fisher AL, et al. Determination of MK-0431 in human plasma using high turbulence liquid chromatography online extraction and tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2006;20:1169-1175.

3. Dethy JM, Ackermann BL, Delatour C, et al. Demonstration of direct bioanalysis of drugs in plasma using nanoelectrospray infusion from a silicon chip coupled with tandem mass spectrometry. *Anal Chem*. 2003;75:805-811.

4. Chang MS, Kim EJ, El-Shourbagy TA. A novel approach for in-process monitoring and managing cross-contamination in a high-throughput high-performance liquid chromatography assay with tandem mass spectrometric detection. *Rapid Commun Mass Spectrom*. 2006;20:2190-2200.

5. Chang MS, Kim EJ, El-Shourbagy TA. Managing bioanalytical cross-contamination. Available at: http://www.americanpharmaceuticalreview.com/articleDetail.asp?SID=02F0ABB61C2C47B6BAC99C9C6A00A84D&ArticleID=483. Accessed October 16, 2007.

6. Mannonen S, Syrja, K. Liquid handling application notes. Available at: http://www.biohit.com/pdf/app11.pdf. Accessed October 16, 2007.

7. Chang MS, Kim EJ, El-Shourbagy TA. Evaluation of 384-well formatted sample preparation technologies for regulated bioanalysis. *Rapid Commun Mass Spectrom*. 2007;21:64-72.

8. Lambert W. Analytical Pitfalls and Trends in Clinical and Forensic Toxicology. Joint GTFCh/TIAFT Symposium of the Analytical Conference; May 12, 2004; Munich, Germany. 2004; Available at: http://www.gtfch.org/tk/tk71_2/Lambert.pdf. Accessed November 1, 2007.

9. Hedgepeth W, Steinike S. High Sensitivity MS Determination of Carryover in a New Autosampler Design. Application Note. Available at: http://www.ssi.shimadzu.com/products/pdfs/hplc/Shimadzu-ABI_Carryove_PC2006.pdf. Accessed November 1, 2007.

10. Elmashni D. HPLC Carryover----Decreased Sample Carryover Using the Finnigan Surveyor Autosampler. *Application Note 330*. Thermo Electron Corporation.

11. Dolan JW. Autosampler Carryover. 2006; Available at: http://www.lcgceurope.com/lcgceurope/LC+Troubleshooting/Autosampler-Carryover/ArticleStandard/Article/detail/377221. Accessed October 31, 2007.

12. Tobien T, Bethem R. Problems and Solutions for Carry-Over on High-Throughput LC Autosamplers. Proceedings of the 50th ASMS Conference on Mass Spectrometry and Allied Topics; June 2-6, 2002; Orlando, Florida. Available at: http://www.asms.org/asms02pdf/1891.pdf. Accessed November 1, 2007.

13. Watanabe J, Hike H, Koura Y, Cueni H, Bando Y. Method Development of Minimizing Carry Over in LC/MS/MS-Novel Cleaning Technique and Device. Paper presented at: Proceedings of the 54th ASMS Conference on Mass Spectrometry and Allied Topics; May 28-June 1, 2006; Seattle, Washington. 2006: Poster TP0118.

14. Dolan JW. Autosampler Carryover. 2001; Available at: http://www.lcgceurope.com/lcgceurope/data/articlestandard//lcgceurope/042002/7657/article.pdf. Accessed October 31, 2007.

15. Yamagishi Y. Utility of H-SRM to Reduce Matrix Interference in Food Residue Analysis of Pesticides by LC/MS/MS Using the Finnigan TSQ Quantum Discovery. Application Note 355. Thermo Scientific. 2005; Available at: http://www.thermo.com/eThermo/CMA/PDFs/Articles/articlesFile_28817.pdf. Accessed November 1, 2007.

16. A Simple and Rapid LC-MS/MS Method for the Simultaneous Determination of Nine Antiretroviral Drugs Commonly Used in Europe (Protease Inhibitors and Non-Nucleoside Reverse Transcriptase Inhibitors). Application Note. Applied Biosystems. 2005; Available at: http://www3.appliedbiosystems.com/cms/groups/psm_marketing/documents/generaldocuments/cms_042226.pdf. Accessed November 1, 2007.

17. Organ A. Mux for Parallel LC-MS. CPSA Digest. New Technologies and Approaches for Increasing Drug Candidate

Survivability: Lead Identification to Lead Optimization. October, 2001. Available at: http://www.milestonedevelopment.com/CPSA/2001/day1oc3.html. Accessed October 31, 2007.

18. Yang L, Mann TD, Little D, et al. Evaluation of a four-channel multiplexed electrospray triple quadrupole mass spectrometer for the simultaneous validation of LC/MS/MS methods in four different preclinical matrixes. *Anal Chem.* 2001;73: 1740-1747.

19. Thurmond M, Hodgin JC. Chromatographic factors affecting sample carry-over in the LC/MS analyses of small molecules pharmaceuticals. Available at: http://www.nocarryover.com/Chromatographic-Tabloid.pdf. Accessed October 15, 2007.

20. Rutters H, Mohring T, Rullkotter J, et al. The persistent memory effect of triethylamine in the analysis of phospholipids by liquid chromatography/mass spectrometry. *Rapid Commun Mass Spectrom.* 2000;14:122-123.

21. Wickremsinhe ER, Ackermann BL, Chadhary AK. Validating regulatory-compliant wide dynamic range bioanalytical assays using chip-based nanoelectrospray tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2005;19:47-56.

# Appendix C

Elliott et al., Current Trends in Quantitative Proteomics. J. Mass. Spectrom., 44 (12): 1637–1660 (2009)

Special Feature: Tutorial

MASS SPECTROMETRY

Received: 10 July 2009 Accepted: 6 October 2009 Published online in Wiley Interscience: 2 December 2009

(www.interscience.com) DOI 10.1002/jms.1692

# Current trends in quantitative proteomics

Monica H. Elliott,[a] Derek S. Smith,[a] Carol E. Parker[a] and Christoph Borchers[a,b]*

**It was inevitable that as soon as mass spectrometrists were able to tell biologists which proteins were in their samples, the next question would be how much of these proteins were present. This has turned out to be a much more challenging question. In this review, we describe the multiple ways that mass spectrometry has attempted to address this issue, both for relative quantitation and for absolute quantitation of proteins. There is no single method that will work for every problem or for every sample. What we present here is a variety of techniques, with guidelines that we hope will assist the researcher in selecting the most appropriate technique for the particular biological problem that needs to be addressed. We need to emphasize that this is a very active area of proteomics research – new quantitative methods are continuously being introduced and some 'pitfalls' of older methods are just being discovered. However, even though there is no perfect technique – and a better technique may be developed tomorrow – valuable information on biomarkers and pathways can be obtained using these currently available methods.** 

*Supporting information may be found in the online version of this article.*

**Keywords:** proteomics; plasma; quantitation; iTRAQ; SISCAPA; MRM; ICAT; iMALDI; SILAC; Label-free; DIGE, $^{15}N$; $^{18}O$

## Introduction

The ability to accurately quantitate changes in protein expression in response to a variety of changes is one of the most important goals of proteomics.[1] The development of methods for accurate protein quantitation is currently one of the most challenging – and rapidly changing – areas of proteomics. Whether observing protein expression during stem cell differentiation,[2] or investigating changes in host protein concentrations during a viral infection, researchers need accurate and quantitative proteomics techniques in order to answer biological and biomedical questions. The currently available methods which are best suited for a particular project depend on multiple factors including the source of the samples, the number of samples, the number of treatments being compared, the type of equipment available, and, of course, the expense and time required.

Quantitative proteomics can be categorized into two types: absolute and relative. Absolute quantitation determines changes in protein expression in terms of an exact amount or concentration (e.g. ng or nmoles per gram of tissue, or ng or nmoles/ml of plasma) of each protein present, whereas relative quantitation determines the up- or down-regulation of a protein relative to the control sample, and the results are expressed as 'fold' increases or decreases. The decision to use absolute or relative quantitation will depend on whether it is necessary to find the exact concentration values or just the 'differences' between the various treatments. It should be noted, of course, that even 'absolute' quantitation is relative – relative to an internal standard.

From a practical point of view, to determine 'differences', one sample can simply be compared to another. For absolute quantitation, however, the absolute amount of material in the reference sample must be known. This means that standard known amounts of the target proteins or peptides must be available for all of the targeted analytes. This is not a trivial requirement, as will be clear from the discussions below.

## Targeted Absolute Quantitation

### Multiple reaction monitoring

In its simplest form, quantitation can be done by comparing key characteristic ions from the labeled and unlabeled standard with respect to the peak heights or peak areas of labeled versus unlabeled analytes. Online separation can be added to increase the specificity of the assay. Targeted quantitation of analytes based on isotopically labeled internal standards has been used for GC/MS since 1981,[3] and for LC/MS since 1987 or earlier.[4,5] To improve the sensitivity of the assay, instead of scanning all of the possible *m/z* values, only selected ions can be measured. This selected ion monitoring (SIM) technique can improve the limits of detection for an analyte by several orders of magnitude. Monitoring selected parent or fragment ions compared with corresponding ions from isotopically labeled standards has been used for at least 40 years. In 1996, Barr *et al.* measured the amount of apolipoprotein A-1 in, a purified protein reference standard by synthesizing deuterium-labeled peptides corresponding to their native counterparts formed by proteolysis.[6] An example of a quantitation study based on SIM with labeled internal standards is shown in Fig. 1 below.

The development of tandem mass spectrometry greatly enhanced the accuracy and specificity of analyte quantitation through the use of selected *reaction* monitoring (SRM). In this form of quantitation, precursor/fragment ion *pairs* are monitored.

* *Correspondence to: Christoph Borchers, University of Victoria, Biochemistry & Microbiology, 4464 Markham Street, Suite 3101, Victoria, British Columbia, V8Z7X8, Canada. E-mail: christoph@proteincentre.com*

a *University of Victoria Genome BC Proteomics Centre, British Columbia, V8Z 7X8, Canada*

b *University of Victoria, Department of Biochemistry & Microbiology, British Columbia, V8Z7X8, Canada*

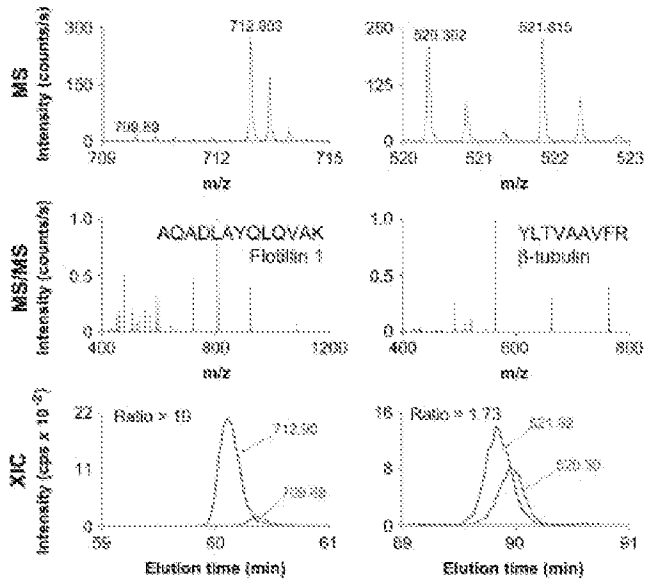



**Figure 1.** MS, MS/MS and extracted ion chromatograms (XIC) for labeled and unlabeled peptides from flotillin 1 and $\beta$-tubulin. Reprinted from Ref. [7], with permission.

This type of analysis has been performed on triple-quadrupole instruments for more than 30 years and has become commonplace in both clinical and industrial settings. It is now routinely used for proteomics and metabolomics.[8]

In a triple-quadrupole mass spectrometer (Fig. 2), high sensitivity and specificity are achieved by only allowing a selected peptide to pass through the first quadrupole and enter the collision cell (Q2). Inside the collision cell, the peptide dissociates into fragments specific to the amino acid sequence of the precursor peptide. A second stage of specificity is added in the second MS (Q3), and only a specific fragment is allowed to pass through and strike the detector. By repeatedly cycling through a list of SRM ion pairs associated with a set of specific retention times, multiple peptides can be targeted in a single *multiple* reaction monitoring (MRM) experiment.

MRM quantitation is ideal for proteomics because the MRM scan type allows for targeted detection of analytes from a complex background.[9] Typically, in an MRM experiment, a stable isotope labeled analog to the target analyte is spiked into the sample matrix and the mass spectrometer is set to scan only for the diagnostic fragment ions of both the analyte and the spiked compound. Quantitation is determined using the relative response of the spiked compound to the target analyte.

The application of the MRM technique for the absolute quantitation (AQUA) of proteins was first introduced by Steven Gygi in 2003.[10] Gygi *et al.* proposed that protein concentrations could be determined by quantitating their component tryptic peptides. This means that instead of *protein* quantitation, one is actually performing *peptide* quantitation. A known amount of isotopically labeled tryptic peptides is added to a tryptic digest of a complex protein mixture in order to monitor changes in protein expression. Now protein quantitation can be performed in the same way that small molecules have been quantitated for decades. A general approach to this technique can be found in Ref. [11].

Absolute quantitation workflows include stable isotope labeled ($^2$H, $^{13}$C, or $^{15}$N) versions of the targeted peptides. The endogenous and isotopically labeled peptides co-elute (or nearly co-elute in the case of deuterated standards) from reversed-phase high-performance liquid chromatography (HPLC) columns, and will behave identically in the mass spectrometer with the exception of the introduced mass shift. Both versions are monitored and quantified, and their concentrations are determined from the relative responses (peak heights or peak areas) of the spiked-in compound to the target analyte.

This approach is usually limited to a small number of preselected proteins because suitable internal standards need to be identified and synthesized.[1] However, a method has recently been developed in our laboratory for the MRM-based quantitation of 45 proteins in human plasma (Fig. 3).[12]

There are three steps to designing an MRM assay for proteomics biomarker studies: First, the target proteins have to be selected, e.g. from the set of differentially expressed proteins found in the biomarker 'discovery' experiments.

Second, peptides have to be selected which have good MS responses and uniquely identify the target protein. Finally, the fragment ions for each peptide that provide the best signal intensity and discriminate the targeted peptide from other species present in the sample have to be selected. Even after the targeted proteins have been selected, developing an MRM method for quantifying their peptides in a plasma digest, which may contain hundreds of thousands of proteins, is a daunting task. This process is described in Ref. [12] and is diagrammed in Fig. 4.

As mass spectrometry-based proteomics moves from the 'discovery' stage to the verification and validation stages, new bioinformatics software packages are being developed to assist in MRM-based assay development. There are already 'libraries' of MRM transitions currently in use for the targeted analysis of drugs and small molecules.[13] There is new software for predicting 'proteotypic peptides' (i.e. peptides that are consistently observed from a given protein),[14,15] calculation of unique peptide sequences for target proteins,[16] predicting peptide intensities in MALDI[17] and ESI[18] and new libraries of MRM pairs for target proteins (cancer biomarker peptide libraries,[19] yeast,[20] and other libraries generated from data in the literature or the Global Proteome Machine Database[21]). There are several new software packages that have been developed to calculate MRM transitions.



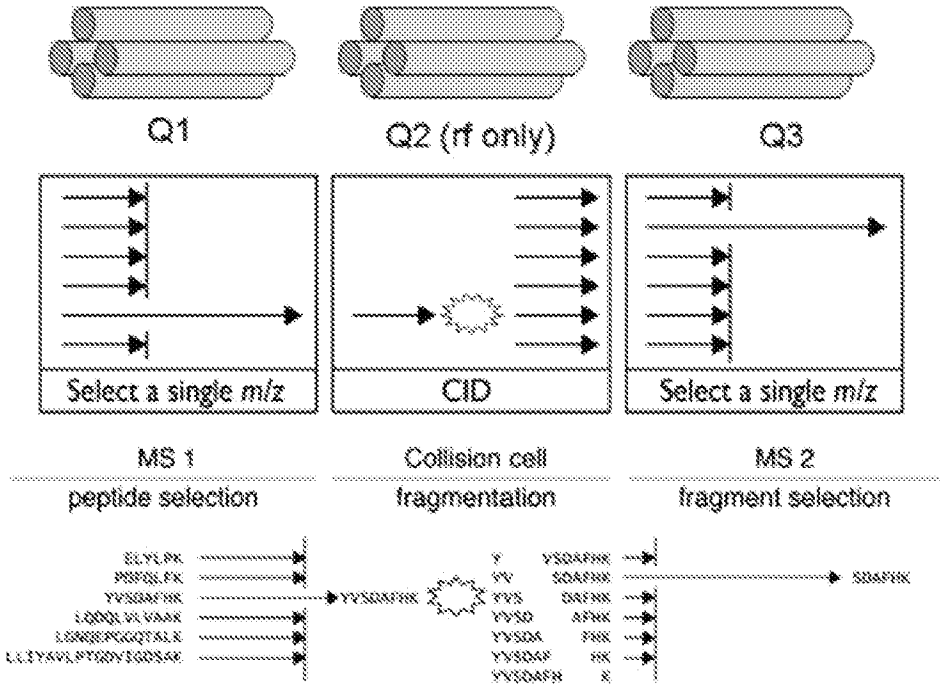



**Figure 2.** Schematic of ion selection and fragmentation in a triple-quadrupole mass spectrometer, and the principle behind MRM analysis.

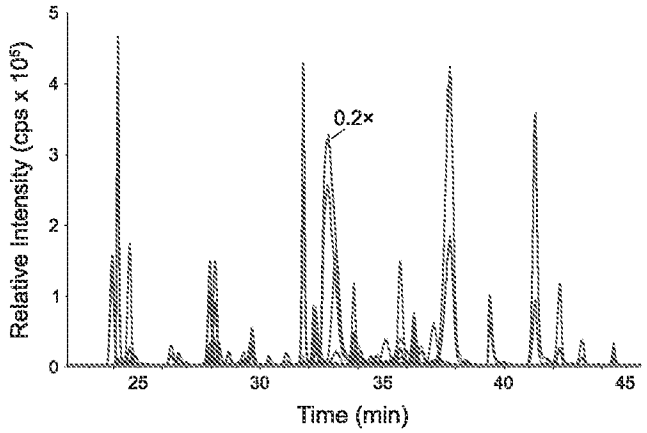



**Figure 3.** Extracted ion chromatograms (XICs) of MRM Q1/Q3 ion pairs for 45 endogenous peptides and 45 isotopically labeled standards analyzed in a single 60 min LC-MS/MS MRM analysis. Each blue peak represents the signal observed for a natural peptide, and the red peaks reflect the signal of the internal standard peptides.

These software packages include Skyline SRM/MRM builder,[22–24] AB's MRMpilot,[25] Agilent's Peptide Optimizer and Dynamic MRM software[26] and MaRiMba,[27] MRMaid[28] and MRMer.[29] There are also several libraries of MRM ion pairs currently being developed from literature data. These include new libraries of MRM pairs for target proteins (cancer biomarker peptide libraries,[19] yeast,[20] as well as other libraries generated from data in the literature or the Global Proteome Machine Database.[21]

Once the transitions have been selected, there are software packages designed to streamline the tuning process. PROMS was developed in 2001 for a PE Sciex API 2000 and automatically tunes the instrument to optimize the instrument conditions for each selected precursor/product ion pair.[30] Although this software was designed for small-molecule analysis, with each MRM optimization taking 2–3 min per compound, it should be possible to use this software for peptides as well. One instrument company (PE Sciex) has similar software for the ABI 2000 and ABI 3000 ion traps to select and optimize tuning conditions for MRM transitions (Automaton)[31] Waters Corporation's QuanOptimize software performs a similar function for the Micromass triple quadrupole.[32] We have found, however, that empirically determining the optimum instrument conditions can result in a factor of 10 increase in sensitivity compared to using literature-based settings.

*Advantages and disadvantages*

Because the internal standard is always present in the sample, its presence provides a useful confirmation that the instrument is functioning properly. The internal standard, because it is present in every sample, means that more accurate expression ratios can be calculated. It is not possible to calculate an accurate 'treated/control' expression ratio if there is no peptide detected in the control sample. With an internal standard, if no signal is detected in the sample, and the internal standard IS detected, then you can be certain that the level of that peptide in the sample is below the detection limit.

For an experiment with multiple treatments, the use of an internal standard ensures that all of the treatments can be compared with each other (e.g. treatment 1/control, treatment 2/control, allows the calculation of treatment1/treatment 2). In label-free quantitation, the control peptide may not always be present or may not be detected. The internal standard also corrects for suppression effects or irreproducibility in sample processing – at least when these problems occur after the point where the standard is added to the sample.

We have found that if isotopically labeled peptides are added prior to digestion, this sometimes produces elevated and unpredictable results. Therefore, stable isotope standard (SIS) peptides should be added post-digestion.[12] Because the internal standard

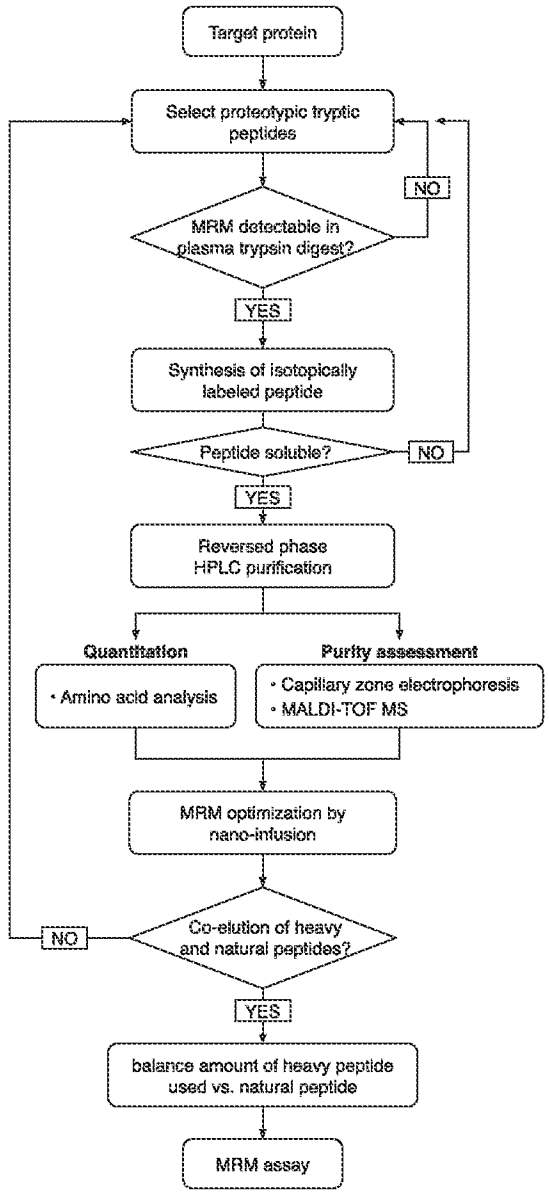



**Figure 4.** Schematic representation of selection process for MRM transitions. Reprinted from Ref. [12] with permission.

is added after the digestion step, this method of quantitation cannot correct for variable and unpredictable losses that occur during the sample digestion/processing steps that occur prior to the addition of the reference standards.[1] However despite this flaw, 'this highly accurate method of quantification has become the gold standard of MS-based proteomic quantification'.[1] Applications of this rapid and selective form of quantitative mass spectrometry include the analysis and validation of potential biomarkers in a large number of clinical samples or for measuring the levels of particular peptide modifications, such as ubiquitination.[11,33] A multi-laboratory study has recently been completed evaluating the reproducibility of MRM-based quantitation of proteins in plasma.[34]

MRM analysis with labeled internal standards is capable of sensitive (attomole) and absolute determination of peptide concentrations across a wide dynamic range of $10^3$–$10^4$.[11,34–39] One of the strengths of this quantitation method is the low coefficients of variation (CVs) which can be obtained. The amount of internal standard added for each peptide can be optimized to obtain low CVs (5–10%).[12] This amount will be different for different proteins, and is a function of the protein's individual relative abundances within the sample. This precision and accuracy makes this method attractive to researchers in clinical or pharmacological laboratories.

Because of the use of a labeled internal standard, ion suppression and matrix effects are less of a problem than in the label-free methods described later, because the standard and the native forms will be suppressed to the same extent. However, sample complexity can limit the sensitivity of the MRM assay for the detection of low-abundance proteins in the presence of high-abundance proteins. Using online HPLC coupled to a mass spectrometer further complicates the detection of low-abundance proteins, because an HPLC column can only be loaded with a fixed amount (typically 1 $\mu$g on 75 $\mu$ i.d. nanoflow column) of protein. Thus, the amount injected may not contain enough peptide from a low-abundance protein for it to be detected by the mass spectrometer. Plasma, for example, may contain more than 10 000 protein species,[40] which can have a dynamic range in concentration that spans 10 orders of magnitude, with 20 proteins making up 95% of the total protein mass.[41] The most abundant proteins in serum and plasma can be depleted using an immunoaffinity column, usually the 6, 12 or 14 most-abundant proteins, depending on the manufacturer of the column. Depletion suffers from two problems, however. The first problem is that the columns only remove 80–90% of the targeted proteins – it is possible, therefore, to *still* have $10^9$ excess of the abundant proteins. The second problem is that non-targeted proteins which bind to the targeted proteins can be removed as well. This can introduce an unwanted source of variation. The third problem, especially for tissue samples, is that the antibody columns recognize the proteins in their native conformations, while the protocols to extract the proteins from the sample may denature the proteins. To date, immunoaffinity depletion is unable to provide the sensitivity required to discover novel proteins found at the lowest levels in plasma.[42]

### MRM combined with immunoenrichment

An alternative to immunodepletion is to use antibodies to *enrich* the sample in target proteins or peptides. Two methodologies that have shown promise in improving the sensitivity of the MRM assay by enriching the target from a complex background are Stable Isotope Capture by Anti-Peptide antibodies (SISCAPA)[42] and immuno-MALDI (iMALDI).[43,44] Both methods use similar sample preparation steps, in that isotopically labeled peptides are added to the tryptic digest of the sample, followed by the addition of anti-peptide antibodies covalently bound to beads (magnetic or non-magnetic). The background, unbound peptides are washed away and the quantitation is performed by mass spectrometry.

*SISCAPA:* In SISCAPA, the sample is eluted from the beads prior to MRM analysis using electrospray ionization (ESI) on a triple-quadrupole mass spectrometer.[42] SISCAPA has been shown to increase the sensitivity of MRM by 3–4 orders of magnitude.[45]

*iMALDI:* In iMALDI, the beads with the affinity-bound peptides still attached are placed directly on a MALDI mass spectrometer, and the MALDI matrix solvent elutes the peptides from the beads. The presence of the peptide, and its peak height or peak area, is then determined from an MS spectrum.[43,44] In iMALDI, the peptides are *confirmed* with MS/MS, but the quantitation is

normally performed in the MS mode. (SISCAPA, in contrast, is based on MRM.) This means that, in principle, you could do iMALDI even with only a MALDI-MS instrument. iMALDI can also be used in the MRM mode on MALDI-MS/MS mode – this is called iMALDI$^{+}$.

*Advantages and disadvantages*

An advantage of SISCAPA over iMALDI is that the dynamic range of a triple quadrupole mass spectrometer is higher than that of a MALDI-TOF/TOF. In addition, a 'screening' experiment for determining sensitive peptides (i.e. potential SISCAPA target peptides) from a complex mixture can be carried out with LC/ESI-MS/MS. This screening could, of course, be performed by LC-MALDI for potential iMALDI target peptides, but LC-MALDI is still less commonly available than LC-ESI/MS/MS. An advantage of iMALDI is that it requires less sample handling and is more robust and faster (because SISCAPA involves the use of LC). Because SISCAPA is LC-MS based, it is also susceptible to cross-contamination during the analysis; iMALDI is not.

iMALDI is also somewhat less targeted than SISCAPA. We have recently demonstrated that iMALDI is able to detect variants of the target peptide which can still be captured by the antibody.[46] However, the design of a few MALDI instruments (e.g. the AB 4700 MALDI-TOF/TOF) makes them more susceptible to possible contamination by dislodged beads, so this should be checked with the manufacturer before performing iMALDI experiments.

## Relative Quantitation

Relative quantitation is based on the introduction of a chemically equivalent differential mass tag that allows the comparative quantitation of proteins in one sample to another. The labels change the mass of a protein or peptide without affecting the analytical or biochemical properties.[47] Differential isotopic labels can be introduced metabolically, enzymatically or chemically and – depending on the method used – at either the peptide or the protein level. A strategy involving reversing the labeled and unlabeled sample to assist in detection of changes between treatments was developed by Wang *et al.* for $^{18}$O labeling,[48] and then was extended to other labeling techniques, including $^{15}$N labeling[49] and iCAT.[50]

### Metabolic labeling

Metabolic *or in vivo* labeling involves the incorporation of stable isotopes during protein biosynthesis. During cell culture, cells are provided with media containing isotopically labeled amino acids which will be incorporated into the proteome during normal cell growth and division. The advantage is that samples grown with different labeled amino acids can be pooled prior to the sample preparation and analysis steps. This technique, however, is not applicable to all sample types but is limited to situations where the cells are metabolically active. As a result, metabolic labeling cannot be used on tissue samples or biofluids.[47] Metabolic labeling has been used for relative protein expression studies since at least the early 1980s, but this work usually involved $^{14}$C, $^{3}$H, $^{32}$P or $^{35}$S with separation by isoelectric focusing (IEF) and polyacrylamide gel electrophoresis (PAGE) gels, and quantitation by autoradiography or radioactivity detection. An early application of stable isotopic labeling with mass spectrometric detection for protein analysis was done by Raap *et al.*, in 1990, using GC/MS detection.[51]

#### $^{15}$N

The first type of metabolic labeling applied to MS-based proteomic analysis was $^{15}$N labeling. Metabolic labeling of yeast using $^{15}$N-labeled growth medium, followed by in-gel digestion and LC-ESI-MS/MS identification of the proteins and determination of the ratios was performed by the Chait group in 1999.[52] This study involved baker's yeast grown in medium that was more than 96% enriched in $^{15}$N, which results in the replacement of each $^{14}$N with $^{15}$N. In 2000, metabolic $^{15}$N labeling of yeast was done by the Smith group using LC-FTICR,[53] with peptide identification and expression ratios determined by using the accurate molecular weight's (mw) of the peptides as 'accurate mass tags'.[54]

*Advantages and disadvantages*

$^{15}$N labeling can be a good choice for autotrophic organisms such as plants and bacteria. Because these organisms can synthesize their own amino acids, the incorporation of isotopically labeled amino acids from the medium (as in Stable Isotope Labeling by Amino Acids [SILAC]) would not be complete. In *Arabidopsis thaliana*, 75% incorporation was reported using isotopically labeled amino acids, whereas >95% incorporation was reported for $^{15}$N.[55] Therefore, $^{15}$N labeling is often used for autotrophic organisms. One drawback to this approach is that only two samples can be compared within a single experiment.

This method of quantitation is complicated by the fact that the number of replaced nitrogen atoms can vary from peptide to peptide making the exact mass shift unpredictable.[56] In addition, finding software that can deal with this may also prove to be difficult.[47] Mascot (through the Mascot Distiller software package) has recently included the searching and quantitation of $^{15}$N-labeled samples as one of their supported quantitation methods.[57]

Media enriched with $^{15}$N has been used for quantitative proteomics in yeast,[52] *Caenorhabditis elegans* and *melanogaster*,[58] and *Arabidopsis thaliana*.[59] Yates and collaborators also showed that metabolic labeling of an entire organism was possible in mammals (in this case, a rat[60]), but this is usually prohibitively expensive.[49]

#### SILAC

There appear to be many patents for various aspects of SILAC, but the earliest patent for the method using isotopically labeled amino acids seems to be held by Franza and Rochon, and was filed in 2001 (Ref. [30] issued 2003). In any event, this method was popularized by Mann and co-workers in 2002.[61] The SILAC technique relies on the incorporation of isotopically labeled amino acids into proteins formed by the growing organism. Isotopically labeled amino acids are usually added to the growth medium,[61,62] or the labeled amino acids can be generated by the organism through the addition of isotopically labeled *salts* to the growth medium.[52] Lysine and arginine are the two most commonly used labeled amino acids which are added to the medium, because each tryptic peptide from proteins synthesized from these amino acids will contain an isotopically labeled lysine or arginine, which increases the quantitative coverage of the experiment.

Arginine, however, can be converted into proline during cell division, leading to potential inaccuracies and complicating the final quantitation. One strategy has been to reduce the concentration of arginine in the media to a level where conversion to proline would be unfavorable.[63] However, although this does

act to reduce the conversion, it does not completely prevent it.[64] Lajoie's group has recently reported that arginine to proline conversion can be eliminated by adding sufficient amounts of proline to the media as well as ensuring that sufficient arginine is present to prevent the back-conversion of proline to arginine.[65]

*Advantages and disadvantages*

Unlike $^{15}N$, SILAC allows for more comparisons within a single experiment due to the availability of several labels. In addition, SILAC has the advantage of a predictable mass shift. However, complete incorporation of isotopic amino acids is not the same for all cell lines.[66] SILAC cannot be used on cell types that are unable to incorporate certain amino acids. For example, SILAC works well for mammalian cell lines because of their inability to synthesize all of their amino acids, but does not work as well for plants, due to their autotrophic nature. In addition to this, some cells are harder to grow in the dialyzed serum required for SILAC due to the loss of essential growth factors.[1]

The isotope distributions resulting from the combination of $^{13}C$ and $^{15}N$ can be difficult to interpret. However, software for calculating theoretical isotope distributions is available to assist in the interpretation of the complex isotope distributions that can be produced.[67,68]

### Proteolytic labeling ($^{18}O$)

$^{18}O$ is a technique where proteolytic labeling and stable isotope incorporation occurs simultaneously during digestion.[69,70] Samples are digested, usually with trypsin, in the presence of $H_2{}^{18}O$ resulting in a 2–4 Da mass shift from the incorporation of one or two $^{18}O$ atoms on the carboxy terminus of each peptide. The presence of the label on the carboxyl termini of peptides is advantageous because it facilitates the assignment of 'y' ions in the spectra.[56] In order to prevent contamination with $^{16}O$ water, samples must be completely dry prior to labeling.

*Advantages and disadvantages*

Labeled and unlabeled samples are combined post-digestion, so protein losses that occur during sample preparation would not be compensated for and would affect experiment reproducibility. Another problem with sample pooling is that trypsin-mediated back exchange (the replacement of $^{18}O$ with $^{16}O$), can occur in solvents containing natural water.[71] The use of immobilized trypsin may prevent this type of back exchange because the enzyme would no longer be present at the time of pooling.[72] There are also only two forms of this label.

One, two, or even three carboxyl oxygens can be incorporated (Fig. 5), and this can lead to variability in the quantitation. If one oxygen is exchanged, the mass shift is only 2 Da, and a mass difference of less than 4 Da can complicate quantitation, especially when low resolution mass spectrometers are used, because there may be an overlap of isotopic envelopes,[47,73,74] although software is under development to help address this problem.[75] Another drawback is that different peptides incorporate the label at different rates.[69,72] For example, the exchange seems to be slow and incomplete for highly acidic peptides.[71] Immobilized trypsin may increase the efficiency of this exchange due to a higher concentration of enzyme at the time of labeling, although complete labeling is rarely achieved.[76]

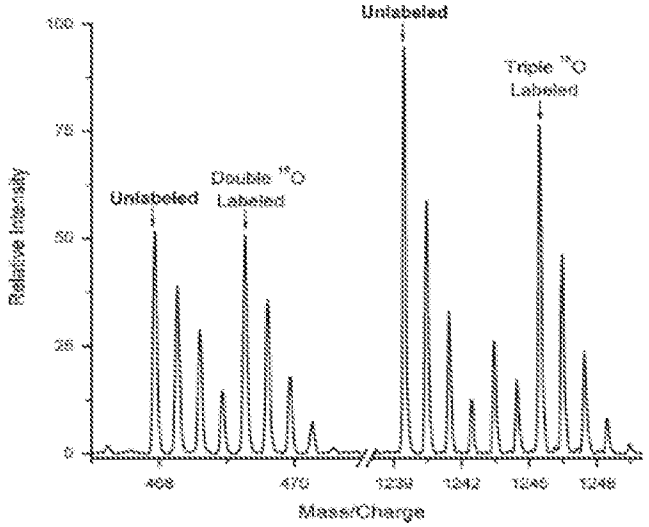



**Figure 5.** Spectrum of $^{18}O$ double- and triple-labeled peptides. Reprinted from Ref. [72] with permission.

## Chemical Labeling

Most protein-based labeling strategies are hindered by the conditions needed to solubilize proteins which may interfere with the labeling process. However, guanidine can be used as a denaturant, which should solubilize most proteins prior to labeling. Because labeling is done prior to digestion, variations in protein losses are also minimized; however, incomplete denaturation could lead to errors in quantitation.

### ICPL

The isotope-coded protein label (ICPL) technique is based on isotopic labeling of all of the free amino groups on the proteins. N-terminal nicotinylation was first introduced by Peter James in 2000,[77] who developed a method for protein quantitation by labeling the N-terminus of peptides using $D_4$ or $H_4$ forms of Nicotinyl-N-Hydroxysuccinimide (Nic-NHS). $H_4$ or $D_4$-labeled N-nicotinoyloxy-succinimide was used to label free amine groups on peptides for relative quantitation by Lottspeich and co-workers (Fig. 6),[78] and produced a mass difference of 4 Da per label. The ICPL reagent is now marketed by Bruker Daltonics. The original Bruker reagent produced a mass difference of 6 Da; in December 2008, Bruker introduced a 4-plex version of this reagent, the Serva ICPL 4plex Kit. This means that there are actually 4 closely related labels: (1) all natural isotopes, (2) X = D, shift of 4 Da to relative to all natural isotopes, (3) X = H, $6x^{13}C$, shift of 6 Da relative to natural, (4) X = D, $6x^{13}C$, shift of 10 Da relative to all natural isotopes.

### $H_3/D_3$ Acetylation or $H_5/D_5$ propionylation

This is probably the oldest and least expensive method of differentially labeling two sets of proteins. It is similar in concept to the ICPL technique, except that the proteins are labeled with $H_6$- or $D_6$-acetic anhydride or $H_{10}$- or $D_{10}$-propionic anhydride.[79] If the proteins are not denatured prior to labeling, this technique can be used to probe surface-accessible residues.[80,81] An example of a comparison of two samples differentially labeled with $H_{10}/D_{10}$ propionic anhydride is shown below (Fig. 7). Difficulties can arise due to protein precipitation, and only two treatments can be

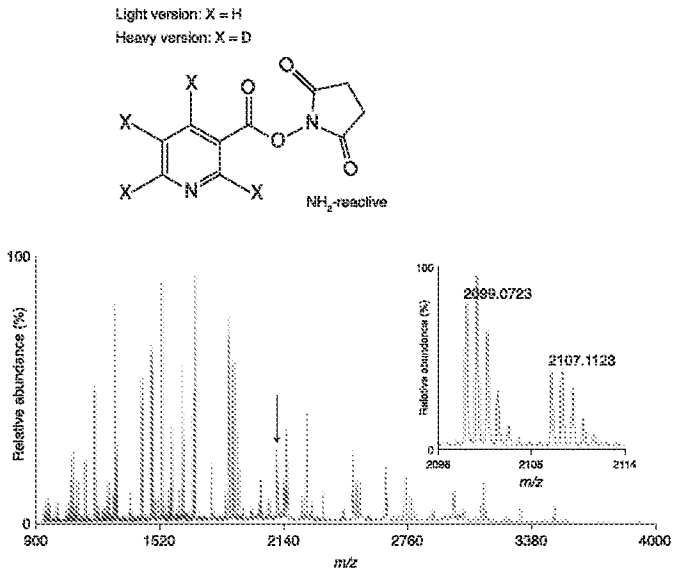



**Figure 6.** Structure and spectra resulting from the Lottespeich ICPL label. Reprinted from Ref. [78] with permission.

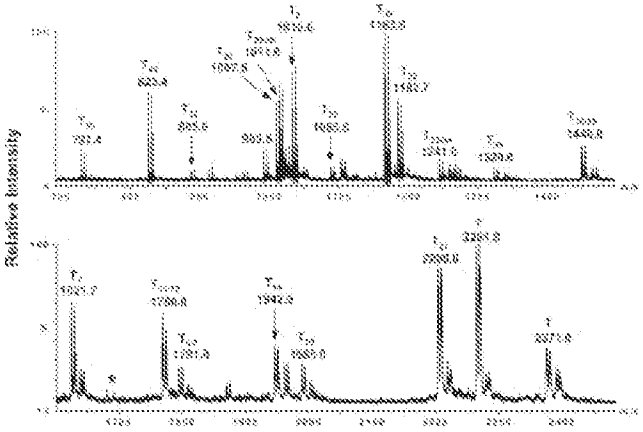



**Figure 7.** Example of 5H/5D-labeled MAPK peptides. Reprinted from Ref. [79] with permission.

compared. For complex samples, such as plasma, the resolution of an FTICR-MS may be needed to separate overlapping isotope clusters (Fig. 8).

## ICAT

ICAT (Isotope Coded Affinity Tag), introduced by Gygi and Aebersold in 1999, is one of the earliest chemical reagents introduced for quantitative proteomics.[82] The original ICAT label included a thiol-specific reactive group (iodoacetamide), biotin, and either 0 or 8 deuterium atoms, leading to a difference in mw of 8 Da between the two different forms of the tag (Fig. 9). The biotin group allowed for avidin affinity purification before the proteins are eluted. A new cleavable version of this reagent was developed in 2003.[83,84] This cICAT reagent contained nine $^{13}C$ instead of eight deuteriums, and an acid-cleavable biotin moiety. This new form removed the potential confusion of a double ICAT label and oxidation, both of which would have led to a +16 Da mass shift. Deuterium causes a slight shift in retention time in reversed-phase HPLC, with the heavy form eluting slightly earlier than the light form.[85] The use of the $^{13}C$ stable label removed this retention time shift. Cleavage of the biotin moiety after affinity purification and before the MS analysis also improved the quality of the CID spectra,[84] leading to the identification of a larger number of proteins.

Proteins are labeled using a thiol-specific reagent (iodoacetamide) containing a biotin moiety (Fig. 9). The biotin group allows for affinity purification with avidin or streptavidin beads following tryptic digestion. Labeling is done prior to digestion decreasing the chance of differential protein losses and therefore, increasing the chance of experiment reproducibility. However, high amounts of detergent interfere with the labeling process, so 6 M urea is used to solubilize samples. However, this may not be sufficient for insoluble samples, such as membrane proteins, and also increases the chances of protein carbamylation.

One advantage of ICAT is the reduction in sample complexity because the label specifically targets cysteines, a relatively rare amino acid making up only 1.42% of all amino acids.[71] However, this simultaneously reduces the reliability of the quantification as the experiment is based on a limited number of peptides per protein. It also makes it impossible to detect changes in the ~10–13% of proteins that do not contain cysteine residues.[50,86,87]

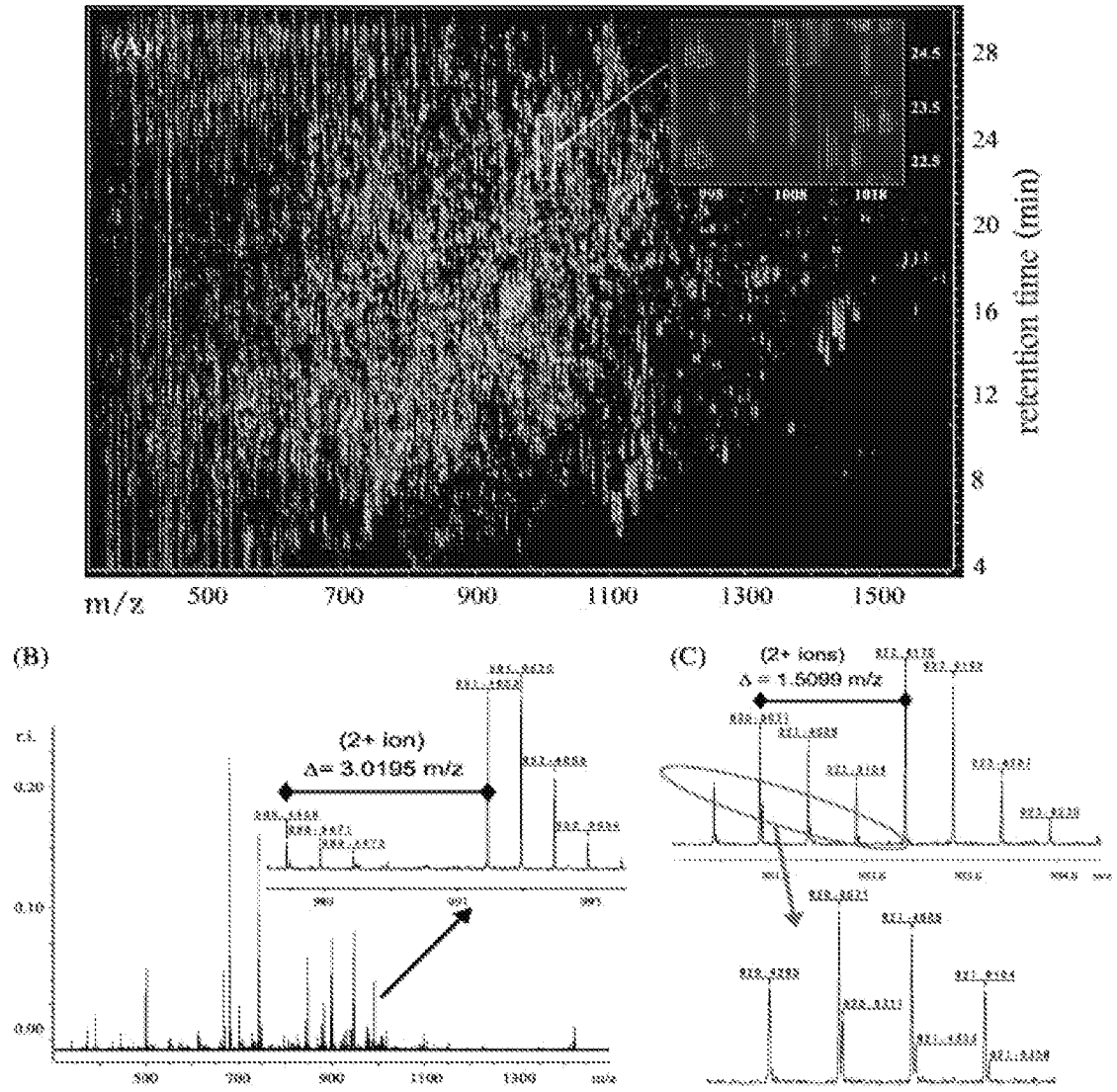



**Figure 8.** Relative quantitation of protein expression in human plasma using $H_6$ *versus* $D_6$ acetic anhydride. (A) Heat map of the LC/MS analysis (the inset shows doublets due to the differential labeling), (B) FTICR MS spectra demonstrating the resolution required for accurate determination of the overlapping isotopically labeled doublets.

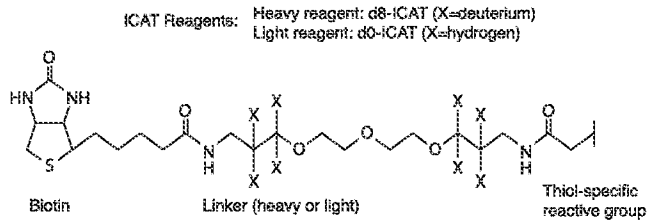



**Figure 9.** Chemical structure of the original ICAT label. Reprinted from Ref. [82] with permission.

### mTRAQ

This is a new amine-reactive isotope-coded reagent recently introduced by Applied Biosystems. It comes in three forms, and is not an isobaric tag like iTRAQ (see below). It is designed to be used after the biomarker discovery stage, during the validation stage of biomarker project as an alternative to having to synthesize deuterated standard peptides for MRM-based quantitation.[88] The precursor-ion mass shifts of these $^{13}C$ and $^{15}N$-labeled reagents are 0, 4 and 8 Da for arginine and 0, 8 and 16 Da for lysine C-terminated peptides (Fig. 10). The reagent is designed to be used in MRM assays, using a 'Global Internal Standard' methodology, where relative expression ratios relative to a 'pooled' sample are obtained. Thus, this is a *relative* quantitation method, in contrast to the standard MRM assay based on known amounts of isotopically labeled internal standard peptides which is an *absolute* quantitation method. However, the use of this label is allows the multiplexing of MRM assays – 'the triplex [version of the label] will double sample throughput'.[89] The two-label form of this new mTRAQ technique has already been successfully applied by the Siu laboratory to the quantitation of endogenous levels of a potential cancer biomarker in endometrial tissues.[90] Interestingly, in this study, the expression ratios found by mTRAQ were significantly higher than those found in the iTRAQ-based discovery phase. This provides more evidence of the 'compression' of iTRAQ-determined expression ratios.

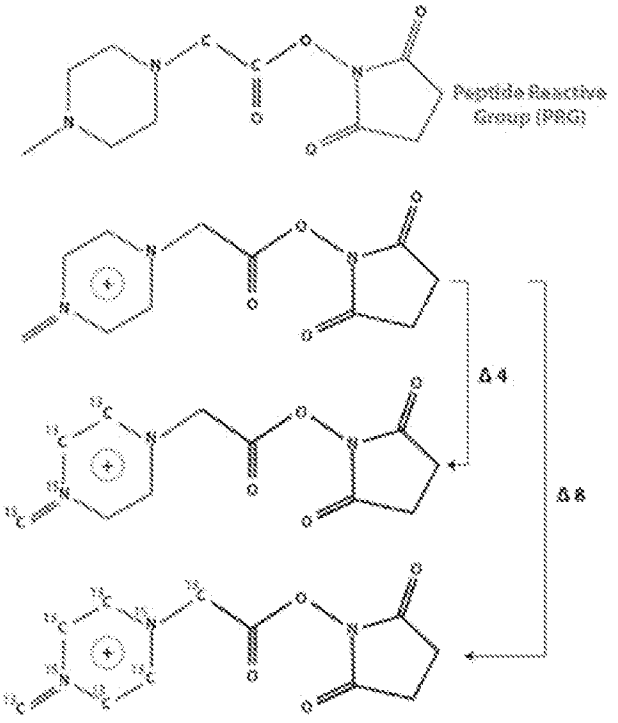



**Figure 10.** Structures of the mTRAQ reagents.

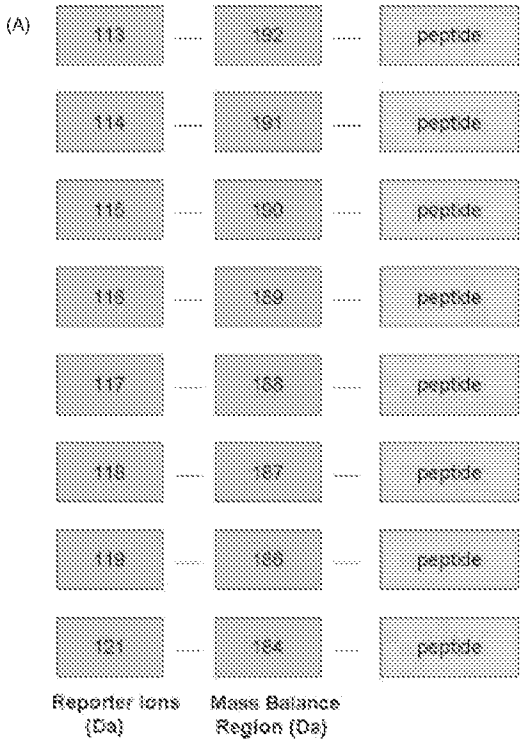



**Figure 11.** (a) Schematic representation of the mass-balanced iTRAQ 8-plex reagent, (b) an MS/MS spectrum showing the reporter region of an iTRAQ-labeled peptide.

### iTRAQ (isobaric tags for relative and absolute quantification)

One limitation of ICAT is that there are only two labels available. This could result in multiple experiments if more than two treatments need to be compared, and would increase the cost accordingly. The need for comparisons of larger numbers of treatments led to the development of the 2- or 4-plex ICPL, the 4- or 8-plex iTRAQ, and the 2- or 6-plex Tandem Mass Tag (TMT) labeling techniques, which can compare up to four, eight, or six samples in a single analysis, respectively. Unlike the previously described labeling techniques, which use the *parent ion* peak heights or peak areas from the MS spectra, the iTRAQ labels from AB and the TMT labels from Thermo Fisher[91] are currently the *only* tagging technologies commercially available where quantitation is carried out in the MS/MS mode. A similar reagent, also relying on reporter ion abundances in the MS/MS spectrum was PerkinElmer's Exactag reagent, which allowed up to ten multiplexed assays in a single analysis.[92] At this time, however, this reagent is not being marketed.

The iTRAQ technique was first described by Ross *et al.* in 2004,[93] and was subsequently commercialized by Applied Biosystems. Since then, iTRAQ has been used to examine a variety of cell systems and biological fluids and to analyze proteins expression changes related to cancer, Alzheimer's disease and other diseases.

The iTRAQ label is an isobaric tagging compound consisting of a reporter group (variable mass of 114–117 Da or 113–121 Da), a balance group and an amino-reactive group that introduces a highly basic group at lysine side chains and at peptide N-termini (Fig. 11). During the initial MS scan, labeled peptides appear as a single peak due to the isobaric masses. The isobaric nature of iTRAQ-labeled peptides allows the signal from all peptides to be summed in both MS and MS/MS modes thus enhancing the sensitivity of detection. During MS/MS, the label releases the reporter group as a singly charged ion of masses at *m/z* 114–117 (4-plex) or *m/z* 113, 114, 115, 116, 117, 118, 119 and 121 (8-plex). iTRAQ can be used to analyze up to four different samples using the 4-plex kit, or up to eight samples using the new 8-plex kit. The TMTs are based on a similar principle, with up to six possible labels (Fig. 12). These isobaric tagging techniques can also be manipulated to perform absolute quantitation by adding SIS peptides. This then involves comparing peptides from a target protein to a known amount of labeled standard peptide spiked into the sample.[74]

#### *Advantages and disadvantages*

In iTRAQ and similar labels, the labeling is done at the peptide level, and, because every tryptic peptide should be labeled, multiple peptides can be detected for the same protein, thus giving multiple quantitation measurements per protein and increasing the confidence of protein identification. This potential benefit – to identify and quantify low-abundance proteins in complex samples – coupled with the ability to multiplex up to eight samples in parallel suggests that iTRAQ and similar mass-balanced labels holds the most promise for quantitative biomarker discovery.[76] Initially, only AB's ProteinPilot software could be

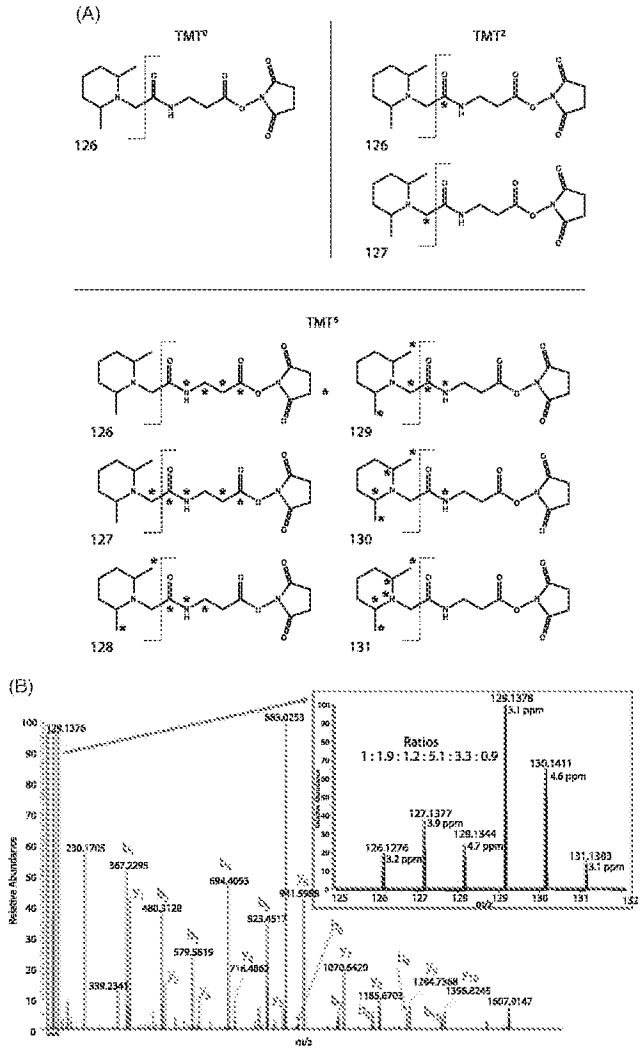



**Figure 12.** (a) Structures of the Thermo Fisher TMT family of labels. (b) MS/MS spectrum of TMT-labeled peptide, showing reporter region. Reprinted from Ref. [91] with permission from Thermo Fisher.

used to analyze the data from iTRAQ-labeled samples. Now, other commercially available software packages (Mascot,[94] Thermo Fisher's Protein Discoverer, etc.) can be used to analyze iTRAQ and TMT data, which opens up this technique to analyses performed on other manufacturers' instruments.

Disadvantages of the technique are the variability in labeling efficiencies, the variability in the initial protein digestion and the expense of the reagents. This is not an easy technique – these isobaric-tagging protocols involve multiple steps. In our laboratory, more than 40 standard operating protocols (SOPs) had to be developed in order to achieve reproducible and reliable results with iTRAQ.[95] Current studies also indicate that there is an inherent dynamic range limitation in iTRAQ results, so that the observed expression ratios are compressed.[96,97] The maximum expression ratios that can be observed in the discovery experiments using iTRAQ seem to be only approximately 3–5, whereas larger differences are observed in the validation experiments using mTRAQ or isotopically labeled standard peptides.[12] This most likely affects the other mass-balanced labeling methods as well.

Because of the expense of these labeling reagents, as well as the difficulties of the labeling protocols, these techniques are better suited for biomarker discovery rather than for biomarker verification or validation, or for clinical studies.

## Label-Free Quantitation

Because of the expense and effort required for use of some of the labeling techniques, there now seems to a trend away from isotopic labeling, toward methods of label-free quantitation. Several types of label-free quantitation methods are currently in use.

Label-free quantitation is exciting because it holds the promise of 'shotgun' quantitation. Label-free quantitation seems at first to be 'easy', requiring no labeling steps, only 'standard' LC/MS or LC/MS/MS as is typically done for the identification of mixtures of proteins. It has the advantage of eliminating the need for costly labeling reagents, and does not require the multi-step labeling protocols which can lead to irreproducibility and loss of target peptides. It also eliminates the need for the synthesis of expensive isotopically labeled reference peptides.

However, on closer examination, we find that label-free quantitation is not so easy. There is a greater need for reducing potential interferences that could lead to suppression effects – effects that could have been compensated for by using isotopically labeled peptide standards. This leads to an increased need for high-resolution or multidimensional chromatography to reduce suppression effects and to allow the detection of low-abundance proteins. There is a useful empirically derived determination of the number of MudPIT fractions needed, depending on the expected number of proteins in the sample.[98] However, sometimes peptides do not behave as anticipated, for example in a study by Old *et al.*, where 90% of the peptides eluted in a single SCX fraction.[99]

In addition, there is the challenge of normalizing the data so that accurate quantitation can be done across multiple samples and multiple analyses. Two approaches are being used to compare samples: normalizing and aligning the *chromatograms* (and using the peptide MS data), or using the MS and/or MS/MS *spectra* to normalize the data.

## Normalization based on chromatograms and retention time

One problem in normalizing based on chromatograms that may affect quantitation more than identification based on MudPIT data is need for high reproducibility in the fractionation (i.e. the first dimension in a MudPIT experiment) as well as high reproducibility in the online chromatography. In a standard MudPIT analysis, the data from all of the fractions are usually combined and searched together, and even then, there have been questions about the reproducibility in MudPIT separations for protein identification.[100] Quantitation based on alignment of chromatograms could potentially be even more adversely affected if a peptide is 'split' between fractions, because the fractions will be aligned separately. This problem has been noted by Wolters *et al.*,[101] and by Old *et al.* who reported that variability in SCX fractionation may result in peptides not always eluting in the same SCX fraction in replicate analyses.[99]

Even in a 1D LC/MS experiment, it has been reported that only 66% of peptides which were present in one analysis were also present in the second LC/MS run, and that ten replicate analyses were needed before no new peptides were detected.[102] A similar

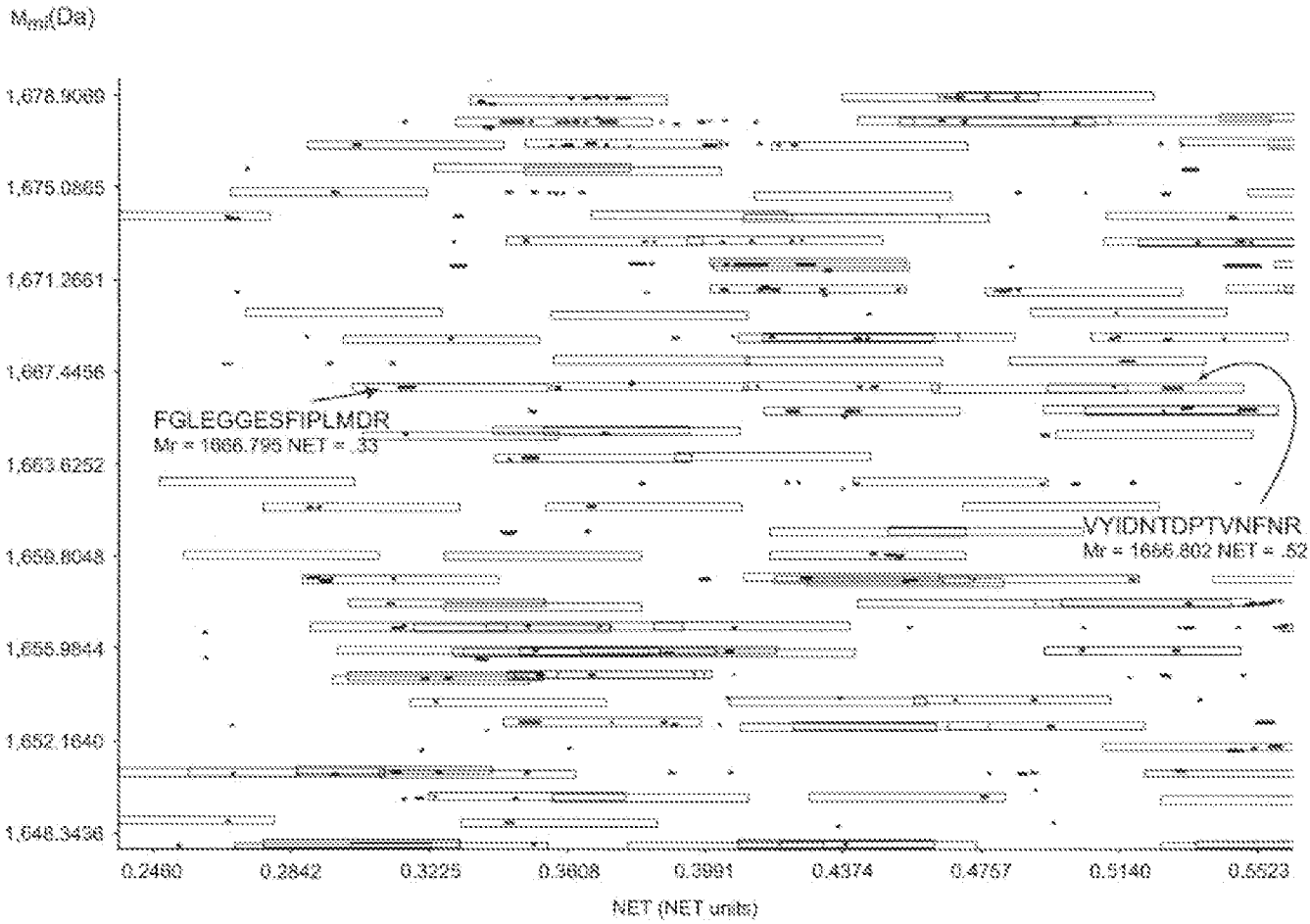



**Figure 13.** Example of the normalized elution time (NET) window approach for aligning and assigning peptides, based on data in the AMT database. The boxes show where a peptide is expected to elute, assuming a 100 ppm mass accuracy. Reprinted from Ref. [115] with permission.

number was found in a study that compared the number of reacquired and non-reacquired MS/MS spectra (i.e. the number of MS/MS spectra that were found *again*) when the analysis was repeated.[99] In this study, twice as many reacquired spectra were obtained in replicate LC/MS/MS-based MudPIT experiments than non-reacquired MS/MS spectra.[99] Unfortunately, a large number of replicate analyses is too costly in terms of expense and time (and even of sample) to be done for most projects. However, the ProQuant software package was specifically designed to address the problem of 'missing peptides'.[103] In many software packages, including ABI's ProteinPilot and GE Healthcare's DecyderMS, and Thermo Fisher's Proteome Discoverer, the user can select the number of replicates where a peptide has to appear in order for it to be considered significant, and can manually decide whether to include a particular peptide in the quantitation results.

Several software packages are available for normalizing and aligning the chromatograms,[104–107] and can be used for stable isotope labeled experiments as well as label-free quantitation. A table showing more than 20 software packages for chromatographic alignment is given in a recent review of label-free quantitation.[108] Two of these are the commercially available DecyderMS from GE Healthcare, and the accurate mass tag (AMT) approach, which are described below. Label-free quantitation has also led to a renewed interest in LC/FTICR-MS methods where the high mass-accuracy and high resolution can be used to help ensure that one is comparing the same peptide in different samples.[105,109–113]

### AMT

The original AMT was developed in 2000 for protein identification and metabolic labeling studies, taking advantage of the high resolution of the FTICR.[54] By 2003, this had evolved into the AMT ['accurate mass and (retention) time'] tag method, and was being applied, not only for label-free quantitation on FT-MS instruments, but also for a study of microbial digests using a Micromass Q-tof with 10 ppm mass accuracy. Figure 13 shows the expected elution time window, normalized elution time (NET[114]), for the peptides.[115] A software package, called Visual Inspection of Peak/Elution Relationships (VIPER) is available as a free download, and can be used to assist in managing large amounts of accurate mass and retention time data from high-throughput proteomics experiments.[116] An interesting approach being the creation of a set of $^{18}$O-labeled reference peptides for plasma proteins for use with LC/FT-MS.[117] A combination of the AMT approach, $^{18}$O-labeling, and VIPER software was used to quantitate 312 proteins from 18 depleted plasma samples, where an $^{18}$O-labeled reference digest was added to each sample to provide a set of labeled internal standards.[118]

### DecyderMS

This software is marketed by GE Healthcare, and turns an LC/MS/MS analysis into a virtual 2D gel, with retention time on the *x*-axis, *m/z*

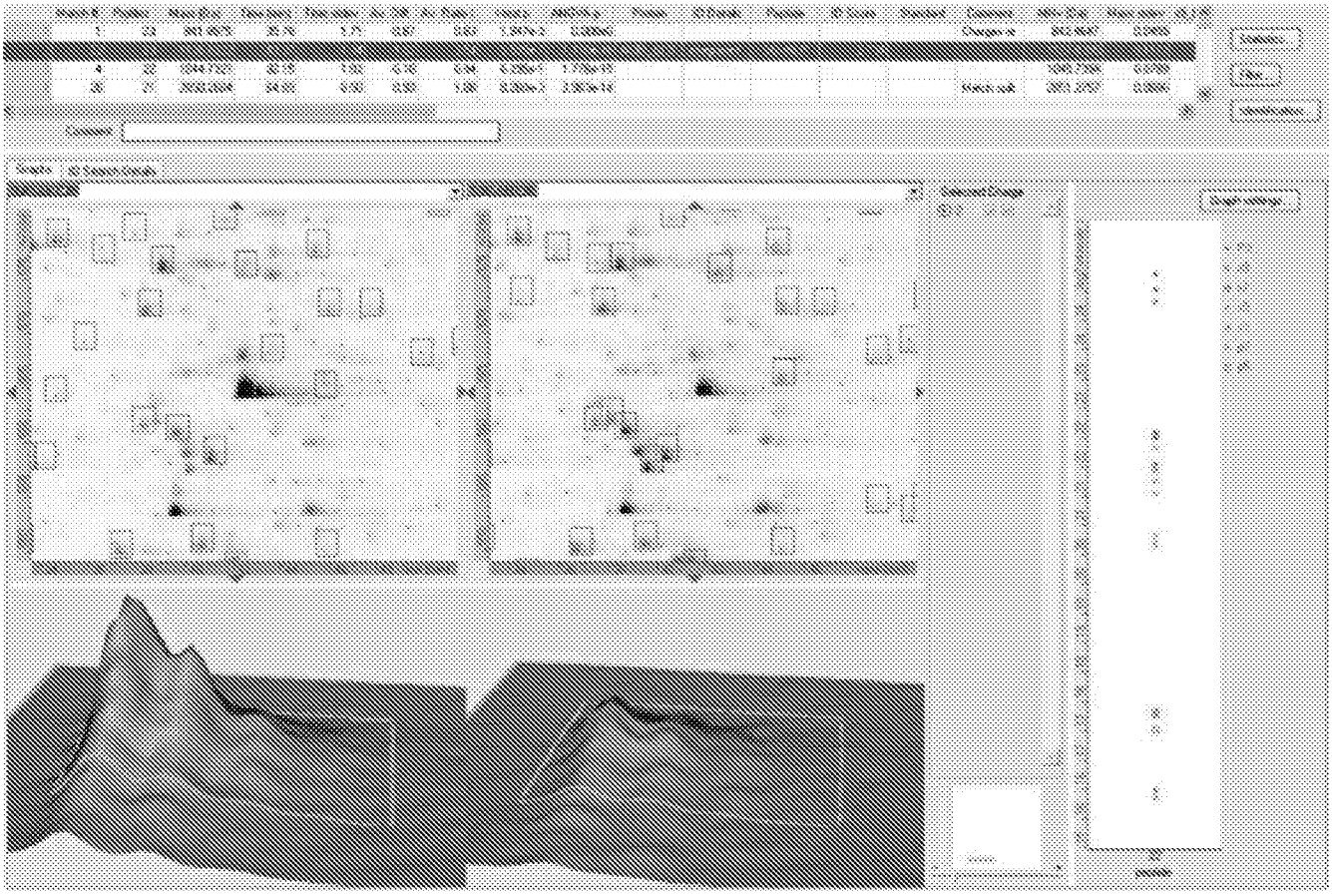


**Figure 14.** A screen-capture of the PepMatch module of DeCyder MS. Reprinted from Ref. [119], with permission.

on the *y* axis, and abundance on the *z*-axis (Fig. 14). The software allows the alignment of pairs of LC/MS/MS chromatograms, and determines the expression ratios of the two samples by comparing the peak volumes of the corresponding 'spots'. The DecyderMS output shows the ratios of the peptides or proteins detected. The software exports the data to database search engines and imports the results to give relative peptide and protein abundance ratios of identified proteins (and unidentified peptides).

*Advantages and disadvantages.* There is no question that label-free quantitation is easier and less expensive than using the isotopic labeling techniques. It also allows the comparison of an unlimited number of experiments and can be used on any type of biological material, and does not appear to have artificial limits on expression ratios.

One disadvantage is that label-free approaches do not allow for sample multiplexing – each sample has to be analyzed individually. Moreover, there are still controversies on the reliability of label-free quantitative analysis.[74] In particular, there are concerns over the possibility of suppression effects which might be different between the samples being compared. We have observed that if the biological matrix is very similar – for example, for tissues from inbred mice – then the 'noise' is very similar and quantitation is quite accurate. However, suppression effects can be a serious problem if you are analyzing human plasma samples, and might adversely affect the results. For this reason, label-free techniques are generally considered inferior in their quantification accuracy when compared to methods relying on stable isotopes, and the accuracy and linearity of label-free techniques is still in question.[76]

## Normalization based only on spectra (no retention time considerations)

It seems intuitively obvious that the more abundant proteins will produce a larger number of spectra. This should be reflected both in increased sequence coverage and in the number redundant scans per peptide. Certainly, the extreme case is true: if there is no protein in the sample, then one should not be able to detect any peptides from that protein. The correlation of abundance with the number of spectra observed, the % sequence coverage, and the number of peptides identified per protein was experimentally validated by the Yates group in 2004.[98]

Several commonly used methods of quantitation based on features in the spectra are described below:

### Empai score

This method to determine the protein abundance was originally developed as the protein abundance index (PAI) method.[120] It is based on the number of peptides observed from a particular protein, compared to the number of peptides that it would be possible to observe, taking into account the mass range of the instrument, and the hydrophobicities of the peptides. Peptides too large, too small, too hydrophobic, or too hydrophilic are excluded.





When it was observed that the PAI score varied as a function of concentration in LC/MS/MS experiment, the Exponentially Modified Protein Abundance Index (emPAI) method for absolute quantitation was developed by converting the PAI value to exponentially modified PAI (emPAI), equal to $10^{PAI}$ minus one, and by determining the total amount of protein in the sample.[121] In this report, the emPAI score was found to be accurate within 63% of the results from protein staining. The emPAI score is reported as part of the standard Mascot search result.

### Peptide ion intensity counting

In this technique, the peak height or the area of a peak at a selected mass to charge ratio is obtained by counting the number of ions. For a complex digested mixture, peptides are separated by a chromatographic gradient followed with a mass spec analysis.[74] One disadvantage to the peptide ion intensity method is that the counting of the peptide ion chromatograms is a tedious process and, in the absence of adequate internal controls, it is usually regarded as unreliable, especially for small molecules.[56]

### Ion accounting

This technique is based on the observation that 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%.[122] This means that intensities of the three most sensitive peptides from each protein can be used as a measure of the protein abundance. A universal signal factor (counts/mol) was determined to be the same for all proteins within a factor of 15%, and in a dilution study, the linearity of the peptide response *vs.* concentration curve was linear with $R^2 = 0.99$. Quantitation software based on this algorithm is marketed by Waters Corporation for use on $MS^E$ instruments, as part of their ProteinLynx Global Server software package.

### Spectral counting method

This label-free quantitation method is based on counting the number of MS/MS obtained from a specific protein. It assumes that the more abundant the protein, the more abundant the peptide, and the more likely it will be selected for MS/MS analysis.[74] Spectral counting has been shown to have a linear dynamic range of 2 orders of magnitude in yeast.[98]

One disadvantage to spectral counting is that high-abundance proteins can mask low-abundance proteins,[74] and it assumes that the linearity of response is the same for every protein. Saturation of the detector can occur at higher spectral counts, and at different levels for different proteins, which may lead to potential problems with dynamic range.[76] Current software packages featuring spectral counting-based quantitation are Scaffold and ProteoIQ. Other software packages that have been developed to determine the *significance* of spectral-counting data include QSpec[123] and the Spectral Index.[124]

### Spectra TIC method

In this method, the TICs of the MS/MS spectra from each protein are averaged and used for quantitation.[125] The Thermo Fisher software, for example, can perform this summation for each protein. The average TIC for each peptide MS/MS spectrum from a particular protein is then calculated. Using the average, rather than the sum of the TIC values, helps normalize for proteins of different sizes. The authors report that this method suffers less from 'compression' than the spectral counting method.

### Absolute protein expression (APEX)

APEX is an open-source software program that uses a 'learning machine' approach, and combines some of the features of spectral counting and EmPAI in that the observed peptide counts are modified by the expectation of observing each peptide in the experiment (which includes the number of unique peptides in the protein), the total sampling depth and the confidence in the protein identification.[126,127] In a study on yeast and *Escherichia coli*, the dynamic range was 2–4 orders of magnitude, and the results were found to agree with Western blot, and 2D gel studies.[127] In a more recent study on *Shigella*, the APEX modified spectral counting method was found to have a low false-positive rate (%5), and a linear dynamic range 10x wider than density-based Coomassie 2D gel quantitation.[128]

### Other methods

The peptide ion current area (PICA) method of label-free quantitation is based on the area under reconstructed ion chromatograms. This method was able to detect spiked-in standard proteins at a 90% confidence level, and the PICA analysis results agreed with ICAT and spectral counting on a mixture of known proteins of varying concentrations with a mean squared error of 0.09.[129] The Serac software,[99] developed by the Ahn group, is designed to normalize and background subtract the peak intensity and spectral count data, as well as to compensate for cases where the spectral count is very low or zero in one of the samples being compared. The ProteinQuant Suite,[130] developed by the Novotny group, filters the MS/MS datafiles for Mascot searching using ProtParser, and calculates the relative concentration based on the peak area under the RIC traces for each peptide in the MS mode, correcting for the 'gaps' due to the data-dependent acquisition of MS/MS data. The false-positive rate of this technique was reported to be <4.9%.

## Fluorescent Labeling with Mass-Spectrometric Identification

### Differential gel electrophoresis (DIGE)

This is a relative protein quantitation technique developed by GE Healthcare that does not quite fit the above categories. This quantitation technique is done on intact proteins, not their peptides, and the differential expression determination is not based on mass spectrometry, but is based on fluorescence. In this technique, three different fluorescent labels with different absorbance and emission characteristics, are used to label a small portion of the proteins in each of three different samples. These samples are then combined and separated on a 2D gel. The relative expression ratios are determined from the fluorescence spectra by GE Healthcare's DeCyder software, or another image analysis software package such as ProGenesis, or on a per-gel basis from Ludesi. If DeCyder is used, the output resembles the Decyder MS output shown in Fig. 14. Only those spots which show differential expression are excised (based on their x–y coordinates), and identified using standard in-gel protein identification techniques. The small percentage of the protein that is labeled does not affect the protein identification part of the experiment.

*Advantages and disadvantages*

This is a fairly expensive technique because of the fluorescent labels which have a limited shelf-life. However, because it is a multiplex technique it may be only slightly more expensive

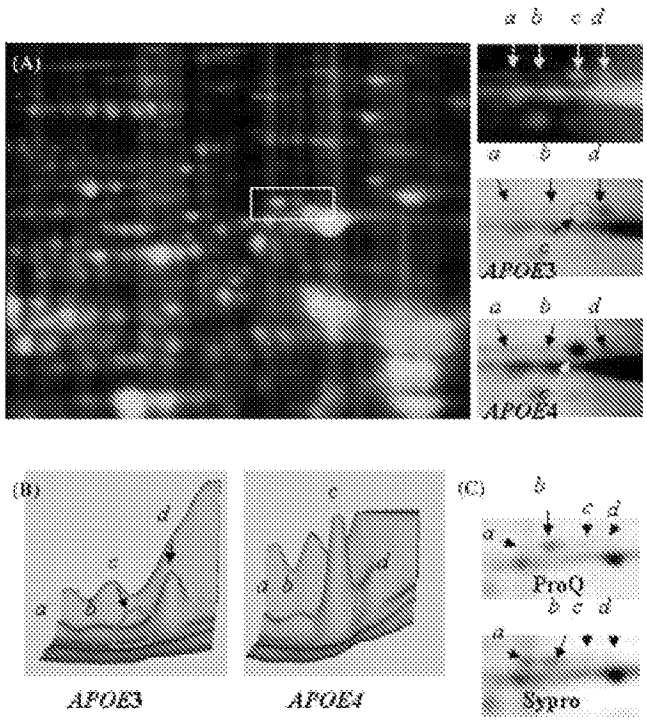



**Figure 15.** DIGE gel showing differential expression of mortalin isoforms. 2D-DiGE and phosphostain analysis of mouse hippocampus. Spots *a*, *b*, *c* and *d* are mortalin. (B) Three-dimensional visualization of protein spots *a*, *b*, *c* and *d*, indicated in the rectangle in panel A. (C) Phosphostaining of phosphorylated mortalin. The arrows show mortalin. Protein isoform *b* is highly phosphorylated, while protein isoforms *a*, and *c* are slightly stained. Isoform *d* is not detected with the Pro-Q phosphostain. Reprinted from Ref. [131] with permission.

than the cost of three separate label-free analyses. Whether it is more or less expensive than iTRAQ depends on the number of samples being compared. One challenge for mass spectrometry is that fluorescence detection is very sensitive – sometimes the differentially expressed protein in the spot cannot be detected by mass spectrometry. In this case, a 'prep gel' with a higher sample loading must be run.

This technique is probably more sensitive to protein degradation in the sample than are the techniques based on peptide detection, because truncated proteins would show up as separate spots on the 2D gel. Conversely, this technique is probably better at detecting isoforms than iTRAQ or the other peptide-based techniques, because these would probably be reported as a single protein (see Fig. 15).

## Discussion

### Assumptions in all peptide-based protein quantitation methods

In all peptide-based quantitation techniques, we are making the assumption that the protein can be digested into peptides, and that these peptides are detectible by the mass spectrometric technique used for the analysis. These assumptions are not unique to label-free quantitation, but are also made in 'shotgun' proteomics in general. In other words, for 'shotgun' proteomics to be able to detect *every* protein present in the sample, *every* protein must be able to be digested by the enzyme used, and *every* protein must produce peptides that are detectible by mass spectrometry, in terms of sensitivity and size. Some of these assumptions were validated by a study by Liu *et al.*,[98] where factors including the size of the protein and the production of detectible ions affected the mass-spectrometry-based quantitation results. However, even if detecting *every* protein is not actually achievable, it is a 'goal', and we need to keep in mind the assumptions that are being made.

Any technique that assumes that peptide quantitation reflects protein concentration is also making several additional assumptions. These are (1) that the digestion is complete and (2) that it is reproducible (one peptide = one protein). Actually, completeness of the digestion is less critical than the reproducibility, because reproducibly incomplete digestion can be empirically corrected for, but reproducibility is essential for accurate peptide-based protein quantitation. Optimizing protein digestion conditions for reproducibility and completeness is still an active area of research.

One peptide = one protein is obviously not true for proteins with repeat units. New software packages are being developed to assist in peptide-based quantitation and dealing with problems relating to these assumptions. For example, software has been designed to address the problem of redundant peptides (i.e. peptides that appear in more than one protein).[132]

### Additional challenges of label-free quantitation

In both the alignment and the spectra-based label-free approaches, it is necessary to compare several different runs. One issue that has to be considered is the reproducibility of replicate label-free quantitation experiments.[98,108,109,111,133,134] One way would be to compare methods using a 'standard' set of proteins. This standard set of proteins has been developed and has been proposed as a way to evaluate reproducibility.[135]

As part of the reproducibility question, one also has to consider the *amount* of each sample to inject, and the effect of sample loading on protein identification and quantitation. This led to the

observation that matrix effects are the dominant factor in peptide detectibility, and that there is a linear range for good peptide quantitation.[136]

Usually, one tries to inject the same total amount of protein. Conceptually, this would work if there were only small changes between samples, but how well would it work if there were changes in *abundant* proteins? One solution is to use internal reference standards for alignment or for normalizing the different runs.[137,138]

Mass-spectrometry based quantitation poses significant statistical challenges. Because of the high cost per analysis, and the often limited amounts of sample, there are very few studies that have addressed the issues of biological variability (samples from different patients or animals), and technical variability (the same biological sample, split and processed independently), and experimental variability (different analyses of the same processed sample). Notable exceptions are the papers by Gan *et al.* for iTRAQ[139] and a recent paper by Li *et al.*[140] on label-free quantitation, where the authors developed a method to determine statistical significance and false positives using AMT-based label-free data. A 'false positive' is defined as a misassignment of differential expression. These authors also discuss the challenges of performing statistics on label-free analysis, as well as the added difficulties of performing statistical analysis on peptide-based quantitation data in general, because different numbers of measurements are made on different proteins, which is not the case for microarray data. A fold-change cutoff was not found to be sufficient – an additional statistical test, performed at the peptide level, was found to be necessary.[141] This method, however, was still not sufficient to determine the false discovery rate (FDR) and statistical significance of relative expression data from label-free experiments. Although the authors recognize that often only one analysis per biological sample is normally performed, they found that the minimum number of analyses required obtaining these statistics was two LC/MS analyses (i.e. two experimental replicates), spiked with the same level of $^{15}$N-labeled internal standard. By quantitating the labeled and unlabeled sample separately, they were able to produce four possible pairings. Three parameters were used to determine differential expression: fold-change, the *t*-test or Wilcoxon ranksum test, and a minimum number of permuted statistical pairings (MPSPs). Using the internal standard as the control, an unlabeled protein found to be differentially labeled was considered to be a positive, while the labeled internal standard found to be differentially labeled, was considered to be a false positive. Interestingly, at a confidence level of 95%, a critical fold-change was found below which there was a drop in the number of positives, while the number of false positives stayed constant which was dependent on the number of analyses (MPSPs). This critical fold-change was 2.75, 2.5, 2.5 and 2.0 for MPSPs of 1, 2, 3 and 4, respectively, which corresponded to FDRs of 22, 15, 8.7 and 4.2%.[140]

### Metabolic versus non-metabolic labeling

Chemical labeling can be applied to any source of biological material. Metabolic labeling using SILAC can be used for cell culture, but it is not effective for autotrophic organisms such as plants or bacteria. For these autotrophic organisms, $^{15}$N labeling is preferred. SILAC typically works well for mammalian cell lines, which do not synthesize all of the amino acids, and so incorporation of the labeled amino acids from the growth medium is more comprehensive.[47]

SILAC has been used for one study involving plant cell culture (*A. thaliana*), and although the average incorporation of $^{13}C_6$-arginine was only 75%, it allowed the study of differential expression of glutathione S-transferase in response to sialic acid treatment.[55] To our knowledge, this is the only report about the use of SILAC in plant proteomics. In contrast, virtually complete labeling (95% and higher) of proteins in both suspension cultures and entire plants of *A. thaliana* has been attained using $^{15}$N isotopes.[55]

There have been a few reports on stable isotope labeled protein quantitation in 'unusual' organisms. Drosophila was also the first multicellular model organism subjected to labeling with $^{15}$N – Heck and collaborators labeled *D. melanogaster* and *C. elegans* with $^{15}$N.[58] Proteomic studies in *Drosophila* are rare, and only a few quantitative proteomics studies have been performed. Aebersold and co-workers have used 4-plex iTRAQ and protein phosphatase treatment for specific substrates in Drosophila cell lines.[142] Our laboratory was involved in an iTRAQ study on *Leishmania*, in which 21% of the proteome was identified and quantified over seven timepoints.[143] Siuzdak and co-workers used stable isotope labeling to monitor the expression kinetics of viral proteins, changes in the expression levels of cellular proteins, and fluctuations in metabolites in response to Flock House Virus (FHV) viral infection.[144]

Yates and collaborators applied the $^{15}$N metabolic labeling technique to *Rattus norvegicus* by feeding them a $^{15}$N-enriched diet. The strategy was employed to generate internal standards to quantify proteins in mammalian tissues. This work provided the proof-of-principle that metabolic labeling of whole organisms is feasible in mammals, as had already been demonstrated for worms and flies, and opened up new possibilities for similar applications.

Mann and co-workers have established protocols for the SILAC-labeling of mice.[145] Labeling of whole animals is based on a special diet containing either the natural or the $^{13}C_6$-substituted version of lysine. Labeling was carried out over four generations, with no effect on development, growth or behavior. Full incorporation of SILAC amino acids was achieved for all organs in the F2 generation animals. However, metabolic labeling strategies for animals are often impractical, due to the high cost of the diet and the long time required for labeling (full incorporation is typically not achieved in the first or even the second generation of animals).

Because of the high cost of isotopically labeled materials, metabolic labeling studies tend to be used for pathway determination. iTRAQ and other chemical-based isotopic-labeling methods are used for biomarker discovery, and MRM methods are used for biomarker verification or validation.

Label-free methods for biomarker discovery are currently receiving a lot of attention, because of their simplicity and low cost. However, the lack of labeled internal standards makes them susceptible to suppression effects from other components in the sample. Of course, at the 'discovery stage' one cannot add internal standards because one does not yet know what standards you will need. This conundrum reflects the current status of protein quantitation.

Also, at this point, there is no single method that will identify and quantitate all of the proteins in the sample – different techniques will find different proteins. Several studies illustrate this point. In the first example, a study of insect salivary gland extracts, the iTRAQ technique identified 43 proteins not observed using in the LC/MS/MS analysis of salivary gland extracts from insects of the same age. This result is consistent with the previous observations that better fragmentation is obtained using this technology, giving more peptides per protein and allowing the identification of less-abundant proteins. iTRAQ labeling led to the identification of 78 proteins, 39 of which were not identified by in a standard LC/MS/MS

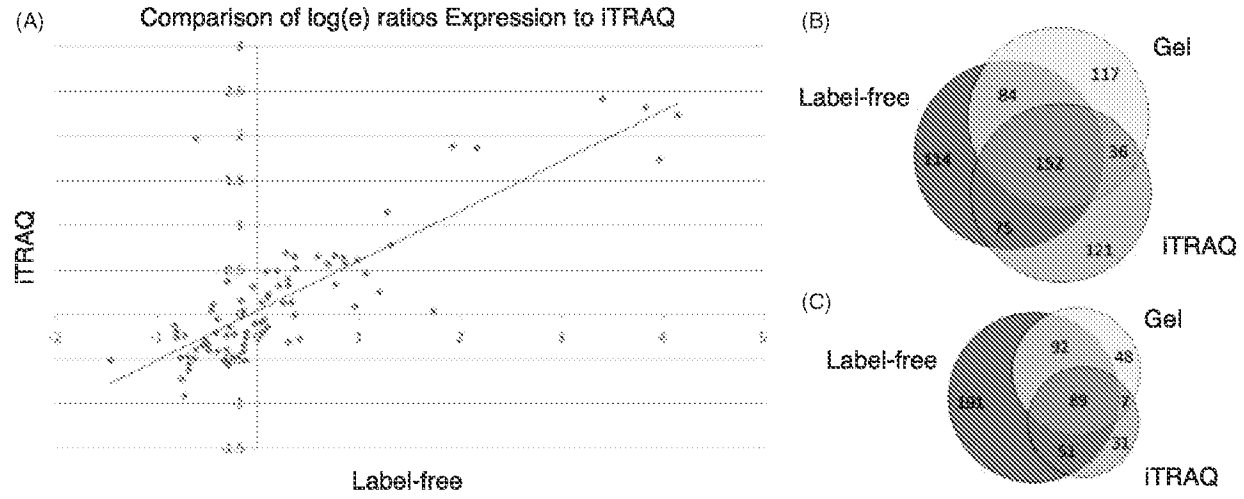



**Figure 16.** A). The correlation between the iTRAQ results and the label-free results, B). Venn diagram of proteins identified by the three techniques. C). Venn diagram of proteins identified by the three techniques, requiring at least 2 peptides for an identification. Reprinted from [146], with permission.

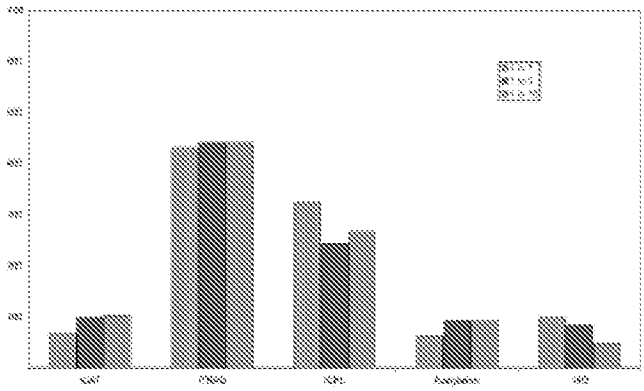


**Figure 17.** The total number of high-confidence proteins identified by MASCOT in each of the five labeling experiments. MS/MS data were searched against an *E. coli* database. The 1 : 25 ratio gave anomalous results for the ICPL labeling, so the 1 : 25 ratio is not included.

analysis, illustrating the value of using both two technologies in parallel for maximum proteome coverage.

### Comparison of methods

In a very recent comparison of iTRAQ, label-free (ion accounting), and geLC by the Patel *et al.*,[146] the expression ratios were higher for the label-free analyses than for iTRAQ, as had been noted in other studies (note the slope of the line). The Venn diagram in Fig. 16 shows the number of proteins identified by the three techniques (including identifications based on a single peptide). It is clear from these studies that, at this point, there is no one technique that can quantitate – or even detect – every protein.

In a recent study in our laboratory, five of the most common labeling techniques – ICPL, cICAT (cleavable ICAT), iTRAQ, $^{18}O$, and acetylation – were compared on an *E. coli* tryptic digest to determine the method that identifies the highest number of proteins and provides the most accurate quantitation. In this study, the highest number of proteins was identified with the iTRAQ labeling system, followed by ICPL. The peptides in these two methods, however, were separated by 2D-LC unlike the other experiments which were done using 1D-LC, thus demonstrating the advantages for prefractionation of peptides in complex samples. The three other labeling systems ($^{18}O$, acetylation and cICAT) resulted in approximately the same number of protein identifications (Fig. 17).

Peptides labeled at 1 : 1 and 1 : 3 ratios with cICAT, iTRAQ and acetylation were quantified with reasonable accuracy. However, only the highest-confidence proteins in iTRAQ-labeled samples resulted in an acceptable amount of variation when labeled at a ratio of 1 : 10 (Fig. 18). We were unable to analyze the $^{18}O$ and ICPL data as we could not find or modify any of our software to accept these labels with the QStar data files. The variation observed in these experiments clearly demonstrates the need for both technical and biological replicates.

The advantages and disadvantages of each procedure, as found in our study, are compared in Table 1.

There have been several other comparisons of different label and label-free methods, where the same samples were analyzed through various quantitation techniques. The results of these comparisons are shown in given in the table in the Supporting Information. In iTRAQ, where the protein identification is done on the same set of labeled peptides as the quantitation, there is an inverse relationship between the confidence of the identification and the number of proteins on which quantitative data can be obtained. In a set of ten experiments on the reproducibility of iTRAQ analyses, Gan *et al.* found that allowing a ±50% change in expression ratio between *biological* replicates





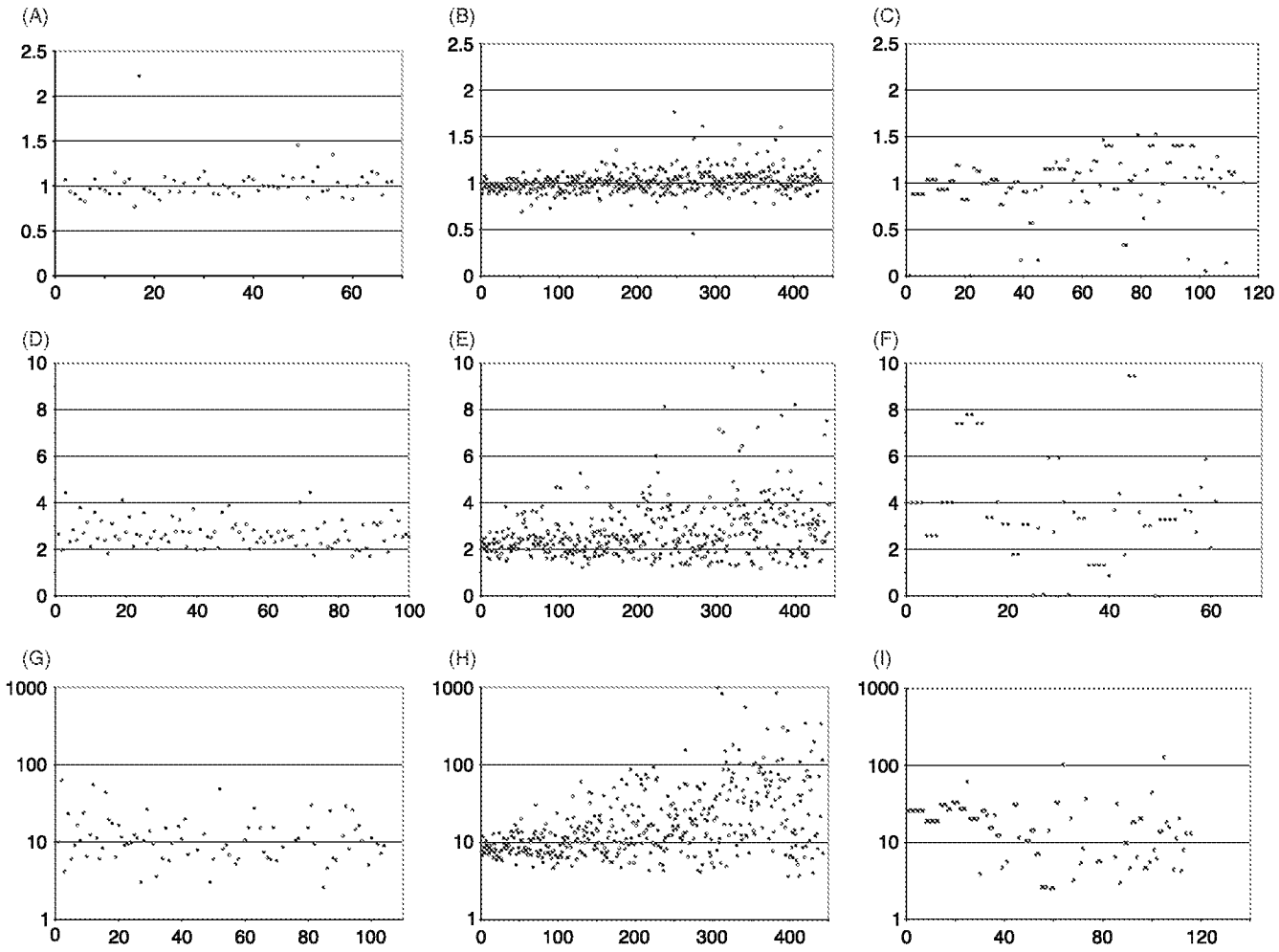



**Figure 18.** The experimental average ratio of cICAT (A), iTRAQ (B), and acetyl (C) labeled *E. coli* proteins at 1 : 1 ratios were calculated by ProteinPilot (ICAT, iTRAQ) and MSQUANT (acetyl). The experimental average ratio of cICAT (D), iTRAQ (E) and acetyl (F) labeled *E. coli* proteins at 1 : 3 ratios were calculated by ProteinPilot (ICAT, iTRAQ) and MSQUANT (acetyl). The experimental average ratio of cICAT G), iTRAQ (H) and acetyl (I) labeled *E. coli* proteins at 1 : 10 ratios were calculated by ProteinPilot (ICAT, iTRAQ) and MSQUANT (acetyl).

resulted in 88% protein coverage, allowing a ±30% change in expression ratio between *technical* replicates resulted in 95% protein coverage, with only a ±0.1% variance coming from the MS.[138] Liu *et al.*[98] found that the correlation of abundance with the number of spectra observed, was better than that based on % sequence coverage or the number of peptides identified per protein. In a comparison of spectral counting *versus* peptide ion intensities, Xia *et al.*[147] found that spectral counting gave better agreement with the true protein ratios. A recent comparison of studies using the spectral counting and ion intensity-based methods of label-free quantitation, with respect to dynamic range of quantitation and dynamic range of protein detection was done by Wong *et al.*[148] Both methods were able to detect changes in protein levels of approximately 2.5. However, machine learning methods and methods using peptide ion intensity were computationally more difficult. This study concluded, however, that these label-free approaches were complementary, and recommended using both for increased confidence in the results.

Some label-free methods, however, were found to underestimate expression ratios if the true ratios were >2.5.[149] One reason that has been proposed for the lower range of ratios produced by peak area intensity measurements compared to spectral counting is the restriction by different software packages on the number of peptides required for a protein to be considered 'detected'.[99] A requirement for a larger number of peptides per protein discriminates against lower-abundance proteins, thus removing the larger expression differences. The larger the difference in abundance ratios, the more reliably this difference could be detected through label-free techniques. The study by Liu *et al.* reported that the *number* of spectra produced was a reliable indication of expression ratio if the concentration difference was >5.[98] Other factors include the size of the protein, the number of tryptic cleavage sites and the amount of protein that can be loaded onto a capillary LC column.[149]

The challenge of comparing quantitation methods is shown in Fig. 19.[125] In this study, three different methods were compared, and all three gave different expression rations – which one is correct? In this particular study, the authors were able to experimentally confirm that the spectral TIC method (using the average MS/MS TIC) was correct, and they attributed the lower ratios obtained from SILAC and spectral counting to compression. However, most studies are performed without this validation

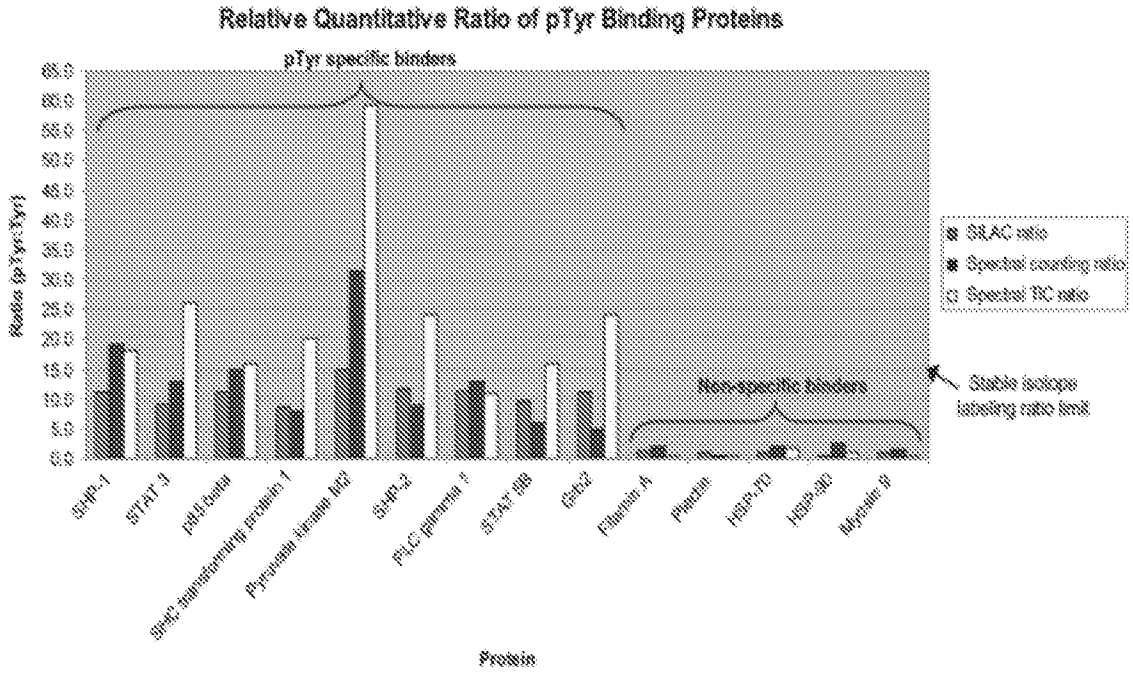



**Figure 19.** Relative expression ratios of pTyrosine-binding proteins using three different quantitation methods. Reprinted from Ref. [125] with permission.

**Table 1.** Comparison of five quantitation methods on an *E. coli* tryptic digest

| | cICAT | iTRAQ | ICPL | $^{18}O$ | Acetylation |
|---|---|---|---|---|---|
| Advantages | Complex samples are simplified | Large number of peptides per protein identified increase confidence and quantitation statistics | Short digest time | Guanidine can be used for solubilization | Inexpensive |
| | MS-based quantitation can reduce costly analysis time | Multiplex with 4/8 labels | Lysis buffer contains guanidine which is excellent for solubilization | MS-based quantitation can reduce costly analysis time | Short preparation time |
| | | | MS-based quantitation can reduce costly analysis time | | MS-based quantitation can reduce costly analysis time |
| Disadvantages | Very few peptides identified per protein leads to low confidence and poor statistics | Fractionation required | Software available only for certain instruments | Long incubation/digestion times | Requires fractionation |
| | Not all proteins contain cysteines | MS/MS quantitation is costly in terms of MS time | Requires fractionation | Pooled sample is only stable for 1 h | Availability of software for data analysis |
| | | Expensive | Expensive | Availability of software for data analysis | Acetic anhydride is a controlled substance |

of true protein concentrations. Clearly, more work needs to be done to validate different quantitation methods using mixtures of proteins with known concentration ratios.

Although there is general agreement that comparison with a labeled standard peptide is still the 'gold standard',[98,112] it is not a global technique. However, many studies have concluded that the *trends* observed in label-free experiments are valid (i.e. overexpression *versus* underexpression) even if the actual observed expression ratios may be compressed.[149]

## Software Considerations

Before you start a project, it is important to be certain that the software 'matches' the label you are trying to use. Certain software packages, for example, can only handle certain types of labeled amino acids for SILAC. Certain companies make proprietary labels, which may only be able to be analyzed using their own software. However, there seems to be a welcome change toward 'openness' now, and iTRAQ-based quantitation, for example, can be analyzed using an increasing number of instruments and software.





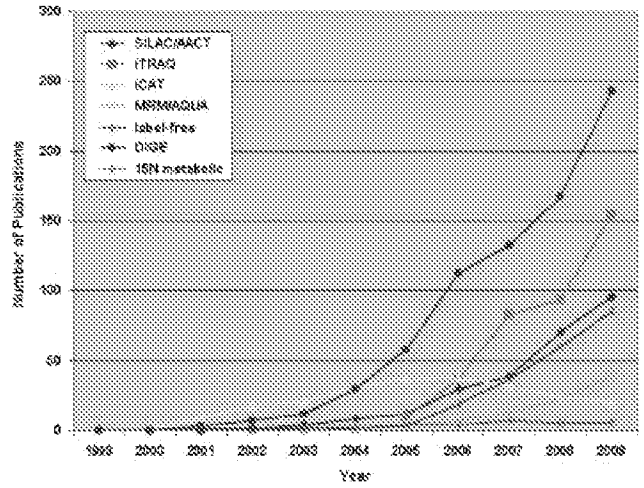



**Figure 20.** Trends in mass-spectrometry-based quantitation. Publications per year, based on a keyword search in SciFinder Scholar™. Note that comparisons over time are probably more accurate than comparisons *between* techniques, due to the difficulties in finding keywords to catch every reference. The number of 2009 publications was extrapolated from the number published by August 2009.

Mascot,[94] for example, can perform quantitation of a variety of 'precursor' methods (where the quantitation based on shifts in the molecular weight), as well as TMT and iTRAQ-type labels where the quantitation is based on MS/MS reporter ions.[150] However, additional software packages, such as Mascot Distiller may be required. They specifically note that for AB MALDI-TOF/TOF data, there is a special software available (TS2Mascot) that should be used because the standard GPS explorer Mascot data is de-isotoped. $^{15}$N-labeling may pose a particular problem because every amino acid will get labeled, but new software (QuantiSpec) has recently been written to enable interpretation and quantitation of $^{15}$N-labeled mass spectra.[151] In general, it is still prudent to make certain to select a label that your data-processing software can handle.

Multifunctional software packages are being developed to handle data from label-free and stable-labeled samples, and from a variety of instrument platforms. These include the ProteinQuant Suite,[130] developed by the Novotny group, Census software[152] developed by the Yates group, the PatternLab software[153] also developed by the Yates group for normalizing spectral count data, and the Corra software developed by the Aebersold group, which in addition, produces protein interaction networks from the differentially expression data.[154,155] MS-Biomarker Discovery Platform (MS-BID),[156] from the Aebersold group, is designed for determining peptides that discriminate between treatment groups. Markerview software (Applied Biosystems) is also designed to facilitate detection of biomarker peptides that correlate with treatment.[157]

## Conclusions

It would have been nice to be able to end this article with a recommendation for a single method. However, as is clear from the above data and discussion, there is really no one single method that will solve all of the analytical problems associated with protein quantitation. This is partly because 'quantitation' means so many different things – global or targeted, absolute or relative.

The ICPL and iTRAQ methods (from our study) and the ion-accounting label-free method (from the Patel study) seemed to identify and quantify significantly more proteins than the other methods in these studies. It should be remembered, however, that the variability of the enzymatic digestion step can affect all of the chemical labeling techniques and the label-free methods by leading to analytical variability. We and others are actively exploring solutions to this problem, including the use of microwave digestion, detergents, pressure, and a variety of solvents, chaotropic agents and denaturants.[158] We are confident, therefore, that this problem will be able to be solved (or at least reduced) in the near future. Another significant source of variability comes from depletion steps. This variability, however, can be reduced by the stringent use of well-developed SOPs.

Label-free methods are based on less-rigorous mass spectrometry, with more reliance on bioinformatics and separation techniques. Factors that have to be considered when selecting a method are the number of treatments, the cost of the experiment, the complexity of the sample, the biological source of the sample and whether the experiment will be done in cell culture. These ultimately will be the determining factors in choosing the appropriate quantitation method.

For this review, we used SciFinder Scholar[159] to count the number of publications per year using these various types of

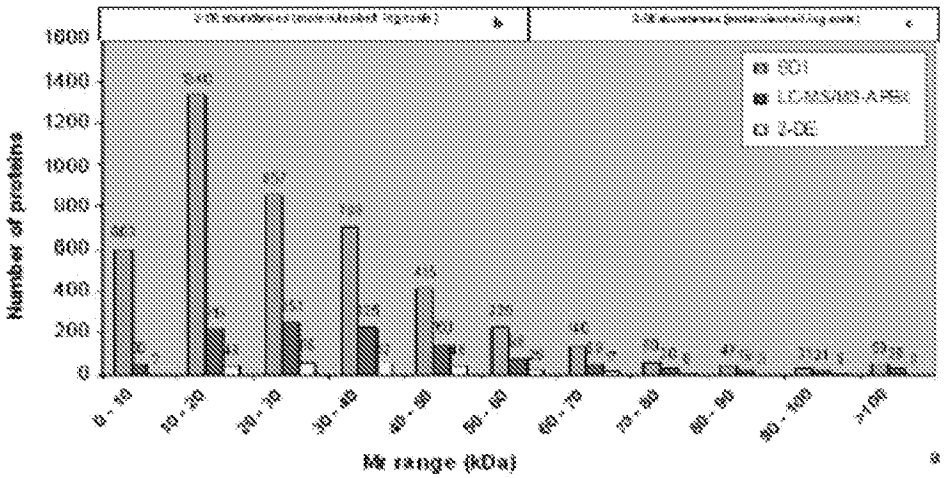



**Figure 21.** Distribution of proteins predicted from the *Shigella dysenteriae* genome (SD1, blue), and found by either 2D gel electrophoresis (2-DE, yellow) or LC/MS/MS (blue). Reprinted from Ref. [128] with permission.

quantitative proteomics (Fig. 20). Although it was impossible to select 'key words' to include every relevant publication, it was clear that all of the methods described above are still in use. Moreover, with the possible exception of $^{15}N$ metabolic labeling and ICAT, the use of most of these methods is increasing each year. In fact, the numbers of publications using DIGE, iTRAQ, label-free methods are all increasing at approximately the same rate – on the order of 50% per year.

In conclusion, even though there is, as yet, no 'one perfect method', this does not mean that there are no *usable* methods. It is still possible to obtain useful relative and/or absolute quantitative data by matching the biological problem to the proper quantitative proteomics approach. Furthermore, mass spectrometers continue to be improved with respect to sensitivity, dynamic range, mass accuracy and scan rate. We are also certain that new multiplex labeling techniques will continue to be developed, and that better SOPs will bring improvements in reproducibility. This will certainly lead to changes and significant improvements in the field of quantitative proteomics.

We should remember, too, that mass spectrometric coverage of the proteome still needs to be improved. As can be seen from Fig. 21, mass spectrometry is still not 'catching' large numbers of lower-molecular weight proteins which are predicted from the genome, and these proteins are underrepresented in current proteomics studies. This might be because insufficient numbers of tryptic peptides of the appropriate size are produced from these smaller proteins, as discussed above. Whether 'top-down' techniques or other methods, such as the use of proteotypic synthetic peptide standards and MRM techniques, as proposed by Aebersold as part of the PeptideAtlas project can fill this gap still remains to be seen. It is important, however, to remember that mass-spectrometry-based protein quantitation still requires mass spectrometric peptide or protein detection, and much work remains to be done in this area as well.

### Acknowledgements

This work was funded by a platform grant from Genome Canada and Genome British Columbia.

### Supporting information

Supporting information may be found in the online version of this article.

## References

[1] S. Ong, M. Mann. Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology* **2005**, *1*, 252.
[2] T. A. Prokhorova, K. T. G. Rigbolt, P. T. Johansen, J. Henningsen, I. Kratchmarova, M. Kassem, B. Blagoev. Stable isotope labeling by amino acids in cell culture (SILAC) and quantitative comparison of the membrane proteomes of self-renewing and differentiating human embryonic stem cells. *Molecular and Cellular Proteomics* **2009**, *8*, 959.
[3] A. Van Langenhove, C. E. Costello, J. E. Biller, K. Biemann, T. R. Browne. A gas chromatographic/mass spectrometric method for the simultaneous quantitation of 5,5-diphenylhydantoin (phenytoin), its para-hydroxylated metabolite and their stable isotope labelled analogs. *Clinica Chimica Acta, International Journal of Clinical Chemistry* **1981**, *115*, 263.
[4] S. J. Gaskell, K. Rollins, R. W. Smith, C. E. Parker. Determination of serum cortisol by thermospray liquid chromatography/mass spectrometry: comparison with gas chromatography/mass spectrometry. *Biomedical and Environmental Mass Spectrometry* **1987**, *14*, 717.
[5] A. L. Yergey, N. V. Esteban, D. J. Liberato. Metabolic kinetics and quantitative analysis by isotope dilution thermospray LC/MS. *Biomedical and Environmental Mass Spectrometry* **1987**, *14*, 623.
[6] J. R. Barr, V. L. Maggio, D. G. J. Patterson, *et al.* Isotope dilution–mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I.1676-82. *Clinical Chemistry* **1996**, *42*, 1676.
[7] L. J. Foster, C. L. de Hoog, M. Mann. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 5813.
[8] N. R. Kitteringham, R. E. Jenkins, C. S. Lane, V. L. Elliott, B. K. Park. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2009**, *877*, 1229.
[9] V. Lange, P. Picotti, B. Domon, R. Aebersold. Selected monitoring monitoring for quantitative proteomics: a tutorial. *Molecular Systems Biology* **2008**, *4*, 1.
[10] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, S. P. Gygi. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 6940.
[11] D. S. Kirkpatrick, S. A. Gerber, S. P. Gygi. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **2005**, *35*, 265.
[12] M. Kuzyk, D. Smith, J. Yang, *et al.* MRM-based, Multiplexed, Absolute Quantitation of 45 proteins in human plasma. *Molecular and Cellular Proteomics* **2009**, *8*, 1860.
[13] C. A. Mueller, W. Weinmann, S. Dresen, A. Schreiber, M. Gergov. Development of a multi-target screening analysis for 301 drugs using a QTrap liquid chromatography/tandem mass spectrometry system and automated library searching. *Rapid Communications in Mass Spectrometry* **2005**, *19*, 1332.
[14] P. Mallick, M. Schirle, S. S. Chen, *et al.* Computational prediction of proteotypic peptides for quantitative proteomics. *Nature Biotechnology* **2007**, *12*, 125.
[15] PeptideAtlas. www.peptideatlas.org. [Last accessed: 2009].
[16] M. Kohl, G. Redlich, M. Eisenacher, *et al.* Automated calculation of unique peptide sequences for unambiguous identification of highly homologous proteins by mass spectrometry. *Journal of Proteomics and Bioinformatics* **2008**, *1*, 006.
[17] W. Timm, A. Scherbart, S. Böcker, O. Kohlbacher, T. W. Nattkemper. Peak intensity prediction in MALDI-TOF mass spectrometry: A machine learning study to support quantitative proteomics. *BMC Bioinformatics* **2008**, *8*, 443.
[18] A. M. Frank. Predicting intensity ranks of peptide fragment ions. *Journal of Proteome Research* **2009**, *8*, 2226.
[19] Y. Xu, I. M. Lazar. MRM screening/biomarker discovery with linear ion trap MS: a library of human cancer-specific peptides. *BMC Cancer* **2009**, *9*, 96.
[20] P. Picotti, H. Lam, D. Campbell, *et al.* A database of mass spectrometric assays for the yeast proteome. *Nature Methods* **2008**, *5*, 913.
[21] G. M. Walsh, S. Lin, D. M. Evans, A. Khosrovi-Eghbal, R. C. Beavis, J. Kast. Implementation of a data repository-driven approach for targeted proteomics experiments by multiple reaction monitoring. *Journal of Proteomics* **2009**, *72*, 838.
[22] Skyline_SRM/MRM_Builder. **2009**, https://brendanx-uw1.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=default.
[23] B. MacLean, D. Tomazela, G. Finney, *et al.* Automated creation and refinement of complex scheduled srm methods for targeted proteomics. In *Presented at the 57th ASMS Conference on Mass Spectrometry and Allied Topics*. Philadelphia, PA, May 31 – Jun 4, **2009**.
[24] A. Prakash, D. M. Tomazela, B. Frewen, *et al.* Expediting the development of targeted SRM assays: using data from shotgun proteomics to automate method development. *Journal of Proteome Research* **2009**, *8*, 2733.
[25] MRMpilot. https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=605354&tab=DetailInfo.
[26] Agilent_MRM. www.chem-agilent.com/pdf/QQQ_peptide_quant_on_eseminar.pdf.

[27] C. Sherwood, A. Eastham, A. Peterson, *et al.* MaRiMba: a software application for spectral library-based MRM transition list assembly. *Journal of Proteome Research* **2009**, *8*, 4396.
[28] J. A. Mead, L. Bianco, V. Ottone, *et al.* MRMaid, the web-based tool for designing multiple reaction monitoring (MRM) transitions. *Molecular and Cellular Proteomics* **2009**, *8*, 696.
[29] D. B. Martin, T. Holzman, D. May, *et al.* MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments. *Molecular and Cellular Proteomics* **2009**, *7*, 2270.
[30] B. R. Franza Jr, Y. P. Rochon. Stable isotope metabolic labeling for analysis of biopolymers. Patent No. 6653076. Assignees: The Regents of the University of Washington, filed 2001, issued 2003.
[31] Automaton. http://www3.appliedbiosystems.com/cms/groups/psm_marketing/documents/generaldocuments/cms 039798.pdf. [Last accessed: 2009].
[32] M. Fonsi, F. Fiore, P. Jones, *et al.* Metabolism-related liabilities of a potent histone deacetylase (HDAC) inhibitor and relevance of the route of administration on its metabolic fate. *Xenobiotica* **2009**, 1.
[33] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry* **2007**, *389*, 1017.
[34] T. A. Addona, S. E. Abbatiello, B. Schilling, *et al.* Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nature Biotechnology* **2009**, *27*, 633.
[35] B. L. Ackermann, M. J. Berna. Coupling immunoaffinity techniques with MS for quantitative analysis of low-abundance protein biomarkers. *Expert Review of Proteomics* **2007**, *4*, 175.
[36] J. Kamiie, S. Ohtsuki, R. Iwase, *et al.* Quantitative atlas of membrane transporter proteins: development and application of a highly sensitive simultaneous LC/MS/MS method combined with novel in-silico peptide selection criteria. *Pharmaceutical Research* **2008**, *25*, 1469.
[37] D. R. Barnidge, M. K. Goodmanson, G. G. Klee, D. C. Muddiman. Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *Journal of Proteome Research* **2004**, *3*, 644.
[38] E. Kuhn, J. Wu, J. Karl, H. Liao, W. Zolg, B. Guild. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. *Proteomics* **2004**, *4*, 1175.
[39] F. Zhang, M. J. Bartels, W. T. Stott. Quantitation of human glutathione S-transferases in complex matrices by liquid chromatography/tandem mass spectrometry with signature peptides. *Rapid Communications in Mass Spectrometry* **2004**, *18*, 491.
[40] N. L. Anderson, N. G. Anderson. The human plasma proteome : history, character, and diagnostic prospects. *Molecular and Cellular Proteomics* **2002**, *1*, 845.
[41] M. J. McKay, J. Sherman, M. T. Laver, M. S. Baker, S. J. Clarke, M. P. Molloy. The development of multiple reaction monitoring assays for liver-derived plasma proteins. *Proteomics: Clinical Applications* **2007**, *1*, 1570.
[42] N. L. Anderson, N. G. Anderson, L. R. Haines, D. B. Hardie, R. W. Olafson, T. W. Pearson. Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). *Journal of Proteome Research* **2004**, *3*, 235.
[43] J. Jiang, C. E. Parker, J. R. Fuller, T. H. Kawula, C. H. Borchers. An immunoaffinity tandem mass spectrometry (iMALDI) assay for detection of Francisella tularensis. *Analytica Chimica Acta* **2007**, *605*, 70.
[44] J. Jiang, C. E. Parker, K. A. Hoadley, C. M. Perou, G. Boysen, C. H. Borchers. Development of an immuno tandem mass spectrometry (iMALDI) assay for EGFR diagnosis. *Proteomics: Clinical Applications* **2007**, *1*, 1651.
[45] N. L. Anderson, A. Jackson, D. Smith, D. Hardie, C. Borchers, T. W. Pearson. SISCAPA peptide enrichment on magnetic beads using an in-line bead trap device. *Molecular and Cellular Proteomics* **2009**, *8*, 995.
[46] B. Shah, C. H. Borchers, iMALDI: a targeted proteomics approach to the differentiation of EGFR and its isoforms. In *Presented at the 57th ASMS Conference on Mass Spectrometry and Allied Topics*. Philadelphia, PA, May 31 – Jun 4, **2009**.
[47] A. Iliuk, J. Galan, W. A. Tao. Playing tag with quantitative proteomics. *Analytical and Bioanalytical Chemistry* **2009**, *393*, 503.
[48] Y. K. Wang, Z. Ma, D. F. Quinn, E. W. Fu. Inverse 18O labeling mass spectrometry for the rapid identification of marker/target proteins. *Analytical Chemistry* **2001**, *73*, 3742.
[49] Y. K. Wang, Z. Ma, D. F. Quinn, E. W. Fu. Inverse 15N -metabolic labeling/mass spectrometry for comparative proteomics and rapid identification of protein markers/targets. *Rapid Communications in Mass Spectrometry* **2002**, *16*, 1389.
[50] Y. K. Wang, D. F. Quinn, Z. Ma, E. W. Fu. Inverse labeling-mass spectrometry for the rapid identification of differentially expressed protein markers/targets. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, *782*, 291.
[51] J. Raap, C. Winkel, A. H. de Wit, A. H. van Houten, A. J. Hoff, J. Lugtenburg. Mass spectrometric determination of isotopically labeled tyrosines and tryptophans in photosynthetic reaction centers of Rhodobacter sphaeroides R-26. *Analytical Biochemistry* **1990**, *191*, 9.
[52] Y. Oda, K. Huang, F. R. Cross, D. Cowburn, B. T. Chait. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 6591.
[53] H. Gao, Y. Shen, T. D. Veenstra, *et al.* Two-dimensional electrophoretic/chromatographic separations combined with electrospray ionization FTICR mass spectrometry for high throughput proteome analysis. *Journal of Microcolumn Separations* **2000**, *12*, 383.
[54] T. P. Conrads, G. A. Anderson, T. D. Veenstra, L. Pasa-Tolic, R. D. Smith. Utility of accurate mass tags for proteome-wide protein identification. *Analytical Chemistry* **2000**, *72*, 3349.
[55] S. Oeljeklaus, J. Barbour, H. E. Meyer, B. Warscheid. Mass spectrometry-driven approaches to quantitative proteomics and beyond. *Comprehensive Analytical Chemistry* **2009**, *52*, 411.
[56] A. Bachi, T. Bonaldi. Quantitative proteomics as a new piece of the systems biology puzzle. *Journal of Proteomics* **2008**, *71*, 357.
[57] Mascot quantitation. http://www.matrixscience.com/help/quant overview help. html. [Last accessed: 2009].
[58] J. Krijgsveld, R. F. Ketting, T. Mahmoudi, *et al.* Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. *Nature Biotechnology* **2003**, *21*, 927.
[59] C. J. Nelson, E. L. Huttlin, A. D. Hegeman, A. C. Harms, M. R. Sussman. Implications of 15N-metabolic labeling for automated peptide identification in Arabidopsis thaliana. *Proteomics* **2007**, *7*, 1279.
[60] C. C. Wu, M. J. MacCoss, K. E. Howell, D. E. Matthews, J. R. Yates III. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Analytical Chemistry* **2004**, *76*, 4951.
[61] S.-E. Ong, B. Blagoev, I. Kratchmarova, *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular and Cellular Proteomics* **2002**, *1*, 376.
[62] X. Chen, L. M. Smith, M. E. Bradbury. Site-specific mass tagging with stable isotopes inproteins for accurate and efficient protein identification. *Analytical Chemistry* **2000**, *72*, 1134.
[63] B. Blagoev, M. Mann. Quantitative proteomics to study mitogen-activated protein kinases. *Methods* **2006**, *40*, 243.
[64] S.-I. Hwang, D. H. Lundgren, V. Mayya, *et al.* Systematic characterization of nuclear proteome during apoptosis: a quantitative proteomic study by differential extraction and stable isotope labeling. *Molecular and Cellular Proteomics* **2006**, *5*, 1131.
[65] S. C. Bendall, C. Hughes, M. H. Stewart, B. Doble, M. Bhatia, G. A. Lajoie. Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Molecular and Cellular Proteomics* **2008**, *7*, 1587.
[66] H. C. Harsha, H. Molina, A. Pandey. Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nature Protocols* **2008**, *3*, 505.
[67] M. T. Sykes, J. R. Williamson. Envelope: interactive software for modeling and fitting complex isotope distributions. *BMC Bioinformatics* **2008**, *9*, 446.

[68] E. Sperling, A. E. Bunner, M. T. Sykes, J. R. Williamson. Quantitative analysis of isotope distributions in proteomic mass spectrometry using least-squares fourier transform convolution. *Analytical Chemistry* **2008**, *80*, 4906.
[69] X. Yao, A. Freas, J. Ramirez, P. A. Demirev, C. Fenselau. Proteolytic 18O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Analytical Chemistry* **2001**, *73*, 2836.
[70] J. Sakai, S. Kojima, K. Yanagi, M. Kanaoka. 18O-labeling quantitative proteomics using an ion trap mass spectrometer. *Proteomics* **2005**, *5*, 16.
[71] K. Gevaert, F. Impens, B. Ghesquiere, P. Van Damme, A. Lambrechts, J. Vandekerckhove. Stable isotopic labeling in proteomics. *Proteomics* **2008**, *8*, 4873.
[72] K. J. Reynolds, X. Yao, C. Fenselau. Proteolytic 18O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *Journal of Proteome Research* **2002**, *1*, 27.
[73] K. Gevaert, J. VandeKerckhove. Protein identification methods in proteomics. *Electrophoresis* **2000**, *21*, 1145.
[74] C.-H. Chen. Review of a current role of mass spectrometry for proteome research. *Analytica Chimica Acta* **2008**, *624*, 16.
[75] S. Dasari, P. A. Wilmarth, A. P. Reddy, L. J. G. Robertson, S. R. Nagalla, L. L. David. Quantification of isotopically overlapping deamidated and 18O-labeled peptides using isotopic envelope mixture modeling. *Journal of Proteome Research* **2009**, *8*, 1263.
[76] K. L. Simpson, A. D. Whetton, C. Dive. Quantitative mass spectrometry-based techniques for clinical use: biomarker identification and quantification. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2009**, *877*, 1240.
[77] M. Munchbach, M. Quadroni, G. Miotto, P. James. Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. *Analytical Chemistry* **2000**, *72*, 4047.
[78] A. Schmidt, J. Kellermann, F. Lottspeich. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **2005**, *5*, 4.
[79] X. Zhang, Q. K. Jin, S. A. Carr, R. S. Annan. N-Terminal peptide labeling strategy for incorporation of isotopic tags: a method for the determination of site-specific absolute phosphorylation stoichiometry. *Rapid Communications in Mass Spectrometry* **2002**, *16*, 2325.
[80] E. O. Hochleitner, M. K. Gorny, S. Zolla-Pazner, K. B. Tomer. Mass spectrometric characterization of a discontinuous epitope of the HIV envelope protein HIV-gp120 recognized by the human monoclonal antibody 1331A. *Journal of Immunology* **2000**, *164*, 4156.
[81] K. Ishikawa. Analysis of molecular surface structure of HIV proteins by deuterium labeling and peptide mapping. *Journal of the Mass Spectrometry Society of Japan* **1999**, *47*, 397.
[82] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology* **1999**, *17*, 994.
[83] J. Li, H. Steen, P. K. C. GS. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. *Molecular and Cellular Proteomics* **2003**, *2*, 1198.
[84] K.C.Hansen, G. Schmitt-Ulms, R. J. Chalkley, J. Hirsch, M. A. Baldwin, A. L. Burlingame. Mass spectrometric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and multidimensional chromatography. *Molecular and Cellular Proteomics* **2003**, *2*, 299.
[85] R. Zhang, C. S. Sioma, R. A. Thompson, L. Xiong, F. E. Regnier. Controlling deuterium isotope effects in comparative proteomics. *Analytical Chemistry* **2002**, *74*, 3662.
[86] A. Leitner, W. Lindner. Current chemical tagging strategies for proteome analysis by mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2004**, *813*, 1.
[87] C. P. Vaughn, D. K. Crockett, M. S. Lim, K. S. J. Elenitoba-Johnson. Analytical characteristics of cleavable isotope-coded affinity tag-LC-tandem mass spectrometry for quantitative proteomic studies. *Journal of Molecular Diagnostics* **2006**, *8*, 513.
[88] mTRAQ. **2009**, https://www3.appliedbiosystems.com/cms/groups/global_marketing_group/documents/web_content/cms_064662.jpg.
[89] mTRAQ. **2009**, https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=605417&tab=DetailInfo.
[90] L. V. De Souza, A. M. Taylor, W. Li, *et al.* Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *Journal of Proteome Research* **2008**, *7*, 3525.
[91] Tandem Mass Tag (TMT). **2008**, http://www.piercenet.com/products/browse.cfm?fldID=C5A5F2FF-5056-8A76-4E4C-971BF496D176.
[92] Exactag. http://las.perkinelmer.com/Content/RelatedMaterials/Brochures/sht_affinity-tools.pdf.
[93] P. L. Ross, Y. N. Huang, J. N. Marchese, *et al.* Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular and Cellular Proteomics* **2004**, *3*, 1154.
[94] Mascot. www.matrixscience.com [Last accessed: 2009].
[95] UVic Proteomics Center. **2009**, http://www.proteincentre.com/services/quantitative-proteomics.
[96] L. V. DeSouza, J. Grigull, S. Ghanny, *et al.* Endometrial carcinoma biomarker discovery and verification using differentially tagged clinical samples with multidimensional liquid chromatography and tandem mass spectrometry. *Molecular and Cellular Proteomics* **2007**, *6*, 1170.
[97] L. V. DeSouza, A. D. Romaschin, T. J. Colgan, K. W. M. Siu. Absolute quantification of potential cancer markers in clinical tissue homogenates using multiple reaction monitoring on a hybrid triple quadrupole/linear ion trap tandem mass spectrometer. *Analytical Chemistry* **2009**, *81*, 3462.
[98] H. Liu, R. G. Sadygov, J. R. Yates, III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry* **2004**, *76*, 4193.
[99] W. M. Old, K. Meyer-Arendt, L. Aveline-Wolf, *et al.* Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular and Cellular Proteomics* **2005**, *4*(10), 1487.
[100] D. Stalder, A. Haeberli, M. Heller. Evaluation of reproducibility of protein identification results after multidimensional human serum protein separation. *Proteomics* **2008**, *8*, 414.
[101] D. A. Wolters, M. P. Washburn, J. R. r. Yates. An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry* **2001**, *73*, 5683.
[102] E. Durr, J. Yu, K. M. Krasinska, *et al.* Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. *Nature Biotechnology* **2004**, *22*, 985.
[103] S. M. Bridges, G. B. Magee, N. Wang, W. P. Williams, S. C. Burgess, B. Nanduri. ProtQuant: a tool for the label-free quantification of MudPIT proteomics data. *BMC Bioinformatics* **2007**, *8*, S24.
[104] B. Fischer, J. Grossmann, V. Roth, W. Gruissem, S. Baginsky, J. M. Buhmann. Semi-supervised LC/MS alignment for differential proteomics. *Bioinformatics* **2006**, *22*, e132.
[105] V. P. Andreev, L. Li, L. Cao, *et al.* A new algorithm using cross-assignment for label-free quantitation with LC-LTQ-FT MS. *Journal of Proteome Research* **2007**, *6*, 2186.
[106] G. L. Finney, A. R. Blackler, M. R. Hoopmann, J. D. Canterbury, C. C. Wu, M. J. MacCoss. Label-free comparative analysis of proteomics mixtures using chromatographic alignment of high-resolution microLC-MS data. *Analytical Chemistry* **2008**, *80*, 961.
[107] A. S. Haqqani, J. F. Kelly, D. B. Stanimirovic. Quantitative protein profiling by mass spectrometry using label – free proteomics. *Methods in Molecular Biology* **2009**, *439*, (Genomics Protocols (2nd Edition)): 241.
[108] A. H. P. America, J. H. G. Cordewener. Comparative LC-MS: a landscape of peaks and valleys. *Proteomics* **2008**, *8*, 731.
[109] J. Sutton, T. Richmond, X. Shi, *et al.* Performance characteristics of an FT MS-based workflow for label-free differential MS analysis of human plasma: standards, reproducibility, targeted feature investigation, and application to a model of controlled myocardial infarction. *Proteomics: Clinical Applications* **2008**, *2*, 862.
[110] S. J. Callister, R. C. Barry, J. N. Adkins, *et al.* Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *Journal of Proteome Research* **2006**, *5*, 277.
[111] G. Wang, W. W. Wu, W. Zeng, C.-L. Chou, R.-F. Shen. Label-free protein quantification using LC-coupled ion trap or FT mass

spectrometry: reproducibility, linearity, and application with complex proteomes. *Journal of Proteome Research* **2006**, *5*, 1214.
[112] M. Mann, N. L. Kelleher. Precision proteomics: the case for high resolution and high mass accuracy. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18132.
[113] L. N. Mueller, O. Rinner, A. Schmidt, *et al.* SuperHirn – a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* **2007**, *7*, 3470.
[114] K. Petritis, L. J. Kangas, P. L. Ferguson, *et al.* Use of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analyses. *Analytical Chemistry* **2002**, *75*, 1039.
[115] E. F. Strittmatter, P. L. Ferguson, K. Tang, R. D. Smith. Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2003**, *14*, 980.
[116] M. E. Monroe, N. Tolic', N. Jaitly, J. L. Shaw, J. N. Adkins, R. D. Smith. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinformatics* **2007**, *23*, 2021.
[117] W.-J. Qian, M. E. Monroe, T. Liu, *et al.* Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 16O/18O labeling and the accurate mass and time tag approach. *Molecular and Cellular Proteomics* **2005**, *4*, 700.
[118] W.-J. Qian, T. Liu, V. A. Petyuk, *et al.* Large-scale multiplexed quantitative discovery proteomics enabled by the use of an 18O-labeled "Universal" reference sample. *Journal of Proteome Research* **2009**, *8*, 290.
[119] C. Johansson, J. Samskog, L. Sundstroem, H. Wadensten, L. Bjoerkesten, J. Flensburg. Differential expression analysis of Escherichia coli proteins using a novel software for relative quantitation of LC-MS/MS data. *Proteomics* **2006**, *6*, 4475.
[120] J. Rappsilber, U. Ryder, A. I. Lamond, M. Mann. Large-scale proteomic analysis of the human spliceosome. *Analytical Chemistry* **2002**, *12*, 1231.
[121] Y. Ishihama, Y. Oda, T. Tabata, *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular and Cellular Proteomics* **2005**, *4*(9), 1265.
[122] J. C. Silva, M. V. Gorenstein, G.-Z. Li, J. P. C. Vissers, S. J. Geromanos. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular and Cellular Proteomics* **2006**, *5*, 144.
[123] H. Choi, D. Fermin, A. I. Nesvizhskii. Significance analysis of spectral count data in label-free shotgunproteomics. *Molecular and Cellular Proteomics* **2008**, *7*, 2373.
[124] X. Fu, S. A. Gharib, P. S. Green, *et al.* Spectral index for assessment of differential protein expression in shotgun proteomics. *Journal of Proteome Research* **2008**, *7*, 845.
[125] J. M. Asara, H. R. Christofk, L. M. Freimark, L. C. Cantley. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics* **2008**, *8*, 994.
[126] J. C. Braisted, S. Kuntumalla, C. Vogel, *et al.* The APEX quantitative proteomics tool: generating protein quantitation estimates from LC-MS/MS proteomics results. *BMC Bioinformatics* **2008**, *9*, 529.
[127] P. Lu, C. Vogel, R. Wang, X. Yao, E. M. Marcotte. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology* **2007**, *25*, 118.
[128] S. Kuntumalla, J. C. Braisted, S.-T. Huang, *et al.* Comparison of two label – free global quantitation methods, APEX and 2D gel electrophoresis, applied to the Shigella dysenteriae proteome. *BMC Bioinformatics* **2009**, *10*, 43.
[129] S. Ryu, B. Gallis, Y. A. Goo, S. A. Shaffer, D. Radulovic, D. R. Goodlett. Comparison of a label-free quantitative proteomic method based on peptide ion current area to the isotope coded affinity tag method. *Cancer Informatics* **2008**, *6*, 243.
[130] B. Mann, M. Madera, Q. Sheng, H. Tang, Y. Mechref, M. V. Novotny. ProteinQuant Suite: a bundle of automated software tools for label-free quantitative proteomics. *Rapid Communications in Mass Spectrometry* **2008**, *22*, 3823.
[131] C. Osorio, P. M. Sullivan, D. N. He, *et al.* Mortalin is regulated by APOE in hippocampus of AD patients and by human APOE in TR mice. *Neurobiology of Aging* **2007**, *28*, 1853.
[132] S. Jin, D. S. Daly, D. L. Springer, J. H. Miller. The effects of shared peptides on protein quantitation in label-free proteomics by LC/MS/MS. *Journal of Proteome Research* **2008**, *7*, 164.
[133] J. A. Koziol, A. C. Feng, J. Yu, N. M. Griffin, J. E. Schnitzer. Range charts for agreement in measurement comparison studies, with application to replicate mass spectrometry experiments. *Journal of Proteomics and Bioinformatics* **2008**, *1*, 287.
[134] Y. J. Kim, P. Zhan, B. Feild, S. M. Ruben, T. He. Reproducibility assessment of relative quantitation strategies for LC-MS based proteomics. *Analytical Chemistry* **2007**, *79*, 5651.
[135] E. Kolker, J. M. Hogan, R. Higdon, *et al.* Development of BIATECH-54 standard mixtures for assessment of protein identification and relative expression. *Proteomics* **2007**, *7*, 3693.
[136] K. Liu, J. Zhang, J. Wang, L. Zhao, X. Peng, W. Jia, W. Ying, Y. Zhu, H. Xie, F. He, X. Qian. Relationship between Sample Loading Amount and Peptide Identification and Its Effects on Quantitative Proteomics. *Anal. Chem.* **2009**, *81*, 1307.
[137] F. Brambilla, D. Resta, I. Isak, M. Zanotti, A. Arnoldi. A label-free internal standard method for the differential analysis of bioactive lupin proteins using nano HPLC-Chip coupled with ion trap mass spectrometry. *Proteomics* **2009**, *9*, 272.
[138] T. Tabata, T. Sato, J. Kuromitsu, Y. Oda. Pseudo internal standard approach for label-free quantitative proteomics. *Analytical Chemistry* **2007**, *79*, 8440.
[139] C. S. Gan, P. K. Chong, T. K. Pham, P. C. Wright. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *Journal of Proteome Research* **2007**, *6*, 821.
[140] Q. Li, B. Roxas. An assessment of false discovery rates and statistical significance in label-free quantitative proteomics with combined filters. *BMC Bioinformatics* **2009**, *10*, 43.
[141] B. A. Roxas, Q. Li. Significance analysis of microarray for relative quantitation of LC/MS data in proteomics. *BMC Bioinformatics* **2008**, *9*, 187.
[142] B. Bodenmiller, J. Malmstrom, B. Gerrits, *et al.* PhosphoPep--a phosphoproteome resource for systems biology research in Drosophila Kc167 cells. *Molecular Systems Biology* **2007**, *3*, 139.
[143] D. Rosenzweig, D. Smith, F. Opperdoes, S. Stern, R. W. Olafson, D. Zilberstein. Retooling Leishmania metabolism: from sand fly gut to human macrophage. *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* **2008**, *22*, 590.
[144] E. P. Go, W. R. Wikoff, Z. Shen, *et al.* Mass spectrometry reveals specific and global molecular transformations during viral infection. *Journal of Proteome Research* **2006**, *5*, 2405.
[145] M. Krüger, M. Moser, S. Ussar, *et al.* SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* **2008**, *134*, 353.
[146] V. J. Patel, K. Thalassinos, S. E. Slade, *et al.* A comparison of labelling and label-free mass spectrometry-based proteomics approaches. *Journal of Proteome Research* **2009**, *8*, 3752.
[147] Q. Xia, T. Wang, F. Taub, *et al.* Quantitative proteomics of intracellular porphyromonas gingivalis. *Proteomics* **2007**, *7*, 4323.
[148] J. W. H. Wong, M. J. Sullivan, G. Cagney. Computational methods for the comparative quantification of proteins in label-free LCn-MS experiments. *Briefings in Bioinformatics* **2007**, *9*, 156.
[149] Q. Xia, E. L. Hendrickson, T. Wang, R. J. Lamont, J. A. Leight, M. Hackett. Protein abundance ratios for global studies of prokaryotes. *Proteomics* **2007**, *7*, 2904.
[150] Mascot Quantitation FAQ. **2009**, http://www.matrixscience.com/pdf/2009WKSHP3.pdf.
[151] K. Haegler, N. S. Mueller, G. Maccarrone, *et al.* QuantiSpec – quantitative mass spectrometry data analysis of 15N-metabolically labeled proteins. *Journal of Proteomics* **2009**, *71*, 601.
[152] S. K. Park, J. D. Venable, T. Xu, J. R. I. Yates. A quantitative analysis software tool for mass spectrometry–based proteomics. *Nature Methods* **2008**, *5*, 319.
[153] P. C. Carvalho, J. S. G. Fischer, E. I. Chen, J. R. r. Yates, V. C. Barbosa. PatternLab for proteomics: a tool for differential shotgun proteomics. *BMC Bioinformatics* **2008**, *9*, 316.
[154] M.-Y. Brusniak, B. Bodenmiller, D. Campbell, *et al.* Corra: computational framework and tools for LC-MS discovery and targeted mass spectrometry-based proteomics. *BMC Bioinformatics* **2008**, *9*, 542.

[155] S. Letarte, M.-Y. Brusniak, D. Campbell, *et al.* Differential plasma glycoproteome of p19ARF skin cancer mouse model using the corra label-free LC-MS proteomics platform. *Clinical Proteomic* **2008**, *4*, 105.
[156] D. Hwang, N. Zhang, H. Lee, *et al.* MS-BID: a Java package for label-free LC-MS-based comparative proteomic analysis. *Bioinformatics* **2008**, *24*, 2641.
[157] Markerview. https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&catID=601523&tab=DetailInfo.
[158] M. A. Kuzyk, J. Proc, D. B. Hardie, *et al.* Quantitative study of chaotropic and solvent effects on the digestion efficiency of human plasma proteins by trypsin. Mansucript in preparation. **2009**.
[159] SciFinder. https://scifinder.cas.org. [Last accessed: 2009].

# Appendix D

St. Jean, "Krill oil production according to the Beaudoin patent," Notebook page, Neptune Technologies & Bioresources (2009)

Report R & D activities in November 09
Project: *New marine biomass & other products*
Made by: Pierre St-Jean

**Krill oil production according to the Beaudoin patent**

**Methodology**

**Extraction according to the Beaudoin patent**

**Step 1**

3 kg of krill split into 6 X 500g
Add 6 volumes of acetone
Grind for 20 min.
Extract for 120 min. without agitation
Filter
Wash solid with 2 volumes of acetone
Combine the two filtrates
Evaporate the acetone
Separate the fat from the water by decantation after addition of 1 volume of acetone
Recover oil
Evaporate under vacuum
Put aside

**Step 2**

Extract from solid with 2 volumes of ethanol 30 min at 4°C
Filter
Combine with the filtrate to the first oil extraction
Evaporate under vacuum
Heat at 125°C

**Results**

The yield of oil production was more than 260g. The oil is reddish brown and it is very odorous. Extraction of krill in the first extraction gave a stable emulsion that we had to break by adding acetone to reduce the volume of water associated with the fat.

# Production d'huile de krill selon le brevet de Beaudoin

## Méthodologie

### Extraction selon brevet Beaudoin

#### Étape 1

3 kg de krill KII RI fractionnés en 6 X 500g
Ajouter 6 volumes d'acétone
Broyer pendant 20 min.
Extraire pendant 120 min. sans agitation
Filtrer
Laver l'aquatéine avec 2 volumes d'acétone
Combiner les deux filtrats
Évaporer l'acétone
Séparer le gras de l'eau par décantation après ajout de 1 volume d'acétone
Récupérer l'huile
Évaporer sous vide
Réserver

#### Étape 2

Extraire l'aquatéine avec 2 volumes d'éthanol 30 min à 4°C
Filtrer
Joindre le filtrat à l'huile de la première extraction
Évaporer sous vide
Chauffer à 125°C

## Résultats

Le rendement de production d'huile a été de plus de 260g. L'huile est de couleur brun-rouge et elle est très odorante. L'extraction du KII RI a donné en première extraction une émulsion stable qu'on a dû briser par ajout d'acétone pour réduire le volume d'eau associé au gras.

# EXHIBIT E

Herman and Groves
"The Influence of Free Fatty Acid Formation on the pH of Phospholipid-Stabilized Triglyceride Emulsions"
*Pharmaceutical Research* 10(5): 774-776 (1993))

# The Influence of Free Fatty Acid Formation on the pH of Phospholipid-Stabilized Triglyceride Emulsions

**Clifford J. Herman[1,2] and Michael J. Groves[1,3]**

*Received November 2, 1992; accepted December 1, 1992*
KEY WORDS: phospholipids; thermal degradation; triglyceride emulsions; pH; free fatty acids.

## INTRODUCTION

The intravenous administration of triglyceride emulsions stabilized with phospholipid emulsifiers has been employed for parenteral nutrition for over 30 years (1). Terminally heat sterilized, these systems are required to be physically and chemically stable in order to avoid harming the patient (2). Nevertheless, slow hydrolysis of the phospholipids is known to occur after the initial sterilization-induced degradation (3).

It is implicitly assumed that the subsequent fall of product pH is due to the formation of free fatty acids (4,5). We have recently demonstrated that the initial hydrolysis of the phospholipids during the heat sterilization process paradoxically promotes physical stabilization of the emulsion system, most probably because of the formation of liquid crystalline structures at the oil/water interface (3,6). The principal degradation process is due to the hydrolysis of the diacylphosphatidylcholines and diacylphosphatidylethanolamines to their corresponding monoacyl (lyso-) derivatives and free fatty acid (FFA) moieties. In turn, the lyso derivatives can degrade to the corresponding glycerophosphoryl compounds, with the formation of additional FFA. FFA can also be formed by the hydrolysis of emulsified triglycerides to the corresponding mono- and diglycerides, although this reaction is believed to be relatively slow compared to the breakdown of the diacylphosphatidyl derivatives (3).

The emulsion systems are unbuffered and the formation of FFA will inevitably lower the pH from the initial value of 8.0 (1) over a period of time poststerilization. Håkansson (2) demonstrated that the degradation rate decreases until pH 6.5 is reached, after which there is again an acceleration of the degradation process. This effect has been confirmed by Grit *et al.* (7). Stabilization may, therefore, be improved by the addition of extraneous FFA, a suggestion made by Washington and Davis (4), who evaluated the effect induced by the addition of oleic acid to their emulsion systems. However, the value of this suggestion is unclear since, during the phospholipid hydrolysis process, lyso compounds are produced in addition to FFA and these materials also contribute to the emulsion stabilization process (3,6).

Measurement of pH could, therefore, provide an indirect method of determining the FFA content of phospholipid-stabilized emulsions. This concept was evaluated measuring the FFA content directly by potassium hydroxide titration of the degrading emulsion at the same time as taking the pH of the system with a glass electrode.

## MATERIALS AND METHODS

### Materials

Purified egg phospholipid, Asahi Injectable grade (lot 900-80201), was received as a gift from Austin Chemical, Chicago, IL. Dipalmitoyl phosphatidylethanolamine (Lot 1 60PE-45) (DPPE) and hydrogenated egg phosphatidylcholine (Lot HEPC – 44) (HEPC) were purchased from Avanti Polar Lipids, Birmingham, AL. Pharmaceutical-grade and "superrefined"-grade soybean oils USP were received as gifts from Croda Inc., Edison, NJ. Glycerol, sodium hydroxide, potassium hydroxide, and potassium hydrogen phthalate were all used as received from Fisher Scientific, Itasca, IL. A Milli-Q ion-exchange water system was used.

### Model Emulsion

The model emulsion was made to the following formula: soybean oil, 20 g; egg phospholipid, 1.2 g; glycerol, 2.25 g; and water to 100 mL.

The egg phospholipid was dispersed in the glycerol and about 95% of the water at 70°C, the oil mixed in, and the coarse emulsion passed through a Microfluidics Model 110T homogenizer at a pressure of 10,000 psig for a total of 10 times to ensure minimal particle size (8). The system was washed through the homogenizer and made up to volume with water, and the pH adjusted to 8.0 with 0.01 *N* sodium hydroxide and packed and sealed in 2-mL volumes in 2-mL glass ampoules. Gas sparging with nitrogen or oxygen was carried out for 15 min when required. Hydrogenated phospholipid-stabilized emulsions were prepared by the same method, using 0.24 g DPPE and 0.96 g HEPC instead of the 1.2 g egg phospholipids.

Samples were stressed and analyzed in replicate; $n = 4$.

### pH Measurement

An Orion Model 811 glass electrode pH meter was used. Potassium chloride was added to counteract the ion adsorbing effects of charged droplets (9), a process also used in the USP XX11 for the measurement of the pH of Dextrose solutions. After experimentation to determine conditions required to give reproducible results, measurements were made following the addition of 50 μL of a saturated aqueous potassium chloride solution to 2 mL of emulsion sample.

### Free Fatty Acid Measurement

FFA measurement of a solution of the emulsion sample

[1] Institute for Tuberculosis Research, College of Pharmacy, University of Illinois at Chicago (M/C 964), 840 West Taylor (2014 SEL), Chicago, Illinois 60607.
[2] Present address: Mallinckrodt Specialty Chemicals Company, P.O. Box 5439, St. Louis, Missouri 63147.
[3] To whom correspondence should be addressed.

0724-8741/93/0500-0774$07.00/0 

was by direct titration to neutrality with 0.01 *N* potassium hydroxide solution standardized using potassium hydrogen phthalate, with phenolphthalein as indicator (USP XXII). The IUPAC method (10) specifies ethanol:diethyl ether (95:5) as the solvent of choice. However, for safety reasons, the solvent was changed to ethanol:chloroform (2:1), each titration being preceded by bringing the solvent to neutrality with 0.01 *N* potassium hydroxide solution.

### Thermal Stress

Thermal stress was applied by filling heating block chambers (Dry Baths, Fisher Scientific, Itasca, IL; 60 chambers per block, each 12 mm diameter and 50 mm deep) with oil and immersing the 2-mL ampoules containing the emulsion at the desired temperature, covering the blocks with aluminum foil to minimize thermal fluctuation. Temperatures were determined with calibrated mercury-in-glass thermometers placed at random in the block chambers.

In general, all ampouled emulsion samples were initially sterilized at 121°C ($F_o$ = 18) using a Getinge BioF$_a$OE autoclave. Unautoclaved controls were stored at 5°C prior to evaluation.

## RESULTS AND DISCUSSION

The rate of pH change and the rate of FFA formation are compared in Fig. 1 for a model emulsion prepared with the pharmaceutical grade of soy oil. Results obtained using the "superrefined" grade of oil were similar. It is evident that the rate of formation of FFA is slower than the rate at which the pH drops, suggesting that some other factors are involved. As shown in Fig. 2, emulsions prepared with unsaturated or saturated acyl groups on the phospholipid moieties and sparged with nitrogen prior to sterilization and storage changed pH at almost identical rates. However, it should be noted that these two emulsions had markedly different hydrolysis rates, (3), which suggested that the lowering of

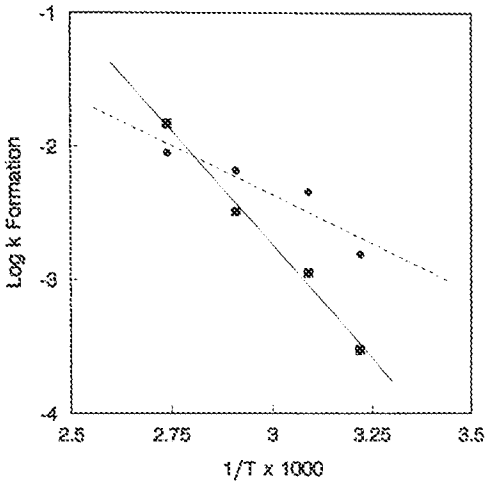



Fig. 1. The rate of FFA formation (by KOH titration) and the fall of pH (glass electrode) in a 20% soy oil emulsion stabilized with egg lecithin after sterilization and storage over the range 25–90°C. pH (slope = −3.42), ·····▩·····; FFA (slope = −1.47), —◉—. *K*, rate of formation (slope of change against time); T, absolute temperature (kelvins).

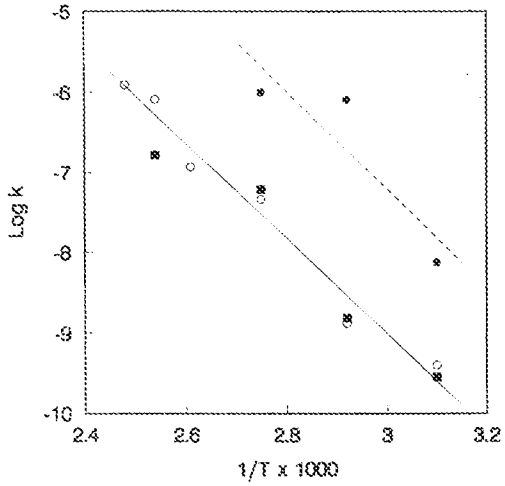



Fig. 2. The rate of pH change in a 20% soy oil emulsion stabilized with egg lecithin or saturated phospholipids and sparged with nitrogen or oxygen before sterilization and storage over the range 25–90°C. Control: egg phospholipids, nitrogen sparged (estimated slope = −5.93), ——○——; saturated phospholipids, nitrogen sparged (estimated slope = −5.30), ▩; egg phospholipids, oxygen sparged (estimated slope = −6.11), ·—◉—·.

pH was due to some other factor not necessarily associated with phospholipid hydrolysis. As noted earlier, free fatty acids can also arise by hydrolysis of triglycerides but the total (titratable) FFA in the system, irrespective of its source, is clearly unable to account for the relatively rapid lowering of pH.

When sparged with oxygen, the pH change in the control emulsion was considerably increased (Fig. 2), and this observation suggests that the effect of oxygen on the emulsion pH was more pronounced than that produced by the heat-induced hydrolytic degradation reaction.

The solubility of gases in liquids is described by Henry's law (11):

$$P_A = X_A K_a$$

where $P_A$ is the vapor pressure of a solution containing solute A, $X_A$ is the mole fraction of A, and $K_a$ is Henry's law constant. Thus, intuitively, since the $K_a$ for gases in nonaqueous solvents is generally higher than the corresponding value in water, it would appear that the oxygen is likely to preferentially dissolve in the oil phase of the emulsion. Atkins (11) discussed this issue in relation to benzene at standard temperature and pressure and observed that, in all cases, the gas was more soluble in the benzene than in the water.

Based on this consideration, it seems feasible to suggest that residual oxygen may remain dissolved in the triglyceride oil phase after preparation and manipulation of the emulsion. Some of the triglycerides contain unsaturated acyl centers, which could, therefore, become partially oxidized, although not necessarily to the point where they would be titratable with alkali and estimated as FFA. These moieties would affect the pH of the unbuffered system, in addition to the effects produced by the FFA resulting from hydrolysis of phospholipid and triglyceride entities. The evident complexity of the physical and chemical structure of phospholipid-

stabilized emulsions makes this suggestion difficult to confirm at present.

## ACKNOWLEDGMENTS

Our joint thanks are due to Mr. Peter Cade and his colleagues at Croda (US) Inc. for partial support on this project.

## REFERENCES

1. P. K. Hansrani, S. S. Davis, and M. J. Groves. The preparation and properties of sterile intravenous emulsions. *J. Parent. Sci. Technol.* 37:145–150 (1983).
2. I. Håkansson. Physico-chemical changes in artificial fat emulsions during storage. *Acta Chem. Scand.* 20:2267–2281 (1966).
3. C. J. Herman. *The Influence of Thermal Stress on the Properties of Phospholipid Stabilized Emulsions*, Ph.D. thesis, University of Illinois at Chicago, Chicago, 1992.
4. C. Washington and S. S. Davis. Ageing effects in parenteral fat emulsions: The role of fatty acids. *Int. J. Pharm.* 39:33–37 (1987).
5. J. M. A. Kemps and D. J. A. Crommelin. Chemical stability of phospholipids in pharmaceutical preparations: Hydrolysis of phospholipids in an aqueous medium. *Pharm. Weekblad.* 123:355–363 (1988).
6. C. J. Herman and M. J. Groves. Hydrolysis kinetics of phospholipids in thermally stressed intravenous lipid emulsion formulations. *J. Pharm. Pharmacol.* 44:539–542 (1992).
7. M. Grit, J. H. de Smidt, A. Struijke, and D. J. A. Crommelin. Hydrolysis of phosphatidylcholine in aqueous liposome dispersions. *Int. J. Pharm.* 50:1–6 (1989).
8. D. M. Lidgate, R. C. Fu, and J. S. Fleitman. Using a microfluidizer to manufacture emulsions. *BioPharm.* 45:28–33 (1989).
9. G. Lee, D. Dick, E. V. Vasquez, and K. Werner. pH measurements of suspensions. In M. H. Rubenstein (ed.), *Pharmaceutical Technology: Drug Stability*, Ellis Horwood, Chichester, UK, pp. 113–117.
10. C. Paquot. Determination of the acid value and the acidity. In *Standard Methods for the Analysis of Oils, Fat, and Derivatives*, Pergamon, New York, 1979, pp. 52–55.
11. P. W. Atkins. *Physical Chemistry*, W. H. Freeman, New York, 1986.

# EXHIBIT F

Winther *et al.*,
Elucidation of Phosphatidylcholine Composition in Krill Oil Extracted from *Euphausia superba*
*Lipids* 46 (1): 25-36 (2011)

Lipids (2011) 46:25–36
DOI 10.1007/s11745-010-3472-6

ORIGINAL ARTICLE

# Elucidation of Phosphatidylcholine Composition in Krill Oil Extracted from *Euphausia superba*

**Bjørn Winther · Nils Hoem · Kjetil Berge · Léon Reubsaet**

Received: 9 June 2010 / Accepted: 30 August 2010 / Published online: 17 September 2010


**Abstract** High performance liquid chromatography-electrospray tandem mass spectrometry was used to elucidate the phospholipids in krill oil extracted from *Euphausia superba*, an emerging source for human nutritional supplements. The study was carried out in order to map the species of the choline-containing phospholipid classes: phosphatidylcholine and lyso-phosphatidylcholine. In addition, the prevalent phosphatidylcholine class was quantified and the results compared with prior analysis. The qualification was performed with separation on a reverse phase chromatography column, while the quantification was obtained with class separation on a normal phase chromatography column. An Orbitrap system was used for the detection, and pulsed-Q dissociation fragmentation was utilized for the identification of the species. An asymmetrical exclusion list was applied for detection of phospholipid species of lower concentration, significantly improving the number of species observed. A total of 69 choline-containing phospholipids were detected, whereof 60 phosphatidylcholine substances, among others seven with probable omega-3 fatty acids in both *sn*-1 and *sn*-2. The phosphatidylcholine concentration was estimated to be $34 \pm 5$ g/100 g oil ($n = 5$). These results confirm the complexity of the phospholipid composition of krill oil, and the presence of long chained, heavily unsaturated fatty acids.

B. Winther · L. Reubsaet (✉)
Department of Pharmaceutical Chemistry,
School of Pharmacy, University of Oslo, Oslo, Norway
e-mail: j.l.reubsaet@farmasi.uio.no

N. Hoem · K. Berge
Aker BioMarine ASA, Fjordalléen 16, Vika,
P.O. Box 1423, 0115 Oslo, Norway

**Keywords** Fish oil · Krill oil · Mass spectrometry · Omega-3 · Phosphatidylcholine · Phospholipid

**Abbreviations**

| | |
|---|---|
| EPA | Eicosapentaenoic acid |
| DHA | Docosahexaenoic acid |
| lyso-PtdCho | Lyso-phosphatidylcholine |
| NPLC | Normal phase liquid chromatography |
| PtdCho | Phosphatidylcholine |
| PtdEtn | Phosphatidylethanolamine |
| PtdIns | Phosphatidylinositol |
| PtdSer | Phosphatidylserine |
| PL | Phospholipid |
| RPLC | Reverse phase liquid chromatography |

## Introduction

Krill oil has emerged as an important source of omega-3 fatty acids for human consumption during the last decade, and the amount sold on the world market is rapidly increasing. In contrast to traditional omega-3 supplements on today's market, which are based on omega-3 fatty acids bound to triglycerides (such as cod liver oil and fish oil) or bound as ethyl esters (Omacor/Lovaza), krill oil contains a high proportion of omega-3 fatty acids bound to phospholipids.

Krill oil has been investigated in several preclinical and clinical studies [1–4], and there is growing evidence that the molecular form of the omega-3 fatty acids (i.e. triglycerides, ethyl-esters, phospholipids) might be of importance for their biological effect as well as distribution of the omega-3 fatty acids in the body. In one animal study,

it was demonstrated that when krill oil and fish oil were administrated to Zucker rats with an equimolar dose eicosapentaenoic acid (EPA) + docosahexaenoic acid (DHA), krill oil had stronger and in some instances different effects than fish oil on specific parameters related to the metabolic syndrome [1]. The lipid level in both heart and liver was significantly lower in rats treated with krill oil, when compared to rats fed the fish oil diet. The authors suggest that this difference may be linked to differences in the incorporation of omega-3 fatty acids into membranes, and consequently a reduction of inflammatory molecules and endocannabinoids, which might be relevant for the differences observed between fish oil and krill oil. Further, in the same study, it was demonstrated that the level of DHA in the brain increased significantly after krill oil administration, but not after fish oil administration, when compared to control animals [2]. Thus, omega-3 fatty acids linked to phospholipids may be differently distributed in the body compared to omega-3 fatty acids in other molecular forms. Moreover, in a clinical safety study, the presence of EPA and DHA in the blood plasma was determined after daily administration of 2 g krill oil or 2 g menhaden oil for 4 weeks [3]. The authors concluded that EPA and DHA from krill oil are absorbed at least as well as that from menhaden oil.

The aim of the current study was to characterize the phospholipids in krill oil in more detail to evaluate the composition of the fatty acids present in the phospholipids. The composition was determined using LC/ESI–MS(/MS), a technique which has lately played an important role in characterization of the lipidome in tissues and organisms [5]. An inherent limitation in the use of ESI for the ionization of long chained fatty acids has been described by Koivusalo et al. [6]. The study showed that the instrument response is affected by the acyl chain length. This is a consideration which is important particularly in the quantification of the lipids.

The elucidation of the phospholipid species is often performed either by doing a precursor ion scan or a neutral loss scan with triple quadrupole instrumentation [7–9], or with $MS^n$ fragmentation with systems based on ion traps [10–13]. Normal phase liquid chromatography (NPLC) and reverse phase liquid chromatography (RPLC) are both frequently used for the separation of the components [10, 14–17]. Of these two separation techniques, RPLC has been shown to be more suitable for species separation and characterization [8].

Different ionization and fragmentation techniques can be used for the evaluation of phospholipids. Ionization of the phospholipids may be performed in negative- and positive-ionization mode. In general, fragmentation of phospholipids in the positive mode provides information about the phospholipid head group, while fragmentation in the negative mode is the source of structural information. For phospholipids containing choline-headgroups, the choline-specific fragment *m/z* 184 has been used in precursor ion scanning operating in the positive ionization mode for class determination [7, 18]. Also in the negative mode, class-specific fragments may be used in the characterization. All phospholipid classes, except those containing choline, yield molecular ions $[M-H]^-$ when a formate-based mobile phase is used. On the other hand, the choline-containing classes form stable adducts with formic acid in the mobile phase, yielding $[M + FA-H]^-$ ions ($m/z$ = M + 45) [19, 20]. With fragmentation, this adduct dissociates with the loss of ($HCOO + CH_3$) into the fragment ion $[M-CH_3]^-$. This is in particularly useful in methods utilizing RPLC for separation. Although the chromatographic class information is lost in such setups, the class-specific fragments may be used in the characterization of the species [11].

Two ion activation techniques may be used for MS analysis utilizing ion traps: collision-induced dissociation (CID) and pulsed-Q dissociation (PQD) techniques. While CID has a low mass cut off below 28% of the *m/z* for the precursor ion, the novel PQD technique eliminates the potential loss of low mass fragments [21, 22]. This difference could be crucial in the fragmentation of larger molecules into low mass, specific fragments, as shown with detection of iTRAQ fragments with a linear ion trap [23].

The fatty acid composition of phosphatidylcholine (PtdCho) from krill oil has previously been investigated by Le Grandois et al. [24]. This study was performed with a method based on the ESI operated in the positive mode with triple quadrupole detection of lithium adduct ions, and showed the presence of a higher number of PtdCho species with long chained unsaturated fatty acids, than seen in egg yolk, ox liver and soy.

We believe the current study verifies previously presented findings and offer new insights into the composition of krill oil. In addition; it shows the advantage of performing an additional fragmentation using an exclusion list in the identification of low prevalent species.

## Experimental Procedures

### Chemicals

Phospholipid standards of lyso-phosphatidylcholine (lyso-PtdCho), PtdCho, phosphatidylethanolamine (PtdEtn), phosphatidylinositol (PtdIns) and phosphatidylserine (PtdSer) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Lyso-PtdCho, PtdCho and PtdEtn were lyophilized powders obtained from egg yolk, whereas the PtdIns source was *glycine max* and the PtdSer source was

bovine brain. EPAX 6000 TG® fish oil was donated by EPAX (Ålesund, Norway), and Superba™ krill oil was obtained from Aker BioMarine (Oslo, Norway). All other chemicals were of MS grade.

### Instrumentation

The chromatography was carried out on a Dionex system consisting of an Ultimate 3000 pump, an Ultimate 3000 RS autosampler, and an Ultimate 3000 flow manager. Detection was obtained using a linear ion trap LTQ XL coupled to an Orbitrap Discovery, LC-operation, data acquisition and processing were carried out using Chromelion SDK 6.80 SP2 Build 2327 and Xcalibur version 2.0.7 coupled with DCMS$^{Link}$ 2.5 (all Instrument-Teknikk AS, Østerås, Norway).

### Mass Spectrometry

The LTQ Orbitrap system was operated with a spray voltage of 5.00 kV, nitrogen as the sheath gas with flow rate set to 30 arbitrary units, and helium as the collision gas. The quantification of the PtdCho class was performed with a scan from $m/z = 400$ to $m/z = 1{,}000$ operated in negative ionization mode. $MS^n$ experiments for identification of the choline-containing phospholipids were performed using data dependent PQD for the first fragmentation step. The molecular ion selected for each fragmentation in this step, was the most intense ion detected by the Orbitrap analyzer with target mass resolution of 30,000 and a scan window from $m/z = 400$ to $m/z = 1{,}000$. The normalized collision energy was 200 and the isolation width 2.00 Da. Subsequently, the most intense fragment ion detected was further fragmented using CID, with the normalized collision energy at 35 and an isolation width of 2.00 Da ($MS^3$). The LTQ was utilized for the detection of the fragments and the $m/z$ range was relative to the $m/z$ of the molecular ion. An alternative method was used in order to be able to observe species that were not selected for fragmentation in this way. The overall setup of this method was as described above, with the distinction of adding an asymmetric exclusion list. The exclusion list was generated with the purpose of the LTQ to ignore already identified substances. The list was based on the $m/z$ of the molecular ions, with an exclusion window from this mass-to-charge ratio, up to $m/z + 1$. The width of the exclusion window was selected in order to diminish the detection of isotopes of the molecular ions.

### Chromatographic Conditions RPLC

Chromatographic separation was performed on a ZORBAX Eclipse Plus C18 column with particle diameter of 5 μm and the column dimensions were 150 × 2.1 mm i.d.

The mobile phase A consisted of 90 parts 1% TEA and 0.2% formic acid in water, and 10 parts mobile phase B (*v/v*). Mobile phase B consisted of 1% TEA and 0.2% formic acid in 60 parts methanol and 40 parts acetonitrile (*v/v*).

A linear gradient was used for the separation. The system was first kept isocratic at 65% mobile phase B for 5 min after injection of sample. The gradient was then run from 65 to 100% mobile phase B in 5 min and was kept isocratic at 100% mobile phase B for 20 min, before it was returned to the initial condition in 0.1 min. The column was regenerated with 65% mobile phase B for 16 min. The mobile phase flow was set to 0.2 mL/min and the injection volume was 20 μL throughout the study.

### Chromatographic Conditions NPLC

NPLC was performed on a HiCHROM LiChrospher 100 DIOL column with a particle diameter of 5 μm and column dimensions of 250 × 2 mm i.d. Mobile phase C was 100% chloroform, and the mobile phase D consisted of 0.05% TEA, 0.05% ammonia and 0.1% formic acid in methanol (*v/v*). For the class separation of the phospholipids, a linear gradient was used. The gradient was run from 5 to 27.5% mobile phase D in 15 min, followed by a rise to 80% in 2 min to flush the column. This concentration was kept isocratic for 4 min, before it was returned to the initial condition in 2 min. The column was regenerated with 5% mobile phase D for 12 min. The mobile phase flow was set to 0.3 mL/min, and the injection volume was 20 μL throughout the study.

### Sample Preparation

Samples of krill oil and stock solutions of standards were prepared by dissolving the lipids in a mixture of chloroform and methanol at a ratio of 2:1. These solutions were stored at −32°C and excessive heating cycles were avoided. Samples were prepared by further dilution with solvents compatible with the mobile phases used. For NPLC, this was achieved with chloroform:MeOH 95:5, while it was attained by dilution in mobile phase A for RPLC.

### Calibration Curve

For the quantification, a calibration curve was established with samples of PL free fish oil (EPAX®) spiked with a PtdCho standard purified from egg yolk to concentrations of 100 μg/mL. The spiking of PL free fish oil was performed in order to produce comparable matrixes in the standards and the krill oil samples. Stock solutions were made by dissolving PtdCho standard, PL free fish oil and krill oil separately in mixtures of chloroform:MeOH 2:1. The concentration of krill oil and PtdCho standard was

1 mg/mL and for PL free fish oil 10 mg/mL for these solutions. Respectively, 100 µL of PtdCho standards and of krill oil was added to 900 µL of the PL free fish oil, producing samples with concentrations of 100 µg/mL. For the calibration curve, the PtdCho standard samples were consecutively diluted to the desired concentrations of 10.0, 5.00, 2.50, 1.00, 0.50, 0.25, and 0.10 µg/mL ($n = 5$) with a mixture of chloroform:MeOH 95:5. The krill oil samples were diluted in the same way to a concentration of 1.00 µg/mL in order to measure the PtdCho content within the linear area of response of the calibration curve.

## Results

### Selection of MS-Mode for PtdCho-Classification

Initial experiments with standards of PtdEtn, PtdIns, PtdSer, PtdCho and lyso-PtdCho, were performed in both positive- and negative- ion mode using the LTQ Orbitrap as a detector. Separation of these substances on a NPLC column yielded class-specific baseline separation (data not shown). The results of these tests indicated a minor difference in the signal intensities obtained between the two settings, with a slightly higher response in negative mode.

### Identification of Choline-Containing Phospholipids in Krill Oil

Since krill oil, as established in Fig. 1, is predominantly composed of choline-containing phospholipids, the emphasis of the work was focused on the elucidation of the species in the PtdCho and lyso-PtdCho classes. In NPLC, class separation of the phospholipids is achieved. A clear tendency of the elution order from the column was seen from high *m/z* to lower *m/z*. PtdCho class species eluted from 6.5 to 8 min, and the elution of the lyso-PtdCho class occurred between 10.5 and 12 min. Some species separation

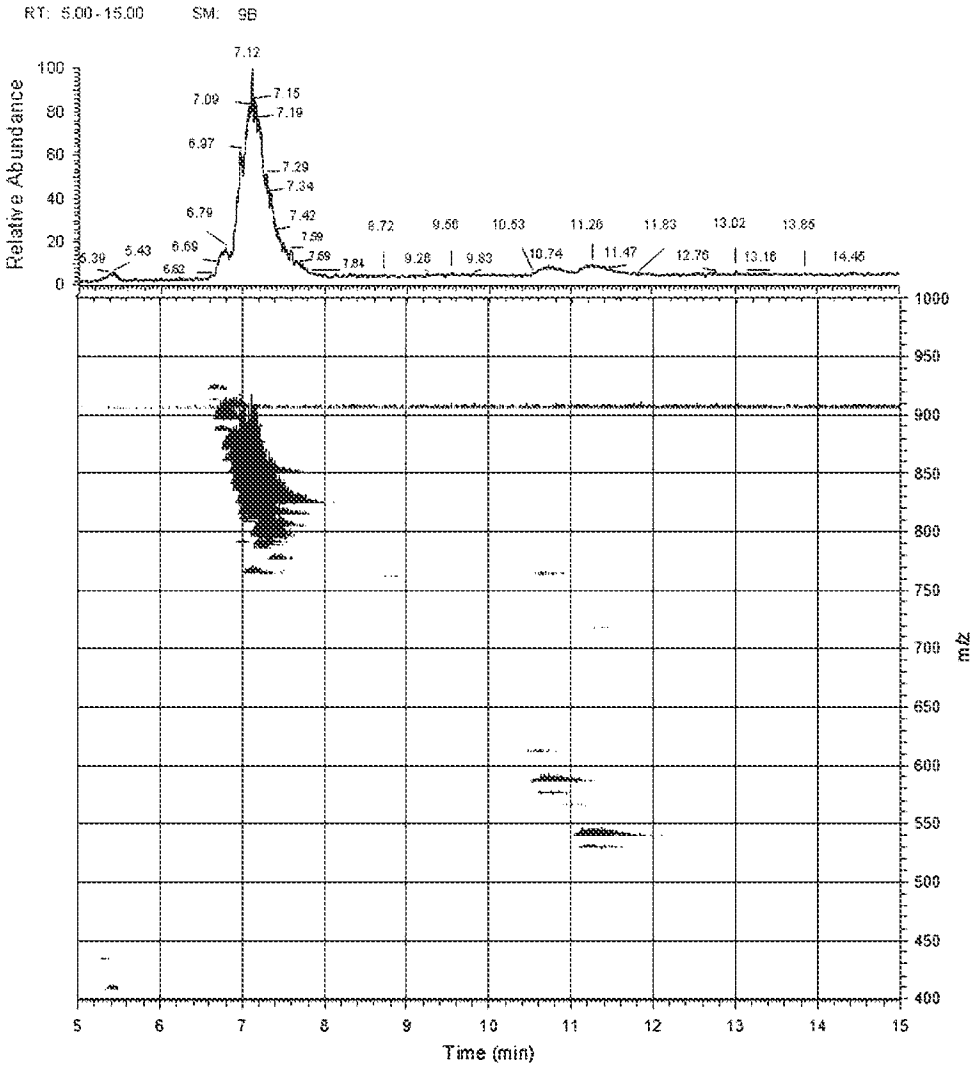



**Fig. 1** Base peak chromatogram and three dimensional (*3D*) map of phospholipid class separation of krill oil, performed with a linear gradient NPLC/ESI–MS on a HiCHROM LiChrospher 100 DIOL column (250 × 2 mm i.d., 5 µm). MS was operated in the negative ionization mode and set to scan *m/z* 400–1,000. Only the relevant part of the chromatogram is shown (5–15 min). For the major part of the species, the map show adduct ions in the form of $[M + FA–H]^-$

was seen within the PtdCho class, however, this was not adequate for identification of the diverse species within the classes.

As the identification of the components is performed with data dependent fragmentation, chromatographic separations of the substances are critical for detection of the less prominent species. Hence, the separation for characterization of the species was performed utilizing a RPLC C18 column as described under "Experimental Procedures". This improved the chromatographic performance for species separation compared to NPLC (Fig. 2). As lyso-PtdCho only carry one fatty acyl group, these components elute earlier in the chromatogram than the PtdCho species. Lyso-PtdCho dominate the region between 12.0 min and 14.5 min in the chromatogram, while the heavier PtdCho components dominate the chromatogram after 14.5 min.

The identification of the species was performed utilizing a $MS^3$ data dependent fragmentation method, with an initial PQD fragmentation resulting in the loss of methyl formate, followed by a CID fragmentation. Analysis of the fragmentation spectra obtained typically revealed the identity of the substances without ambiguity. However, co-elution of isobaric compounds could potentially complicate the interpretation of the spectra. This challenge is minimized by applying a set of criteria for reliable identification. The following criteria were applied for reliable interpretation of a choline-containing phospholipid: based on the mass-to-charge ratio of the molecular ion, it is likely to be a choline-containing phospholipid (i.e. *m/z* being an even number). Following the first CID with PQD, a daughter ion should be produced by the loss of ($HCOO + CH_3$) as 60 Da. Further fragmentation with CID of the resulting product should produce specific fragments revealing the nature of the fatty acyl groups in both the *sn*-1 and the *sn*-2 position, either by the occurrence of the fragment for the fatty acyl group itself, or by the presence of the corresponding fragments of the lyso-compound. The sum of the fatty acyl groups elucidated in this matter should yield a mass matching the initial molecular mass. This is illustrated in Fig. 3, showing the elucidated fragment identity for the fragmentation of the 20:5–22:6 diacyl PtdCho.

The spectra were generally dominated by fatty acyl fragments from both the *sn*-1 and the *sn*-2 positions in addition to their corresponding fragments of the lyso-compound, ensuring identification of the species. The lyso-PtdChoand PtdCho substances identified by applying this method are presented in Tables 1 and 2. The relative intensity of the molecular ions is also presented.

As described earlier, the use of signal intensities in MS, for concentration comparison of the different substances, is only semi-quantitative. However, it provides a valuable indication of the composition of the PtdCho and lyso-PtdChoclasses. Chromatograms and fragmentation patterns are presented in Figs. 4 and 5 for the 10 foremost substances characterized from the PtdCho class.

As data dependent fragmentation methods are, by nature, biased in the selection of the most prevalent substances, the experiments where repeated with the use of an asymmetric exclusion list added to the MS-method. The exclusion list was based on the data sets obtained with the initial settings (i.e. Tables 1, 2). This method allowed the detection and identification of the additional substances presented in Table 3.

### Quantification of PtdCho-Class in Krill Oil

Krill oil predominantly contains phospholipids from the PtdCho class (Fig. 1). It was therefore attempted to quantify the absolute concentration of this class by use of class separation with NPLC. Quantification of the PtdCho class was performed with a method developed "in-house" with a LTQ Orbitrap mass spectrometer for the detection.

In the construction of the calibration curve, PtdCho concentrations above 1.00 μg/mL resulted in a relative decrease in the MS signal response, yielding a quadratic polynomial curve ($y = ax^2 + bx + c$) where $a = -9{,}766$, $b = 330{,}360$ and $c = 4763.8$ with $r^2 = 0.9995$. From 0.10 μg/mL to 1.00 μg/mL, the calibration curve showed a high degree of linearity ($r^2 = 0.9995$) with a linear regression curve ($y = bx + c$) where $b = 368{,}737$ and $c = -16{,}547$.

The latter area was chosen for quantification purpose. From this, the PtdCho content of the undiluted krill oil was determined to be 34 ± 5% (*w/w*) ($n = 5$). Comparisons of the mass spectrum of the sample with the spectrum of the PtdCho standard indicated a difference in PtdCho class composition (Fig. 6). From the results, the average acyl chain lengths appear to be higher in krill oil than in egg yolk. This has also previously been shown by others [24]. As mentioned above, instrument response is affected by the acyl chain length of the PtdCho. These differences in chain lengths could therefore influence the quantification of the PtdCho class as discussed later. The quantitative results were compared with an earlier analysis of the krill oil, performed by the accredited analytical company Nofima (Bergen, Norway). They reported the PtdCho concentration in the krill oil sample to be 35 g/100 g oil.

## Discussion

The fact that fatty acyl chain lengths of the PtdCho species are relatively long, affects both the choice of fragmentation technique and the effect of standards used for quantification purposes. Due to the low mass cut-off limit at 28% of the molecular ion mass with CID fragmentation in ion

traps, fragmentation utilizing PQD in the positive ionization mode was chosen as the first fragmentation step. In addition, utilizing PQD in the first fragmentation step of a $MS^3$ method, operating with negative mode ionization, also yields class elucidation of phospholipids with choline head groups. This is achieved by the detection of the $[M-CH_3]^-$ fragment formed after ($HCOO + CH_3$) loss from formate-molecular ion adducts. Consequently, performing a mode shift from the positive to the negative ionization mode is not necessary for the overall identification of lyso-PtdCho and PtdCho class phospholipids. PQD with a normalized collision energy at 200 produced the $[M-CH_3]^-$ fragments, without extensive fragmentation to secondary fragment ions. Higher collision energies produced secondary fragments that could be used as a source of structural information; however, this was better achieved with CID as a second fragmentation step. The CID was operated with a normalized collision energy of 35, for further fragmentation of the $[M-CH_3]^-$ ion. This value was not optimized for the individual species, and additional structural information could potentially have been achieved by specie specific optimization of this setting.

With the described method for separation and fragmentation of the phospholipids, typical fragmentation patterns were obtained, as shown in Fig. 3. The spectra were dominated by fatty acyl fragment ions originating

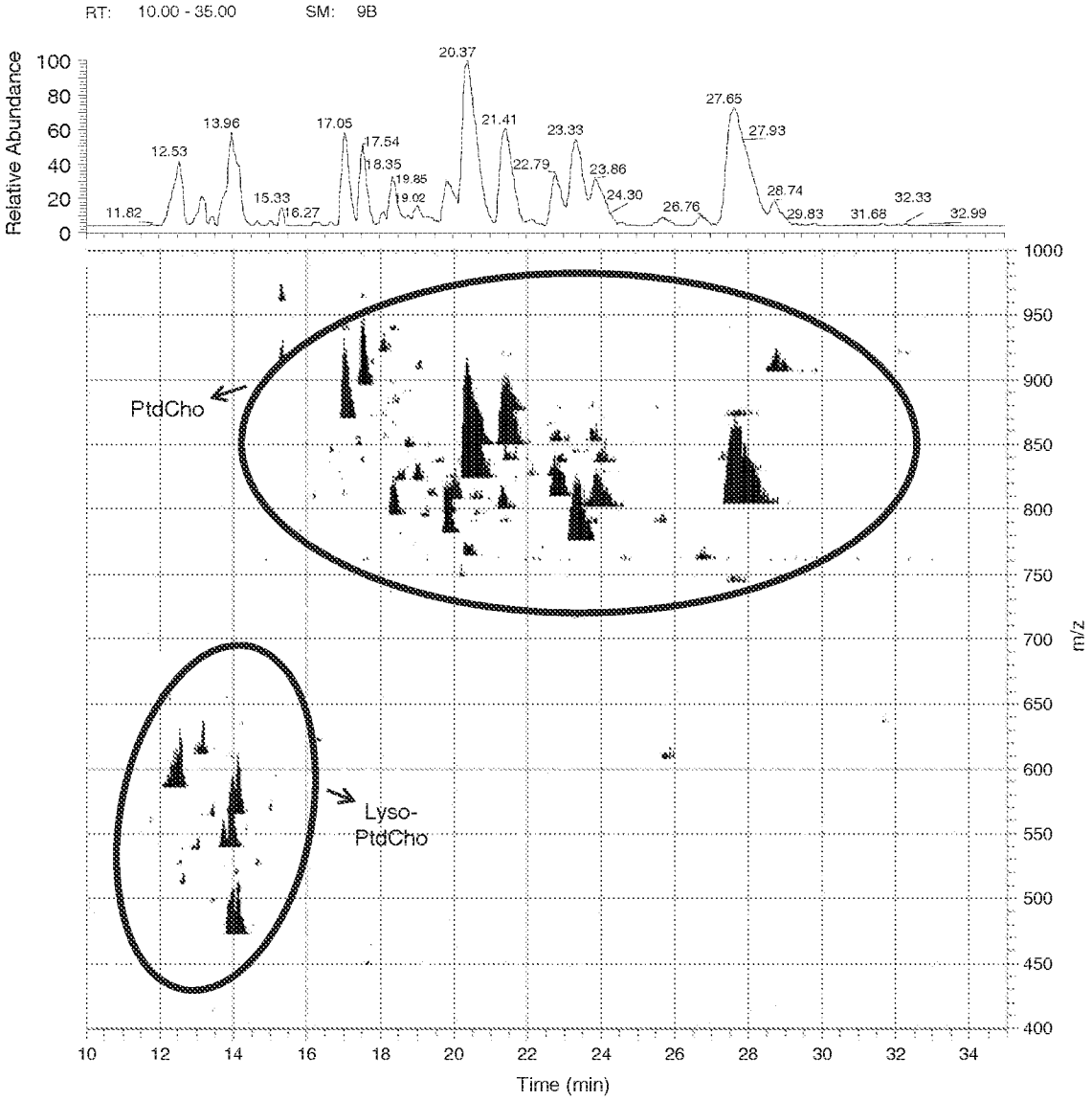



**Fig. 2** Base peak chromatogram and three dimensional (*3D*) map of phospholipid specie separation of krill oil, performed with RPLC/ESI–MS on a ZORBAX Eclipse Plus C18 column (150 × 2.1 mm i.d., 5 μm). MS was run in negative the ionization mode and set to scan from *m/z* 400 to *m/z* 1,000. Only the most relevant part of the chromatogram is shown (10–35 min). Adduct ions in the form of $[M + FA-H]^-$ is seen throughout the map

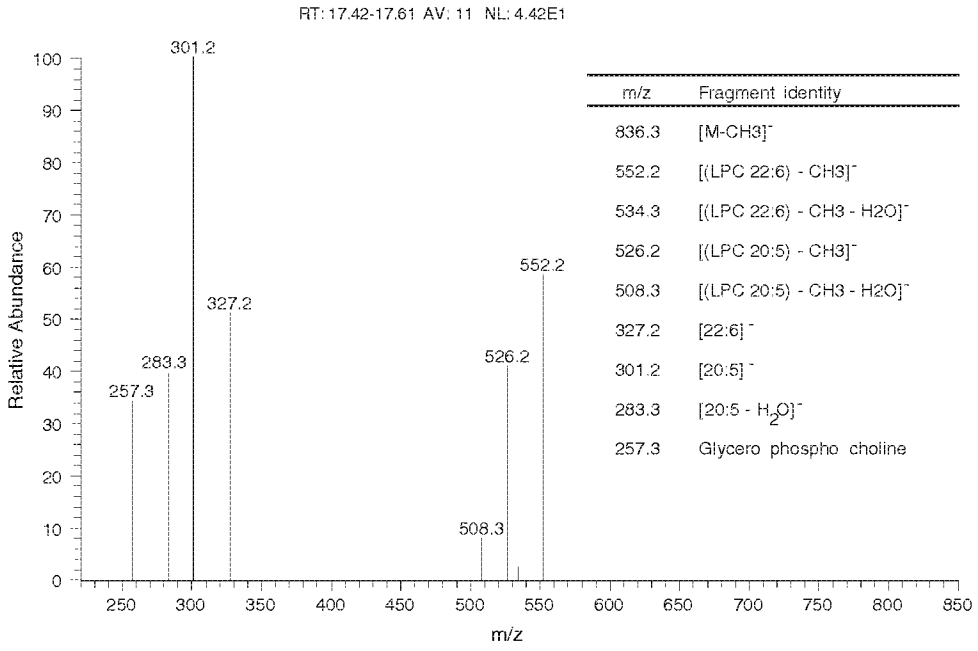



**Fig. 3** $MS^3$ product ion spectrum of 22:6–20:5 diacyl PtdCho obtained in negative ionization mode. The molecular ion of *m/z* 896.6 was selected for PQD fragmentation; this yielded a fragment ion of *m/z* 836.3 which was further fragmented with CID resulting in the presented spectrum. Fragment identity is explained in the table on the *right hand side*, indicating the high level of certainty in the characterization

**Table 1** Identified lyso-phospholipid species with choline head group in krill oil

| Class | Mass | *m/z*[a] | Molecular specie | Relative intensity |
|---|---|---|---|---|
| Lyso-PtdCho | 493.4 | 538.4 | 16:1 | 6.01 |
| Lyso-PtdCho | 495.4 | 540.4 | 16:0 | 32.42 |
| Lyso-PtdCho | 509.4 | 554.4 | 17:0 | 4.06 |
| Lyso-PtdCho | 521.4 | 566.4 | 18:1 | 23.21 |
| Lyso-PtdCho | 541.4 | 586.4 | 20:5 | 31.31 |
| Lyso-PtdCho | 563.4 | 608.6 | 21:1 | 16.70 |
| Lyso-PtdCho | 567.4 | 612.4 | 22:6 | 12.19 |

Data were obtained with RPLC/ESI–$MS^3$ operated in negative ionization mode and with data dependent fragmentation without exclusion list

[a] *m/z* for $[M + FA–H]^-$ adduct

from both *sn*-1 and *sn*-2, except in the fragmentation of alkyl–acyl species, where only a single fatty acyl fragment ion was observed. In addition, the corresponding fragment ions of the lyso-PtdCho compound were prevalent, confirming the characterization of the species. The fragments associated with the lyso-PtdCho compounds were either in the form of $[lyso\text{-}PtdCho–CH_3]^-$ or $[lyso\text{-}PtdCho–CH_3–H_2O]^-$. Furthermore, fragments specific for phospholipids carrying a choline head group were often registered. These dissociation products were *m/z* 257, *m/z* 242 and *m/z* 223, representing [Glycero phospho choline]$^-$, $[Glycero\ phospho\ choline–CH_3]^-$ and $[Glycero\ phospho\ choline–CH_3–H_2O]^-$, respectively. The identification of the phospholipid class was made based on the loss of 60 Da in the PQD fragmentation step. However, fragments specific for phospholipids carrying a choline head group affirms this interpretation. In spectra obtained from the dissociation of species carrying a 20:5 fatty acyl group, an ambiguous $[20:5–H2O]^-$ fragment ion with *m/z* 283 can often be detected. This ion can potentially be misinterpreted as the fatty acyl fragment $[18:0]^-$, but meeting the criteria for reliable identification will rule out this erroneous conclusion.

In some incidences, a molecular ion could be explained by either being a diacyl-, or an alkyl–acyl-compound. In these cases, the possible identification of alkyl–acyl PtdCho species was based on the presence of a relatively high MS-signal for a single fatty acyl fragment ion and the corresponding lyso-PtdCho compound. In addition, there should be a total absence of signals (i. e. both fatty acyl-, and lyso-PtdCho-ions) potentially explained by fragmentation of an ester bond in the opposite *sn*-position. An example of this is *m/z* 764.6 (spectrum shown in Fig. 5f) which could originate from both O16:0–20:5 alkyl–acyl PtdCho and 15:0–20:5 diacyl PtdCho. The presence *m/z* 466 $[(lyso\text{-}PtdCho\ O16:0/15:0)–CH_3]^-$, 448 $[(lyso\text{-}PtdCho\ O16:0–15:0)–CH_3–H_2O]^-$ and 301 $[20:5]^-$ indicate a fatty acyl group of 20:5, while there is no fragment indicating a fatty acyl group of 15:0 in the opposite *sn*-position. This is therefore assumed to be a PtdCho-specie with an alkyl-acyl composition of O16:0–20:5. Altogether, seven different potential alkyl-acyl PtdCho species were characterized. For all these species, the fatty alkyl chains were either hexadecanoic or octadecanoic, and either saturated or with a single double bond.

No further attempt was made to clarify the stereospecificity of the species. It is important to keep in mind that

**Table 2** Identified phospholipid species with choline head group in krill oil

| Class | Mass | *m/z*[a] | Molecular specie | Relative intensity |
|---|---|---|---|---|
| PtdCho | 703.6 | 748.6 | 14:0–16:1 | 5.33 |
| PtdCho | 717.7 | 762.7 | 15:0–16:1 | 0.90 |
| PtdCho | 717.7 | 762.7 | 13:0–18:1 | 21.08 |
| PtdCho | 731.6 | 776.6 | 14:0–18:1 | 13.75 |
| PtdCho | 731.6 | 776.6 | 16:0–16:1 | 27.48 |
| PtdCho | 737.6 | 782.6 | 13:0–20:5 | 17.38 |
| PtdCho | 745.6 | 790.6 | 15:0–18:1 | 6.71 |
| PtdCho | 745.6 | 790.6 | 16:0–17:1 | 6.80 |
| PtdCho | 749.6 | 794.6 | 14:1–20:5 | 6.94 |
| PtdCho | 751.6 | 796.6 | 14:0–20:5 | 17.62 |
| PtdCho | 753.6 | 798.6 | 14:0–20:4 | 0.26 |
| PtdCho | 753.6 | 798.6 | 16:0–18:4 | 14.88 |
| PtdCho | 755.6 | 800.6 | 16:0–18:3 | 13.30 |
| PtdCho | 757.6 | 802.7 | 16:1–18:1 | 1.41 |
| PtdCho | 757.6 | 802.7 | 16:0–18:2 | 32.42 |
| PtdCho | 759.7 | 804.7 | 16:1–18:0 | 0.10 |
| PtdCho | 759.7 | 804.7 | 16:0–18:1 | 100.00 |
| PtdCho | 761.7 | 806.7 | 16:0–18:0 | 15.83 |
| PtdCho | 763.6 | 808.6 | 13:0–22:6 | 0.30 |
| PtdCho | 763.6 | 808.6 | O16:1–20:5 | 20.17 |
| PtdCho | 765.6 | 810.7 | O16:0–20:5 | 28.24 |
| PtdCho | 777.6 | 822.6 | 18:3–18:3 | <0.01 |
| PtdCho | 777.6 | 822.6 | 12:4–24:2 | 0.03 |
| PtdCho | 777.6 | 822.6 | 18:1–18:5 | 0.07 |
| PtdCho | 777.6 | 822.6 | 16:1–20:5 | 4.85 |
| PtdCho | 777.6 | 822.6 | 14:0–22:6 | 12.25 |
| PtdCho | 779.6 | 824.6 | 18:1–18:4 | 0.10 |
| PtdCho | 779.6 | 824.6 | 16:0–20:5 | 96.57 |
| PtdCho | 781.6 | 826.6 | 16:0–20:4 | 24.23 |
| PtdCho | 783.7 | 828.7 | 18:1–18:2 | 7.69 |
| PtdCho | 785.6 | 830.7 | 18:1–18:1 | 14.69 |
| PtdCho | 789.6 | 834.7 | O16:1–22:6 | 3.50 |
| PtdCho | 789.6 | 834.7 | 17:2–20:5 | 3.85 |
| PtdCho | 791.7 | 836.7 | 17:1–20:5 | 15.64 |
| PtdCho | 791.7 | 836.7 | O16:0–22:6 | 22.97 |
| PtdCho | 793.6 | 838.6 | O18:0–20:5 | 15.01 |
| PtdCho | 799.6 | 844.6 | 18:4–20:5 | 8.17 |
| PtdCho | 803.6 | 848.7 | 18:2–20:5 | 9.28 |
| PtdCho | 805.6 | 850.6 | 18:1–20:5 | 3.51 |
| PtdCho | 805.6 | 850.6 | 16:0–22:6 | 76.34 |
| PtdCho | 807.6 | 852.6 | 18:0–20:5 | 23.92 |
| PtdCho | 825.6 | 870.6 | 18:4–22:6 | 0.15 |
| PtdCho | 825.6 | 870.6 | 20:5–20:5 | 30.31 |
| PtdCho | 827.6 | 872.6 | 20:4–20:5 | 16.18 |
| PtdCho | 831.7 | 876.7 | 18:1–22:6 | 17.10 |
| PtdCho | 833.7 | 878.7 | 20:1–22:6 | <0.01 |
| PtdCho | 851.6 | 896.6 | 20:5–22:6 | 24.80 |
| PtdCho | 861.7 | 906.7 | 20:5–22:1 | 20.80 |
| PtdCho | 867.6 | 912.7 | 20:5–23:5 | <0.01 |
| PtdCho | 875.7 | 920.7 | 20:5–23:1 | 5.75 |
| PtdCho | 877.6 | 922.6 | 22:6–22:6 | 7.49 |

Data were obtained with RPLC/ESI-$MS^3$ operated in negative ionization mode and with data dependent fragmentation without exclusion list

[a] *m/z* for $[M + FA–H]^-$ adduct

stereoisomers would be difficult to separate and specifically identify. Therefore, the relative intensity values presented in Table 2 will in many cases be the sum of the signal intensities of the two stereoisomers. The ratio between the stereoisomers will vary among the different species. However, an interesting observation was that several of the *n*-3 acylated species appear to carry the *n*-3 fatty acyl group in the *sn*-1 position, based on the fragment ratios in the fragmentation spectra (e.g. 22:6–18:1 diacyl PtdCho).

In total, 58 species were characterized without the use of the exclusion list, whereof seven were from the lyso-PtdCho class and 51 from the PtdCho class. An additional 11 species were detected by applying the asymmetric exclusion list. Of these latter, two were identified as lyso-PtdCho and nine as PtdCho compounds, yielding an overall total of nine lyso-PtdCho class species and 60 PtdCho class species elucidated. Seven species yielded signals for highly probable fatty acyl *n*-3 groups in both the *sn*-1 and *sn*-2 positions (i.e. the diacyl PtdCho species 18:4–20:5, 18:4–22:6, 20:5–20:5, 20:5–22:6, 20:5–23:5, 22:6–22:6 and 20:5–22:5). Those species and the detection of more exotic species such as 22:6–23:5 and 20:5–26:4 diacyl PtdCho, show the complexity of krill oil.

Prior to analysis of the krill oil, the total fatty acid composition, wherein information on the concentration of individual fatty acids and their *n*-3 content was obtained and provided by Nofima, with the method AOCS Ce 1b-89 (data not shown). The sum of polyunsaturated (*n*-3) fatty acids was reported to be 18.5 g/100 g oil. The assumed homologous distribution of the fatty acid composition between triacylglycerols, free fatty acids, and the lyso-PtdCho- and PtdCho-classes, combined with the described relative intensities of the known species (Tables 1, 2), makes it possible to estimate the prevalence of *n*-3 fatty acids in one or both *sn* positions of the PtdCho species. For the PtdCho class, approximately 58% of the components contained a single *n*-3 fatty acid, and 10% held an *n*-3 fatty acid in both *sn*-1 and *sn*-2 positions. Of the species in the lyso-PtdCho class, approximately 35% contained an *n*-3

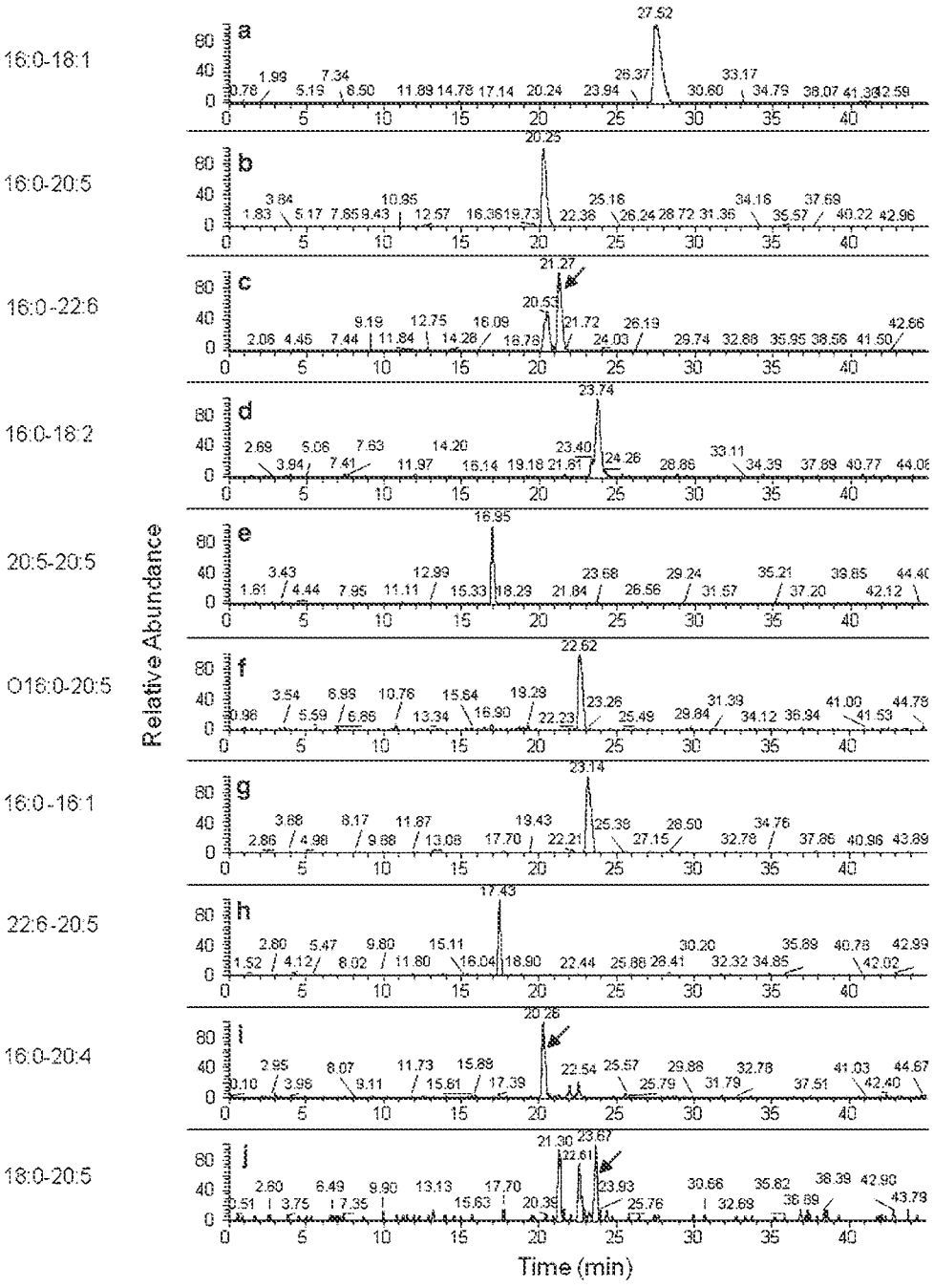



**Fig. 4** Reconstructed ion chromatograms obtained for the 10 species with the highest relative intensities in falling order from **a** to **j**. The following *m/z* values for the adducts $[M + FA-H]^-$ were used for the reconstruction: **a** 804.6 **b** 824.6 **c** 850.6 **d** 802.6 **e** 870.6 **f** 810.6 **g** 776.6 **h** 896.6 **i** 826.6 **j** 852.6

fatty acid. The phospholipid composition might potentially give better insight into the mechanism and distribution of krill oil in the body.

In the analysis performed by Nofima, the PtdCho concentration of the same krill oil as was used in the current study was reported to be 35 g/100 g oil (Nofima internal method N A88, based on [25, 26]). With the method developed for this work, the PtdCho concentration was estimated to be 34 ± 5 g/100 g oil. This shows that, in spite of the difference between the methods, the correlation of the values obtained with the two methods is relatively high. Dissimilarity between the standards used for calibration and the actual composition of the phospholipid classes in the sample could influence the quantification. In this work, we utilized a PtdCho standard originating from egg yolk. The difference in the PtdCho profile between this standard and the krill oil (Fig. 6) could potentially result in an underestimation of the PtdCho content of the sample. This is a result of krill oil containing long chained fatty acids in the phospholipid components, which are less prone

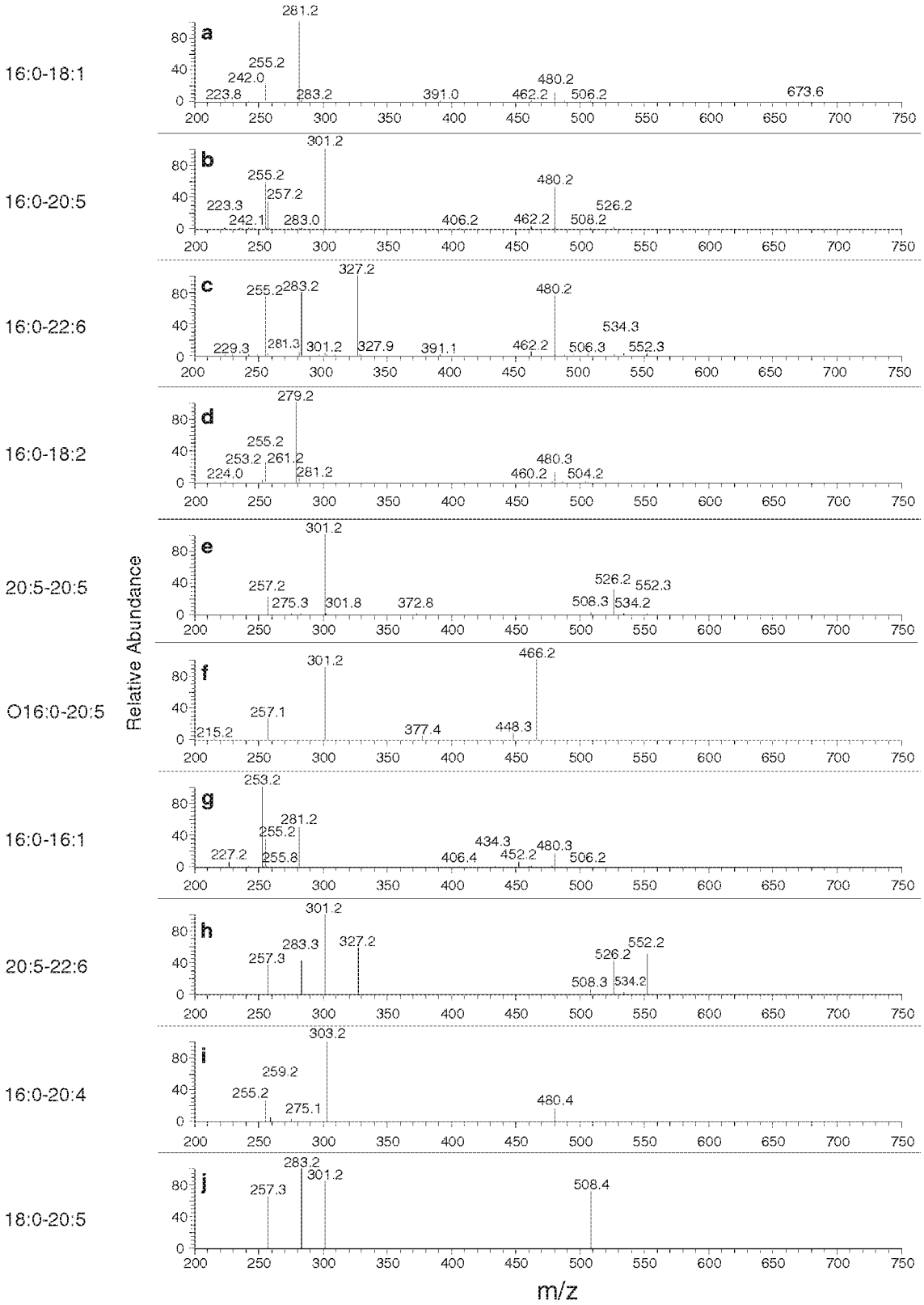



**Fig. 5** $MS^3$ product ion spectrum for the 10 species with the highest relative intensities, obtained using RPLC for separation with data dependant fragmentation. Fig. 5 **a–j** corresponds to Fig. 4 **a–j**, respectively

**Table 3** Identified lyso-phospholipid and phospholipid species with choline head group in krill oil

| Class | Mass | *m/z*[a] | Molecular specie |
|---|---|---|---|
| Lyso-PtdCho | 549.4 | 594.4 | 20:1 |
| Lyso-PtdCho | 569.4 | 614.4 | 22:5 |
| PtdCho | 771.3 | 816.3 | O18:1–18:1 |
| PtdCho | 795.6 | 840.6 | 16:0–21:4 |
| PtdCho | 819.7 | 864.7 | O18:0–22:6 |
| PtdCho | 829.5 | 874.5 | 18:2–22:6 |
| PtdCho | 845.6 | 890.6 | 20:5–21:2 |
| PtdCho | 853.2 | 898.2 | 22:6–23:5 |
| PtdCho | 853.2 | 898.2 | 20:5–22:5 |
| PtdCho | 893.6 | 938.6 | 20:4–22:6 |
| PtdCho | 911.6 | 956.6 | 20:5–26:4 |

Data were obtained with RPLC/ESI-MS$^3$ operated in negative ionization mode and with data dependent fragmentation with asymmetric exclusion list. Relative intensities for the two different methods cannot be compared

[a] *m/z* for $[M + FA–H]^-$ adduct

to ionization with ESI as described earlier. Hence, the use of a PtdCho standard originating from krill oil would have been preferred. However, this type of standard was not commercially available when the study was performed.

Krill oil has been analyzed by a novel method and numerous molecular species were found that were previously not known to be present in krill oil. To our knowledge, there has previously only been one report showing the fatty acid composition in PtdCho from krill oil [24] The method used in that study was comparable to the platform used in the current study, however, their detection was based on the formation of lithium adducts, detected with a triple quadrupole. The krill oil used in that study was delivered by Nestec SA (Lausanne, Switzerland), and the composition of the two krill oils might therefore not be directly comparable in respect to composition. However, in both studies the following three PtdCho species were among the five most prevalent: (16:0–20:5) (16:0–22:6), and (16:0–18:1). On the other hand, they detected high levels of other PtdCho species that are low or absent in our study, such as: (18:1–20:5) and (18:0–18:2). Le Grandois et al. detected neither fatty acids longer than 22 carbons nor ether-linked fatty acids. Moreover, in our study, a higher number of PtdCho species was detected. Whether these differences are due to a different composition of the two krill oils, or the methodology is not known, but might indicate that the detection method used in the current study is more sensitive.

The method used for quantification purposes in this study was optimized for specific quantification of the PtdCho class, and quantification of other PL classes was not attempted. However, it would be of great interests and an aid in the field of lipidomics, to develop standardized and validated LC–MS methods that meet the need for quantification of all phospholipid classes in complex matrixes like krill oil. One of the potential challenges in the development of such strategies would be the choice of suitable standards for the calibration curves.

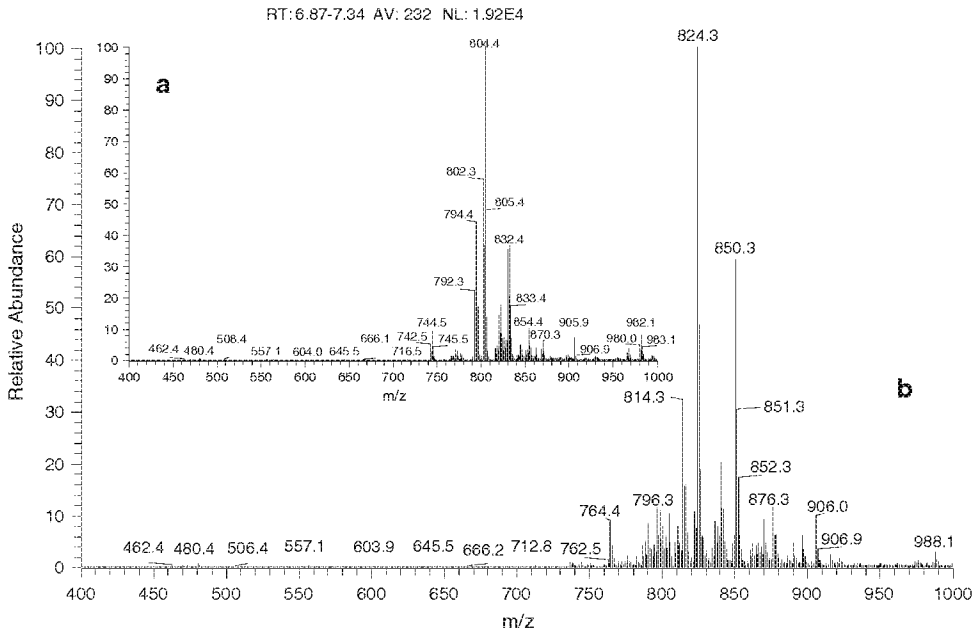



**Fig. 6** **a** Mass spectrum of PtdCho class standard used as calibrant for quantification. **b** Mass spectrum of the krill PtdCho class. Both spectra were obtained with NPLC/ESI–MS operated with negative mode ionization, and show $[M + FA–H]^-$ adduct ions. The *mass spectra* indicate a higher occurrence of long chained fatty acyl groups in PtdCho species originating from krill than in the PtdCho standard from egg yolk

**Acknowledgments** The authors would like to thank Dr. Inge Bruheim and Dr. Åsmund Larsen for their helpful input and valuable discussions during the method development. This work was partially funded by Aker BioMarine ASA.



## References

1. Batetta B, Griinari M, Carta G, Murru E, Ligresti A, Cordeddu L, Giordano E, Sanna F, Bisogno T, Uda S, Collu M, Bruheim I, Di Marzo V, Banni S (2009) Endocannabinoids may mediate the ability of (n-3) fatty acids to reduce ectopic fat and inflammatory mediators in obese Zucker rats. J Nutr 139:1495–1501
2. Di Marzo V, Griinari M, Carta G, Murru E, Ligresti A, Cordeddu L, Giordano E, Bisogno T, Collu M, Batetta B, Sanna F, Uda S, Berge K, Banni S (2010) Dietary krill oil increases docosahexaenoic acid and reduces 2-arachidonoylglycerol but not *N*-acylethanolamine levels in the brain of obese Zucker rats. Int Dairy J 20:231–235
3. Maki KC, Reeves MS, Farmer M, Griinari M, Berge K, Vik H, Hubacher R, Rains TM (2009) Krill oil supplementation increases plasma concentrations of eicosapentaenoic and docosahexaenoic acids in overweight and obese men and women. Nutr Res 29:609–615
4. Tandy S, Chung RWS, Wat E, Kamili A, Berge K, Griinari M, Cohn JS (2009) Dietary krill oil supplementation reduces hepatic steatosis, glycemia and hypercholesterolemia in high-fat fed mice. J Agric Food Chem 57:9339–9345
5. Han X, Gross RW (2005) Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. Mass Spectrom Rev 24:367–412
6. Koivusalo M, Haimi P, Heikinheimo L, Kostiainen R, Somerharju P (2001) Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response. J Lipid Res 42:663–672
7. Bruegger B, Erben G, Sandhoff FT, Wieland WD, Lehmann WD (1997) Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. Proc Natl Acad Sci USA 94:2339–2344
8. Houjou T, Yamatani K, Imagawa M, Shimizu T, Taguchi R (2005) A shotgun tandem mass spectrometric analysis of phospholipids with normal-phase and/or reverse-phase liquid chromatography/electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom 19:654–666
9. Lehmann WD, Koester M, Erben G, Keppler D (1997) Characterization and quantification of rat bile phosphatidylcholine by electrospray-tandem mass spectrometry. Anal Biochem 246:102–110
10. Beermann C, Mobius M, Winterling N, Schmitt JJ, Boehm G (2005) *sn*-Position determination of phospholipid-linked fatty acids derived from erythrocytes by liquid chromatography electrospray ionization ion-trap mass spectrometry. Lipids 40:211–218
11. Houjou T, Yamatani K, Nakanishi H, Imagawa M, Shimizu T, Taguchi R (2004) Rapid and selective identification of molecular species in phosphatidylcholine and sphingomyelin by conditional neutral loss scanning and MS3. Rapid Commun Mass Spectrom 18:3123–3130
12. Larsen A, Uran S, Jacobsen PB, Skotland T (2001) Collision-induced dissociation of glycero phospholipids using electrospray ion-trap mass spectrometry. Rapid Commun Mass Spectrom 15:2393–2398
13. Schwudke D, Hannich JT, Surendranath V, Grimard V, Moehring T, Burton L, Kurzchalia T, Shevchenko A (2007) Top-down lipidomic screens by multivariate analysis of high-resolution survey mass spectra. Anal Chem 79:4083–4093
14. Hvattum E, Hagelin G, Larsen A (1998) Study of mechanisms involved in the collision-induced dissociation of carboxylate anions from glycerophospholipids using negative ion electrospray tandem quadrupole mass spectrometry. Rapid Commun Mass Spectrom 12:1405–1409
15. Karlsson AA, Michelsen P, Larsen A, Odham G (1996) Normal-phase liquid-chromatography class separation and species determination of phospholipids utilizing electrospray-mass spectrometry tandem-mass spectrometry. Rapid Commun Mass Spectrom 10:775–780
16. Larsen A, Mokastet E, Lundanes E, Hvattum E (2002) Separation and identification of phosphatidylserine molecular species using reversed-phase high-performance liquid chromatography with evaporative light scattering and mass spectrometric detection. J Chromatogr B Analyt Technol Biomed Life Sci 774:115–120
17. Uran S, Larsen A, Jacobsen PB, Skotland T (2001) Analysis of phospholipid species in human blood using normal-phase liquid chromatography coupled with electrospray ionization ion-trap tandem mass spectrometry. J Chromatogr B Biomed Sci Appl 758:265–275
18. Ekroos K, Chernushevich IV, Simons K, Shevchenko A (2002) Quantitative profiling of phospholipids by multiple precursor ion scanning on a hybrid quadrupole time-of-flight mass spectrometer. Anal Chem 74:941–949
19. Harrison KA, Murphy RC (1995) Negative electrospray ionization of glycerophosphocholine lipids: formation of [M-15] ions occurs via collisional decomposition of adduct ions. J Mass Spectrom 30:1772–1773
20. Khaselev N, Murphy RC (2000) Structural characterization of oxidized phospholipid products derived from arachidonate-containing plasmenyl glycerophosphocholine. J Lipid Res 41: 564–572
21. Schwartz JC, Syka JP, Quarmby ST (2005) Improving the fundamentals of Msn on 2D ion traps: new ion activation and isolation techniques. In: 53rd ASMS Conference on Mass Spectrometry, San Antonio, Texas
22. Tyurin VA, Tyurina YY, Feng W, Mnuskin A, Jiang J, Tang M, Zhang X, Zhao Q, Kochanek PM, Clark RS, Bayir H, Kagan VE (2008) Mass-spectrometric characterization of phospholipids and their primary peroxidation products in rat cortical neurons during staurosporine-induced apoptosis. J Neurochem 107:1614–1633
23. Griffin TJ, Xie H, Bandhakavi S, Popko J, Mohan A, Carlis JV, Higgins L (2007) iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. J Proteome Res 6:4200–4209
24. Le Grandois J, Marchioni E, Zhao M, Giuffrida F, Ennahar S, Bindler F (2009) Investigation of natural phosphatidylcholine sources: separation and identification by liquid chromatography-electrospray ionization-tandem mass spectrometry (LC–ESI–MS2) of molecular species. J Agric Food Chem 57:6014–6020
25. Homan R, Anderson MK (1998) Rapid separation and quantitation of combined neutral and polar lipid classes by high-performance liquid chromatography and evaporative light-scattering mass detection. J Chromatogr B Biomed Sci Appl 708:21–26
26. Moreau RA (2006) The analysis of lipids via HPLC with a charged aerosol detector. Lipids 41:727–734