Author's response to reviews

Title: A high confidence, manually validated human blood plasma protein reference set

Authors:

Susann Schenk (sschenk@bmb.sdu.dk)
Gary J Schoenhals (gary@bmb.sdu.dk)
Gustavo de Souza (gso041@uib.no)
Matthias Mann (mmann@biochem.mpg.de)

Version: 4 **Date:** 27 May 2008

Author's response to reviews: see over

Susann Schenk, Ph.D.
Associate Professor
Department of Biochemistry and Molecular Biology
University of Southern Denmark
Campusvej 55
5230 Odense M
Denmark
Tel: +45 6550 2361
E-mail: sschenk@bmb.sdu.dk

Ms. Ashleigh Manning
Assistant Editor
BMC-series journals
Tel: +44 (0)20 7631 9921
Fax: +44 (0)20 7631 9923
E-mail: editorial@biomedcentral.com

May 28, 2008

RE: MS: 1691558822166885
A high confidence, manually validated human blood plasma protein reference set
Susann Schenk, Gary J Schoenhals, Gustavo de Souza and Matthias Mann

Dear Ms. Ashleigh Manning,

Thank you for sending us the new comments from the reviewer Tim Griffin.

We are pleased to read that Dr. Griffin is satisfied with the way that we addressed his concerns from January 5th 2008.

As previously we have taken the report from Tim Griffin and broken it into sections according to the individual points that he feels need to be addressed and we have left the order of the referee's comments intact. We have then placed our response to each of the referee's comments below the text written by the referee.

Tim Griffin, referee

The authors have done a good job in addressing the main concerns from the first review. The manuscript is generally improved. Before publication I would suggest the following revisions:

Points that should be addressed before publication:

1. Although the manuscript has been shortened, both in terms of text length and number of figures, one final suggestion to further shorten the manuscript would be to consider **removing Figure 5.**

> We removed Figure 05 concerning the signal peptide analysis. The manuscript now has a total of 10 figures.

2. **On the issue of using MS3 in order to increase confidence in peptide identifications, some confusion still exists.** The authors have added in a better explanation of how their **MS3 experiment works (loss of phosphate)**, **however this is cause for some questions.** For one, there is no reason presented to the reader **as to why they decided to include this operating mode in their experiments?** One would assume that they are working on the hypothesis that there are a high number of phosphorylated peptides within the plasma which would benefit from an MS3 scan? Some **rationale should be given in the Results section to clue the reader into why these experiments were included.** Also, if the reason is to analyze phosphopeptides, does this mean that all the peptides that were identified with the aid of MS3 were phosphorylated peptides? Or did the MS3 scan randomly sequence fragments from non-phosphorylated peptides as well which still helped to identify these?

> This comment is not entirely clear to us and we assume that there must be some kind of misunderstanding. However, we have tried to clarify this issue as best we can.

a) References to MS3 as an additional validation requirement for single peptide identification and the advantage of using MS3 in improving the confidence level of peptide identification is mentioned throughout the manuscript. Examples:

   Page 02: "The combination of MS3 with very high mass accuracy in the parent peptide allows peptide identification with orders of magnitude more confidence than that typically achieved."

   Page 05: "……, and to further increase the reliability of our data, we employed MS3".

   Page 09: "…..according to an algorithm that assigns MS3 spectra to peptide fragment sequences [7]". (Reference 7 points to a paper by Jesper Olsen and Matthias Mann on the improved peptide identification by additional MS3, as well as a very detailed explanation on how MS3 is performed.)

   Page 13: "….Briefly, proteins identified with one single peptide were required to have an MS3 spectrum, an MS3 score, and a total score (also known as a 'Mascot peptide score plus MS3 score') ≥ 42, which assured with 99.9% confidence that this was a correct identification. Proteins identified with one single peptide but without an available MS3 scan were discarded regardless of their Mascot peptide score."

   Page 21: "It was identified with 4 valid, non-redundant peptides of 10, 13, 17 and 20 amino acids in length. The three peptides having lengths of 10, 13 and 17 residues possess an MS3 spectrum in addition to their MS2 Mascot scores, giving rise to scores of 70, 131 and 176 (summed

score of 377), respectively". It is unlikely that these scores would be achieved without the use of MS3.
Page 23: "Note that single peptide identifications do not appear in 01_MS2 and 04_MS2_prec because our validation criteria demanded an MS3 spectrum, which is precluded in these experiments which were only MS2-based."
Page 29: "We therefore discarded all protein identifications based on 1 peptide unless a given peptide was validated by a MS3 spectrum, significantly increasing confidence in peptide identification [7]".

b) MS3 in a neutral loss-dependent mode was applied only to experiment 10 as mentioned on page 11: "…in 10_Alb_depl_NL, MS3 acquisition was only done in a neutral loss-dependent fashion, in order to detect possible phosphopeptides."

   Additionally we added on page 24 the following (see addition in brackets): "(neutral loss-dependent MS3 was used in this study in experiment 10_Alb_depl_NL only)."

c) The reason why we included MS3 neutral loss-dependent fashion in experiment 10 is now mentioned and cited on page 11 "Also, in 10_Alb_depl_NL, MS3 acquisition was only done in a neutral loss-dependent fashion, in order to detect possible phosphopeptides. As early as in the pilot phase of the HUPO project [3][4][5], the possible determination of posttranslational modifications of plasma proteins such as protein phosphorylation was mentioned as an important issue for the comprehensive analysis of the protein constituents of human plasma as well as the identification of biomarkers".

3. The inclusion of the discussion on peptide length seems warranted based upon the authors explanation. One suggestion would be **to include a brief explanation of why peptide length is an important parameter, perhaps with a citation to the Adamski et al** publication where this issue is also discussed.

On page 29, where we mention our minimum required peptide length of 7 amino acids, we added a brief explanation and the citation of Adamski et al. "Because the accuracy of protein identification increases with peptide length, we excluded, for example, all peptides that did not have the minimum required length of 7 amino acids, while Adamski et al. eliminated all peptides shorter than 6 amino acids from further analysis [14]".

4. Throughout the manuscript **"Orbitrap" should be capitalized** because it's a product name

We capitalized “Orbitrap” throughout the manuscript.

5. On p. 28 the authors explain that all “identifications based on 1 peptide unless a given peptide was validated by a third MS spectrum (MS3).**” What is meant by “third MS spectrum”? What are the other two spectra?**

The sentence on page 29 “We therefore discarded all protein identifications based on 1 peptide unless a given peptide was validated by a third MS3 spectrum” was modified and supplemented to **“**We therefore discarded all protein identifications based on 1 peptide unless a given peptide was validated by a MS3 spectrum, significantly increasing confidence in peptide identification [7]”.

As an additional explanation, with 'third' MS spectrum we referred to the additional stages of mass spectromic fragmentation (MSn). Just as one stage of MS2 provides more information than the molecular mass alone, additional stages provide further information on the fragments. In brief, the mass spectrometer was operated in a data-dependent mode to automatically switch between MS, MS2 and MS3 acquisition. For example, survey full scan MS spectra (high dynamic range survey of the total mass range) were acquired in the FT-ICR (MS1 full scan). The three most intense ions were sequentially isolated for accurate mass measurements by a FT-ICR selected ion monitoring scan (SIM). These precursor ions were then fragmented in the linear ion-trap to obtain sequence information (MS2 scan). Up to three ions in each MS2 spectrum (the most intense ions) were further isolated and fragmented (MS3).

We hope that you will find our responses to the referee’s comments to be satisfactory and that our manuscript is now suitable for publication. Thank you very much for your kind assistance.

With kind regards,

Susann Schenk