University of South Bohemia České Budějovice
Faculty of Science

Branišovská 31, 370 05 České Budějovice, Czech Republic

Opponent Report on

Bachelor Thesis of Tamara Buchberger

The Bachelor thesis of Ms. Tamara Buchberger is a review of current information on protein investigation by mass spectrometry (MS). The thesis has 33 pages and is divided into six sections. After a short Introduction, there are four sections on the history of MS, basic information on MS, but also on new technologies.

Generally, the thesis is well written. However, the quality of the language is not sufficient especially in the section 5. The student is totally missing the importance of commas and thus the reader has to think hard to get the meaning of some sentences. The reader should understand the whole text and all its parts without any thinking about what the author wanted to say. As an example, here is one sentence, where one comma would be enough to clarify the meaning:

"Depending on the composition of the matrix molecules with different molecular weight are dissolved in the matrix forming matrix-analyte crystals".

Corrections by native speaker (or a more fluent English speaker) would be beneficial. Anyway, most of these mistakes could be corrected by the student herself – therefore, I suggest her to read the written text several times, if possible wait some time between two corrections. It is possible then to see the mistakes more clearly and also to enhance the readability and understandability of the text. Also, it is good to change the language in the software, while writing in English – then we will not find "und" instead of "and". As this thesis was prepared in two years, there was enough time to do it.

Citations are another problem. In the text it is OK, but if you copy WHOLE figures from books or other sources, they HAVE TO BE CITED. Otherwise, it can be considered as plagiarism.

Some of the mistakes and incomprehensible sentences:

For this method peptides are dissolved in a solvent that absorbs UV … (page 9, line 17),

The matrix usually contains aromatic acids or aromatic carbonyl compounds with a high absorption coefficient for the laser and to have the ability to convert the laser energy into heat (page 10, line 3),

In common MALDI experiments concerning proteome research are applied directly to the output of tryptic digests (page 10, line 11),

The analysis benefits from all advantages from the Q-TOF that have already been mentioned before (page 15, line 4) – the problem is they were not,

Further it is useful in the analysis of large protons … (page 19, line 30),

In Table 1 additionally the masses of the appropriate immonium ions are listed. They provide further information about the content of amino acids but not about their position (page 24, line 2),

There exist several different procedures … (page 27, line 11) … In principle there exist two methods ...(page 27, the same paragraph, line 16),

In former days clinical application was limited as a analysis of a small tissue micro array would last seven days and produce around 50 GB of data (page 27, line 30).

Page 21 is interesting by itself. In the second paragraph (line 15 and further) the student describes peptide-fragmentation. As she states, she is describing it using the singly charged peptide ATSFYL in Figure 12 (can be found in Reference 1, page 70). There is just one problem – she

copied the whole figure from a book and she copied the wrong figure – the doubly charged peptide ATSFYK (can be found in Reference 1, page 71). I am not sure what does it mean – lack of knowledge, interest? Student defending a bachelor thesis in chemistry should be able to differentiate between lysine and leucine and to see two charges instead of one … Also in this case, citation of the source in the figure legend is lacking.

Student should answer the following question during the defense:

1) You write that it is possible to analyze several proteins per hour (page 1, "Speed"). Did you mean proteins or samples?

2) On page 14 you describe the function of the middle quadrupole in the triple quadrupole instrument. What are the functions of the first and the third quadrupoles?

3) On page 23, you mention tryptic digests. What are these?

4) On page 24, you write about the analyzing of the fragmentation spectra. Why are the b and their corresponding a ions separated by 28?

5) On page 26, you write about destruction of proteolytic enzymes (Imaging Mass Spectrometry). How is it possible to destroy these enzymes? Is it possible to destroy proteolytic enzymes but not to harm the sample?

6) On page 26, you write about removing hemoglobin from tissue samples by washing (Imaging Mass Spectrometry). How is it possible to remove only one specific protein (hemoglobin) and not other proteins, too?

7) You are writing about the importance of post-translational modifications (PTMs) in diseases but not under other conditions. Are the PTMs important also in healthy organisms?

8) On page 28 you state, that 50% of all proteins are phosphorylated. Is this true for all the proteins in all organisms?

9) The same you state with glycosylation. Again, is this true?

10) How is it possible to introduce ^{18}O enzymatically into proteins? Which enzymes and which reactions can be used and what is the product of such reactions?

In conclusion, the thesis fulfills the requirements for bachelor theses at our faculty and I recommend the thesis for defense. Once, I will have the answers to my questions, I will decide on the mark.

In České Budějovice, June 1st 2010

.....

Mgr. Ján Štěrba

1) You write that it is possible to analyze several proteins per hour (page 1, "Speed"). Did you mean proteins or samples?

I meant proteins. One must take into account that the analysis also includes fragmentation and isolation of fragments. Therefore a "speed" of several proteins per hour is common.

2) On page 14 you describe the function of the middle quadrupole in the triple quadrupole instrument. What are the functions of the first and the third quadrupoles?

The first and the third quadrupole can have two possible modes. One is the so-called SIM mode and the other one is the scan mode. The scan mode allows a measurement of a range of m/z (usually from m/z from 50 to 2000). The SIM mode (single ion monitoring) allows the selection of one or several specific m/z ratios so that only these selected m/z walk through the quadrupole. The SIM mode has therefore an increased sensitivity over the scan mode and quantitative precision and accuracy are improved due to a improved peak shape. The SIM mode is of special interest for the first quadrupole as it ensures the fragmentation of one peptide (with a specific m/z) which makes the interpretation of the obtained fragment easier.

3) On page 23, you mention tryptic digests. What are these?

A digest is a hydrolysis by a protease, usually trypsin is used (tryptic digest). It usually cuts the amino bonds of the amino acids lysine or arginine. The derived peptides are specific for the protein like a fingerprint. Digests are pretreatment methods that can also be done in the separation gel.

4) On page 24, you write about the analyzing of the fragmentation spectra. Why are the b and their corresponding a ions separated by 28?

The difference between the b and its corresponding a ion is due to a loss of CO_2.

5) On page 26, you write about destruction of proteolytic enzymes (Imaging Mass Spectrometry). How is it possible to destroy these enzymes? Is it possible to destroy proteolytic enzymes but not to harm the sample?

One possibility is to heat the tissue sample immediately after extraction with a microwave. The microwave ensures a rapid temperature increase which degrades the enzymes and prevents postmorten degradation of peptides and proteins without loss of morphology.

6) On page 26, you write about removing hemoglobin from tissue samples by washing (Imaging Mass Spectrometry). How is it possible to remove only one specific protein (hemoglobin) and not other proteins, too?

It means that you remove an excessive amount of haemoglobin. This washing step is before the washing step where you destroy the matrix/ lipids. So the removal of proteins is reduced. By the way the removal of wanted proteins has to be checked by analysis of washing media. Further it is important that the different washing steps have to be adjusted to the tissue because depending on the tissue sample washing steps might be different.

7) You are writing about the importance of post-translational modifications (PTMs) in diseases but not under other conditions. Are the PTMs important also in healthy organisms?

PTMs change several properties of a protein, such as lifetime, localisation or solubility. They are also present in healthy organisms. In organisms with a disease these PTMs seem to have an influence as their abundance is different than in healthy organisms. For the identification of diseases and their treatment ist is important to know about PTMs.

8) On page 28 you state, that 50% of all proteins are phosphorylated. Is this true for all the proteins in all organisms?

In literature it is stated that this is true for mammalians. It is only an estimation because of probably 100,000 potential phosphorylation sites (again an estimation) in human proteome fewer than 2,000 are currently known. (Dario E Kalume, Henrik Molina and Akhilesh Pandey, Tackling the phosphoproteome: tools and strategies, *Current Opinion in Chemical Biology* **7/1** (2003) 64-69)

9) The same you state with glycosylation. Again, is this true?

In literature it is stated that more than 50 % of all proteins are glycosylated but it is not proven yet. This is the result from an extrapolation from the values of a database. (Rolf Apweiler, Henning Hermjakob and Nathan Sharon, On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database, *Biochimica et Biophysica Acta (BBA) - General Subjects* **1473/1** (1999) 4-8)

10) How is it possible to introduce ^{18}O enzymatically into proteins? Which enzymes and which reactions can be used and what is the product of such reactions?

The introduction of ^{18}O takes for example place during a digest (hydrolysis). The surrounding medium contains $H_2{}^{18}O$ which enables the introduction of ^{18}O into the peptide. The introduction into a C-termini leads to a mass shift of 4 Da as two ^{18}O are introduced. The result is an ^{18}O-containing carboxyl group. Examples for such enzymes would be trypsin or Glu-C. On the other hand, enzymes like chymotrypsin and Lys-N introduce only one ^{18}O.