Opponent´s review of Jan Hájek´s master thesis „Identification of products of tetrapyrrole pathway"

The work of Jan Hájek is dealing with a very demanding topic, which is searching for novel molecules in non-standard natural sources. Even mere mastering of their isolation, analysis and identification alone require substantial deal of work.

The presented work contains 60 pages. Introduction review encompasses the knowledge of cyanobacteria, their morphology, metabolism, and photosynthesis with the particular attention to their tetrapyrrole pigments. Figures, which are adopted from scientific papers, contain correct references to the original sources. Page 21 contains aims of the work, which are purification, characterization and structure determination of unknown yellow pigment excreted by *Synechocystis* cell culture.

The main focus of the author´s work is primarily thorough investigation of isolation conditions of the new metabolite and development of analytical and preparative methods. Several methods were tested for isolation and finally attempts were made to elucidate the structure by MS and NMR. Although there are no doubts about the substantial amount of hand-work that was actually carried out, I have a number of comments and questions to be addressed:

1. Majority of figures describe UV spectra, using one instrument (page 22). Some spectra contain switch between halogen and deuterium lamp at 310 nm (Fig´s 11, 12), some at 270 nm (Fig 14), some without (Fig 18) on apparently different instrument. Please explain. Y-axis is described as absorbance units, which is in several figures 400 (Fig. 20) or even 2995 and 4000 (Figure 24), frankly, hard to believe. At figure 13, negative absorbance is observed.
2. Isolation. What is the yield of compound isolated mg/%? Chromatogram showing the final purity?
3. MS characterization. LC/MS is described on Figure 22, where are the conditions used? Ion current at a) for time frame 6-20 min is clearly $0.2 . 10^9$, however b) is one order higher, why? There is not likely unknown compound at till 6 min, it is basically salt (p. 34).
4. NMR is highly sophisticated technique, but the description seems to be incomplete. Figure 28, singlets at 1.8 and 2.0, are likely methyls. COSY and HETCOR spectra are for small areas and the complete listing is absent. Figure 37 should describe typical splitting of CH_2 group. Figure 37 shows spectra of 2nd order without the possibility to identify neighbouring groups, without any proof, if it is a signal of one or more groups and their assignment to CH or CH_2. Spectra provide proof of 3 nitrogen atoms and ^{13}C NMR provides 3 signals, which could be carbonyl groups. Spectral predictions at p. 47 are not consistent with this finding.

MALDI spectrum at p. 58 provides and interesting possibility that the yellow pigment may originate from TES. Such result seems to be very interesting, since TES is considered as basically chemically inert buffer. Even in spite of some reservations I have to the work, finally the result is interesting, and I recommend it for the defence, where the questions and comments should be clarified.

In České Budějovice, 28th May 2013

Alexandr Jegorov
Teva Czech Industries
Branišovská 31
České Budějovice

UNIVERSITY OF SOUTH BOHEMIA IN ČESKÉ BUDĚJOVICE

Faculty of Science

STATEMENT OF THE ~~BACHELOR~~/DIPLOMA* THESIS REVIEWER

Name of the student: Jan Hájek

Thesis title: Identification of Products of Tetrapyrrole Pathway
Supervisor: prof. RNDr. Josef Komenda, CSc.

Reviewer: Ing. David Kahoun, Ph.D.
Reviewer`s affiliation: University of South Bohemia – Faculty of Science

	Point scale[1]	Points
(1) FORMAL REQUIREMENTS		
Extent of the thesis (for bachelor theses min. 18 pages, for masters theses min. 25 pages), **balanced extents of the thesis divisions** (recommended extent of the theoretical part is max. 1/3 of the total extent), **logical structure of the thesis**	0-3	3
Quality of the theoretical part (review) (number and relevancy of the references, recency of the references)	0-3	2
Accuracy in citing of the references (presence of uncited sources, uniform style of the references, use of correct journal titles and abbreviations)	0-3	2
Graphic layout of the text and of the figures/tables	0-3	2
Adequacy and clarity of the results and conclusions	0-3	2
Quality of the annotation	0-3	3
Language and stylistics, complying with the valid terminology	0-3	2
Accuracy and completeness of figures/tables legends (clarity even without reading the rest of the text, explanation of the symbols and labeling, indicating the units)	0-3	2
Formal requirements – points in total		18
(2) PRACTICAL REQUIREMENTS		
Clarity of the aims	0-3	3
Fulfillment of the aims	0-3	2
Discussion quality – interpretation of results and their discussion with the literature	0-3	2
Logic in the course of the experimental work	0-3	2
Completeness of the description of the used techniques	0-3	2
Experimental difficulty of the thesis	0-3	3

* Choose one
[1] Mark as: 0-unsatisfactory, 1-satisfactory, 2-average, 3-excellent.

Quality of experimental data presentation	0-3	2
The use of up-to-date techniques	0-3	3
Contribution of the thesis to the knowledge in the filed and possibility to publish the results (after eventual supplementary experiments)	0-3	2
Formal requirements – points in total		21
POINTS IN TOTAL (MAX/AWARDED)	**51**	**39**[2]

Suggestions and questions, to which the student has to answer during the defense:

1) Pages 22 – 24, Chapter 2.1: How did you find out SPE and HPLC chromatography conditions?

2) Page 25, Chapter 2.1.6: Why did the author use two the same NMR spectrometers? Which spectra were measured using Linz equipment and which spectra were measured using Prague equipment? This information is not given in following chapters.

3) Page 34, Chapter 3.3, Figure 22: How did you decide, that the peak which you marked belongs to the 414 nm compound? Figures 17a – 17c show, that the peak of this compound is approx. 2 minutes wide. All chromatographic conditions for LC/MS analysis, except flow rate which was decreased of 16.7 % (peak becomes more wide and its retention time increased), were not modified and so under these new chromatographic conditions the peak of 414 nm compound should have been started at 10^{th} minute and ended at 12^{th} minute. As the Figure 22 shows, there are lots of peaks between 10^{th} minute and 12^{th} minute but you marked just the only this peak. Why did you make this decision?

4) Page 35, Chapter 3.3: "*The 414 nm compound eluted at 11^{th} min showed the m/z value 383 and gradually fragmented into m/z 355 and m/z 124 (Fig. 23), which corresponds to previously obtained highly resolved MS spectra (Komenda, unpublished, see Attachment 9.1)*". From my point of view the MS spectrum shown in Figure 23a (Page 35) does not correspond to the spectra in Chapter 9.1. Figure 23a (Page 58) but it probably belongs to the compound which the author mentions in the following paragraph.

   Electron spray ionization (ESI) is a soft ionization technique which forms mostly even-electron molecular adducts $[M+H]^+$ (or $[M+NH_4]^+$ and $[M+H+Methanol]^+$ in this case). The spectrum on the Page 58 shows lots of peaks. Why did you decide that m/z 383 is m/z of the 414 nm compound? Why did you decide, that the full scan MS spectrum contains only two compounds (molecular ion at m/z 383 and molecular ion at m/z 351?

5) Page 37, Chapter 3.4: "*Thought previous purification on C30 column did not show a presence of large amount of the UV-absorbing substances in the 11^{th} min fraction, its separation on HILIC column surprisingly resulted in two large UV-absorbing fractions eluted at around 8^{th} and between 12^{th} and 18^{th} min (its spectrum see Fig 25).*". Why is this fact surprising?

[2] Enter the number of points awarded.

6) Page 46, Chapter 4: Author speculated that the reaction of an unknown compound with TES was catalyzed by an unknown enzyme? Why did author not verify this speculation?

7) Page 46, Chapter 4: *"... but the absence of aromatic carbons (= carbons with an NMR spectra shift between 100 and 150 ppm) excluded its identity as a simple derivative of tetrapyrrole*. Where did author obtain this information? From my point of view the information is not correct because I found some publications where ^{13}C chemical shifts of some tetrapyrrole derivatives came up to 180 ppm. The information is also wrong due to the fact that some simple aromatic compounds provide ^{13}C chemical shifts higher than 150 ppm because ^{13}C chemical shifts of aromatic carbons (128.5 ppm) strongly depend on the substituent (+/- XY ppm) e.g. nitrosobenzene (128.5 + 37.4 = 165.9), methoxybenzene (128.5 + 30.2 = 158.7), N-methylaniline (128.5 + 21.7 = 150.2) and others.

8) Page 49, Chapter 4: "*This hydrogen has relatively high correlation with carbon with chemical shift 150 ppm.*". Do the NMR ^{13}C spectra of the 414 nm compound contain carbon with chemical shift 150 nm or not? On Page 46 author denies presence of this chemical shift but on Page 49 the presence of this chemical shift is discussed. This chemical shift is also shown in Figure 29 and Figure 32.

9) Page 58, Chapter 9.1: Which MS detector was used? Was the spectrum obtained using MALDI-TOF detector (see the heading of the chapter) or using ORBITRAP detector (see the title of the picture).

Eventual additional comments of the supervisor on the student and the thesis:

This part does not have to be read during the master's thesis defence due to time reasons. These comments are made especially for the author.

1) Do not leave one-letter words at the end of a line and keep the number and the unit on the same line. Do not start a sentence with an abbreviation.

2) Some figures (including letters, numbers and symbols used in these figures) should be larger and sharper (e.g. Fig. 3a, Fig. 4, etc.).

3) Some paragraphs are not fully justified (e.g. Chapter 1.2.1.3.2, Chapter 1.2.2, etc.).

4) A lot of abbreviations (e.g. NMR, DNA, ATP, NADPH, etc.) are not listed in the list of abbreviations.

5) Some chapters (especially results, discussion and conclusion) are usually written using past tense, e.g. "data were examined ...". Notice that it is written in the passive form as well - this is used so that focus falls on what was examined and not on who did the examining.

6) The thesis contains some typing errors or grammar mistakes, e.g.:

   - Page 9, Chapter 1.2.1: Chapter name contains the word "*grown*" instead of the word "growth".

   - Page 23, Chapter 2.1.4: The first sentence of the chapter contains the word "*termostatted*" instead of the word "*termostated*".

- Page 28, Chapter 3.2: Do not use preposition "by Fig. 13" but use preposition "in Fig. 13".

7) Page 23, Chapter 2.1.3: What volume of the sample solution was passed through SPE column? What weight of the adsorbent [mg/column] was used for SPE?

8) Page 23, Chapter 2.1.3, Table 2: The volume of each solvent should be also specified. All columns in the table should be named.

9) Page 23, Chapter 2.1.4.1: The term "*solvent system*" should be replaced with more appropriate term "mobile phases".

10) Page 24, Chapter 2.1.4.1, Table 3: The term "*Timetable of solvent*" should be replaced with more appropriate term "Gradient program". The unit "min" in the fourth row is not given.

11) Page 26, Chapter 3.1: Figure 10 is placed before Figure 9.

12) Page 27, Chapter 3.1, Figure 11 and Figure 12: Values of the y-axis have too many decimal places.

13) Page 28, Chapter 3.2, Figure 13: The absorption curve enters to the area of the x-axis values and so y-axis range should be extended to lower values in this case.

14) Page 30, Chapter 3.2, Figure 15 and Figure 16: y-axis contains too many values.

15) Page 31, Chapter 3.3, Figure 17a etc.: The term "Retention Time" used for x-axis is more appropriate and more correct than the term "*Time*".

16) Page 32, Chapter 3.3: The term "death volume peak" is more appropriate and more correct than the term "*injection peak*".

17) Page 33, Chapter 3.3: "... *but there was no peak between 9th and 15th min typical for 414 nm compound*". Why did you evaluate such a long range of retention time? Chromatographic conditions were not modified so was no reason for any change of retention time of this 414 nm compound.

18) Page 33, Chapter 3.3, Figure 21: How did you obtain this spectrum? Corresponding chromatograms shown in Figure 20 do not contain any peak at this retention time.

19) Page 34, Chapter 3.3: It is appropriate to mention, that the flow rate decreasing from 0.6 ml/min to 0.5 min caused the retention time of the 414 nm compound to increase from 9th minute to 11th minute.

20) Page 34, Chapter 3.3, Figure 22 etc.: The y-axis has not appropriate title. It should represents the signal intensity of ions (expressed as "intensity" or "relative intensity") and this intensity of ions do not equal to the number of ions.

21) Page 36, Chapter 3.4, Figure 24: Does this figure really show chromatograms of the 11th min C30 fraction which was prepared without TES?

22) Page 38, Chapter 3.4: Compounds eluted at 22.5th min do not have absorption maximum at 450 nm. As is shown in Figure 27, absorption maximum of this spectrum is approx. at 210 nm. It is obvious that this wavelength is not specific and there is not peak, but

owing to the definition of the absorption maximum there is no doubt the absorption maximum of the spectra is here. Of course, if it is possible, absorption maxima of absorption peak (the highest peak in a spectrum) and/or other smaller absorption peaks are usually used for comparison but not even 450 nm is the correct wavelength for the absorption maxima peak. In the spectrum there are three absorption peaks – the highest peak is approx. at 270 nm, smaller peak is approx. at 295 nm and the smallest peak is approx. at 425 nm. No peak at 450 nm was found.

23) Page 40 – 44, chapter 3.5: All the NMR spectra and the enclosed tables are only mentioned but not explained at all.

24) Page 39, Chapter 3.5.1, Figure 28: The enclosed table headings are not written consistently because only some of them start with capital letter. The first column does not contain any heading.

25) Page 40, Chapter 3.5.3, Figure 29: The first and the third column do not contain any heading. The second and the fourth column (^{13}C chemical shifts) do not contain values "[ppm]".

26) Page 41 – 44, Chapter 3.5.3 – 3.5.6, Figures 30 – 33: Axes are labelled with undefined titles "*F1 Chemical Shift*" and "*F2 Chemical Shift*".

27) Page 45, Chapter 3.6, Figure 34: Values of the axes have too many decimals.

28) Page 46, Chapter 4: "*In the HMBC (^{1}H-^{13}C) spectrum (Fig. 32), we distinguished 3 different nitrogen atoms...*" How can the author distinguish 3 different nitrogen atoms from Figure 32 – from the HMBC (^{1}H-^{13}C) spectrum?

29) Page 58, Chapter 9.1: The enclosed table does not correspond with the spectrum shown on this but with the spectrum on the following page.

Conclusion:

In conclusion, I

recommend / ~~**do not recommend***~~

the thesis for the defense and I suggest the grade 2 .[3]

In **České Budějovice** date **May 24, 2013**

.....

Signature

[3] You can suggest a grade, which can be modified during the defense based on the presentation. However, if the reviewer is not present at the defense, the grade will not be counted.