Oponent's assessment of bachelor thesis

First name and surname of student: Natália Komorníková

Thesis name: Development and analysis of a database of reactions catalyzed by cytochrome P450 enzymes for machine learning applications

A. Point assessment of individual aspects of the work (choose only one option)

1. S	1. Scale of BT and its structure	
A	A -excellent	
	B – very good - it is however not balanced – scale of some parts does not reflect their importance	
	C – good, but some parts are not covered sufficiently	
	N - unacceptable	

2. I	2. Language level		
A	A - excellent		
	B – very good, only a few minor errors or typos found		
	C - good, minor errors are common		
	N – unacceptable, many serious mistakes and errors		

3. V	3. Work with literarure		
A	A – excellent – citations are relevant and appear at relevant places		
	B – very good – minor objection to a number or a placement of citations		
	C – good, wrong format or missing citations do appear		
	N – unacceptable – a lot of missing citations		

4. (. Clarity of the work	
A	A – excellent – the thoughts are well-formulated, it is a pleasure to read	
	B- very good – there are a few clumsy expressions or unclear formulations	
	C - good, clumsy expressions or unclear formulations are more common	
	N – unacceptable, the the work is difficult to understand	

5. F	5. Formal and graphical level of the work		
A	A - excellent		
	B – very good, but individual missing links to figures, shortcuts etc. appear,		
	C - good, bigger errors like missing pages appear		
	N – unacceptable, many serious errors		

Please add any comment to the above (if you feel like doing so):

Thesis about establishment of cytochrome P450 (CYPs) reactions database is well written and documented.

If the work contains results of the authors (it is not mandatory) please comment the following:

Are the aims of the work clearly stated? **YES**

Is the amount of the experimental work equivalent to the aims? YES

Are the results well documented? YES, GitHub and GoogleDrive is nice way to present data

although Github repository is not public (yet?) and personal

GoogleDrive is not really permanent repository.

Are the results discussed with existing literature? YES

Please add any other comment to the work of student:

Thesis starts with short but concise description of cytochrome P450 nomenclature and chemistry, followed by the description of available reaction databases and a way how to gather the data from them. Analysis of collected data is relatively short but it shines lights into the future possibilities on the topic.

B. Defense

Your questions to the author (please have at least one)

- 1. Since you have gathered cytochrome P450 reaction data from Uniprot and Rhea database was it possible to cross-validate them? Was there significant overlap in their sources, and were there any disagreements (e.g., reaction going wildly differently)?
- 2. Humans have 57 known CYP genes are any of them in uncharacterized section of your work or they are completely characterized already?
- 3. In Plantae, it is hypothesized that multiple CYPs are following each other in consecutive reactions channeling metabolites from one CYP to another in the so called "metabolon" complex. Have you found any example of that in your database i.e., product from one reaction being a substrate for another?

C. Final assessment

I do recommend this work to be accepted: YES

Suggested classification: A

Date: 1.6.2023

Name and signature: doc. RNDr. Karel Berka, Ph.D.

Department of Physical chemistry, Palacký University Olomouc

Instruction for filling and sending this form:

- Please use this form for your assessment.
- Please send the filled form to the following address: marian@natur.cuni.cz. We also need an original of signed document either send it to the following address: sekretariát Katedry buněčné biologie PřF UK (p. Růžičková), Viničná 7, 128 44 Praha 2, or being it to the defence.
- Student shall get your assessment at least three days before the defence you can send it yourself or we can do it for you.