



Warsaw, 05 November 2025

Marcin Grynberg, DSc

Institute of Biochemistry and Biophysics,

Polish Academy of Sciences

### **Review of the doctoral dissertation of Andrea Cappannini, MSc**

entitled: *Development of MODOMICS and NACDDB databases and their application to studying the impact of RNA modifications on structure* carried out at the Laboratory of Bioinformatics and Protein Engineering at the International Institute of Molecular and Cell Biology in Warsaw under the supervision of Professor Janusz Bujnicki, DSc.

#### **1. Legal and formal basis for the preparation of the review**

The legal basis for the preparation of this review comprises:

1.1. The doctoral dissertation submitted by Andrea Cappannini, MSc.

1.2. The Act of 14 March 2003 on academic degrees and titles, and degrees and titles in the field of art (Journal of Laws No. 65, item 595, Article 13.1, as amended), as well as the provisions introducing the Act – Law on Higher Education and Science of 3 July 2018 (Journal of Laws of 2018, item 1669), consolidated text: Journal of Laws of 2021, item 478.

This review pertains to the substantive evaluation of the doctoral dissertation (formal criteria are beyond the scope of this assessment). The dissertation should constitute an original solution to a scientific problem, demonstrate the candidate's general theoretical knowledge in the field of natural and exact sciences, and confirm the ability to conduct independent scientific research. These criteria will be taken into account in the evaluation of the dissertation submitted by Andrea Cappannini, MSc.

## 2. Factual basis of the review

This review has been prepared based on the assessment of the doctoral dissertation submitted in the form of a monograph. The dissertation comprises 176 pages, including a table of contents, abstracts in Polish and English, an introduction, aim, results, discussion, conclusions, future perspectives, materials and methods, table of figures, table of tables and a bibliography.

The submitted dissertation is an original work, as confirmed by the similarity indices provided in the anti-plagiarism report. The only identified similarity, at the level of 1%, pertains to phrases matched against the repository of diploma and doctoral theses.

## 3. Objective of the doctoral dissertation

The main aim of the work is to ‘investigate the hypothesis that post-transcriptional RNA modifications influence RNA structure’.

The steps that lead to verification of this question are:

1. Assessment and refinement of RNA modification data: Critical evaluation and validation of RNA modification data deposited in MODOMICS, with a focus on confirming the accuracy and quality of annotations related to experimentally determined RNA structures.
2. Structural and evolutionary analyses of RNA modifications: Analysis of RNA sequence alignments and experimentally determined structures stored within MODOMICS, investigating the evolutionary conservation of modified positions in tRNA and rRNA, and assessing their structural implications.
3. Integration of Circular Dichroism spectroscopy data: Establishment and expansion of the Nucleic Acid Circular Dichroism Database (NACDDB) to provide circular dichroism spectra for RNA molecules, including modified nucleic-acid oligos. Development of computational tools based on Cosine Similarity and Root Mean Square Deviation algorithms enabled quantitative comparisons of nucleic acid structures, further elucidating structural effects induced by specific RNA modifications.

Based on the results presented in the dissertation, it can be concluded that the primary objective of the study has been achieved, and the utility of the newly developed methods has been confirmed. The specific objectives (1) and (3) have been fully accomplished, while objective (2) has been partially fulfilled.

In light of the stated aim and the proposed methodological approach, it is evident that the subject matter of the dissertation is consistent with its content.

## 4. Substantive evaluation of the doctoral dissertation

### 4.1. Assessment of the topic selection and scope of the study

The candidate undertook the analysis of a highly significant yet ambitious subject, namely the modification of the MODOMICS database and the development of the NACDDB database. The necessity and objectives of establishing such publicly accessible resources are self-evident; hence, their creation constitutes a justified and purposeful endeavor.

Both the selection of the research topic and the execution of the computationally demanding project are, in my view, appropriate and well-founded.

#### 4.2. Structure of the doctoral dissertation

The structure of the reviewed dissertation is, in principle, appropriate. It comprises classical parts typical for such works, however, some chapters are quite messy, like the material and methods section or bibliography. For instance, the 7.6 section describes methods used to build NACDDB. Subchapter 7.6.8 describes circular dichroism experiments, which are the prerequisite to the rest of this work. Why at the end? Another issue are the references that are nowhere near a proper list of articles cited: formats vary, no numbers describing articles are present (even if there are 'mysterious' doubles, like Su et al., 2022). Therefore, navigation in this chapter was highly difficult.

#### 4.3. Substantive Evaluation of the Dissertation

The work presented in this dissertation is certainly useful and it must be underlined as a basis of the analysis of this work. Both main works realised by Andrea Cappannini, M.Sc., which are modifications and building of databases represent logical steps in the development of the laboratory of prof. Bujnicki.

What should be underlined here is the introduction, which is well built. To put it in short words, it's a perfect 101 course for students. I appreciate it.

I also like the presence of the 3.1.16 section (Limitations of the analyses). It always gives this impression of a mature researcher.

I appreciate the statement about the help of AI to properly use English.

Even if the structure, parts of data presented in this dissertation are valuable, my main concern with the text is that it's sloppy. I'll name my comments and questions below:

1. In the Results section some figures are barely readable, badly planned, like Figure 12. Why not to put them vertically in order to enlarge descriptions?
2. In most cases figure titles are very modest; they lack colour descriptions, sometimes it is not known what the drawings refer to, e.g. figure 28A: what is the percentage presented? What are the Y axis symbols? Unfortunately, definitions from page 12 do not refer to all symbols in fig. 28A.
3. In the RNA modifications section figures each modification changes colour with each figure. It's very hard to follow.
4. Quite often figures are not in the vicinity of the chapter it refers to. Sometimes two pages away. Why?
5. Surprising were 2-page figures with no explanations it's a 2-page figure.
6. In chapter 3.1.15 you state: '[...] This data supports a model of reinforcing or suppression of evolutionary pressure [...]'. How? What are the rules then?
7. Chapter 3.2.3: You used Kendall's tau. Why? I couldn't find an explanation. Why, for instance, not Spearman's rho or any other method?
8. Discussion section is a gem. I started to understand some elements of your work because they were properly explained only there. For instance, I learned why it was so crucial and what were the links to include the matter of SciModoM, which you do not mention before. Why so?

9. 7.5.1: No idea what methods were used to align sequences. To my knowledge, it is not mentioned in this chapter.
10. 7.6.8: No idea how many times the experiment was repeated per wave length, how many sequence and modification versions were used. You lack data here, even if you did not make it personally.
11. Table of figures and Table of tables lack page numberings, which would help to locate them.
12. References start with names or surnames, no rule. Some cited references do not exist, like Boileau, 2024.

Additionally, I would like to ask the following questions:

1. Can you think of a more ambitious hypothesis for the results you obtained? Any change in any structure usually leads to something. It seems obvious. So what question would you ask now?
2. On page 144 you speak of 'visual assessment'. Can you think of a method to automate this procedure?
3. There are many environmental reasons for RNA structure dynamics, like metal binding that you mentioned in the introduction. How would you measure/add these effects to MODOMICS/NACDDB?

## 5. Summary

The doctoral dissertation submitted for review by Andrea Cappannini, MSc. falls within the scope of fundamental research concerning the structure and function of RNA. The research problem undertaken is of considerable importance, particularly in terms of organizing and systematising the broad field of RNA studies. The drawback is data presentation. I hope these remarks will help the candidate to improve its next writing endeavours.

An analysis of the submitted dissertation indicates that Andrea Cappannini, MSc, has achieved the learning outcomes defined by the Polish Qualifications Framework for level eight (doctoral level).

## 6. Final conclusion

I, the undersigned, hereby state that the doctoral dissertation of Andrea Cappannini meets the requirements specified in Article 187 of the Act of July 20, 2018 – Law on Higher Education and Science (c.t., Journal of Laws of 2024, item 1571, as amended). I hereby recommend to the Doctoral Committee of the International Institute of Molecular and Cell Biology in Warsaw to admit Andrea Cappannini to the subsequent stages of the procedure for the conferment of the doctoral degree in the field of natural sciences, in the discipline of biological sciences.