PhD Thesis Review (Ján Hájek): Structural variability of cyclic cyanobacterial lipopeptides: their biosynthesis and features affecting their bioactivity

Cyanobacteria are ancient and ubiquitous prokyarotes that dedicate a large percentage of their genome to secondary metabolism. The majority of these secondary metabolites are modified peptides or peptide-polyketide hybrids. Most of the biosynthetic hybrids are encoded by nonribosomal peptide synthases (NRPSs) and polyketide synthases (PKSs) whose action is required to assemble the products that we call natural products. They act on various therapeutically relevant targets and possess intriguing bioactivity. One prime example is the marine natural product and NRPS/PKS hybrid dolastatin 10 which provided the basis for several FDA-approved anticancer drugs. Dolastatin 10 is a linear modified peptide; however, a major portion of cyanobacterial compounds are cyclic and additionally possess lipid moiety as two characteristic features. The PhD candidate's thesis centers around cyclic lipopeptides produced by cyanobacteria. Specifically, the structural diversity is discussed, the NRPS- and PKS-mediated biosynthesis, and insight into the cytotoxicity against fungi and mammalian (cancer) cells in a structure-dependent manner.

Part I of the thesis provides a brief background on NPRPS and PKS and mainly a comprehensive review of the literature describing the structural diversity of cyclic peptides with fatty acid chain units. The systematic classification (Figure 2) is very helpful to the field and urgently needed, given the large number of compounds that fall into those categories.

Part II is a discussion of puwainaphycins (PUW) and minutissamides (MIN) and related beta-amino fatty acid lipopeptides, which are the subject of the presented publications in **Part III**. The modular biosynthesis is well described based on the published PUW biosynthetic gene cluster. Unambiguous mass spectrometry-based identification and sequencing of cyclic peptides with fatty acid moiety is highly challenging. The candidate described the successful method development using different collision energies. Evidence is presented that the fatty acid moiety is responsible for the biological effects. Specifically, these compounds interact with biological membranes and therefore induce cytotoxicity. A longer fatty acid chain appears to induce a stronger interaction. Semisynthesis yielded additional analogues to probe the structure-activity relationship for these lipopeptides.

Part III includes four papers in very good international journals, including one first-author publication in *RSC Advances*.

Significance:

The general significance of the topic is high, given the demonstrated therapeutic value of cyanobacterial compounds, warranting in-depth studies and analyses on the genomic and metabolomic level. Specifically, the systematic classification of cyclic lipopeptides is of great value to the field. Also, the developed HPLC and mass spectrometric methods are important for others studying these types of compounds and will aid the field. The variation in fatty acid chain length and distribution among strains is of biosynthetic interest and the effect on biological effects is of interest to chemical biologists.

Overall quality:

The PhD thesis suggests that the candidate has become a well-rounded natural products expert with a multidisciplinary skill set required for modern natural products science. The quality of the document is high and most of the research has already been peer-reviewed and published. Only relatively minor edits are recommended by this reviewer (see below). Based on the written document and publications, the candidate has strong skills in analytical chemistry with emphasis on separation and mass spectrometry.

He also demonstrated knowledge of the biosynthetic machinery and pathways to hybrid peptide-polyketides as well as bioactivity studies and even semisynthesis.

Comments:

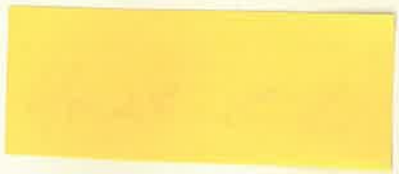
1) Most secondary, or specialized, metabolites from cyanobacteria particularly from the marine environment have been described from *Lyngbya majuscula*. With the availability of genomic data, the taxonomy has changed. I recommend including a statement that the taxonomy is in flux and that the producing genus/species mentioned for each compound in this thesis are based on the original literature data. For example, lyngbyabellins A and B (page 41) and ulongamides A-F (page 48) are now known to be produced by *Moorena bouillonii*; however, these compounds were originally published as *Lyngbya majuscula* metabolites some 20 years ago.
2) NRPS is misspelled at certain places (NPRS). Same for RSC in RSC Advances (RCS).
3) I recommend that the Introduction be published in the scientific literature in form of a Review. Is this the plan?
4) When discussing these structures; however, the configuration should not be ignored, since absolute configuration is critical to bioactivity and also genetically encoded (presence or lack of epimerase).
5) No distinction has been made between marine and terrestrial compounds. Although understandably beyond the scope of the PhD thesis, it would be really interesting to have a comparison of terrestrial and marine metabolites for each of the sub-classes. What is the distribution and what is the overlap?
6) Page 10 footnote suggests that many compounds from *Dolabella auricularia* are products from *Lyngbya majuscula*. While it is true that they are produced by cyanobacteria, they are not necessarily from *Lyngbya majuscula*. This statement here and elsewhere should be generalized (marine cyanobacteria). For example, dolastatin 10 is produced by *Caldora penicillata*, a relatively new genus that was described recently. Originally it was described as a *Symploca* metabolite (see general comment #1). It is unclear if reference 139 is relevant, which should be double checked.
7) The key pharmacophore in the serine protease inhibitors, Cyanopeptolins, is described in most literature as Ahp (not Ahopip). I recommend adding or replacing the current abbreviation. In the text and in Table 2, other related Ahp-containing serine protease inhibitors could be mentioned, including lyngbyastatin 4-10, symplostatins 5-10, loggerpeptins A-C, and tuilamides A-C, some of which have been synthesized and more rigorously biologically investigated.
8) Page 17 statement may need some clarification. While the macrocyclic core of cyanopeptolins is critical for activity, it is insufficient for activity. Recent SAR studies suggest that an additional amino acid unit is necessary. It is correct that the fatty acid does not play a role (*ACS Med. Chem. Lett.* **2020**, *11*, 419–425).
9) Stereocenters are usually not indicated; however, configuration is critical for bioactivity. For most compounds the configuration is known and should either be indicated in the structure or should be mentioned in the corresponding text for each section.
10) Page 44, Figure 13 legend: Should be subclass N-1 and not O-3.
11) Page 46: One needs to be careful generalizing malevamides as cyclic. While true for most malevamides, malevamide D is linear.
12) Page 78 states that PUW/MIN compounds belong to N-3 type; however, page 58/59 indicate that they are N-4 type. Please clarify.

Publications and research output:

The candidate has been highly productive during his PhD. His research has culminated in 15 publications. Although most of them were co-author publications, he also published one high-quality first-author paper. In addition, he presented his work at three national and international conferences.

Recommendation:

I recommend the PhD thesis for defense.



Hendrik Luesch, Ph.D.
Professor and Chair, Department of Medicinal Chemistry
University of Florida