Review of: "Redirecting RiPP biosynthetic enzymes to proteins and backbone-modified substrates"

Sylvie Rebuffat[1]

1 Muséum National d'Histoire Naturelle

Potential competing interests: The author(s) declared that no potential competing interests exist.

The paper from J. Walker et al. describes that the chemical diversity of ribosomally synthesized and posttranslationally modified peptides (RiPPs) can be expanded *in vitro* by including aromatic polyamids (aramids) and beta-amino acids in amino acid sequences, thanks to the ability of cyclodehydratase/heterocyclase and dehydrogenase enzymes from the cyanobactin family to accept such non-proteinogenic and non alpha-amino acid entities. For this, they used the MicD fusion cyclodehydratase/heterocyclase MicD-F (MicD fused to the leader region) from *Microcystis aeruginosa* that is involved in biosynthesis of microcyclamide, and ArtGox, the dehydrogenase from *Arthrospora spirulina* involved in arthrospiramide biosynthesis. Both enzymes were shown previously to accept leaderless peptide substrates and thus can constitute a good model system for the obtention of thiazole/oxazole-containing products. Major results arise from the study: - MicD and ArtGox act in synergy to process amino acid sequences including non alpha-amino acid monomers, including aramids, located at the +1 site (N-terminal to the processing site) and at the -1 site (C-terminal to the processing site) as well, although with a lower tolerance; - MicD-F is not tolerant to non alpha-amino acids at the processing site; - installation of thiazole/oxazole by MicD-F/ArtGox is possible in folded proteins, either alpha-helix bundles or beta-barrels, without perturbing extensively the three-dimensional structures; - introduction of thiazoline/thiazole instead of cysteine in a short model peptide shows a significant reduction of flexibility, as evaluated by molecular dynamics.

This is an in-depth study of the requirements of cyclodehydratase and dehydrogenase enzymes to modify peptide/protein sequences to incorporate thiazol(in)e/oxazol(in)e heterocycles in various contexts, including non canonical amino acids or complex three-dimensional structures. The need for novel biologically active natural products with various major activities including antimicrobial, antiviral or anticancer properties that are critical for biomedical applications, largely justifies the approach.

The study relies on bioorganic chemistry experiments using appropriate and effective methods (presumably as the experimental conditions are not provided...) for expression of the RiPP enzymes and LC-MS for identification of the compounds formed in the enzymatic reactions. The study is well designed, the logic of the paper is clear and the data appear to support the conclusions within the limit of available data (as the 20 supplementary figures cited in the main text are lacking...). The figures are clear, informative and well commented in the legend, providing self-understandable information.

The manuscript provides original information in the domain of RiPPs and especially of cyanobactins taken as model and opens perspectives in the field. The study is self-consistent and the paper will be presumably of high value for a wide range of scientists interested in RiPP biosynthesis, and more largely in chemistry of natural products, medicinal chemistry and pharmacology, provided it is made more easily accessible to non specialists of the RiPP domain and the supplementary data are made available to allow a visualization of many results.
Moreover, a number of points detailed below need clarification or improvement.

Specific comments

1. It is to note that the supplementary figures were not available on the Biorxiv website, and could not be obtained neither by asking the authors for them. In addition, the experimental procedures are not described in the paper. Thus a detailed analysis of the amount of work done and the rigour of the experiments is difficult. Moreover, the lack of many data makes the paper hard to read and the logic more difficult to follow.
   It is a pity that so many interesting and important data that should be presumably provided in Supplementary Figures 1 to 20 are not available to the scientific community.
2. The paper is mainly written for a well-informed community specialized in RiPP biosynthesis and enzymology, and more specifically cyanobactin biosynthesis. Writing of the paper merits being improved to permit the paper concerning a broader scientific community.
   - First, the choice of these specific enzymes from cyanobactin biosynthetic pathway among the known and previously described enzymes that install thiazol(in)e/oxazol(in)e rings has to be justified and argued (for instance accessibility of the enzymes thanks to efficient expression/production procedures, well described biosynthesis pathway, specific characteristics,...).
   - Second, a detailed description of the selected enzymes and especially their origin (genus and species of cyanobacteria), specific characteristics, requirement or not of a leader peptide, processionality, promiscuity,... have to be decribed.
   - Third, the global strategy used in the study to incorporate non-coded, non alpha-amino acid blocks has to be defined in more detail in the text since the beginning of the paper, and not only in the figure legends, although they are very clear. The choice of the CAYDG sequence (including the recognition AYD motif) has to be justified in the description of the global strategy.
   - Moreover, the interest and further consequences/applications of validating the acceptance of thiazole and oxazole rings in beta-barrel (mCherry) or alpha-helical bundle (Rop) protein context has to be announced in the introduction and not only provided in the conclusion.
3. Globally, the experimental procedures are not described. In particular, the description of the conditions used for expression, production and purification of MicD-F and ArtGox has to be provided. Similar, procedures for cloning, expression and purification of mCherry variants (with MCAYDG at the N-terminus, C-terminus or inserted at different positions in the loops) and Rop variants (with CAYD motif

at the N-terminus, C-terminus or central position) have to be provided.

4. The numbering “+1 site” and “-1 site” as regard the reaction site is confusing, as it works in contradiction with the usual numbering used to name the sites before (…, -2, -1) and after (+1, +2,…) the cleavage site between the leader and the core peptides in RiPPsbiosynthesis. This makes the present numbering puzzling.
5. The abbreviation “MicD-F” for MicD fusion is awkward, as it is similar to the usual abbreviations used in the literature for naming enzyme or protein complexes (MicD-F can be understood as a complex MicD/MicE/MicF). Thus MicD-f would be unambiguous and thus more appropriate.

Minor points

1. Page 5 line 5 change “(substrates 1(a-i))” to “(substrates 1(a-i); Figure 2A, B))”
2. Page 8 Figure 3, (D) is lacking on the figure.
3. Page 10, for clarity separate the legend to Figure 4 from the text.
4. Please provide page numbering.