Prof. Matthias Bochtler

International Institute of Molecular and Cell Biology &
Trojdena 4, 02-109 Warsaw, Poland
Institute of Biochemistry and Biophysics
Pawinskiego 5a, 02-106 Warsaw, Poland
Tel: 0048 22 5970732
Fax: 0048 22 5970715
e-mail: mbochtler@iimcb.gov.pl

October 31st 2019

Opinion on the PhD thesis by Mariusz Madej

Mariusz Madej has presented a PhD thesis on the "Structural and functional characterization of (a) unique peptide transporter from Gram-negative bacteria *Porphyromonas gingivalis*". The thesis reports the discovery, based on structural data, that RagAB is involved in peptide uptake, and it clarifies the mechanism, and specificity of this novel uptake pathway.

To put the work into perspective, it is useful to recall that *P. gingivalis* bacteria belong to the Bacteroidetes phylum, one of the two most prominent phyla in the human microbiome. *P. gingivalis* is of interest in particular as "unfriendly" member of the human microbiome, and as a dental pathogen, which primarily causes gum disease, but can also negatively affect the cardiovascular system, and has recently been implicated, albeit not necessary causally, in Alzheimer's disease. Like many commensal bacteria, *P. gingivalis* is metabolically less proficient than many other bacteria. In particular, *P. gingivalis* bacteria are asaccharolytic, i.e. deficient in carbohydrate processing. This makes them heavily dependent on peptides as a source not only of nitrogen, but also of carbon. Prior to Mariusz Madej's work, it was already known that *P. gingivalis* secretes a battery of proteolytic and peptidolytic enzymes, first and foremost the gingipains KgpA, and RgpA and RgpB, to break down host proteins. It was also established that the bacteria can take very short peptides and single amino acids, but it was not clear whether a dedicated import pathway existed.

RagAB, the topic of the PhD thesis, is the protein product of the ragAB locus, which codes for a ~115 kDa RagA protein and a ~55 kDa RagB protein. Prior to Mariusz Madej's work, it was known that RagA was an outer membrane protein, with a controversial role in virulence. Moreover, it was known that several paralogues existed, and that paralogues were present or expressed in a strain-specific manner. Based on sequence comparisons, it was possible to deduce that RagAB belonged to the SusCD/PUL family of proteins, which are believed to be involved in carbohydrate uptake. However, because of the asaccharolytic nature of *P. gingivalis*, it was very unlikely that RagAB had the "canonical" function of other members of the protein complex family. Although the pieces of the puzzle clearly fall into place, this is not a lot start a thesis on.

Prof. Matthias Bochtler

International Institute of Molecular and Cell Biology &
Trojdena 4, 02-109 Warsaw, Poland
Institute of Biochemistry and Biophysics
Pawinskiego 5a, 02-106 Warsaw, Poland
Tel: 0048 22 5970732
Fax: 0048 22 5970715
e-mail: mbochtler@iimcb.gov.pl

Mariusz Madej approached the problem from a structural biologist's perspective. This was a very courageous move, since RagA is membrane embedded, and since membrane proteins are notoriously difficult to handle. RagA's location in the outer membrane (OM) rather than the inner membrane (IM) made success somewhat more likely, but still, it was a very daring project. Adding to the difficulty, the large phylogenetic distance between *Escherichia coli* and *P. gingivalis* meant that "standard" membrane protein expression in *E. coli* was unlikely to work, and that the protein had to be purified from *P. gingivalis* bacteria. Mariusz Madej did this in two different ways, either "classically" without reliance on a protein tag, or by introduction of an affinity tag by homologous recombination, followed by an affinity-based purification. Luckily, the purified RagAB complex turned out to be very stable, and in fact stable enough to obtain a crystal structure of the complex, which could be solved by molecular replacement.

At this point, luck helped greatly, but of course it takes a prepared mind to take advantage of the moment. Mariusz Madej found density for a long peptide in the complex. This of course suggested that RagAB could be the long sought-after importer. And of course, it immediately raised the question how the peptide transport works. Using X-ray crystallography, addressing such questions has taken years, when it has worked at all, in the most favourable cases. Clearly, cryo-EM was the way to go for to separate individual states of the transporter. It's far from trivial that cryo-EM reaches atomic or near atomic resolution, particularly for membrane proteins, even when a high resolution crystal structure is possible. In Mariusz Madej's case, cryo-EM did yield high resolution structures that also were informative about mechanism. This left the question of the mechanism of substrate selection. Using a combination of mass spectrometry and microscale thermophoresis, Mariusz Madej could show that the RagAB proteins have distinct specificities, which differ between paralogues.

Altogether, the breadth of methods that Mariusz Madej has used in his PhD thesis is stunning. He has genetically engineered suitable *P. gingivalis* strains, purified membrane proteins (protein complexes), in some cases without being able to rely on an affinity tag to simplify the task. The work involved crystallizing a membrane protein complex embedded in the outer membrane, and solving the structure. Here, for once, Mariusz Madej had it relatively easy, since the phase problem could be solved by molecular replacement. Had he gotten up to this point only, Mariusz Madej would have more than deserved a PhD. However, we went on to use cryo-electron microscopy to characterize the different states of the reporter, and mass spectrometry and microscale thermophoresis, to characterize cargo specificity. Finally, there is also a considerable and well-written bioinformatics part, which puts the results for the *P. gingivalis* RagAB complex into a broader perspective.

The presentation and editorial standard of the PhD thesis are also very high (repeated use of "clad" instead of "clade" is almost the only error that I found). The presentation is logical and

Prof. Matthias Bochtler

International Institute of Molecular and Cell Biology &
Trojdena 4, 02-109 Warsaw, Poland
Institute of Biochemistry and Biophysics
Pawinskiego 5a, 02-106 Warsaw, Poland
Tel: 0048 22 5970732
Fax: 0048 22 5970715
e-mail: mbochtler@iimcb.gov.pl

clear, and the English is fluent. In some cases, the shear breadth of methods that Mariusz Madej has used in his thesis led him to describe individual methods, particularly those for which he relied on collaboration, such as the cryo-EM work, rather summarily. This is fine, it's the privilege of the very successful PhD students, but I hope that some of the methodological details that I am personally interested in can be clarified at the defence.

For my own education, and not in any way to criticize the work, I would like to discuss the following questions at the thesis:

Technical:

- Why was a detergent exchange required prior to cryo-EM analysis? Why could it not have been done in the same detergent as crystallization?
- How was the number of receptor states in cryo-EM determined? Was this based on manual inspection of micrographs, or was some more objective procedure used?
- What fraction of particles in a micrograph could actually be used? It is common in studies of "flexible" proteins that most images of a complex (sometimes 80-90%) have to be discarded? What fraction of complexes in the mix is described by the analysed states?
- Which criteria were used for validation? The standard correlation criteria measure self-consistency, but not necessarily correctness. Mariusz Madej is in the unusual position of having a reference crystal structure for comparison. Does the comparison support the standard resolution assignment? Have tilt series comparisons been done to check the accuracy of projection angles?

Biological:

- Could it be that other putative carbohydrate channels of the superfamily also are peptide transporters? In particular, how widespread is the asaccharolytic lifestyle in Bacteroidetes that have a SusCD/RagAB homologue?
- Have cases been found whether several RagAB paralogues occur in the same strain? If so, could it be that *P. gingivalis* uses phase variation to switch import on an off? Are there repetitive regions in coding sequence that by expansion or contraction could be involved in such phase variation?

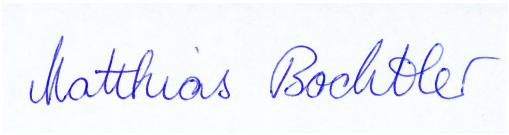
Given the spectacular results of Mariusz Madej's thesis, I obviously support awarding him a PhD and request the Biological Sciences Council of Jagiellonian University to admit Mariusz Madej to further stages of the doctoral dissertation procedure. In Warsaw, the rules require that most of the outcome of the PhD has to be published for the work to be awarded a distinction, which does not seem to be the case. Unless such formalities

Prof. Matthias Bochtler

International Institute of Molecular and Cell Biology &
Trojdena 4, 02-109 Warsaw, Poland
Institute of Biochemistry and Biophysics
Pawinskiego 5a, 02-106 Warsaw, Poland
Tel: 0048 22 5970732
Fax: 0048 22 5970715
e-mail: mbochtler@iimcb.gov.pl

rule out a distinction, **I also strongly encourage awarding him a distinction**. I read in the thesis that the work in under consideration in *Nature Microbiology*, which is of course an excellent journal. I am very optimistic that the work will be accepted, and in fact, I am left wondering why the authors did not try to publish even higher!

With best regards



Matthias Bochtler