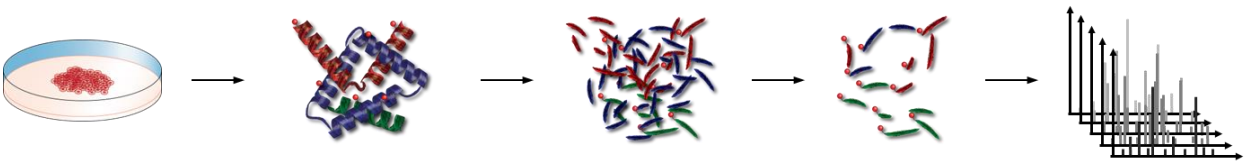
PTM Analysis

A strength of proteomics techniques is the ability to detect post translational modifications (PTMs), major actors of biological processes[1]. PTMs are identified by the mass shift they induce on their site of attachment to a peptide. The modified proteins (and thus the corresponding peptides) are generally present in low abundance compared to non-modified proteins[2]. The mass spectrometer will therefore detect preferentially non-modified peptides, and the modified peptides will most often fall below the limit of detection of the instrument. In order to detect them, it is therefore necessary to filter the peptides for PTMs of interest, a procedure called enrichment. Different protocols have been developed for the detection of phosphorylation, glycosylation, sumoylation, ubiquitination, acetylation, *etc.*, for reviews see[3, 4]. Note that depending on the technique, enrichment can be conducted either prior to or after the digestion of the proteins.



Upon acquisition, the spectra are searched as in the previous chapters, with the inclusion of the PTMs of interest as variable modifications.

What are the consequences of adding these variable modifications? *[1.6a]*

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

In this chapter, we will use a publicly available dataset of mouse brain proteins obtained at different development stages[5, 6]. The samples were analyzed in replicates after fractionation, and the peptides labeled with isobaric reagents prior to enrichment for phosphorylation and glycosylation, as well as acetylation (see the original publication for details on the experimental design). Here, we will solely focus on the identification of peptides from the phosphorylation and glycosylation fraction.

Note that this part of the tutorial has not yet been updated to the most recent version of PeptideShaker. To open the material below please download and use this older version instead:

genesis.ugent.be/maven2/eu/isas/peptideshaker/PeptideShaker/1.16.45/PeptideShaker-1.16.45.zip

To save time, the searching and processing of the data has already been done. In the resources folder, you will find the result file named PTM Tutorial.zip. Open this file in PeptideShaker. You should see the following:

How do the protein results differ compared to a non-enriched sample? [1.6b]

Go to the *Edit* menu and select *Search Settings*.

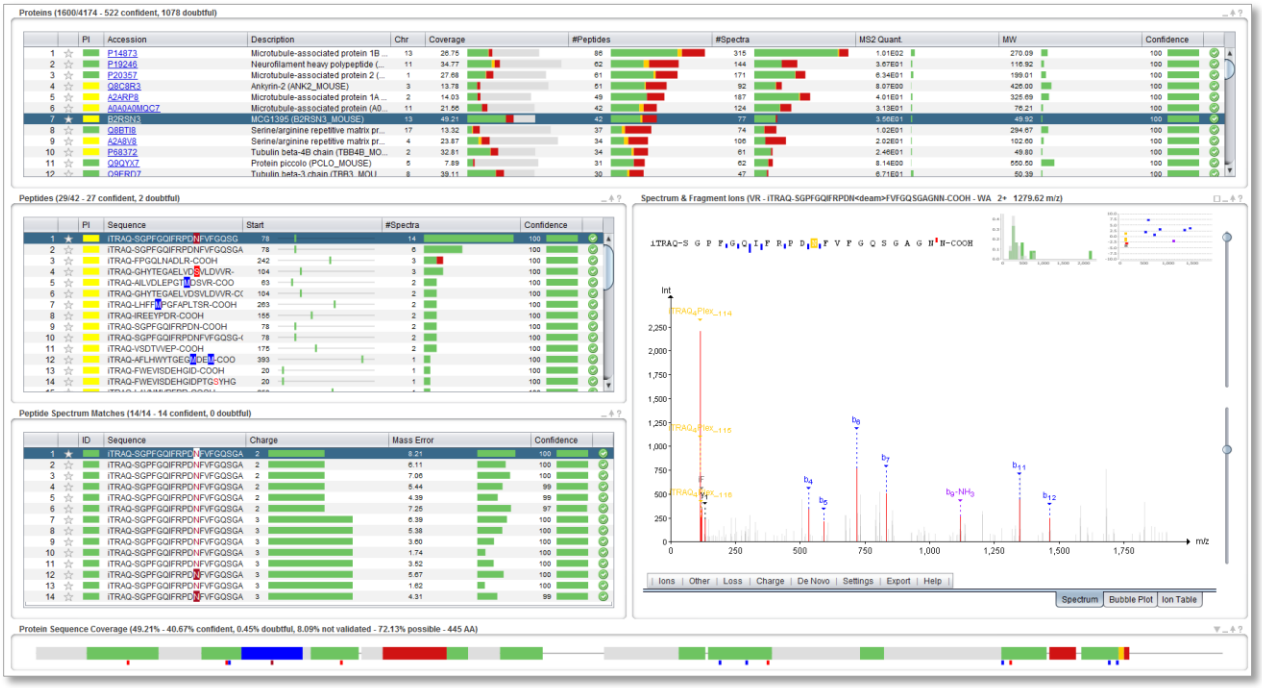
Do you notice any differences compared to the search settings used for non-enriched samples? [1.6c]

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

Note that during the enrichment procedure for glycosylated peptides, the glycan moiety is enzymatically removed, resulting in a mass shift the size of a deamidation at the original glycosylation site. The spectra were therefore inspected for glycosylated peptides using deamidation of N. Note also that glycosylation sites only appear at specific amino acid motifs, namely NX[S or T], where X can be any amino acid except P. Only the motifs presenting a deamidation will thus be retained as potential glycosylation sites.

Is it possible to distinguish glycosylation from deamidation when using the above setup? Why not search glycosylation only at these specific motifs without including deamidation? [1.6d]

Select the protein B2RSN3. You should see the following:



Note how the modifications are color coded in the peptide table and on the protein sequence at the bottom.

How many phosphorylation and glycosylation sites can you find indicated on this protein? How many modified peptides were identified? [1.6e]

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

Select the peptide on line 4 with the sequence GHYTEGAELVDSVLDVVR with a phosphorylation at S12.

Which information was used to assign the phosphorylation to S12 and not Y3 or T4? *[1.6f]*

Next select the peptide at line 17 with the sequence LHFFMPGFAPLTSR with a phosphorylation at S13.

Which peak(s) do we need to decide if this peptide is modified at T12 or S13? *[1.6g]*

As demonstrated with this example, even if the peptide sequence and the modification status are correctly identified, the information required to confidently localize a modification can still be missing. Specific fragment ions are needed to confidently assess the localization of a modification, so-called site determining ions, and these can in many cases either be missing or found with low intensity, often undistinguishable from the noise. As a result, the false localization rate is generally higher than the false identification rate.

How does this affect glycosylation results? Can we inspect every modification site manually? *[1.6h]*

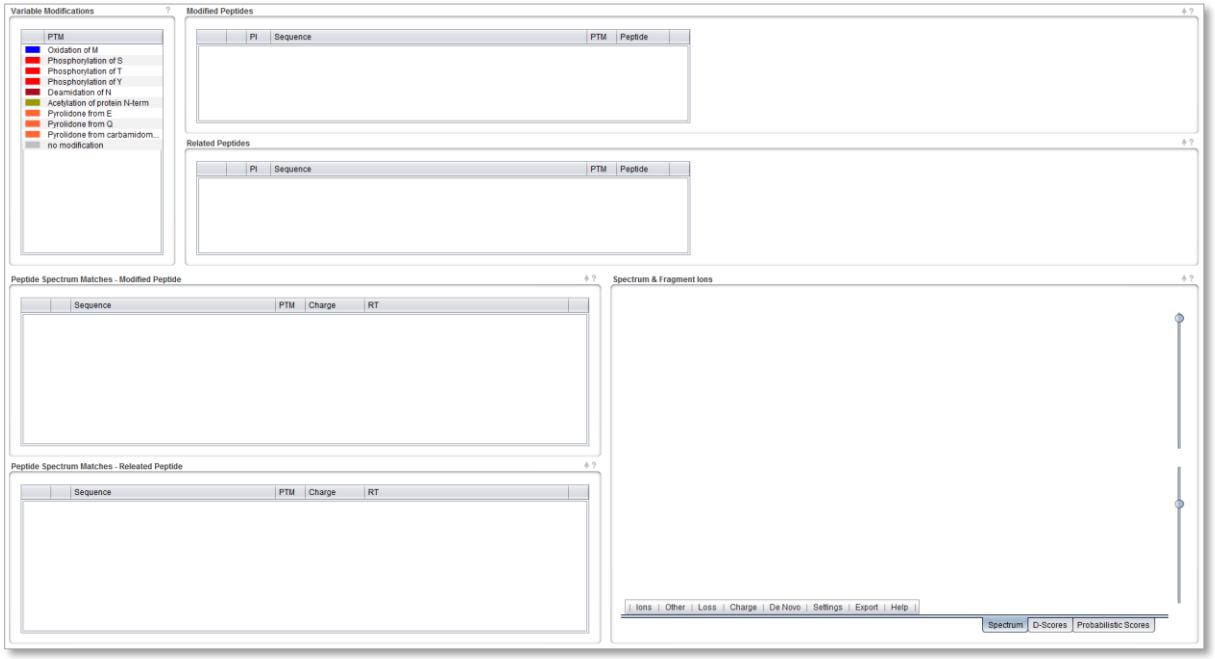
Keen observers will have already noticed that the modification site of the peptide on line 4 is highlighted with a colored background, while a white background with a colored font is used for the peptide line 17. In fact, PeptideShaker has inspected every modified peptide using modification localization scores, and the sites passing a user defined score threshold were considered as significant and highlighted with a colored background. For this purpose two different scores were used: PhosphoRS[7] and D-score[8].

Tip:

When creating a PeptideShaker project, you can change the modification localization scores in the Advanced Settings section of the Identification Settings dialog!

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

Next, select the 'Modifications' tab. You should see the following:



At the top left, you see the list of variable modifications found in the search results.

We will now focus on phosphorylation, so select 'Phosphorylation of S' in the modification table:

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

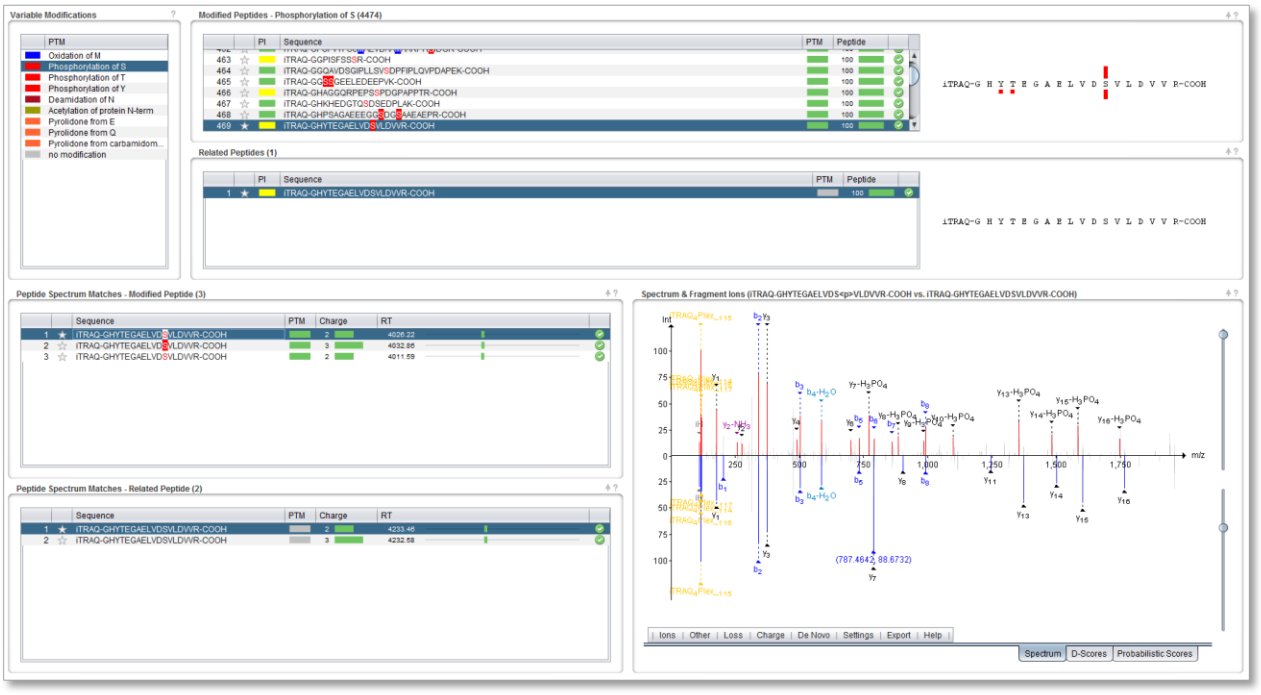
The *Modified Peptides* table at the top lists all the peptides in the dataset carrying this modification. Make sure that the peptide on line 469 with the sequence GHYTEGAELVDSVLDVVR is selected. To the right of the table, you will then see the different scores highlighted on the sequence. Correspondingly, in the *Related Peptides* table, the peptides overlapping with the sequence of the selected peptide is listed.

How do you interpret the relationship between such peptides? How can this be utilized? *[1.6i]*

Select the related peptide. Note that the PSMs found for both of the two peptides are listed in the two tables at the bottom left. In the *Peptide-Spectrum Matches – Modified Peptide* table, you will find three PSMs. Note that you can inspect the D-score and PhosphoRS results for all PSMs in this table in the *D-Score* and *Probabilistic Scores* tabs at the bottom right. For more information about PTM scoring, see the following reviews[3, 9, 10].

For how many of the PSMs from the selected peptide is the modification confidently localized? *[1.6j]*

Next, make sure to have the *Spectrum* tab in the lower right corner selected. Then select the first line in each of the two PSM tables. You should see the following:



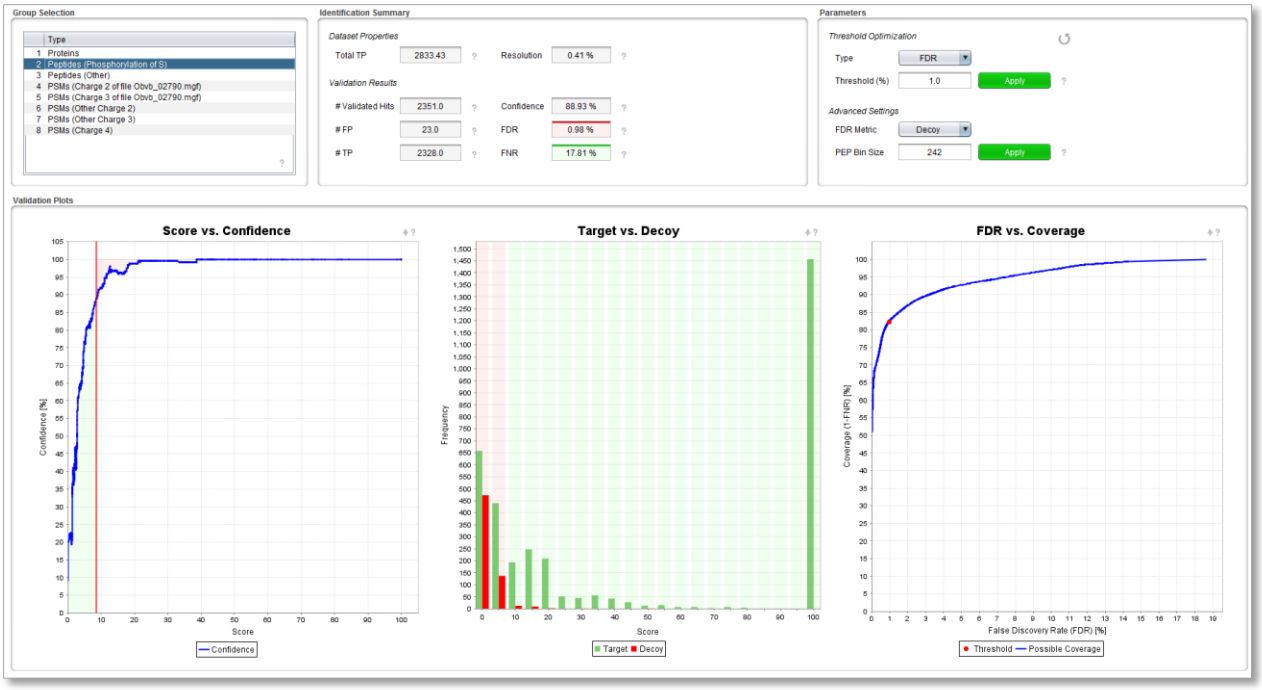
At the bottom right, you can see the spectra of both the modified and unmodified peptides in a mirrored view. *How do the spectra compare for the modified and unmodified peptide?* *[1.6k]*

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

Now select the peptide on line 16 with the sequence AAVGVTGNDITTPPNKEPPPSPEKK.

How many related peptides do you know find? How can these be used? *[1.6l]*

Go to the *Validation* tab in PeptideShaker. At the top left, you will see that both the confidence and the FDR are estimated separately for the phosphorylated peptides. Select the category *Phosphorylation of S*. You should see the following:



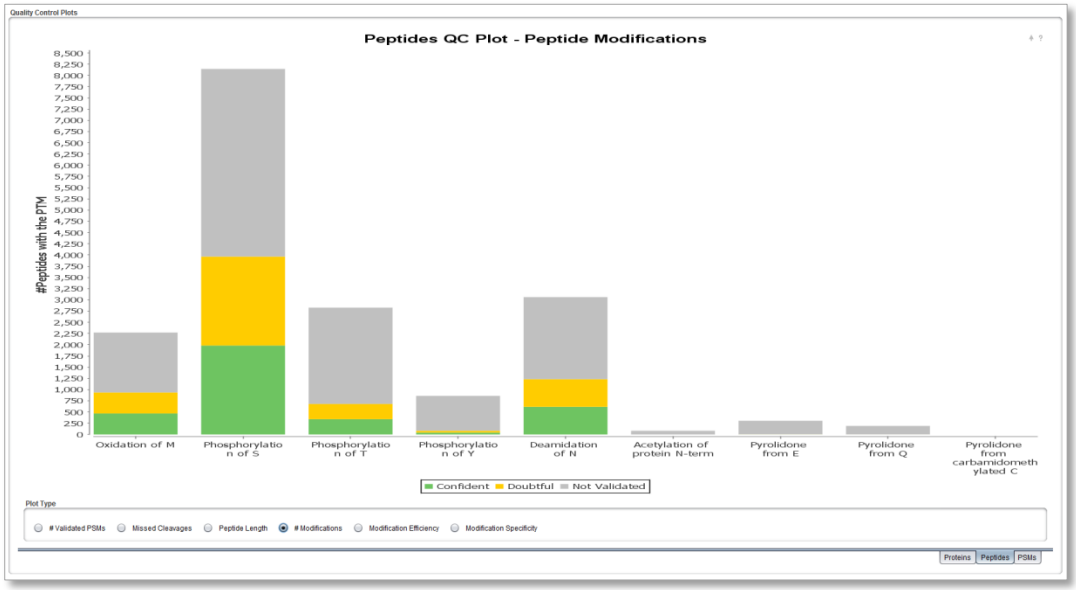
Why are modified peptides treated separately? How do the validation results compare between the modified and unmodified results? Why not group deamidated peptides? *[1.6m]*

Tip:

When creating a PeptideShaker project, you can change the validation parameters in the Advanced Settings section of the Identification Settings dialog!

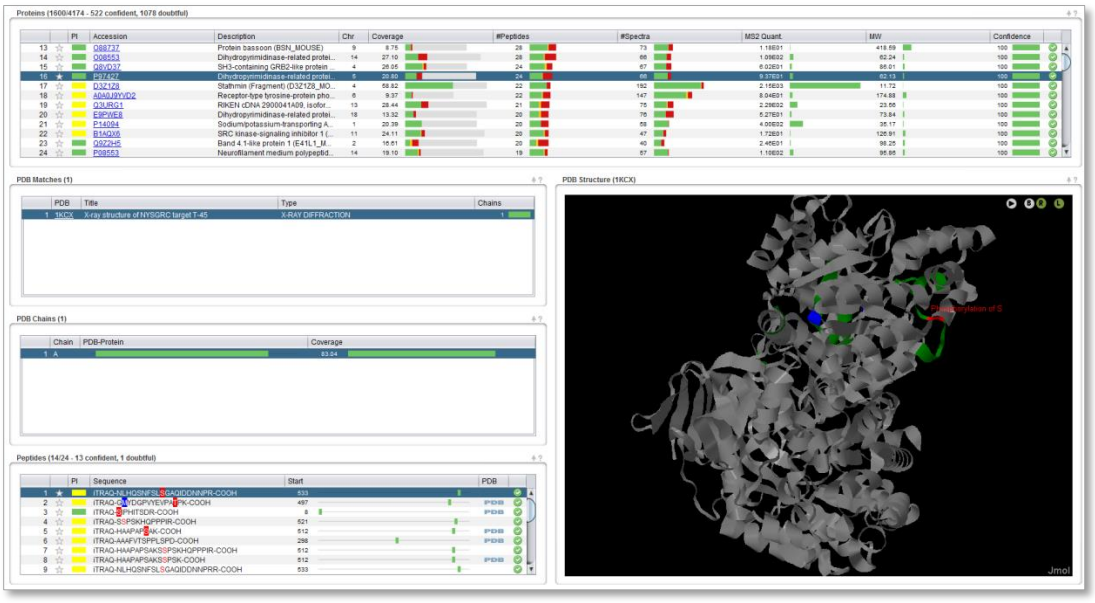
Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

Next, go to the *QC Plots* tab, select *Peptides* and then *# Modifications*:



How can you use the # Modifications, Modification Efficiency, and Modification Specificity plots? [1.6n]

Finally, go to the *3D Structures* tab and select the protein P97427 on line 16. After downloading the structure select the PDB match listed in the *PDB Matches* table:



Where are the modification sites located on the protein structure? Is this as expected? [1.6n]

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)

References

1. Hunter, T. Signaling--2000 and beyond. *Cell* **100**, 113-127 (2000).
2. Solari, F.A., Dell'Aica, M., Sickmann, A. & Zahedi, R.P. Why phosphoproteomics is still a challenge. *Molecular bioSystems* (2015).
3. Olsen, J.V. & Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Molecular & cellular proteomics : MCP* **12**, 3444-3452 (2013).
4. Gevaert, K. et al. A la carte proteomics with an emphasis on gel-free techniques. *Proteomics* **7**, 2698-2718 (2007).
5. Edwards, A.V., Edwards, G.J., Schwammle, V., Saxtorph, H. & Larsen, M.R. Spatial and temporal effects in protein post-translational modification distributions in the developing mouse brain. *Journal of proteome research* **13**, 260-267 (2014).
6. Edwards, A.V., Schwammle, V. & Larsen, M.R. Neuronal process structure and growth proteins are targets of heavy PTM regulation during brain development. *Journal of proteomics* **101**, 77-87 (2014).
7. Taus, T. et al. Universal and confident phosphorylation site localization using phosphoRS. *Journal of proteome research* **10**, 5354-5362 (2011).
8. Vaudel, M. et al. D-score: a search engine independent MD-score. *Proteomics* **13**, 1036-1041 (2013).
9. Chalkley, R.J. & Clauser, K.R. Modification site localization scoring: strategies and performance. *Molecular & cellular proteomics : MCP* **11**, 3-14 (2012).
10. Vaudel, M., Sickmann, A. & Martens, L. Current methods for global proteome identification. *Expert review of proteomics* **9**, 519-532 (2012).

Harald Barsnes (harald.barsnes@uib.no) and Marc Vaudel (marc.vaudel@uib.no)