COMPARISON OF STRATEGIES FOR GLOBAL GLYCOPROTEIN PROFILING BY LC-MS

Javier Beaskoetxea[1], Mikel Azkargorta[1], Alberto Marina[2], Ibon Iloro[1], Iraide Escobes[1], Luca Unione[2], Felix Elortza[1]

[1]Proteomics Platform, CIC bioGUNE, Basque Research and Technology Alliance (BRTA), CIBERehd, 48160 Derio, Bizkaia, Spain. [2]Chemical Biology Lab, CIC bioGUNE, Basque Research and Technology Alliance (BRTA), 48160 Derio, Bizkaia, Spain.

E-mail: jbeaskoetxea@cicbiogune.es

BACKGROUND

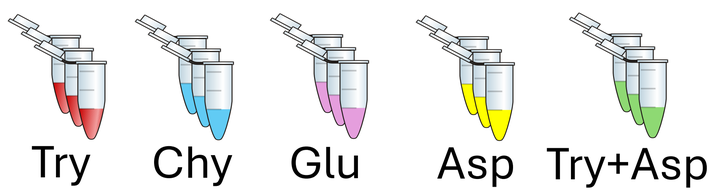
Protein glycosylation is a common post-translational modification, affecting protein folding and stability, trafficking, differentiation, and cell–cell interactions, among others. In addition, aberrant protein glycosylation is a hallmark of cancer, crucial in malignant transformation, tumour development, invasiveness, and metastatic disease.

In this work we have tried to expand and optimize our current protocols for glycoprotein analysis, both at single purified protein and global glycoprotein profiling levels. We have focused on 4 main levels: 1) Sample preparation, 2) Glycopeptide enrichment, 3) Data acquisition and 4) Data analysis.

OPTIMIZATION

1) SAMPLE PREPARATION

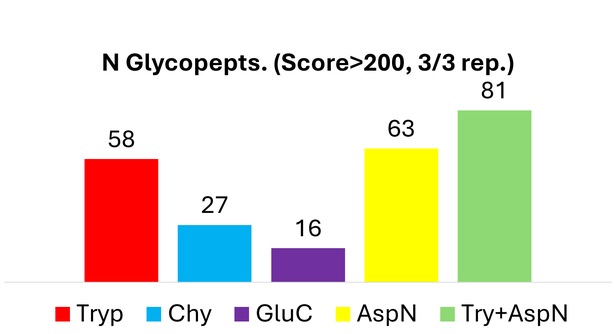
Simple sample - Fetuin (Bovin)

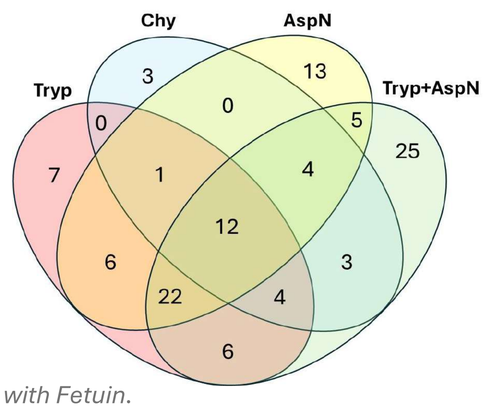


20ug digestions

-We've been using trypsin (**Try**, R/K cleavage) and chymotrypsin (**Chy**, F/W/Y/L/M cleavage) with good complementary results. Yet, still some sites remain undetected for certain proteins.

-What about **adding proteases to the toolbox**? GluC (**Glu**) and AspN (**Asp**) are good options that cleave alternative aminoacids (D/E) – even a combination of them (**Try+Asp**) might be useful.



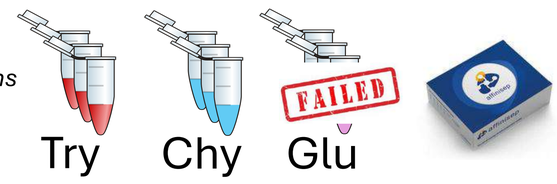


-GluC and Chymo don't seem to work well with Fetuin.
-AspN works well, specially in combination with trypsin.

2) GLYCOPEPTIDE ENRICHMENT

Complex sample - HEK293 cell crude extracts

100/200ug FASP digestions
SepPack cleanup

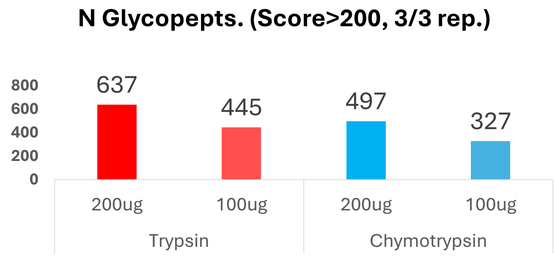


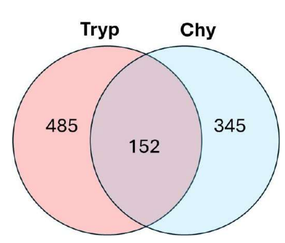
Affinisep Glycaclean Enrichment

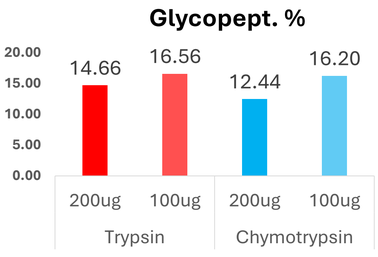
-**Global glycoproteomic profiling** is of special interest in cabcer research.

-Affinisep Glycaclean has rendered good results in our hands but only with **tryptic** simples and at high sample loads (1mg). The goal now is to **downscale** the amount required for the initial peptide load.

-We are also planing to use **alternative sample preparation** kits and check if they provide complementary information. We made a preliminary trial of Thermo Fisher's SOLA-SAX and NEB's N-Glycopepetide Binding Protein, but further optimization is needed (data not shown). GluC also failed to provide useful results.





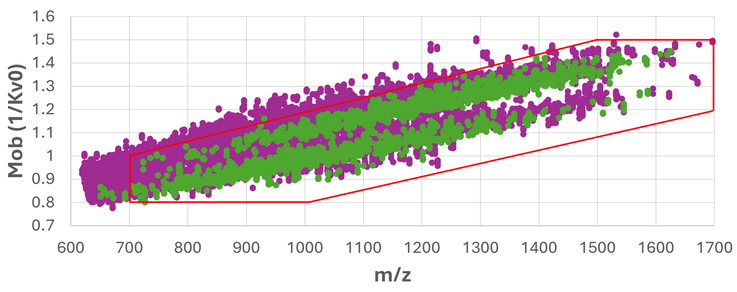


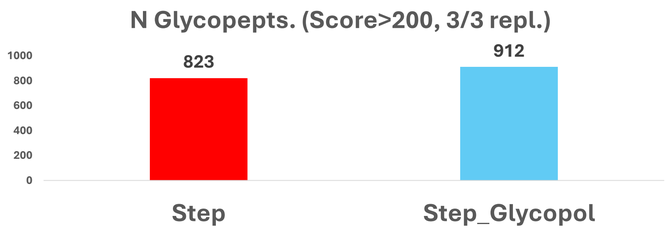
-Affinisep Glycaclean works well with glyco-chymotryptic peptides.
-Chymotrypsin provides complementary results to trypsin.
-200ug loads provide more results than 100ug. However, 100ug loads provide slightly better enrichment.
-Poor overall recovery - room for glycopeptide enrichment improvement.

3) GLYCOPEPTIDE ACQUISITION

-How can we make glycopeptide acquisition **more specific and reduce non-glycopeptide signal** in our samples? The work by Barenefaenger *et al.* [1] defines what they call the "**glycopolygon**" – that is, the m/z and mobility region where most of the glycpopetides lie.

-We defined such glycopolygon using our tryptic and chymotryptic data (in **red** in the following picture) and reacquired the Affinisep tryptic samples (200ug starting material) using both our standard "Stepping" method (defined by Mukherjee *et al.* [2]) as well as that same method plus the glycopolygon applied.





-Overall numbers do not show a big difference (1004 to 1045 increase in average glycopeptides identified when using the Glycopolygon)

-However, the data suggests that using the Glycopolygon provides more reproducible data (912 vs 823 glycopeptide identified in 3/3 replicates)



All samples were acquired in **a timsTOF HT** using the **30 SPD Evosep** method

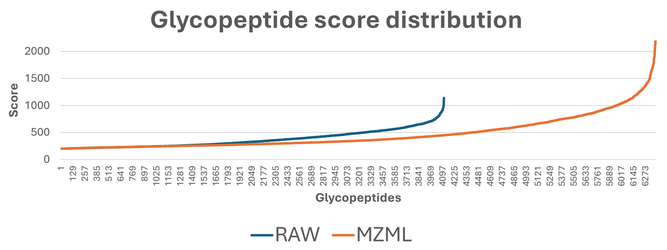
Protein Metrics

All searches were performed using **BYOS** suite by Protein Metrics

4) GLYCOPEPTIDE ANALYSIS

-How can we get the most of the acquired runs and **increase the quality of the search results** and overall output?

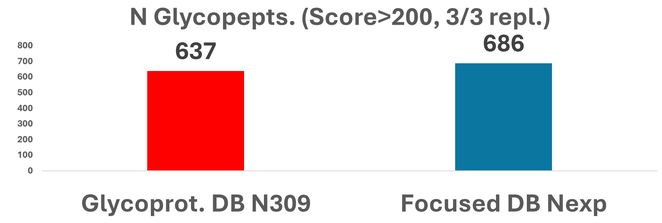
-First, we have used Msconvert to **merge scans** corresponding to the same glycopeptides and generate a **mzml file** to boost the quality of the results:



Using an mzml file with merged scans increases the number and quality of the results

-In addition, we have applied a **dual search** approach to optimize the search times:

Search#1: **broad protein** database (all glycoproteins in Uniprot) / **Restricted Glycan** dabatase (309 N glycans + 78 O glycans)
Search #2: **focused Protein** databse (proteins providing hits in the previous search) + **expanded glycan** database (>1000 glycans considered)



There is a gain of around 50 glycopeptides (7% increase) when considering the focused database plus the expnanded N glycans

CONCLUSIONS

-Using **alternative proteases** might be **helpful** for glycoprotein characterization. GluC doesn't seem to be a good option, but we must consider the use of AspN for such purposes. However, AspN is quite expensive, so its use might be quite limited.

-We can perform **glycopeptide purification** starting from **200ug or even 100ug**, which is a more suitable amount than 1mg. However, there is still room for improvement specificity wise. Increasing wash steps might be useful in this regard.

-**Glycopolygon** enhances the signal corresponding to glycopeptides in the samples analysed and **improves overall results**.

-Our workflow for **data analysis** consisting of **scan merging** and **a two step search** provides **best results** when characterizing complex glyco samples.

FUTURE STEPS

-Give *SOLA-SAX* and *NEB glycoprotein kits* **another try**.

-**Optimize GluC digestion** or add **another protease** that works well to the portfolio.

-With the aim to analyze specifically proteins that are located at the extracelular part of the cell and directly involved in cell-cell interaction, **enrich** the sample in **plasma membrane proteins** prior to digestion and purification of glycopeptides,

References

[1] Baerenfaenger M, Post MA, Zijlstra F, van Gool AJ, Lefeber DJ, Wessels HJCT. Maximizing Glycoproteomics Results through an Integrated Parallel Accumulation Serial Fragmentation Workflow. ***Anal Chem***. 2024 Jun 4;96(22):8956-8964. doi: 10.1021/acs.analchem.3c05874. Epub 2024 May 22. PMID: 38776126; PMCID: PMC11154686.

[2] Mukherjee S, Jankevics A, Busch F, Lubeck M, Zou Y, Kruppa G, Heck AJR, Scheltema RA, Reiding KR. Oxonium Ion-Guided Optimization of Ion Mobility-Assisted Glycoproteomics on the timsTOF Pro. ***Mol Cell Proteomics***. 2023 Feb;22(2):100486. doi: 10.1016/j.mcpro.2022.100486. Epub 2022 Dec 19. PMID: 36549589; PMCID: PMC9853368.