Glycoproteomics-based signatures of cancer cell lines

Pirro, M.

Summary

Cancer-associated alterations in protein glycosylation are commonly known as tumor-associated carbohydrate antigens (TACA), because they are lower abundant or absent on normal human cells. TACA recognition by endogenous lectins expressed by immune cells can have diverse effects on tumor progression and metastasis. In order to characterize the specific nature of the glycan present on glycoproteins of cancer cells, mass spectrometry (MS)-based methods have become increasingly important. **Chapter 1** describes aberrant glycosylation in cancer, the role of lectins, and glycomics and (glyco)proteomics approaches to study (changes in) protein glycosylation.

Due to its ability of binding different TACAs presenting a terminal GalNAc residue, the Macrophage Galactose-type Lectin (MGL) has been used as tool for glycoprofiling various adenocarcinomas. The higher expression of MGL ligands on tumor tissues is correlated with better prognosis in breast cancer patients, whereas it is associated to a more aggressive tumor phenotype and worse prognosis in patients with (adeno-)squamous cervical cancer and colorectal cancer (CRC). MGL is expressed on the surface of immature dendritic cells (DCs) and macrophages. The first MGL binding protein identified in humans was the tyrosine phosphatase CD45. MGL-CD45 interaction results in a decreased proliferation and increased cell death in effector T cells as well as an acute T cell leukemia cell model (Jurkat T cells). However, TACA expression also drastically influences the immune response against tumors through binding to MGL expressed on DCs. However, whether this interaction may facilitate immune evasion or stimulate an anti-tumor response by the immune system depends on cancer type and ligand binding. Hence, the comprehensive characterization of glycoproteins binding to MGL can aid to unravel the intracellular mechanisms activated in cancer by lectin recognition. In this thesis, well-characterized cell lines (Jurkat and CRC cell lines (HT29, HCT116 and LS174T)) were used to gain a broader understanding of glycoproteins binding to MGL. For this purpose, a variety of (glyco)proteomic approaches were used. Moreover, the thesis presents a method to discriminate different populations of glycopeptides in complex quantitative proteomics datasets.

In **Chapter 2**, the optimization of a lectin pull-down assay in combination with mass spectrometric analyses is presented, which allowed the identification of previously

unexplored ligands expressed by Jurkat cells, as well as new insights into site-specific, and density of, *O*-linked glycosylation of known ligands, such as CD43 and CD45.

The characterization of MGL binding proteins from CRC cell lines is presented in **Chapter 3**. The results showed a diverse degree of binding and ligand specificity depending on the CRC cell line. In fact, while HCT116 and HT29 cells bind MGL at high levels, although through a diverse set of ligands, LS174T appeared to be a low MGL binder. The identification of several glycopeptides carrying a LacdiNAc epitope suggested that, in addition to the Tn antigen, *N*-linked glycosylation also plays a role in the binding to MGL.

The involvement of *N*-glycans in binding to MGL was further studied in **Chapter 4** with the use of *N*-glycans release, prior to lectin enrichment. Overall, the results showed a previously underestimated role of *N*-glycosylation in MGL-binding. For example, one of the main MGL-ligands in HT29 and HCT116 cells, Hepatocyte growth factor receptor (c-MET), lost binding to MGL after removal of its *N*-glycans. Moreover, for this chapter we performed quantitative mass spectrometry-based proteomics using the CRC cell lines, in an attempt to explain the differential binding of MGL. Although the expression of MGL binding glycans/proteins could not be assigned to a single mechanism in CRC cell lines, results showed the possible involvement of GALNT3 in the higher MGL binding to HT29, while excluding the overall abundance of MGL-binding proteins as driving factor behind the differences in MGL binding. Moreover, a role for mucins could be excluded because they were more abundant in the low MGL-binding cell line (LS174T).

Chapter 5 describes an optimized workflow for the isolation and purification of secreted proteins in serum-deprived conditioned medium from CRC cell lines. Intriguingly, using the resulting CRC cell line secretomes for MGL binding investigation, we observed that the low MGL binder cell line LS174T secretes glycoproteins carrying MGL-glycan epitopes. This evidence highlights the importance of studying different glycoprotein sources for these type of studies.

In **Chapter 6**, we present the application of a method to unequivocally determine *O*-GalNAcylation from confounding *O*-GlcNAcylation within a complex mixture of (glyco)peptides, especially for peptides containing a single HexNAc. This method could become useful to glyco-scientists for confident identification of such isobaric structures without investing in time-consuming enrichment workflows, tedious method optimization

and demanding expert supervision of manual spectra annotation. Applying this workflow to our CRC quantitative proteomic study, we confidently discriminated between two groups of glycosylated peptides with a single HexNAc. Notably, one of the peptides with a single GalNAc was from anterior gradient protein 2 (AGR2), a hitherto relatively understudied glycoprotein. Using a combination of experiments, we could subsequently determine its *O*-glycosylation in a site-specific manner and show different *O*-glycosylation on intracellular *versus* secreted AGR2.

Finally, **Chapter 7** provides a discussion of the methodological challenges encountered during the study and the complementary approaches that were used or are available for future experiments. Moreover, it discusses the obtained results in the context of literature regarding the suspected mechanisms involved in MGL ligand expression in cancer and relevance of the results for potential clinical applications.