Journal of Chromatography A, xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Mining the acidic serum proteome utilizing off-gel isoelectric focusing and label free quantitative liquid chromatography mass spectrometry

Josh Smith[a,b], Gavin Davey[b], Karol Polom[c,d], Franco Roviello[e], Jonathan Bones[a,f,*]

[a] Characterisation and Comparability Laboratory, The National Institute for Bioprocessing Research and Training, Foster Avenue, Mount Merrion, Co. Dublin, A94 X099, Ireland
[b] School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin 2, D02 R590, Ireland
[c] Department of General Surgery and Surgical Oncology, University of Siena, Siena, Italy
[d] Department of Surgical Oncology, Medical University of Gdansk, Gdansk, Poland
[e] Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy
[f] School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, D04 V1W8, Ireland

ARTICLE INFO

Article history:
Received 22 March 2018
Received in revised form 17 June 2018
Accepted 18 June 2018
Available online xxx

Keywords:
Glycosylation
Sialic acid
Proteomics
Mass spectrometry
Liquid chromatography
Isoelectric focusing

ABSTRACT

Serum remains an attractive source for the discovery of disease related biomarkers due to its intimate interaction with the majority of tissues within the body. The dynamic range of proteins present within serum has long complicated the ability to detect low level tissue leakage proteins that offer more promise as potential biomarkers due to their higher degree of specificity. Depletion strategies, using affinity based sorbents, to remove the most abundant serum proteins are routinely used for matrix simplification during discovery strategies focused on the serum proteome or glycoproteome. Glycoproteins bearing glycans with terminal sialic acid residues have been shown to be differentially expressed in a variety of cancers and are of interest due to the varied involvement of sialic acid in biological functions related to cancer development and metastasis. Herein, we describe the use of 14P serum depletion with subsequent off-gel isoelectric focusing using narrow pH range acidic strips, to facilitate simplification and enrichment of the acidic proteome and glycoproteome prior to label free liquid chromatography mass spectrometry (LC–MS) based proteomic analysis. The developed platform enabled the detection of proteins present within serum at sub-nanomolar concentrations while data analysis using a variety of bioinformatics resources suggested that many of the detected proteins were tissue leakage proteins or proteins associated with specific cellular compartments, rather than traditional secreted proteins. Comparison with theoretical isoelectric point (pI) values also revealed that some detected proteins had higher pI values than the separation range applied during the IEF step, suggesting that their presence may be due to glycans with high degrees of sialylation being attached to the protein backbone, as well as the presence of other post translational modifications. Such findings agree with offline glycomic profiling of the IEF fractions wherein increasing degrees of sialylation were detected across the pH strip from the basic towards the acidic end. Application of the platform using pooled serum from patients bearing gastric cancer, both before and after surgical intervention for tumour removal, revealed a number of differentially expressed proteins associated with cell recognition and cell signalling. The developed platform enabled excellent sensitivity and offers strong potential for application within biomarker discovery studies focused on the acidic proteome.

1. Introduction

Protein glycosylation contributes to a variety of biological functions including involvement in protein folding and protein structure, extracellular matrix and membrane organization, antigen recognition, uptake and processing, cell to matrix interactions, and intercellular signalling and adhesion [1,2]. Aberrant glycosylation is associated with many pathological conditions including cancer, and, as understanding of protein glycosylation expands, greater interest in glycoproteins as disease biomarkers has developed [3–13]. Glycoproteins containing glycans with sialylated and fucosylated features are of increasing interest as alterations in such glycans are well known to be cancer associated, and have been reported in a variety of cancers including breast, liver, pancreatic,

* Corresponding author at: Characterisation and Comparability Laboratory, The National Institute for Bioprocessing Research and Training, Foster Avenue, Mount Merrion, Co. Dublin, Ireland, A94 X099.
E-mail address: jonathan.bones@nibrt.ie (J. Bones).

https://doi.org/10.1016/j.chroma.2018.06.044
0021-9673/

prostate, lung and gastric [8,14,15]. Glycoproteins with sialylated glycans are of particular interest due to their active role in many biological functions, such as those described above, and as such are increasingly becoming targets in cancer biomarker studies [16]. In many cancers increased levels of sialylation have been shown to be associated with disease progression, poor prognosis and tumour burden [17–19]. Because of such findings, glycoproteins containing sialylated glycans are targets of increasing interest in cancer biomarker studies.

Human serum is commonly utilized in biomarker studies because it contains proteins associated with a variety of activities taking place within the body, including tissue leakage proteins which are of great interest as potential disease biomarkers [20]. However, tissue leakage proteins are of low abundance in serum, and the high dynamic range of human serum necessitates the use of serum depletion, protein enrichment and fractionation strategies to identify these low level proteins [20–22]. A variety of protein affinity depletion strategies exist to remove high abundant serum proteins, ranging from affinity ligands for single protein depletion to columns containing ligands to facilitate removal of the twenty most abundant serum proteins [20,23]. However, even with the depletion of these abundant proteins, the most prominent proteins remaining are classical serum proteins and as such, strategies to further mine the serum glycome and glycoproteome are useful when trying to detect low level glycoproteins. For glycoprotein and glycopeptide studies, a variety of enrichment strategies have been applied including the use of hydrazide chemistries [24–27], boronic acid chemistries [28–30], multi-lectin affinity chromatography (M-LAC) [31,32] and hydrophilic interaction liquid chromatography (HILIC) [33–35]. Lectins are often used for the targeted enrichment of glycoproteins containing specific glycan features, because the wide variety of available lectins provides the ability to target specific types of glycosylation [36–41], while titanium dioxide enrichment has proven successful in the enrichment of sialylated glycoproteins [42–44].

Off-gel isoelectric focusing (off-gel IEF) is an alternative platform that has the potential to be utilized for protein enrichment and fractionation. Capable of fractionation on the protein level, off-gel IEF has the same separation capabilities as standard IEF, but is performed in the liquid phase, thus reducing sample handling and simplifying further downstream analysis [45,46]. The availability of different pH gradient IEF strips enables targeted protein analysis based on a protein's isoelectric point (pI) while the commercial availability of varied strips lengths allows for fractionation to be adapted to one's desired application [45,47,48]. As such, utilizing IEF strips with an acidic pH range on the off-gel IEF platform has the potential to enrich acidic proteins including glycoproteins containing glycans with high degrees of sialylation, as previous studies have illustrated that the presence of sialylation increases a proteins negative charge and thus results in a more acidic protein pI [49].

In this study we paired protein affinity serum depletion with off-gel isoelectric focusing to reduce serum matrix complexity and enrich acidic serum proteins and glycoproteins before performing downstream comparative quantitative proteomic analysis of the gastric cancer (GC) serum acidic proteome and glycoproteome before and after tumour resection. 14P serum depletion was utilized to deplete the most abundant human serum proteins while off-gel IEF was used to enrich the acidic proteome, which contains proteins carrying sialylated glycan features. Proteomic analysis revealed that we were able to identify proteins present in the serum at sub-nanomolar (nmol L^{-1}) concentrations including proteins and glycoproteins originating from various cellular components and contributing to a variety of biological processes including various protein transport and binding processes, cell recognition, secretion, signalling and organization, protein processing and phagocytosis. Overall this platform illustrates the ability to detect and enrich low level acidic serum proteins and glycoproteins and can be utilized as the basis for more in-depth glycoprotein analysis including the study of glycoproteins containing sialylated N-glycans, which are a PTM of interest in cancer studies, as increased levels of sialylation have been shown to be associated with disease progression, poor prognosis and tumour burden [17–19].

2. Materials and methods

2.1. Chemicals and reagents

All reagents and solvents used were ACS reagent grade or better. Acetic acid (HAc, glacial), urea, bromophenol blue (sodium salt, electrophoresis grade), trichloroacetic acid (TCA), iodoacetamide (IAA), ammonium bicarbonate (ABC), formic acid (FA, 98–100%), 2-aminobenzamide (2-AB), sodium cyanoborohydride, dimethyl sulfoxide (DMSO), ammonium acetate, ethanol (for HPLC, ≥99.8%), acetonitrile (ACN – for HPLC, gradient grade, ≥99.9%), methanol (for HPLC, ≥99.9%) and trizma hydrochloride were all obtained from Sigma Aldrich (Arklow, Co. Wicklow, Ireland). 20× modified Dulbecco's phosphate buffered saline (PBS), dithiothreitol (DTT – ultrapure, molecular biology grade, USB), acetone (HPLC grade), water (Optima® LC–MS grade), ACN (Optima® LC–MS grade), water containing 0.1% FA (Optima® LC–MS grade), ACN containing 0.1% FA (Optima® LC–MS grade) and tris base were obtained from Fisher Scientific (Dublin 15, Ireland). Peptide N-Glycosidase F (PNGase F) was obtained from New England Biolabs (Ipswich, MA, USA). Purified water was obtained from an Arium® pro Ultrapure Water System (Sartorius, Göttingen, Germany). 50% glycerol stock solution was obtained from Agilent (Santa Clara, CA, USA). Pharmalyte™ 3–10 for IEF was obtained from GE Healthcare (Uppsala, Sweden). The Hi3 *E.Coli* internal standard and [Glu1]-fibrinopeptide B (Glu-Fib) were obtained from Waters (Milford, MA, USA).

2.2. Serum samples

Cancer patient serum samples were obtained with informed patient consent from individual male and female patients with various stages of GC at the time of tumour resection (pre-op) and six weeks after surgery (post-op) by the Department of Surgical Oncology, University of Siena, Italy. Samples were pooled according to when they were taken from the patient, so that two pools were obtained: 1 pre-op and 1 post-op. These pooled samples were then used for in-depth quantitative comparative proteomic analysis. Further information on the clinical samples and the preparation of the pooled samples can be found in the supplementary information section titled "Clinical samples".

2.3. 14P serum depletion

14P serum depletion was performed on a Waters Alliance 2695 HPLC separations module coupled to a Waters W2489 UV–vis detector and Waters Fraction Collector III (Milford, MA, USA). Instrument settings were as follows: autosampler temperature: 5 °C, column temperature: ambient, detection at 280 nm. The system was controlled using Empower 3 chromatography workstation. A 4.6 mm i.d. × 100 mm Human 14 MARS column (Agilent Technologies, Santa Clara, CA, USA) was used for the depletion of the 14 most abundant human serum proteins. Injections of 200 microliters (μL) were prepared by diluting 40 μL of serum to 200 μL with 1 × PBS. Depletion was performed using a 35 min 14 P depletion method where solvent A was 1 × PBS and solvent B was 0.5 M HAc. Samples were loaded onto the column using 100% A at 0.13 mL min^{-1} for 18 min and then from 0.13 mL min^{-1} to 1.0 mL min^{-1}

Please cite this article in press as: J. Smith, et al., Mining the acidic serum proteome utilizing off-gel isoelectric focusing and label free quantitative liquid chromatography mass spectrometry, J. Chromatogr. A (2018), https://doi.org/10.1016/j.chroma.2018.06.044