Protein paucimannosylation is an enriched *N*-glycosylation signature of human cancers

Sayantani Chatterjee[1], Ling Y. Lee[1,2*#], Rebeca Kawahara[1,3*#], Jodie L. Abrahams[1,4*], Barbara Adamczyk[5], Merrina Anugraham[1,6*], Christopher Ashwood[1,7*], Zeynep Sumer-Bayraktar[1,8*], Matthew T. Briggs[9], Jenny H.L. Chik[1,10*], Arun Everest-Dass[1,4*], Sarah Förster[11], Hannes Hinneburg[1], Katia R.M. Leite[12], Ian Loke[1,13*], Uwe Möginger[14], Edward S.X. Moh[1,23], Miyako Nakano[1,15*], Saulo Recuero[12], Manveen K. Sethi[1,16*], Miguel Srougi[12], Kathrin Stavenhagen[1,17*], Vignesh Venkatakrishnan[18], Katherine Wongtrakul-Kish[1,19*], Simone Diestel[20], Peter Hoffmann[9], Niclas G. Karlsson[5], Daniel Kolarich[4], Mark P. Molloy[1,21*], Michael H. Muders[11], Martin K. Oehler[22], Nicolle H. Packer[1,4,23], Giuseppe Palmisano[3], Morten Thaysen-Andersen[1^]

[1]Department of Molecular Sciences and Biomolecular Discovery and Design Research Centre (BDDRC), Macquarie University, Sydney, Australia
[2]ISGlobal, Barcelona Centre for International Health Research (CRESIB), Hospital Clínic -Universitat de Barcelona, Barcelona, Spain
[3]Department of Parasitology, Institute of Biomedical Sciences, University of Sau Paulo, Brazil.
[4]Institute for Glycomics, Griffith University, Gold Coast, Australia
[5]Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[6]Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan
[7]Department of Biochemistry, Medical College of Wisconsin, Milwaukee, USA
[8]School of Life and Environmental Sciences, Charles Perkins Centre (CPC), The University of Sydney, Sydney, Australia
[9]Future Industries Institute, Mawson Lakes Campus, University of South Australia, Adelaide, Australia
[10]International Collaboration On Repair Discoveries (ICORD), Vancouver Coastal Health Research Institute and Department of Pathology and Laboratory Medicine, The University of British Columbia, Vancouver, Canada
[11]Rudolf-Becker-Laboratory, Institute of Pathology, University Hospital Bonn, Bonn, Germany
[12]Laboratório de Investigação Médica da Disciplina de Urologia (LIM55), Faculdade de Medicina da FMUSP, Universidade de Sao Paulo, Sao Paulo, Brazil
[13]Department of Biological Sciences, National University of Singapore, Singapore
[14]Institute for Biochemistry and Molecular Biology, University of Southern Denmark
[15]Graduate School of Advanced Sciences of Matter, Hiroshima University, Hiroshima, Japan
[16]Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University School of Medicine, Boston University, Boston, MA, USA
[17]Beth Israel Deaconess Medical Center, Department of Surgery and Harvard Medical School Center for Glycoscience, Harvard Medical School, Boston, MA, USA
[18]Department of Rheumatology & Inflammation Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
*[19]Bioprocessing Technology Institute (BTI), A*STAR, Singapore, Singapore,*
[20]Institute of Nutrition and Food Sciences, University of Bonn, Bonn, Germany

Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.

[21]*Sydney Medical School, Bowel Cancer and Biomarker Laboratory, Kolling Institute, The University of Sydney, Sydney, Australia*
[22]*Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, Australia*
[23]*ARC Centre for Nanoscale Biophotonics, Macquarie University, Sydney, Australia*
[*] Current affiliation
[#] Equal contribution

^Correspondence: Dr. Morten Thaysen-Andersen, Department of Molecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia. E-mail: morten.andersen@mq.edu.au. Phone/Fax: +61 2 9850 7487 / +61 2 9850 6192

Keywords: Cancer / glycomics / glycan / protein paucimannosylation / paucimannosidic glycan

Running title: Paucimannosylation is elevated in human cancers

Total number of words: 8,720 (Abstract->Discussion 5,139)

Abbreviations: ALL, acute lymphocytic leukaemia; AML, acute monocytic leukaemia; APL, acute promyelocytic leukaemia; ATCC, American type culture collection; BCC, basal cell carcinoma; BPH, benign prostatic hyperplasia; CBA: Cell Bank Australia; CLS: cell line services; CLL, chronic lymphocyte leukaemia; CRC, colorectal cancer; EGFR, epidermal growth factor receptor; EIC, extracted ion chromatogram; FF, fresh frozen; FFPE, formalin-fixed paraffin-embedded; Fuc, α-L-fucose; GlcNAc, *N*-acetyl-β-D-glucosamine; HCC, hepatocellular carcinoma; IAA, iodoacetamide; Man, α/β-D-mannose; MEM, minimum essential medium; MF, microsomal fraction; PCa, prostate cancer; PGC, porous graphitised carbon; PMG, paucimannosidic glycan; PNGase F, peptide-*N*-glycosidase F; RPMI, Roswell Park Memorial Institute medium; S, secretome; SCC, squamous cell carcinoma; SPE, solid-phase extraction; TNM, tumours/nodes/metastases; WCL, whole cell lysate.

Abstract

While aberrant protein glycosylation is a recognised characteristic of human cancers, advances in glycoanalytics continue to discover new associations between glycoproteins and tumourigenesis. This glycomics-centric study investigates a possible link between protein paucimannosylation, an under-studied class of human *N*-glycosylation [$Man_{1-3}GlcNAc_2Fuc_{0-1}$], and human cancers. The paucimannosidic glycans (PMGs) of 34 cancer cell lines and 133 tissue samples spanning 11 cancer types and matching non-cancerous specimens were profiled from 467 published and unpublished PGC-LC-MS/MS *N*-glycome datasets collected over a decade within our laboratories. PMGs, particularly $Man_{2-3}GlcNAc_2Fuc_1$, were prominent features of 29 cancer cell lines, but the PMG level varied dramatically across and within the investigated cancer types (1.0%-50.2%). Analyses of paired (tumour/non-tumour) and stage-stratified tissues demonstrated that PMGs are significantly

enriched in tumour tissues from several cancer types including liver cancer (p=0.0033) and colorectal cancer (p=0.0017) and is elevated as a result of prostate cancer and chronic lymphocytic leukaemia progression (p<0.05). Surface expression of paucimannosidic epitopes was demonstrated on human glioblastoma cells using immunofluorescence while biosynthetic involvement of *N*-acetyl-β-hexosaminidase was indicated by quantitative proteomics. This intriguing association between protein paucimannosylation and human cancers warrants further exploration to detail the biosynthesis, cellular location(s), protein carriers and functions of paucimannosylation in tumourigenesis and metastasis.

Statement of Significance of Study

This is the first study to systematically investigate and document an association between protein paucimannosylation, a less studied human *N*-glycosylation type, and human cancers. The study provides evidence of an over-representation of paucimannosylation in diverse cancer types, and further points to a stage-specific expression, catalytic involvement of *N*-acetyl-β-hexosaminidase and cancer cell surface expression of paucimannosidic proteins. Our findings expand our knowledge of the glyco-phenotypes associated with human cancer, and, at the same time, open many enticing questions concerning the biogenesis, protein carrier(s), subcellular localisation(s) and function(s) of paucimannosylation in the heterogeneous tumour micro-environment. Thus, this work contributes fundamental knowledge required to improve the diagnosis, treatment, and, eventually, prevention of human cancers.

1. Introduction

Protein glycosylation, the addition of complex carbohydrates (glycans) to polypeptides, plays important roles in diverse intra- and intercellular processes [1]. Unsurprisingly, the glycosylation process of proteins is impacted by the onset and progression of cancer. Aberrant glycosylation in human cancers was reported already several decades ago by the observation of a skewed ABO blood type glycosylation in patients suffering from stomach carcinoma [2] and by an antibody-based detection of aberrant carbohydrates in breast and colorectal carcinoma [3].

Many glycoepitopes serving as tumour antigens have since been structurally elucidated using modern analytical methods for glycan structure determination, for example, by fluorescence HPLC detection of 2-aminobenzoic acid labelled colorectal cancer (CRC) glycans [4], MALDI-TOF-MS detection of permethylated prostate cancer (PCa) glycans [5], and porous graphitised carbon (PGC)-LC-MS/MS detection of underivatised breast cancer glycans [6]. The aberrant protein glycosylation features associated with human cancers have been recently summarised [7]. While it remains difficult to unravel if aberrant protein glycosylation is causative or a consequence of cancer, building evidence demonstrates that altered glycosylation can dramatically impact cancer-related molecular and cellular processes, for example, by modulating the recognition of glycoproteins by cancer-associated glycan-binding proteins [8].

Asparagine (*N*)-linked glycoproteins comprise relatively few monosaccharide building blocks including, but not limited to, α/β-D-mannose (Man), α-L-fucose (Fuc), *N*-acetyl-β-D-glucosamine (GlcNAc) and all share a common trimannosylchitobiose core. In spite of such structural constraints, *N*-glycoproteins often display considerable heterogeneity due to structural variations in the glycan antenna region [9]. Three prominent classes, the oligomannosidic, hybrid and complex type *N*-glycoproteins, dominate the human glycobiological literature [10] and are often studied in cancer research [11-16].

A less studied type of human *N*-glycosylation, known as protein paucimannosylation, is receiving growing attention in human glycobiology [17-19]. Our recent summary of the literature supports that humans, similar to most other eukaryotic species, produce proteins modified by paucimannosidic glycans (PMGs) comprising simple monosaccharide compositions i.e. $Man(M)_{1-3}GlcNAc_2$ with a variable presence of core fucosylation i.e. $Fuc(F)_{0-1}$ albeit in a more tissue-specific and physiology-dependent manner than the lower organisms e.g. invertebrates and plants that constitutively express paucimannosidic proteins [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548]. Our detailed literature survey also supports that protein paucimannosylation plays important functions in the human innate immune system, in infectious processes and cellular development [20], accurately mirroring the biological environments where most human PMG-modified proteins to date have been observed [17, 18, 20-22].

Interestingly, several cancer-focused glycomics and glycoproteomics studies published in recent years have indicated a presence of paucimannosylation suggesting that this unusual type of *N*-glycosylation may be a signature of human cancers [4, 13, 23, 24-26]. While it should be emphasised that most cancer-focused glycomics and glycoproteomics studies have not reported on protein paucimannosylation, this may in some cases be explained by a low consideration to these non-conventional glycosylation features or, alternatively, that the employed methodology did not allow for the detection of truncated low-mass glycans. Hence, it remains to be systematically investigated if protein paucimannosylation is a glyco-signature of human cancers.

Herein, we have interrogated an extensive collection of previously published and unpublished PGC-LC-MS/MS glycomics data generated within our laboratories over a decade from tumourigenic and non-tumourigenic cell lines and tissues spanning a range of human cancer types with the aim to systematically test for an association between protein paucimannosylation and human cancers.

2. Materials & Methods

2.1. Materials

2.1.1. Chemicals and reagents

Ultra-high-quality water was from a Milli-Q system (Merck/Millipore). Dulbecco's Modified Eagle's medium (DMEM), MCDB 105: Medium 199 (1:1, v/v), Minimum Essential Medium (MEM alpha), Roswell Park Memorial Institute medium 1640 (RPMI), and fetal bovine serum (FBS) were from Sigma Aldrich or Thermo Fisher Scientific (Sydney, Australia). *Flavobacterium meningosepticum* and *Elizabethkingia miricola* peptide-*N*-glycosidase F (PNGase F) expressed by recombination in *Escherichia coli* were from Roche Diagnostics (Mannheim, Germany) and Promega (Sydney, Australia), respectively. Other chemicals, reagents and/or proteins were from Sigma Aldrich or Thermo Scientific unless specified.

2.1.2. Origin of human cell lines and patient tissues

Human cell lines were obtained from the American Type Culture Collection (ATCC) (Manassas, VA), Cell Bank Australia (CBA) (Sydney, Australia), Cell Line Services (CLS) (Eppelheim, Germany), Garvan Institute of Medical Research (Sydney, Australia) and from multiple other sources as indicated in **Table 1** (see also **Supporting Table S1** for further details of the investigated cell lines). These cell lines were collected and studied over many years by past and present members and collaborators of our glycoanalytical laboratories. The cell lines were generally stored at -140°C in 90% (v/v) FBS and 10% (v/v) dimethyl sulfoxide (DMSO) until use. All cultured cells were confirmed to be mycoplasma-free prior to the *N*-glycome profiling.

Tissue samples were obtained from multiple cancer patient cohorts or from control individuals suffering from a range of other non-cancer conditions at various hospitals and clinics as described in **Table 2** (see also **Supporting Table S2** for further details of the investigated tissue samples). Many of these tissue samples have been described in recent publications [11, 13-16, 24, 27-30]. In brief,

tumourigenic and non-tumourigenic tissues were collected via surgery by trained clinicians after written and informed patient consent were obtained, see **Table 2** and **Extended Experimental Methods** for details of ethics.

2.2. Methods

Details concerning the cell culture conditions, tissue sample handling, and protein extraction, isolation and denaturation procedures have been exhaustively described in the **Extended Experimental Methods.**

2.2.1. N-glycan release and preparation

The *N*-glycans were released from the protein extracts from the obtained samples and prepared for glycomics as previously described [31]. In brief, 10-20 µg protein extract was immobilised in separate spots (technical replicates) on a primed 0.45 µm polyvinylidene fluoride membrane (Merck-Millipore). The dried spots were stained with Direct Blue, excised, transferred to separate wells in a flat bottom polypropylene 96-well plate (Corning Life Sciences), blocked with 1% (w/v) polyvinylpyrrolidone in 50% (v/v) aqueous methanol and washed with water. De-*N*-glycosylation was performed using 1-2 U PNGase F per 20 µg protein in 10 µl water/well, 16 h, 37°C. The released *N*-glycans were transferred into fresh tubes and hydroxylated by the addition of 100 mM aqueous ammonium acetate, pH 5, 1 h, 20°C. The glycans were reduced using 1 M sodium borohydride in 50 mM aqueous potassium hydroxide, 3 h, 50°C. The reaction was quenched using glacial acetic acid. Dual desalting of the reduced *N*-glycans was performed using first strong cation exchange resin (*N*-glycans collected in the flowthrough) and then porous graphitised carbon (PGC) (where the *N*-glycans were retained) packed as micro-columns on top of C_{18} discs in P10 solid-phase extraction (SPE) formats. The *N*-glycans were eluted from the PGC SPE columns using 0.05% trifluoroacetic acid (TFA): 40% acetonitrile (ACN): 59.95% water (v/v/v), dried and reconstituted in 10-20 µl water, centrifuged at 14,000g for 10 min at 4°C and the supernatant finally transferred into high recovery

glass vials (Waters) for LC-MS/MS analysis. Bovine fetuin was included as a sample handling and LC-MS/MS performance control.

2.2.2. PGC-LC-MS/MS N-glycome profiling

The entire set of *N*-glycome profiling data forming the foundation of this study was uniformly collected using our PGC-LC-MS/MS analytical platform [31, 32]. However, as explained below, minor variations in the acquisition style were used for the individual experiments carried out over a 10-year period. The *N*-glycans were separated on a heated (27°C-40°C) HyperCarb KAPPA PGC LC column (particle/pore size 3-5 µm/200-250 Å, column length 100-300 mm, inner diameter 0.075-0.32 mm, Thermo Hypersil Runcorn, UK) over an 85 min gradient of 0-45% (v/v) ACN (solvent B) in 10 mM aqueous ammonium bicarbonate (solvent A) at a constant flow rate of 0.8-2 µl/min delivered by an Agilent 1260 Infinity HPLC (Sydney, Australia) or a Dionex Ultimate-3000 HPLC (Germany/USA) system. The glycans were ionised using ESI and detected in negative ion polarity mode using similar settings across three different mass spectrometry platforms including a i) 3D ion trap mass spectrometer (LC/MSD Trap XCT Plus Series 1100, Agilent Technologies, Australia) operating with a full scan acquisition range of *m/z* 200-2,200. Smart fragmentation (start/end amplitude 30-200%) at 1.0 V and an isolation window of 4 *m/z* with a maximum accumulation time of 200 ms were applied. Smart ion charge control was enabled with a smart parameter setting target of *m/z* 900. A total of 80,000 ions were targeted for all scan events. ESI was performed using a capillary voltage of +3.2 kV, a nitrogen drying gas flow of 6 l/min at 325°C, and a nitrogen-based nebuliser pressure of 12 psi, ii) LTQ Velos Pro ion trap mass spectrometer (Thermo Scientific, Australia) with a full scan acquisition range of *m/z* 500-2,000, a resolution of *m/z* 0.25 full width at half maximum and a source voltage of +3.2 kV. The automatic gain control for the MS1 scans was set to 5 x 10^4 with a maximum accumulation time of 50 ms. For the MS/MS events, the resolution was set to *m/z* 0.35 full width at half maximum, the automatic gain control was 2 x 10^4 and the maximum accumulation time was 300 ms, and iii) AmaZon speed ion trap mass spectrometer (Bruker, Germany) equipped with a

CaptiveSpray source was used with an MS1 full scan acquisition range of *m/z* 380-1,800. The smart parameter setting was set to *m/z* 900 with an ion charge control target of 40,000 and a maximum accumulation time of 300 ms. For MS/MS, the fragmentation cut-off was set to 27% with 100% amplitude using the "Enhanced SmartFrag" option from 30-120% in 32 ms. The ion charge control target was set to 150,000.

Data-dependent acquisition was enabled for all MS/MS scans; the three most abundant precursors in each MS1 full scan spectrum were selected using resonance activation (ion trap) collision-induced dissociation (CID) at a normalised collision energy of 33%. All MS and CID-MS/MS data were acquired in profile mode, and dynamic exclusion was inactivated. The mass accuracies of the precursor and product ions were typically better than 0.2 Da. The LC-MS/MS instruments were tuned and calibrated, and their performance benchmarked using bovine fetuin *N*-glycans prior to data acquisition.

The raw data files of all 467 LC-MS/MS datasets included in this study were browsed, (re-)interrogated and annotated using the ESI-Compass Data Analysis 4.0 software v1.1 (Bruker Daltonics) or Xcalibur v2.2 (Thermo Scientific) using assisting software i.e. GlycoMod (http://www.expasy.ch/tools/glycomod) and GlycoWorkBench v2.1 as well as manual *de novo* glycan sequencing as previously described [12]. The relative abundances of the individual *N*-glycans were determined from relative area-under-the-curve measurements based on extracted ion chromatograms (EICs) performed exclusively for the monoisotopic precursor ion using RawMeat v2.1 (Vast Scientific, www.vastscientific.com), QuantAnalysis v2.1 (Bruker) and Skyline v4.1 [33]. The nomenclature and depiction of glycans follow the latest symbol nomenclature for graphical representations of glycans[34].

All 467 PGC-LC-MS/MS raw data files are available via the MassIVE Consortium (identifier: MSV000083727). The files can be freely downloaded from: ftp://MSV000083727@massive.ucsd.edu.

See also **Supporting Data 1** and **Supporting Data 2** for tabulated *N*-glycome profile data of all samples.

2.2.3. Statistics

The statistical significance of the individual experiments was assessed using paired and unpaired one- or two-tailed Student's *t*-tests where $p < 0.05$ was chosen as the minimum acceptable level of confidence to support a rejection of various proposed null hypotheses (e.g. *the level of paucimannosylation does not differ between tumour and non-tumour tissue*). For each test, the statistical confidence has been indicated by * ($p < 0.05$; ** ($p < 0.01$) and *** *(*$p < 0.001$). The number of biological samples (*n*) is given for the individual experiments involving tissue samples. The *N*-glycome profiles of all cell lines were generated from at least three technical (sample handling and/or LC-MS/MS) replicates. Data points are generally plotted as the mean and error bars represent their standard deviation (SD) if not specified otherwise.

2.2.4. Immunofluorescence

To assess the cancer cell surface expression of paucimannosidic glyco-epitopes, cultured non-permeabilised human glioblastoma cells (U87MG and U138MG) were fixed on 0.002% (w/v) poly-L-lysine-coated microscopy glass slides with 4% (v/v) paraformaldehyde for 20 min. To prevent non-specific antibody recognition, the fixed cells were treated with a blocking solution containing 5% (v/v) horse serum in PBS for 20 min, incubated with the PMG-recognising mouse IgM "Mannitou" (undiluted supernatant from hybridoma cells, described in [22]) for 30 min at room temperature, and then incubated with a goat anti-mouse IgM Cy2-conjugated IgG antibody for 30 min in the dark (1:500 dilution, Dianova, Hamburg, Germany) [22]. The immunofluorescence of the cells was determined using a Zeiss Axioplan2 Apotome system equipped with a Plan Apochromat 63x/1.40 oil immersion objective (Zeiss, Oberkochen, Germany) as recently described [26].

3. Results

3.1. Quantitative PMG profiling using PGC-LC-MS/MS

In this study, a total of 467 published and unpublished PGC-LC-MS/MS datasets obtained from an array of tumourigenic and non-tumourigenic human cell lines and tissue samples were (re-) interrogated for the presence of paucimannosidic glycans (PMGs). All eight combinatorically possible PMGs were confidently identified from this wealth of *N*-glycome data using orthogonal evidence including monoisotopic mass, relative PGC-LC retention time, and negative ion resonance activation (ion trap) CID-MS/MS, a common strategy for glycan fine structure elucidation [12, 35]. Only six of the eight PMGs, i.e. M1, M2F, M2a, M2Fa, M3 and M3F were consistently identified across most samples, **Figure 1A-F**.

Further support for these six PMGs was obtained by the fact that the CID-MS/MS data matched previously published spectra of human PMG standards, i.e. M1F, M2a, M3 and M3F (spectral comparisons not shown) [18]. The relative PGC-LC retention times of the reported PMGs also agreed with published data [36]. The α1,3-terminating bi-mannosylated PMG isomers (i.e. M2b and M2Fb) were absent or negligible in most datasets whereas the corresponding α1,6-terminating M2a and M2Fa isomers were identified in most samples. The relative level of paucimannosylation and the distribution of the individual PMG species within the *N*-glycome profiles were determined using EIC-based quantitation as is common practice in glycomics [32] (see **Supporting Data 1** and **Supporting Data 2** for tabulated *N*-glycome profile data of all samples).

3.2. Cancer-wide distribution of protein paucimannosylation

N-glycomics data from a total of 11 types of human cancers, some with several subtypes, were investigated for PMG signatures including brain (glioblastoma and neuroblastoma), blood (APL, AML, and CLL), skin (melanoma and non-melanoma), breast, lung, liver, gastric, colorectal, prostate,

ovarian and bladder cancers. As visually presented in **Figure 2A,** PMGs were detected in most cancerous cell lines and tissues (blue boxes) albeit the levels of paucimannosylation were found to vary dramatically across and even within the investigated cancer types. The non-cancerous cells and tissues (green boxes) were in general relatively PMG-poor (statistically compared below), but several exceptions to this trend were observed. Although not the focus of this study, the more elongated oligomannosidic *N*-glycans were also found to be significant features of most studied cancers in accordance with the available literature [37].

Further interrogation of the pan-cancer *N*-glycome data indicated that the α1,6-fucosylated PMGs i.e. M2Fa and M3F and to a lesser extent M3 were highly abundant PMG species across the investigated cancers, **Figure 2B.** The afucosylated and highly truncated PMGs i.e. M1, M2a and M1F were rarely observed above trace levels.

3.2.1. PMG expression in cancer cell lines

Focusing firstly on the cancer cell line glycoprofiling data, we compiled and (re-)analysed data from a total of 34 human cell lines spanning 10 types of human cancers, **Supporting Table S1**. The *N*-glycome datasets for 23 of these human cell lines have been published previously as indicated in **Table 1**. By mapping the *N*-glycome of different protein extracts (MF, S, WCL), a total of 55 different samples were explored from the 34 cell lines, **Table 1.** The total level of paucimannosylation within the *N*-glycome varied extensively for the 29 PMG-expressing cell lines (1.0%-50.2%), **Figure 3** and **Supporting Table S3-S4.** In total, 15 cell lines, with an over-representation of cells derived from CRC as well as from blood and brain cancers were classified as high PMG expressing cells (>10%). The whole cell lysate protein fractions generally displayed higher PMG levels than the microsomal and secretome protein extracts.

PMG expression was devoid or trace (<1%) in five of the investigated cell lines; one of these was a non-cancer breast cell line (HMEC). Several non-cancerous cell lines i.e. the ovarian HOSE 6.3 and HOSE 17.1 cell lines and the bladder UROtsa cell line unexpectedly displayed a low-to-medium PMG-

expression (1.5%, 2.6%, and 2.3-3.0%, respectively), which either indicates that some non-cancer cells are also capable of expressing PMG at significant levels or, alternatively, may be explained by the cellular perturbation caused by the human papillomavirus type 16 and the simian vacuolating virus 40 large T-antigen used to generate these continuously replicating cell lines [38]. Isolated primary adipocytes from healthy adipose tissues displayed <1% paucimannosylation (data not shown) providing an example of a non-cancerous and non-infected human cell type with PMG-poor characteristics.

The PMG-rich glioblastoma cell line i.e. U87MG (MF: 7.02% and WCL: 14.34%) was used to assess the cellular location of the expressed paucimannosidic proteins. Significant cancer cell surface presentation of paucimannosidic epitopes was observed using a PMG-recognising antibody (Mannitou) on non-permeabilised cells, **Figure 3, insert**. Similar cancer cell surface expression of PMGs was detected for a related human glioblastoma cell line (U138MG) (data not shown). Immunofluorescence experiments conducted with these and other human cancer cell lines using membrane permeabilisation prior to Mannitou staining have previously indicated that paucimannosidic glycoepitopes are also located intracellularly [22, 26].

3.2.2 Paucimannosylation of cancerous tissues

N-glycome profile data was also obtained from 133 tumour and non-tumour tissue samples obtained from various patient cohorts spanning seven cancer types, **Supporting Table S2.** In total, 63 of these 133 tissues samples (47%) have previously been used for various publications, see **Table 2**). For all 133 tissue samples, the protein extracts were obtained from both fresh frozen and FFPE tissues, many of which were paired samples obtained from cancerous and matching non-cancerous tissues from the same donor including non-melanoma i.e. BCC, n = 34 and SCC, n = 15; gastric cancer, n = 3; liver cancer, n = 3 and CRC, n = 11, **Table 2**. Importantly, all the paired samples, which are particularly valuable due to the absence of inter-person variation, displayed higher PMG levels in the

tumour tissues relative to the matching non-tumour tissues, **Figure 4A-E**. With the exception of gastric cancer that did not reach statistical significance despite an elevated tumour expression of PMGs (non-tumour: 1.36% ± 0.56%, tumour: 3.35% ± 0.12%, n = 3, p = 0.1232), the other four cancer types displayed significantly higher tumour PMG expression i.e. BCC, p = 0.0145, SCC, p = 0.0062, liver cancer, p = 0.0033, and CRC, p = 0.0017 (all paired one-tailed Student's t-tests).

Cordant with observations made for the cell line *N*-glycome data, the PMG level varied dramatically between the cancerous tissue types and even within patients suffering from the same type of cancer. The average PMG level across the studied cancer tissues was 3-11%, a level also identified in advanced ovarian cancer tissues (4.3% ± 0.5%, n = 4) that was, however, not included above due to the lack of matching non-tumour tissues or access to well-annotated ovarian cancer stages precluding testing via comparative glycoprofiling.

The CLL (n = 8) and the PCa (n = 50) tissues collected from patients presenting with different disease stages allowed for an assessment of a possible stage-specific expression of PMGs. Patients with advanced CLL displayed higher PMG expression levels relative to patients with early stage CLL (early stage: 2.12% ± 0.56%, late stage: 3.63% ± 0.12%, p = 0.0205, n = 4), **Figure 4F**. Enhanced PMG expression was also observed in the PCa tumour tissues relative to tissues from a BPH control group (n = 5). The highest PMG expression in PCa was observed in PCa grade 3 tissues (Gleason score 7), **Figure 4G**. Although trends indicating a PCa grade-specific expression of PMG were observed this could not be statistically supported likely due to a gradual rather than an abrupt aberration in the protein and/or glycosylation machinery during PCa progression as we have reported on before [30].

Finally, the protein levels of subunit α (UniProtKB: P06865) and β (P07686) of the human *N*-acetyl-β-hexosaminidase, a homo- (αα or ββ) and hetero- (αβ) dimeric catabolic enzyme that has been associated with paucimannosidic protein formation in humans and across species within the eukaryotic domain [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548], were determined across the PCa grades and in

BPH using quantitative TMT-based proteomics, **Figure 4H** and **Supporting Table S5**. The stage-specific expression pattern of the *N*-acetyl-β-hexosaminidase subunit α and β resembled the grade-specific PMG expression albeit statistical significance for this relationship could not be obtained from the compiled data.

4. Discussion

Enabled by access to an extensive collection of published and unpublished cancer glycomics data collected over the past decade by past and present members and collaborators of our glycoanalytical laboratories, we here provide evidence that protein paucimannosylation is a significant, but variably expressed, glyco-signature across the investigated human cancer types. Of the 467 PGC-LC-MS/MS datasets forming the foundation of this study, 139 datasets (30%) have not previously been used for any publication. The other 328 datasets (70%), which were previously reported on, were thoroughly re-interrogated with a PMG-centric focus for the purpose of this study. Importantly, all *N*-glycome profile data, which was obtained from a variety of tissue preparations and protein extraction methods, were uniformly acquired using a common *N*-glycomics profiling method and systematically compiled and interrogated by the same analyst for analytical rigor and enhanced data confidence.

Paucimannosylation is a common modification of proteins expressed by lower organisms, but our recent reports indicate that higher organisms including humans also express paucimannosidic proteins albeit in a more tissue-restricted and physiology-dependent manner [20] [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548]. Paucimannosylation appears particularly central to the human innate immune system as supported by our previously published data documenting that PMGs are carried by biologically important proteins still displaying bioactivity and structural integrity expressed by resting and activated human neutrophils and macrophages [17, 18, 21]. The findings presented herein documenting that PMGs also serve as signatures of human cancers add to our knowledge of this

understudied *N*-glycan type, but at the same time prompts several questions, some of which are discussed below to encourage and stimulate further activity in this research field.

Are PMGs carried by intact proteins or are they catabolic products of hydrolase-mediated lysosomal degradation in cancer cells?

Glycoproteomics and intact glycoprotein analysis are examples of analytical methods required to adequately address this question. However, several lines of evidence indicate that the PMGs reported herein are, at least in part, carried by fully processed non-degraded proteins, i.e. 1) the presence of paucimannosidic glycoepitopes on the surface of cancer cells, 2) the PMGs were identified from the proteinaceous fractions obtained using classical protein extraction methods (MF, WCL, S) from biological specimens, 3) the PMGs were detected after PNGase F treatment following immobilisation of the proteinaceous fractions on hydrophobic PVDF membranes, and 4) our recent detection of PMGs decorating a non-degraded form of the very large (628 kDa) neuroblast differentiation-associated protein (UniProtKB: Q09666) from human glioblastoma cells [26]. Further, intact glycopeptides carrying PMGs have previously been identified for several cancer cell lines (e.g. THP-1 [21]) and tumour tissues (e.g. PCa, unpublished observations) building further support for a protein-conjugation of the reported PMGs.

Are the PMGs reported here of cancer and/or immune cell origin?

Our study of 34 cancer cell lines including some isolated and cultured primary cancer cells provides solid evidence to document that protein paucimannosylation is a characteristic of human cancer cells. However, it is likely that the PMGs identified from proteins extracted from the tumour tissues originate from multiple cell types including directly from the cancer cells as well as from the cancer-associated stroma, adipocytes, immune cells, and necrotic cells forming the heterogeneous tumour micro-environment [39]. In support, M2Fa and to a lesser extent M3Fa, which in this study were

found to be key PMGs of the tumour tissues, have repeatedly been reported to be abundant glycans carried by proteins expressed by neutrophils and macrophages [17, 18, 21]. Neutrophils and macrophages are important innate immune cells that are known to reside in and modulate the tumour micro-environment as tumour-associated neutrophils (TANs) and macrophages (TAMs) [40]. Experiments employing laser capture microdissection of histologically different regions of the surgically-removed tumour or MS imaging of released *N*-glycans directly on histologically annotated tumour tissue slices are examples of modern glyco-analytical approaches that may provide insight into the cell-specific expression of PMGs (and other glycan types) in tumour tissues [13, 25].

Can the observed PMGs be artefacts of the employed methodology?

High confidence in the reported PMG structures including their monosaccharide composition and linkage/branching patterns was achieved by thorough *de novo* characterisation based on PGC-LC-MS/MS data and via spectral and PGC-LC retention time matching to previously characterised PMGs and *N*-glycan standards [18, 36].

In contrast to other glycomics methods including MALDI-MS profiling of underivatised glycans, *in source* glycan fragmentation can be ruled out in this study since PGC-LC-MS/MS separates glycan analytes prior to ionisation and detection. Furthermore, the employed PGC-LC-MS/MS method detects reduced, but otherwise native (non-derivatised) glycans, which limits the sample handling steps and, in turn, reduces the risk of perturbing individual *N*-glycan analytes and skewing their composition in mixtures. In further support, our previous study of an important paucimannosidic protein, human neutrophil elastase, showed an agreement between the PGC-LC-MS/MS glycan analysis and MS-based intact glycopeptide and glycoprotein level profiling [17].

A recognised shortcoming of PGC-LC-MS/MS is that the highly truncated chitobiose core type *N*-glycans (e.g. M0 and M0F) are inconsistently detected due to poor PGC-LC retention, non-optimal

MS settings to detect low-mass analytes and, likely, a reduced efficiency of PNGase F to release very small *N*-glycan substrates [41].

What is known of the biosynthetic route(s), (sub)cellular locations and roles of PMGs in the tumour microenvironment?

Much is yet to be learned about the cell-specific formation of paucimannosidic proteins, a fact particularly true for human glycobiology [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548]. Building empirical evidence and inference from simpler eukaryotic model systems support that some human cell types, possibly cancer cells included, utilise *N*-acetyl-β-hexosaminidase and the α1,3-specific lysosomal α-mannosidase (O00754) to produce the abundantly observed M3F and M2Fa (α1,6-terminating isomer), respectively [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548]. The prevalent core fucosylation of the observed PMGs indicates that the paucimannosidic proteins are formed via an *N*-acetyl-β-glucosaminyltransferase I (GnT-I)-dependent pathway since β1,2-GlcNAc-capped glycoprotein intermediates are substrates required for the Golgi-resident α1,6-fucosyltransferase responsible for catalysing *N*-glycoprotein core fucosylation [42].

The lysosomal residence of the *N*-acetyl-β-hexosaminidase and lysosomal α-mannosidase, that display optimal catalytic activity at acidic pH [43] and the lysosomal annotation of many of the PMG-containing proteins identified to date including the human lysosomal associated membrane protein-2, β-glucoronidase, CREG1 and gamma-glutamyl hydrolase [18] indicate a formation and temporal or permanent residence of paucimannosidic proteins in lysosomes or lysosome-like compartments.

Importantly, our knowledge of lysosomes has improved considerably over the past years, and their function as biomolecule-degrading organelles has been nuanced by the discovery that lysosomes and their protein content can be mobilised via regulated exocytosis, and, hence, serve important

functions extracellularly or in other cellular compartments [44]. Human neutrophils hosting many PMG-containing proteins i.e. azurocidin, cathepsin G, myeloperoxidase and neutrophil elastase packaged in the azurophilic granule, a mobile organelle with lysosome-like character, is a well-studied example of human cells utilising regulated exocytosis of PMG-rich compartments for a functional purpose [45].

The functions of PMGs in human glycobiology including in cancer processes are also largely unexplored [Tjondro, Loke, Chatterjee, Thaysen-Andersen. *Biol Rev Camb Philos Soc*, accepted 17 July, 2019, DOI: 10.1111/brv.12548]. However, a growing body of literature supports an involvement of PMGs in innate immunity and in infection processes [20]. Future identification of putative PMG-reactive receptors expressed by cancer cells and/or by various tumour-supporting cell types would represent an important advancement to illuminate their potential roles in cell-cell communication possibly driving pro- or anti-tumourigenic processes in the heterogeneous tumour microenvironment.

How can the knowledge of elevated PMG expression in cancers advance cancer research?

The association between protein paucimannosylation and human cancers expand our knowledge of the glyco-phenotypes displayed by cancer cells. A better understanding of the fundamental glycobiology underpinning cancer development and progression is critical to more rapidly improve the diagnosis, treatment and eventually prevention of these prevalent malignancies. Applications including the use of PMG as potential diagnostic cancer markers, perhaps even reporting on the disease progression stage or differentiating between cancer types, may represent future avenues. Further, the cancer cell surface presentation of PMGs, as indicated in this study, may open for exciting new therapeutic avenues including chimeric antigen receptor T-cell strategies.

In conclusion, this is the first study to provide evidence that protein paucimannosylation is a molecular feature of a wide range of human cancers. This new association has sparked many questions of relevance to fundamental glycobiology and cancer research alike that we, and hopefully colleagues within the community, will explore in the years to come.

Supporting Information

Supporting information of this article can be found on Wiley Online Library. This supporting information contains an Extended Experimental Method section and five Supporting Tables (**Supporting Table S1-S4**), which are all provided in a combined SI document (Word). Further, two Supporting Data sheets (**Supporting Data 1-2**) have been provided separately (Excel). All proteomics LC-MS/MS data files supporting the conclusions presented herein have been deposited to the ProteomeXchange Consortium via the PRIDE [46] partner repository with the dataset identifier PXD014271, see also **Supporting Table S5** for an overview of these files. All raw PGC-LC-MS/MS data files forming the foundation of this glycomics-focused study have been deposited to the MassIVE Consortium (identifier: MSV000083727); these data can also be freely downloaded from: ftp://MSV000083727@massive.ucsd.edu.

Acknowledgements

We thank the Australian Proteomics Analysis Facility (APAF) for access to mass spectrometry instrumentation and other essential infrastructure. Dr. Dana Pascovici is thanked for statistical assistance. Hedda von Quistorp is thanked for assistance with immunofluorescence experiments. MTA was supported by a Macquarie University Research Seeding Grant. SC was supported by an international Macquarie University Excellence Scholarship (iMQRES 2017152). GP was supported by FAPESP (2014/06863-3 and 2018/15549-1) and CNPq ("bolsa de produtividade"). RK was supported by FAPESP (2017/03010-8 and 2015/02866-0) and the Cancer Institute NSW (ECF181259). SD was supported by the German Research Foundation DFG (DI 1189/6-1).

Conflict of Interest Statement

The authors have declared no conflict of interest.

Table Legend

Table 1. Overview of the cultured cell lines investigated in this study. #Initials of the primary contributor(s) responsible for acquiring the PGC-LC-MS/MS data. *Indicates non-cancer cell line; S: secretome, WCL: whole cell lysate, MF: microsomal fraction, FF: fresh frozen, FFPE: formalin-fixed and paraffin-embedded, P: paired tissue sample (tumour/non-tumour reference).

Cancer type	Primary contri-butor#	Cell lines (origin/ identifier)	Culture media and properties	Protein fraction	Ref
1.Brain cancers					
A. Glio-blastoma	CA	i. U87MG (ATCC HTB-14)	DMEM, adherent	WCL	In prep
	IL	i. U87MG[a)] (ATCC HTB-14) ii. A172[a)] (ATCC CRL-1620) iii. U118MG (CLS 300362) iv. U138MG[a)] (CLS 300363)	DMEM, adherent	MF MF, S, WCL MF, WCL MF, WCL	In prep
B. Neuro-blastoma	ZSB	i. SK-N-B(2)[b)] (ATCC CRL-2271)	DMEM, adherent	MF	In prep
2. Blood cancers					
A. APL	IL	i. HL-60 (ATCC CCL-240)	RPMI 1640, suspension	Differentia-ted/undif-ferentiated WCL	-

B. AML	LYL	i. THP1 (ATCC TIB-202)	RPMI 1640, suspension	Uninfected /infected MF, WCL	[21]
C. ALL	MN	i. CCRF-CEM (T-cell) (ATCC CCL-119)	RPMI 1640, suspension	MF	[28]
3. Melanoma	JLA	i. MM253 (CBA-1347)	RPMI 1640, adherent	MF	[29]
4. Breast cancers	LYL	i. HMEC* (ATCC CRL-3243) ii. HS578T (ATCC HTB-126) iii. MCF7 (ATCC HTB-22) iv. MDA-MB157 (ATCC HTB-24) v. MDA-MB231 (ATCC HTB-26) vi. MDA-MB468 (ATCC HTB-132) vii. SKBR3 (ATCC HTB-30)	RPMI 1640, adherent	MF, S	[24]
5. Lung cancer	EM	i. A549 (ATCC CCL-185)	RPMI 1640, adherent	MF, WCL	-
6. Liver cancer	SC	i. HepG2 (ATCC HB-8065)	DMEM, adherent	MF, WCL	-
7. CRC	JHLC	i. SW480 (ATCC CCL-228) ii. SW620 (ATCC CCL-227) iii. SW837 (ATCC CCL-235) iv. SW1116	DMEM, adherent	MF	[15]

		(ATCC CCL-233)			
	MKS	v. LIM1215[c)] vi. LIM1899[c)] vii. LIM2405[c)]	RPMI 1640, adherent	MF	[14]
8. Ovarian cancer	MA	i. A2780 (ATCC CRL-2772) ii. IGROV1 (ATCC TCP-1021) iii. OVCAR3 (ATCC HTB-161) iv. SKOV3 (ATCC HTB-77) v. HOSE 6.3*[d)] vi. HOSE 17.1*[d)]	i-iv) RPMI 1640, adherent v-vi) MCDB 105: medium 199, adherent	MF	[11]
9. Bladder cancer	EM	i. J82 (ATCC HTB-1) ii. RT112 (ATCC TCP-1020) iii. UROtsa*[e)]	a) MEM alpha media b-c) RPMI 1640 adherent	MF, WCL	-

[a)] Gift from Prof Torsten Pietsch, Institute of Neuropathology, Univ. Bonn Medical Center, Bonn, Germany
[b)] Gift from Prof Maria Kavallaris, Children's Cancer Institute, Lowy Cancer Research Centre, Univ. New South Wales, Sydney, NSW, Australia
[c)] Ludwig Institute for Cancer Research, Melbourne, VIC, Australia
[d)] Garvan Institute of Medical Research, Sydney, NSW, Australia
[e)] Gift from Prof Philip Erben, Department of Urology, Medical Faculty Mannheim, Heidelberg Univ., Mannheim, Germany

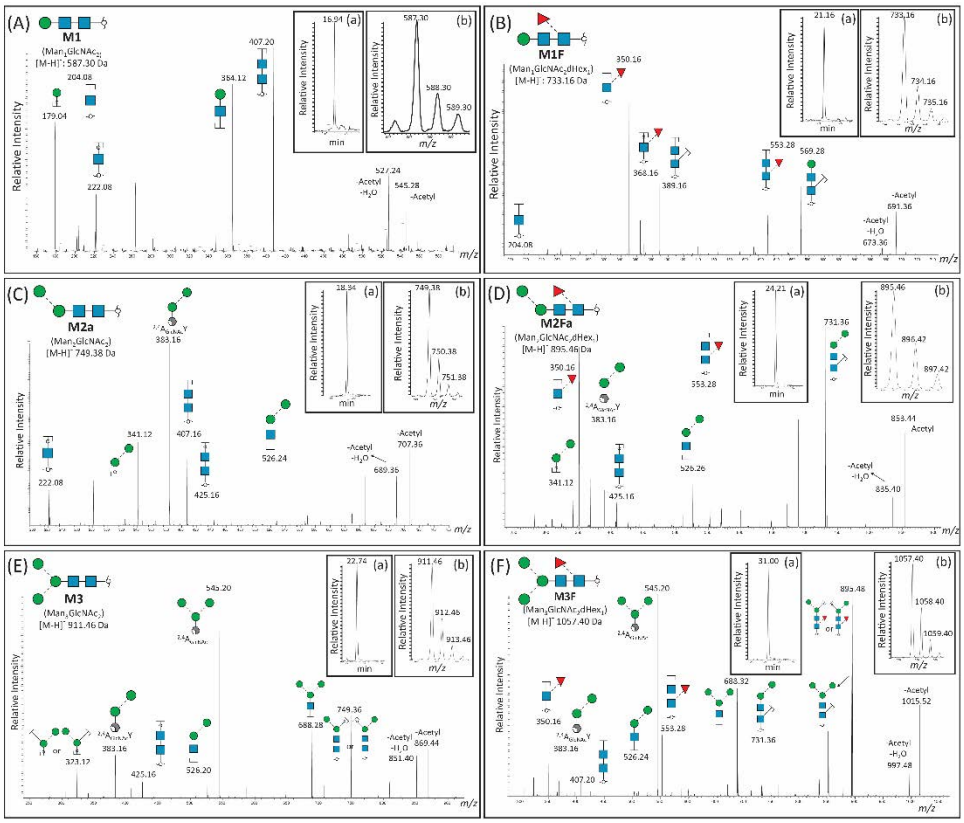
Table 2. Overview of the tissue samples investigated in this study. [#]Initials of the primary contributor responsible for acquiring the PGC-LC-MS/MS data. *Indicates non-cancer cell lines; S: secretome, WCL: whole cell lysate, MF: microsomal fraction, FF: fresh frozen, FFPE: formalin-fixed and paraffin-embedded, P: paired tissue samples (tumour vs non-tumour reference).

Cancer type	Primary contri-butor[#]	Biological replicates (n)	Sample type	Ethics approval	Ref
1. Blood cancer					
A. CLL	KS	i. 8	FF	Project #8935, Univ. Sydney, Sydney, Australia	-
2. Non-melanoma					
A. BCC B. SCC	UM	i. 14 ii. 20 i. 15	FF BCC (P) FFPE BCC (P) FFPE SCC (P)	127-11-18042011 Univ. Leipzig, Leipzig, Germany	[16]
3. Gastric cancer	VV	i. 3	FF (P)	Comitato Etico Regione Toscana (Tuscany, Italy)	-
4. Liver cancer	HH	i. 3	FFPE (P)	Klinieka Bolnica Merkur (Zagreb, Croatia)	[13]
5. CRC	JHLC	i. 6	FF (P)	X08-0614 Sydney South West Area Health Service, Australia	[15]
	MKS	ii. 5	FF (P)	5201100040 Yonsei Univ., Seoul, South Korea	[14]

6. PCa	RKS	i. 50 tumour 5 BPH	FF	nº2695126 Faculdade de Medicina do Estado de São Paulo, Brazil	In prep
7. Ovarian cancer	AD	i. 4	FFPE	140101, Royal Adelaide Hospital Human Ethics Committee	-

Figure Legends

Figure 1. Quantitative PMG profiling using PGC-LC-MS/MS. (**A-F**) PGC-LC-MS/MS-based evidence of six PMGs consistently identified across the cell lines and tissue samples investigated in this study i.e. M1, M1F, M2a, M2Fa, M3 and M3F. Annotated resonance activation (ion trap) CID-MS/MS (-) spectra are shown. Inserts: The elution time, isotopic envelope and charge state of the six precursor ions further document the structural details of the reported PMGs. The α1,3-terminating bi-mannosylated PMGs i.e. M2b and M2Fb were only inconsistently identified at very low levels across the investigated samples and hence left out of this and subsequent figures. The nomenclature and depiction of glycans follow the latest symbol nomenclature for graphical representations of glycans[34]. In short, Blue squares: GlcNAc; green circles: Man; red triangles: Fuc.

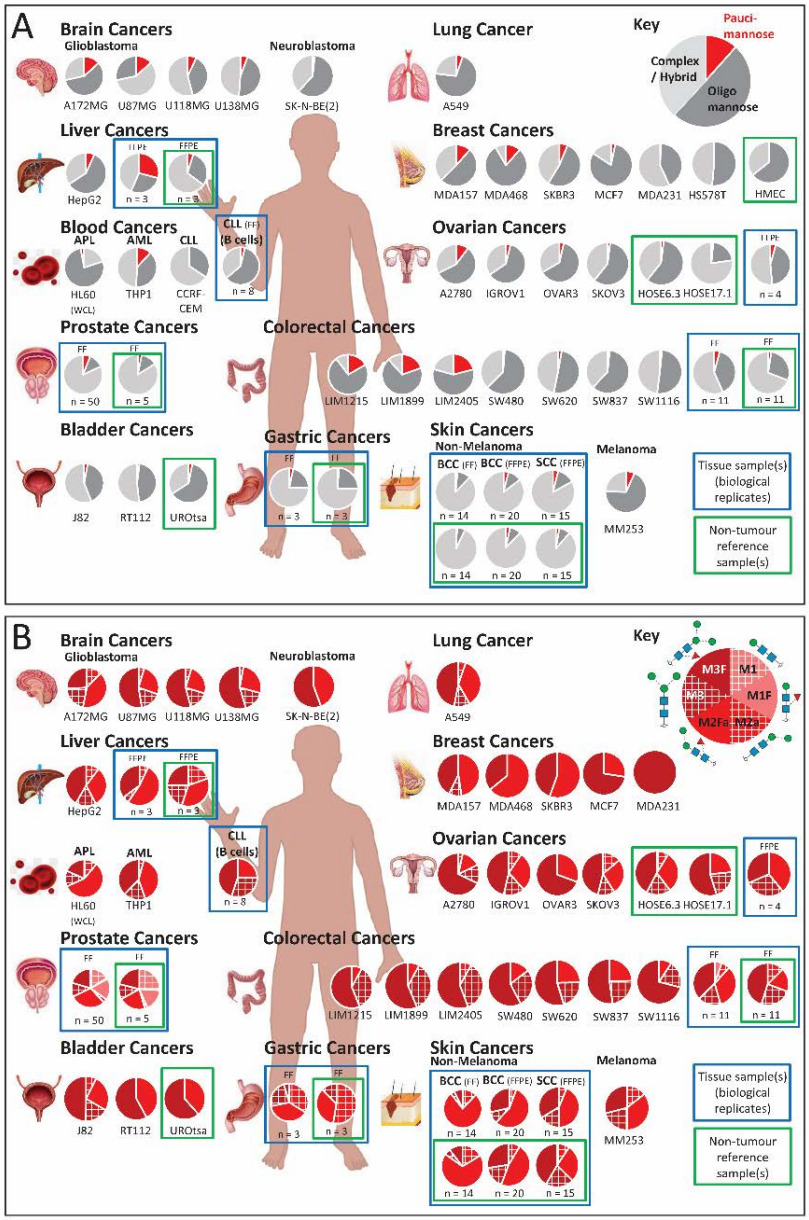


Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.

Figure 2. Pan-cancer map of protein paucimannosylation. (A) Total level of paucimannosylation (red slices) and **(B)** the relative distribution of the individual PMG species (various shades of red, checked/non-checked patterns) within the *N*-glycome of the studied cell lines and tissues (blue boxes). The *N*-glycome data obtained from the microsomal fractions have consistently been shown for all cell lines except the HL-60 cell line where the protein *N*-glycans were only profiled from whole cell lysates protein extracts. For the tissue samples, the *N*-glycome profiles are presented using the fresh frozen (FF) and formalin-fixed paraffin-embedded (FFPE) tissues. Various non-cancerous cell lines and tissues were used as reference samples (green boxes). At least three technical replicates were obtained for each of the cell line *N*-glycome profiles, see Table 1 for experimental details and Supporting Table S3-S4 for tabulated data. The number of biological replicates for the tissue *N*-glycome profiles have been provided, see Table 2 for experiment details and Figure 4 for data. Blue squares: GlcNAc; green circles: Man; red triangles: Fuc.[34]

Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

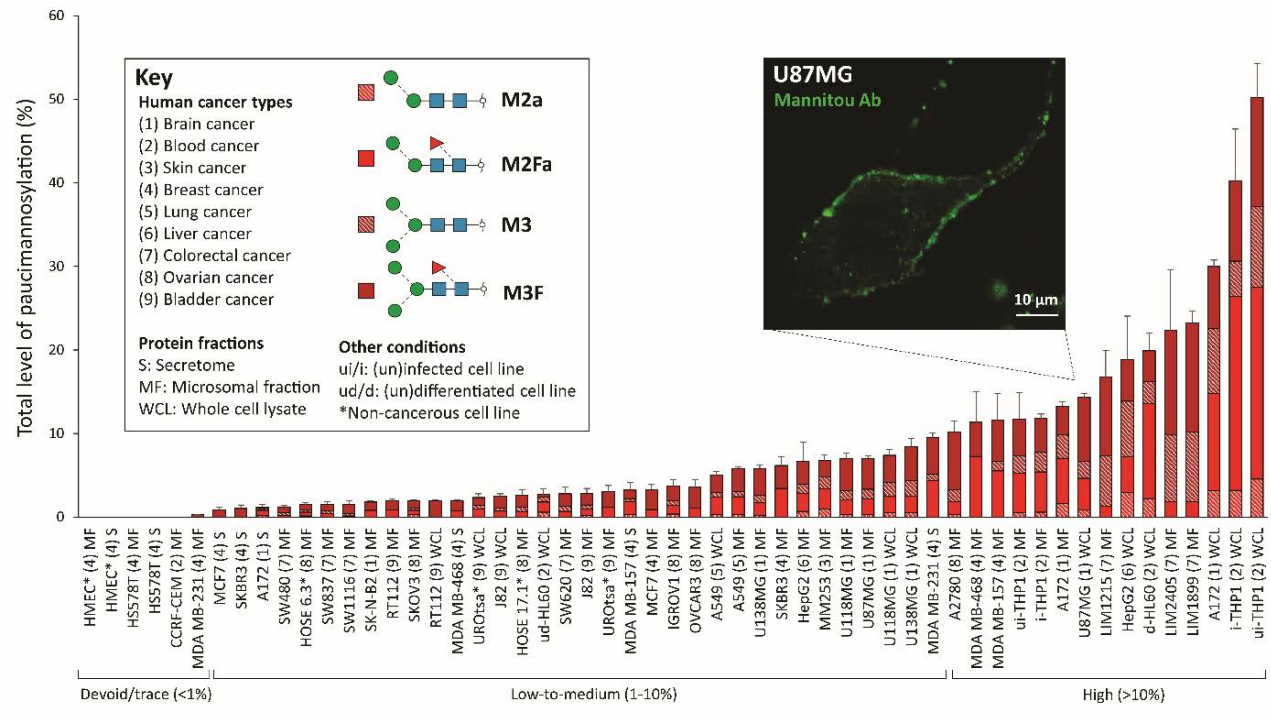


Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

Figure 3. Significant, but non-uniform PMG expression in human cancer cell lines. Total level of paucimannosylation and distribution of four PMG species (M2a, M2Fa, M3 and M3F) observed within the *N*-glycome of three protein fractions (S, MF, and WCL) of all investigated 34 human cell lines comprising nine cancer types. The cell lines, which were glycoprofiled from a minimum of three technical replicates, are grouped according to their PMG expression status i.e. devoid/trace (<1%), low-to-medium (1-10%) and high (>10%) PMG expression levels (see key for details). Insert: Cancer cell surface expression of paucimannosidic epitopes of the PMG-rich U87MG human glioblastoma cell line was indicated by the surface reactivity of a PMG-recognising antibody i.e. Mannitou (green) on non-permeabilised cells, bar = 10 µm. See Table 1 and Supporting Table S1 for experimental details and Supporting Table S3-S4 for tabulated data. Blue squares: GlcNAc; green circles: Man; red triangles: Fuc[34].

Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.



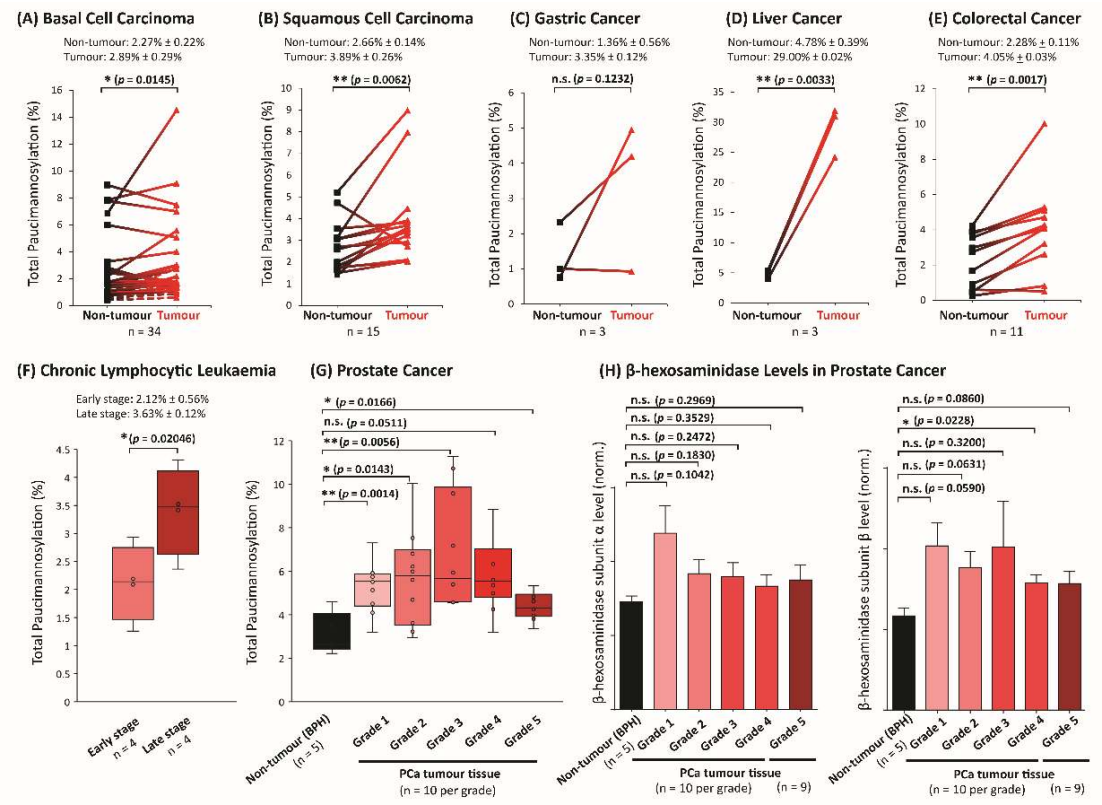
Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.

Figure 4. PMG expression is enriched in tumour tissues and increases with disease progression. Total PMG level within the *N*-glycome of donor-paired tumour and non-tumour tissues obtained from patients suffering from (**A**) basal cell carcinoma, (**B**) squamous cell carcinoma, (**C**) gastric cancer, (**D**) liver cancer, and (**E**) colorectal cancer. Statistical significance was evaluated based on paired one-tailed Student's *t*-tests (n = number of patient samples). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, n.s. not significant ($p \geq 0.05$). The total PMG level was also determined from disease stage-classified tumour tissues (unpaired samples) of patients suffering from (**F**) chronic lymphocytic leukaemia and (**G**) prostate cancer (PCa). Five individuals suffering from benign prostatic hyperplasia (BPH) were included as an inflammatory control group enabling comparisons to the PCa data. Box and whisker plots indicating the mean and spread of the indicated data points are used to plot these data points. Statistical significance was evaluated based on unpaired one-tailed Student's *t*-tests (n = number of patient samples). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, n.s. not significant ($p \geq 0.05$). (**H**) The protein level expression of the *N*-acetyl-β-hexosaminidase subunit α (UniProtKB: P06865) and β (P07686), putatively responsible for the formation of paucimannosidic proteins (see text), were determined for the same BPH and PCa tissue samples using quantitative TMT-based proteomics (BPH, n = 5; PCa grades 1-4, n = 10 per grade, PCa grade 5, n = 9). Statistical analyses were carried out as described above.

Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.



Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.

References

[1] A. Varki, Glycobiology 2017, 27, 3; K. W. Moremen, M. Tiemeyer, A. V. Nairn, Nature reviews Molecular cell biology 2012, 13, 448.
[2] R. P. Ladenson, S. O. Schwartz, A. C. Ivy, The American journal of the medical sciences 1949, 217, 194; S. Callender, M. J. Langman, I. N. Macleod, J. Mosbech, K. R. Nielsen, Gut 1971, 12, 465; P. Meldgaard, E. H. Holmes, E. P. Bennett, H. Clausen, J. Zeuthen, H. Wolf, T. F. Ørntoft, Cancer Research 1994, 54, 2440.
[3] J. Burchell, S. Gendler, J. Taylor-Papadimitriou, A. Girling, A. Lewis, R. Millis, D. Lamport, Cancer Res 1987, 47, 5476; S. F. Kuan, J. C. Byrd, C. Basbaum, Y. S. Kim, The Journal of biological chemistry 1989, 264, 19271; Y. S. Kim, Seminars in cancer biology 1990, 1, 189; T. Feizi, Cancer Surv 1985, 4, 245.
[4] C. I. Balog, K. Stavenhagen, W. L. Fung, C. A. Koeleman, L. A. McDonnell, A. Verhoeven, W. E. Mesker, R. A. Tollenaar, A. M. Deelder, M. Wuhrer, Molecular & cellular proteomics : MCP 2012, 11, 571.
[5] Z. Kyselova, Y. Mechref, M. M. Al Bataineh, L. E. Dobrolecki, R. J. Hickey, J. Vinson, C. J. Sweeney, M. V. Novotny, Journal of proteome research 2007, 6, 1822.
[6] L. Y. Lee, C. H. Lin, S. Fanayan, N. H. Packer, M. Thaysen-Andersen, Front Immunol 2014, 5, 404.
[7] S. R. Stowell, T. Ju, R. D. Cummings, Annual review of pathology 2015, 10, 473; J. Munkley, D. J. Elliott, Oncotarget 2016, 7, 35478; N. Taniguchi, Y. Kizuka, Advances in cancer research 2015, 126, 11; M. N. Christiansen, J. Chik, L. Lee, M. Anugraham, J. L. Abrahams, N. H. Packer, Proteomics 2014, 14, 525.
[8] J. Munkley, D. J. Elliott, Oncotarget 2016, 7, 35478; B. N. Vajaria, P. S. Patel, Glycoconjugate journal 2017, 34, 147.
[9] A. Varki, S. Kornfeld, in *Essentials of Glycobiology*, (Eds: A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, P. H. Seeberger), Cold Spring Harbor (NY) 2017.
[10] C.-Y. Chung, N. I. Majewska, Q. Wang, J. T. Paul, M. J. Betenbaugh, Cell 2017, 171, 258; J. Breitling, M. Aebi, Cold Spring Harbor perspectives in biology 2013, 5, a013359; L. S. Kreisman, B. A. Cobb, Glycobiology 2012, 22, 1019.
[11] M. Anugraham, F. Jacob, S. Nixdorf, A. V. Everest-Dass, V. Heinzelmann-Schwarz, N. H. Packer, Molecular & Cellular Proteomics 2014, 13, 2213.
[12] A. V. Everest-Dass, D. Jin, M. Thaysen-Andersen, H. Nevalainen, D. Kolarich, N. H. Packer, Glycobiology 2012, 22, 1465.
[13] H. Hinneburg, P. Korać, F. Schirmeister, S. Gasparov, P. H. Seeberger, V. Zoldoš, D. Kolarich, Molecular & cellular proteomics 2017, 16, 524.
[14] M. K. Sethi, H. Kim, C. K. Park, M. S. Baker, Y.-K. Paik, N. H. Packer, W. S. Hancock, S. Fanayan, M. Thaysen-Andersen, Glycobiology 2015, 25, 1064.
[15] J. H. Chik, J. Zhou, E. S. Moh, R. Christopherson, S. J. Clarke, M. P. Molloy, N. H. Packer, Journal of proteomics 2014, 108, 146.
[16] U. Möginger, S. Grunewald, R. Hennig, C.-W. Kuo, F. Schirmeister, H. Voth, E. Rapp, K.-H. Khoo, P. H. Seeberger, J. C. Simon, Frontiers in oncology 2018, 8, 70.

Received: 29/04/2019; Revised: 25/07/2019; Accepted: 14/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900010.

[17] I. Loke, O. Ostergaard, N. H. H. Heegaard, N. H. Packer, M. Thaysen-Andersen, Molecular & cellular proteomics : MCP 2017, 16, 1507.
[18] I. Loke, N. H. Packer, M. Thaysen-Andersen, Biomolecules 2015, 5, 1832.
[19] V. Venkatakrishnan, N. H. Packer, M. Thaysen-Andersen, Expert review of respiratory medicine 2013, 7, 553.
[20] I. Loke, D. Kolarich, N. H. Packer, M. Thaysen-Andersen, Molecular aspects of medicine 2016, 51, 31.
[21] N. J. Hare, L. Y. Lee, I. Loke, W. J. Britton, B. M. Saunders, M. Thaysen-Andersen, J Proteome Res 2017, 16, 247.
[22] A. C. Dahmen, M. T. Fergen, C. Laurini, B. Schmitz, I. Loke, M. Thaysen-Andersen, S. Diestel, Glycobiology 2015, 25, 869.
[23] U. Moginger, S. Grunewald, R. Hennig, C. W. Kuo, F. Schirmeister, H. Voth, E. Rapp, K. H. Khoo, P. H. Seeberger, J. C. Simon, D. Kolarich, Front Oncol 2018, 8, 70; H. Chen, Z. Deng, C. Huang, H. Wu, X. Zhao, Y. Li, Tumour Biol 2017, 39, 1010428317716249; S. Hua, M. Saunders, L. M. Dimapasoc, S. H. Jeong, B. J. Kim, S. Kim, M. So, K. S. Lee, J. H. Kim, K. S. Lam, C. B. Lebrilla, H. J. An, J Proteome Res 2014, 13, 961; X. Wang, Z. Deng, C. Huang, T. Zhu, J. Lou, L. Wang, Y. Li, J Proteomics 2018, 172, 1; P. Shah, X. Wang, W. Yang, S. Toghi Eshghi, S. Sun, N. Hoti, L. Chen, S. Yang, J. Pasay, A. Rubin, H. Zhang, Molecular & cellular proteomics : MCP 2015, 14, 2753; S. Holst, A. J. Deuss, G. W. van Pelt, S. J. van Vliet, J. J. Garcia-Vallejo, C. A. Koeleman, A. M. Deelder, W. E. Mesker, R. A. Tollenaar, Y. Rombouts, M. Wuhrer, Molecular & cellular proteomics : MCP 2016, 15, 124.
[24] L. Y. Lee, M. Thaysen-Andersen, M. S. Baker, N. H. Packer, W. S. Hancock, S. Fanayan, J Proteome Res 2014, 13, 4783.
[25] A. V. Everest-Dass, M. T. Briggs, G. Kaur, M. K. Oehler, P. Hoffmann, N. H. Packer, Molecular & cellular proteomics : MCP 2016, 15, 3003.
[26] Y. Becker, S. Forster, G. H. Gielen, I. Loke, M. Thaysen-Andersen, C. Laurini, K. Wehrand, T. Pietsch, S. Diestel, Oncotarget 2019, 10, 4449.
[27] M. K. Sethi, M. Thaysen-Andersen, H. Kim, C. K. Park, M. S. Baker, N. H. Packer, Y.-K. Paik, W. S. Hancock, S. Fanayan, Journal of proteomics 2015, 126, 54.
[28] M. Nakano, R. Saldanha, A. Göbel, M. Kavallaris, N. H. Packer, Molecular & Cellular Proteomics 2011, 10, M111. 009001.
[29] J. Abrahams, N. Packer, M. Campbell, Analyst 2015, 140, 5444.
[30] R. Kawahara, F. Ortega, L. Rosa-Fernandes, V. Guimarães, D. Quina, W. Nahas, V. Schwämmle, M. Srougi, K. R. Leite, M. Thaysen-Andersen, Oncotarget 2018, 9, 33077.
[31] P. H. Jensen, N. G. Karlsson, D. Kolarich, N. H. Packer, Nature protocols 2012, 7, 1299.
[32] H. Hinneburg, S. Chatterjee, F. Schirmeister, T. Nguyen-Khuong, N. H. Packer, E. Rapp, M. J. A. c. Thaysen-Andersen, 2019.
[33] C. Ashwood, C.-H. Lin, M. Thaysen-Andersen, N. H. Packer, Journal of The American Society for Mass Spectrometry 2018, 29, 1194.
[34] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J. J. Prestegard, R. L. Schnaar, H. H. Freeze, J. D. Marth, C. R. Bertozzi, M. E. Etzler, M. Frank, J. F. Vliegenthart, T. Lutteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach, K. F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi, S. Kornfeld, Glycobiology 2015, 25, 1323.
[35] A. V. Everest-Dass, J. L. Abrahams, D. Kolarich, N. H. Packer, M. P. Campbell, J Am Soc Mass Spectrom 2013, 24, 895.
[36] J. L. Abrahams, M. P. Campbell, N. H. J. G. j. Packer, 2018, 35, 15.
[37] M. L. A. de Leoz, L. J. T. Young, H. J. An, S. R. Kronewitter, J. Kim, S. Miyamoto, A. D. Borowsky, H. K. Chew, C. B. Lebrilla, Molecular & cellular proteomics : MCP 2011, 10, M110.002717.
[38] S. W. Tsao, S. C. Mok, E. G. Fey, J. A. Fletcher, T. S. Wan, E. C. Chew, M. G. Muto, R. C. Knapp, R. S. Berkowitz, Experimental cell research 1995, 218, 499.

[39] F. Cavallo, C. De Giovanni, P. Nanni, G. Forni, P.-L. J. C. I. Lollini, Immunotherapy, 2011, 60, 319.
[40] F. Cavallo, R. A. Calogero, G. Forni, Nature reviews. Cancer 2007, 7, 707; M. R. Galdiero, C. Garlanda, S. Jaillon, G. Marone, A. Mantovani, Journal of cellular physiology 2013, 228, 1404.
[41] G. E. Norris, T. J. Stillman, B. F. Anderson, E. N. Baker, Structure 1994, 2, 1049.
[42] T.-H. Tseng, T.-W. Lin, C.-Y. Chen, C.-H. Chen, J.-L. Lin, T.-L. Hsu, C.-H. Wong, Journal of the American Chemical Society 2017, 139, 9431; X. Wang, S. Inoue, J. Gu, E. Miyoshi, K. Noda, W. Li, Y. Mizuno-Horikawa, M. Nakano, M. Asahi, M. Takahashi, Proceedings of the National Academy of Sciences 2005, 102, 15791; Q. Yang, L.-X. Wang, Journal of Biological Chemistry 2016, 291, 11064.
[43] R. DeGasperi, S. Al Daher, P. Daniel, B. Winchester, R. Jeanloz, C. Warren, Journal of Biological Chemistry 1991, 266, 16556; S. Al Daher, R. De Gasperi, P. Daniel, N. Hall, C. Warren, B. Winchester, Biochemical Journal 1991, 277, 743.
[44] S. Ravi, K. A. Peña, C. T. Chu, K. Kiselyov, Cell calcium 2016, 60, 356.
[45] S. Rorvig, O. Ostergaard, N. H. Heegaard, N. Borregaard, J Leukoc Biol 2013, 94, 711; N. Borregaard, Ann N Y Acad Sci 1997, 832, 62; N. Borregaard, O. E. Sorensen, K. Theilgaard-Monch, Trends in immunology 2007, 28, 340.
[46] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, Nucleic acids research 2018, 47, D442.