Clin Proteom (2008) 4:25–36
DOI 10.1007/s12014-008-9006-z

Glycosylation Site Analysis of Human Platelets by Electrostatic Repulsion Hydrophilic Interaction Chromatography

Urs Lewandrowski · Katharina Lohrig · René P. Zahedi · Dirk Wolters · Albert Sickmann

Published online: 12 July 2008
© Humana Press 2008

Abstract
Introduction Glycosylations range among the most common posttranslational modifications with an estimated 50% of all proteins supposed to be glycosylated. These modifications are required for essential cellular processes including cell–cell recognition, protein structure and activity, e.g., of surface receptors, as well as subcellular localization of proteins. Beside the elucidation of the carbohydrate structures, the annotation of glycosylation sites is of primary interest as a basis for subsequent functional characterization. Although mass spectrometry is the method of choice for large-scale analysis of glycosylation sites, it requires initial enrichment of glycopeptides prior mass spectrometric detection in most cases.
Materials and Methods In this paper, we present a novel approach for glycopeptide enrichment by electrostatic repulsion hydrophilic interaction chromatography (ERLIC). Glycopeptides were separated from the bulk of non-modified peptides and gradually eluted from the stationary phase with potential for isoform resolution. Applied to human platelets, 125 glycosylation sites on 66 proteins were identified including major platelet glycoproteins responsible for cellular function.
Conclusion These sites add a major contribution to the now more than 250 glycosylation sites annotated for platelets, which enable the clinically relevant design of quantification assays for platelet glycoproteins.

Keywords Electrostatic repulsion hydrophilic interaction chromatography · Glycosylation · Mass spectrometry · Platelet

Abbreviations

ERLIC	electrostatic repulsion hydrophilic interaction chromatography
PEG	Polyethyleneglycol

Introduction

Glycosylations are among the most common posttranslational modifications to proteins and have been characterized by an increasing number of publications on the glycoproteomic field. Being a very heterogeneous group of modifications, a narrow focus of studies on certain subsets of glycosylation types is required. In this context, the majority of glycosylation site analyses were focused on *N*-glycosylations in the past, demonstrating the strong impact of glycosylations on protein and cellular function. They influence protein structure as well as their subcellular localization, e.g., by controlling correct folding of newly synthesized proteins and transport signals like mannose-6-

Urs Lewandrowski and Katharina Lohrig contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s12014-008-9006-z) contains supplementary material, which is available to authorized users.

U. Lewandrowski · R. P. Zahedi · A. Sickmann (✉)
DFG-Research Center for Experimental Biomedicine,
University Würzburg,
Versbacher Strasse 9,
97078 Würzburg, Germany
e-mail: Albert.Sickmann@virchow.uni-wuerzburg.de

K. Lohrig · D. Wolters
AG Biomolecular Mass Spectrometry,
Analytical Chemistry Department, Ruhr-University Bochum,
Universitätsstrasse 150,
44801 Bochum, Germany

phosphate for lysosomal localization. In addition, protein function can be directly modulated, e.g., in case of glycoprotein VI—the major collagen receptor on platelets, where prevention of *N*-glycosylation leads to reduced binding of collagen [1]. Furthermore, increased levels of terminal beta-*N*-acetylglucosamine on glycoproteins, especially on glycoprotein Ib in cold-storage platelets lead to increased clearance of transfused platelets from the bloodstream by interaction with the macrophage $\alpha_M\beta_2$ integrin receptor [2, 3]. On a systemic level, the impaired synthesis of oligosaccharide side chains of membrane glycoproteins has been shown to be connected to congenital thrombocytopenia [4] as shown for the terminal glycosylation of glycoprotein Ib in platelets [5, 6].

Despite examples of direct clinical relevance for glycosylation related effects, the analysis of glycosylation sites offers a further benefit for proteomic biomarker research. Membrane glycoproteins are the main mediators of extracellular stimuli and thus a primary research target. However, in comparison to non-glycosylated intracellular proteins they comprise only a low proportion of the total cellular protein content and therefore are—among other reasons—underrepresented in many proteomic studies. The same holds true for plasma analysis, where albumin and other abundant non-glycosylated proteins are responsible for up to 95% of the protein content. In contrast, large-scale glycosylation site analysis relies on specific enrichment of glycopeptides before mass spectrometric detection, thus reducing the complexity of the sample mixture by several orders of magnitude. This enables the identification of low abundant glycoproteins for basic research and potential biomarkers with clinical relevance. Stahl-Zheng et al. used this approach for plasma proteins in combination with hybrid-type triple quadrupole instruments and reported potential identification as well as quantification of hundreds of glycoproteins within five orders of magnitude [7]. Therefore, the analysis of glycosylation sites can be considered as beneficial not only for detailed functional protein analysis, but also for the design of assays for clinically relevant biomarkers. As experimental basis, these assays require the identification of accessible glycosylation sites on a wide range of functional proteins. A repository of those sites might therefore enable fast and convenient access to this kind of information without conducting time-consuming basic research for each clinical study anew.

Over the years, many different strategies have been used for glycosylation site elucidation. Nearly all of them employ a three-step procedure based on enrichment of glycopeptides, (partial) removal of the glycan moiety—mostly by enzymatic means—and subsequent mass spectrometric detection and bioinformatic evaluation of datasets. Established approaches mainly differ regarding the primary step for specific isolation of glycopeptides. Lectin-based approaches have been among the first techniques for this purpose, as shown by Kaji et al., using isotope-coded glycosylation-site-specific tagging (IGOT) with Concanavalin A as lectin [8]. Almost simultaneously, Zhang et al. demonstrated the use of hydrazide-based covalent coupling for identification of glycoproteins and their relative quantification in two samples [9]. In addition, approaches based on more general properties of glycopeptides were applied for enrichment by size, hydrophilicity, or charge. Alvarez-Manilla et al. [10] utilized the large size of glycopeptides as selection criterion during size exclusion chromatography, Hagglund et al. [11] employed hydrophilic liquid interaction chromatography to enrich carbohydrate-bearing peptides [12]. Recently, we reported the enrichment of sialic acid bearing peptides by strong cation exchange chromatography exploiting the charge difference of these glycopeptides compared to the bulk of tryptic peptides from platelet membrane digests [13]. It is interesting to note that while mapping glycosylation sites for the platelet proteome project by various methods, a high complementary of glycopeptide enrichment strategies was observed. Whereas subsets of glycosylation sites could be readily detected, e.g., by oxidative hydrazide coupling, the same sites were not accessible via hydrophilic interaction or lectin affinity chromatography and vice versa. The same was previously reported for phosphopeptide analysis as well [14]. To allow for an enhanced survey of certain cell types, tissues or body fluids, several of these methods should therefore be combined to maximize the output of relevant data from highly complex protein mixtures.

However, all generally employed strategies for glycopeptide enrichment show a common limitation: the elution of peptides is achieved as bulk mixture. This can be advantageous with respect to quantitative profiling, but in general, further subfractionation of glycopeptides would be preferable. Already during strong cation enrichment of glycopeptides we observed parallels between glyco- and phosphopeptide behavior. Both types of modified peptides eluted early during the salt gradient or were already found in the flow-through fractions. This parallelism also holds true for enrichment of sialoglycopeptides by TiO_2 as recently demonstrated by Larsen et al. [15]. Therefore, the development of electrostatic repulsion hydrophilic interaction chromatography (ERLIC) for phosphorylation analysis [16] led us to evaluate the potential of glycopeptide enrichment by this novel liquid chromatography technique.

In this study, ERLIC, in combination with plasma membrane preparations based on aqueous two-phase partitioning, allowed for enrichment of carbohydrate-bearing peptides from a large number of platelet glycoproteins being important for platelet function. The method enabled the annotation of 125 glycosylation sites on 66 distinct proteins with a high share of plasma membrane proteins. Moreover, the elution

behavior of the glycopeptides exhibited potential isoform resolution as a result of interaction with the stationary phase. Thereby, this approach adds not only to the growing number of known platelet glycosylation sites supporting biomarker monitoring, but also shows promise for future application of glycopeptide-isoform detection.

Material and Methods

Materials

Chromatographic chemicals used in this study were of analytical or higher grade and were obtained from Sigma, Steinheim, Germany. PNGaseF (*F. meningosepticum*) was from Roche, Mannheim, Germany; sequencing grade trypsin was purchased from Promega, Madison, USA.

Platelet Purification

Human platelets were prepared as described previously [13, 17, 18] from fresh aphaeresis concentrates (leukocyte depleted, $2–4\times10^{11}$ platelets/250 mL, Department of Transfusion Medicine, University Würzburg, Germany). Briefly, concentrates were divided into 50-mL centrifugation tubes and spun twice at 310×*g* for 15 min to remove remaining leukocytes or erythrocytes. The supernatant of the second centrifugation step was spun again at 380×g for 20 min at room temperature. Pellets were washed twice with 10 mM citric acid buffer containing 5 mM KCl, 145 mM NaCl, 14 mM glucose, and 1 mM $MgCl_2$, pH 6.4 to remove residual plasma proteins. Final pellets were directly frozen in liquid nitrogen until further use.

Aqueous Two-Phase Partitioning

Platelet plasma membranes were enriched as previously described by two-phase partitioning [13, 19]. A 200-mL two phase system consisting of 6.3% PEG 3350 and dextran T500 each in 15 mM Tris, pH 7.8 was equilibrated overnight at 6°C. Thereof, a 20-mL system (1:1 *v/v* of the equilibrated PEG and dextran phases) was used for separation of 100-mg (wet weight) platelets. Lysis was facilitated by repetitive ultrasonic bursts and subsequent phase separation was achieved by centrifugation at 500×*g* for 10 min at 6°C. The top PEG phase was extracted with an equal volume of fresh dextran phase. This extraction was repeated once and the final upper PEG phase was diluted 1:1 with water before membranes were pelleted by ultracentrifugation at 100,000×*g* for 1 h at 4°C in a TLA 100.4 rotor (Beckman Coulter, Krefeld, Germany). To remove cytoplasmic and peripheral membrane proteins, the resulting four membrane pellets were twofold extracted in a volume of 3 mL of 100 mM sodium carbonate, pH 11.5 for 0.5 h at 4°C [20]. Pellets were subsequently stored at −80°C until further use.

Proteolytic Digests

A single pellet (or an equivalent amount of whole-cell lysate) was solubilized in 200 μL of 2.5% SDS, 50 M Tris, pH 8.0, briefly heated to 90°C and 1,4-dithiothreitol was added to a concentration of 5 mM before incubation at 57°C for 15 min. After addition of 20 mM iodoacetamide and incubation at 21°C for further 15 min, 1800 μL ethanol (chilled to −30°C) was added and proteins were precipitated at −30°C for 4 h. Proteolytic digest of proteins solubilized in 400 μL 50 mM NH_4HCO_3 was performed with 20 μg trypsin (sequencing grade, Promega, Madison, USA) at 37°C overnight. Samples were acidified by addition of formic acid and dried under vacuum.

Glycopeptide Purification

For isolation of glycopeptides, electrostatic repulsion hydrophilic interaction chromatography as demonstrated by Alpert [16] was used. Therefore, a polyWAX column (4.6 mm inner diameter, 100 mm length, polyWAX LP, 5 μm particle size, 300 Å pore size, Chromatographic Technologies, Basel, Switzerland) was used with a flow rate of 1 mL/min on an inert BioLC HPLC system (Dionex, Idstein, Germany). The binary gradient consisted of 20 mM methylphosphonic acid, 70% acetonitrile for buffer A and 200 mM triethylamine phosphate, 60% acetonitrile for buffer B, respectively. Stock solutions were prepared as 66 mM methylphosphonic acid in water, adjusted to pH 2.0 with sodium hydroxide and 500 mM triethylamine adjusted to pH 2.0 with ortho-phosphoric acid before addition of acetonitrile.

Dried samples were dissolved in 200 μL buffer A and injected to the polyWAX column, which was previously equilibrated for at least 1 h with buffer A. After 10 min at 100% A, a linear gradient was used to 100% B in 45 min for elution of glycopeptides followed by equilibration to 100% A for at least 1 h before the next injection. Elution profiles were monitored by UV absorption at 215 and 295 nm. Fractions were collected in 1-min intervals and acetonitrile was evaporated under vacuum.

Subsequently, peptides were purified using small scale C18 columns (PerfectPure C18, Eppendorf, Hamburg, Germany). Columns were pretreated with acetonitrile, conditioned with water, and samples were repetitively applied. After eight washes with 20 μL of water each, peptides were eluted with 20 μL 60% acetonitrile, 40 mM NH_4HCO_3. Acetonitrile was removed under vacuum and 10 μL 100 mM NH_4HCO_3 were added.

Glycosidase Digest

After C18 microcolumn treatment 0.25 U PNGase F was added to each fraction, followed by incubation at 37°C for 3 h. Samples were acidified with 15 μL of 1% trifluoroacetic acid each and reduced to 15 μL final volume under vacuum.

Mass Spectrometric Detection

Separation of peptide mixtures before mass spectrometric sequencing was performed on a nanoLC–MS/MS system. Briefly, a Famos™, Switchos™, Ultimate™ nanoLC system (Dionex, Idstein, Germany) was set up to trap and desalt peptides from the ERLIC fractions on a custom-made 100 μm ID×2 cm length precolumn (Ace C_{18}, 5 μm particle size, 100 Å pore size; HiChrom Ltd., Berkshire, UK) with 0.1% trifluoroacetic acid (TFA) as loading buffer. Dimensions of custom-made separation columns were 75 μm ID×150 mm length (Ace C_{18}, 3 μm particle size, 100 Å pore size, HiChrom Ltd.). Separations were accomplished at a flow rate of 270 nL/min and gradient slopes of 1% B/min up to 55% buffer B content, followed by 5 min wash at 95% B. Solvent A was 0.1% formic acid in water and solvent B 0.1% formic acid in 84% acetonitrile.

A Qtrap 4000 linear ion trap mass spectrometer (Applied Biosystems, Darmstadt, Germany) was used in the positive ion mode comprising 1) an enhanced multiple-charge scan (380–1,500 amu, three spectra summed at 4,000 amu/s) as survey scan, followed by 2) enhanced resolution scans of three selected precursors (single spectra at 250 amu/s), which were furthermore fragmented by 3) enhanced product ion scans (115–1,500 amu, two spectra summed at 4,000 amu/s). Ion spray voltage was set to 2.3 kV and only ions with charge states 2^+ and 3^+ were chosen for fragmentation. Dynamic exclusion time was set to 22 s after one occurrence of a respective target ion.

Data Evaluation

Mass spectrometric derived datasets were evaluated by searches using Mascot™ (Version 2.1, Matrix Science, London, UK) combined with manual evaluation as well as open mass spectrometry search algorithm (OMSSA) [21] for automated evaluation. In both cases, peak lists were generated from the raw-data format using Analyst 1.4.2 software plug-ins (mascot.dll; Matrix Science/Applied Biosystems). All peaks with intensities below 0.1% of the base peak were omitted, whereas data were centroided in the process. Mass deviance was set to 0.4 Da. A non-redundant human subset of the Swiss-Prot database (version from 11 September 2007, total 566,908 sequences, 34,340 thereof in the human subset, http://www.expasy.ch) was used as a concatenated reversed database for searches with trypsin specified as protease comprising one missed cleavage site. Carbamidomethylation was set as fixed modification and deamidation of asparagine as variable modification. Spectra with a Mascot™ score >34 (significance threshold $p<0.05$, false positive rate for total peptide number ~1.6%) and valid glycosylation consensus sequence were considered for further manual evaluation. OMSSA results were considered to be significant with an e-value below 0.01, which corresponds to a false positive rate <0.1%.

Results

Enrichment of Glycopeptides by ERLIC

Electrostatic repulsion hydrophilic interaction chromatography as presented by Alpert [16] enables the enrichment of phosphopeptides by interaction of the negatively charged phosphate group with the weak anion exchange matrix overlaid with hydrophilic effects caused by the use of 60% to 70% acetonitrile within the buffer system. Glycopeptides have been demonstrated to exhibit a similar behavior as phosphopeptides in other systems such as strong cation exchange [13, 22] and titanium dioxide chromatography [15, 23]. Indeed, we observed distinct enrichment and separation of glycopeptides by the use of ERLIC.

As expected, the bulk of peptides could be observed in the flow-through fractions from 1 min to 12 min (see Fig. 1). They comprise non-glycosylated peptides, whereas high abundant proteins such as glycoprotein IIb/IIIa clearly dominate the analysis of individual fractions. Beginning with fraction 13–14 but more pronounced after fraction 20, less non-glycopeptides were identified (compare Fig. 2). In turn, the number of peptides that exhibits a +1 Da mass shift for the artificial deamidation by PNGaseF deglycosylation increased up to fraction 35–45 where a ratio of roughly 15 for modified versus non-modified peptides could be observed using automated OMSSA evaluation. This is a comparable enrichment effect to that observed by Kaji et al. using a twofold lectin enrichment procedure. In their study, ~16% of the dataset was made up by peptides from highly abundant structural proteins or from unspecific adsorption of peptides to the agarose-based lectin column [8]. Similar shares (~17%) were monitored applying strong cation exchange chromatography to platelet plasma membranes [13] prepared by the same aqueous two-phase partitioning system as used in the current study.

Nevertheless, the main difference between ERLIC and SCX-based enrichment of glycopeptides is the elution of glycopeptides in distinct fractions in the ERLIC mode. Most of the detected glycopeptides elute as broad peaks in

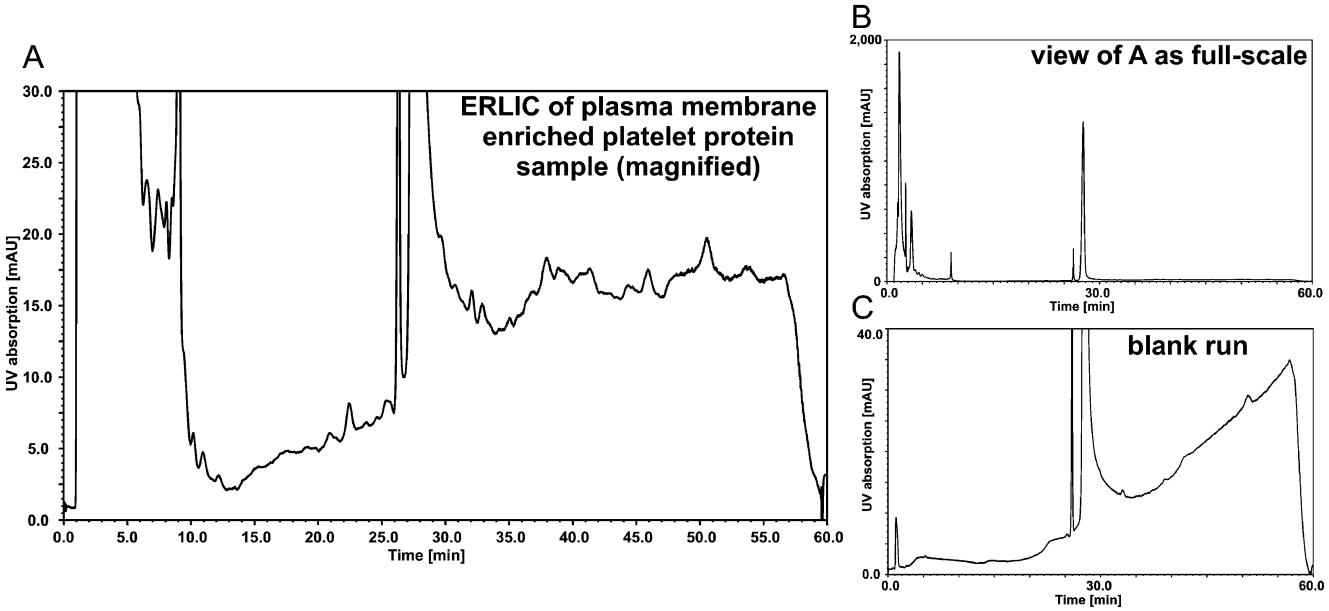

Fig. 1 UV-Chromatogram of ERLIC from a platelet plasma membrane fraction. (**A**, **B**) The majority of peptides reveals only weak or no interaction with the stationary phase and is detected from min 1—12. It is clear that additional peaks in the later fractions after 13 min can be observed in the magnified view (**A**). Comparison with a blank run (**C**) confirms the two major peaks at 26 and 28 min to be artifacts, while mass spectrometry proved the remaining signals to be derived from glycopeptides

time frames over 5–6 min. As could be deduced from the artifact peaks, as well as from minor peaks within the gradient (Fig. 1), this elution window seems rather large. It could either be caused by inferior chromatographic conditions resulting in tailing of peaks; more probably, the presence of glycan isoforms attached to the same peptide moiety contributes to the observed peak broadening. Detailed analysis of ovalbumin, e.g., resulted in the elucidation of 18 glycan isoforms attached to the single glycosylation site N293 by Harvey et al. [24]. Assuming a selective adsorption and desorption of individual glycan isoforms from the polyWAX resin, peak broadening might therefore be a result of oligosaccharide isoform separation. Moreover, the behavior of different glycopeptides (and their potential isoforms) varied to a high degree during gradient elution. Whereas the peptide K.YSL#NSTTWK.V of C-C chemokine receptor type 4 could only be observed in fraction 36, the peptide R.LGGW#NITGPWAK.D of endothelin-converting enzyme 1 was detected in five adjacent fractions (32 to 36) and lastly R.LRPDDSK#NFSIQVR.Q of integrin beta-3 was found throughout the gradient (fractions 21–27, 31, 33, 35–46) (Fig. 3). The extended elution time width of the integrin beta-3 peptide would argue for an extremely heterogeneous glycosylation profile, whereas peptides from endothelin-converting enzyme 1 and C-C chemokine receptor type 4 would support minor heterogeneity of their glycosylation sites. In addition to potential differences caused by the glycan part, the peptide moiety seems to influence the elution behavior as well. In case of integrin beta-1 peptides K.DTCTQECSYF#NITK.V (fractions 38–44) and K.KDTCTQECSYF#NITK.V (fractions 33, 34, 39, 45–47, 49) as well as platelet glycoprotein 4 peptides K.#NCTSYGVLDISK.C (fractions 36,37, 39–42) and K.IISK#NCTSYGVLDISK.C (fractions 40–46, 48) the missed cleavage site resulted in a shift to later retention times of roughly 5 min.

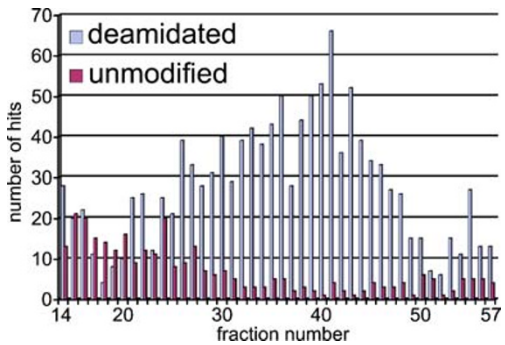

Fig. 2 Distribution of deamidated and unmodified peptides identified during MS-based sequencing of ERLIC fractions from min 14 to min 57

As a conclusion, ERLIC is well suited for the enrichment of glycopeptides, with potential access to the separation of glycopeptide isoforms. This is advantageous when dealing with highly complex glycoprotein samples such as plasma, complete tissue extracts as well as complete organisms, where the expected number of glycopeptides is still higher as in the presented platelet system. Thorough subfractionation would be required in those cases to avoid undersampling during mass spectrometric analysis. ERLIC allows for this fractionation already during the single

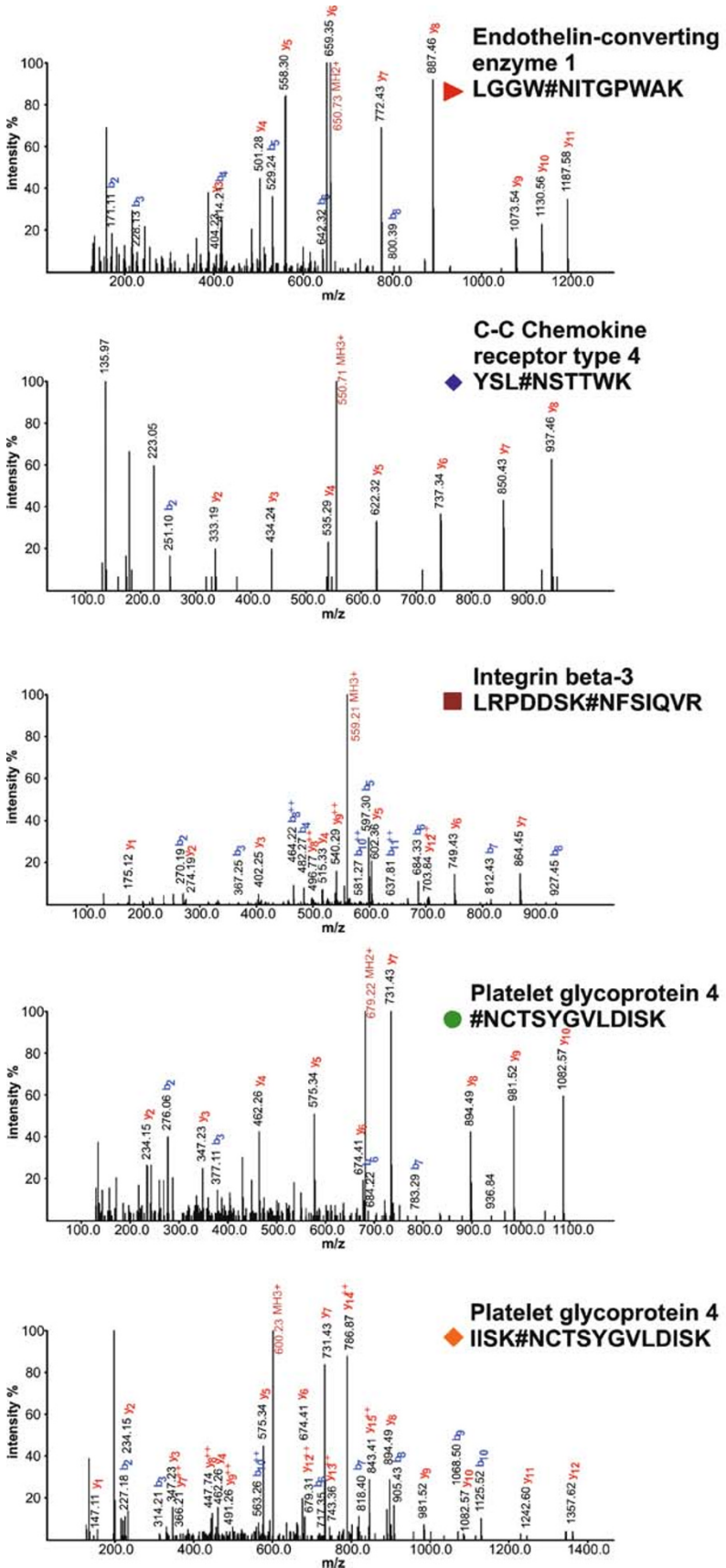

Fig. 3 Examples of glycosylated peptides detected during the ERLIC approach. Elution times are visualized for each peptide to account for possible isoform resolution (right scheme)

enrichment step. However, the exact mechanism of ERLIC retention for glycopeptides remains yet unknown and is the focus of ongoing work in our group. Although the acetonitrile gradient is quite low ranging from 70% to 60%, a hydrophilic retention mechanism is possible (compare Hagglund et al. [11]) in addition to the attractive effects of the polyWAX phase toward negatively charged groups—potentially sialic acid.

Elucidation of Glycosylation Sites

In total, 125 glycosylation sites belonging to 118 distinguishable peptide sequences could be sequenced using the ERLIC approach (compare Table 1; for additional information, see supplemental table). This number is based on Mascot database searches with subsequent manual validation of spectra (for MS/MS spectra compare supplemental section). In addition, 117 of those sites (94%) could also be detected using OMSSA [21] (supplemental table), thereby increasing the confidence in the correctness of these sites. Of 2418 peptides identified by MASCOT under the conditions stated in the methods section, 1577 peptides were annotated as deamidated. Thereof, 616 deamidations (25.5%) were within the NXT and 363 (15.0%) within the NXS consensus sequence. The high number of 598 unspecific deamidations might be explained by an increased retention of this peptide species as a result of the additional carboxy-group of the converted asparagine. In this context, deglycosylation using PNGaseF and $H_2{}^{18}O$ might help to further distinguish between unspecific deamidations, e.g., by natural (+1 Da) and intentionally introduced deamidations (+3 Da). A comparison with two previously published studies on the platelet system [13, 18] results in 63 common sites from the SCX-based study and 20 from the hydrazide/lectin-based approach. Fifty-four sites have not been reported so far for platelets. ERLIC therefore proves to be a complementary approach further increasing the number of known glycosylation sites in human platelets. Novel sites could not only be detected for major glycoproteins such as integrins beta-1 and beta-2, integrin alpha-2 or platelet glycoprotein V but also for less known, e.g., the metal transporter CNNM4, which has so far not been characterized in human platelets. Regarding the function of the identified proteins, three groups can be differentiated: first, 16 proteins exhibit receptor- or receptor-like properties. They comprise integrins as well as C-C Chemokine receptor 4, a prostacyclin receptor or glycoprotein V as component of the glycoproteins Ib-IX–V complex (von Willebrand factor receptor). Second, seven transporter proteins could be identified. Among them are P2X purinoceptor 1 and short transient receptor potential channel 6 as examples for calcium transporters. Third, 19 proteins involved in cell–cell interactions or cellular adhesion are included in the presented dataset. This group again includes integrins responsible for cellular interaction as well as adhesive glycoproteins thrombospondin 1, platelet endothelial cell adhesion molecule or endothelial cell-selective adhesion molecule. Further proteins not included in the groups previously mentioned are, e.g., multimerin 1 as a carrier protein for coagulation factor V/Va or the metalloproteinase ADAM 10.

Using subcellular prefractionation by aqueous two-phase partitioning led to a high share of plasma membrane localized glycoproteins. In this study, 66 proteins could be shown to be *N*-glycosylated. By comparison with the Swiss-Prot database, 60% thereof were annotated as belonging primarily to the plasma membrane, whereas further 17% of the identified proteins were putative membrane proteins with yet unconfirmed localization (Fig. 4). Although this share is lower than in the previous study [13], it still highlights the suitability of ERLIC for the enrichment of glycopeptides from this source. Initial use of whole-platelet lysates in concentrations comparable to the membrane samples only led to a low number of glycopeptide identifications. This is probably because of the lower absolute amount of membrane proteins, which make up the primarily glycosylated species in the analyzed samples. Subfractionation is therefore recommended, although soluble glycoproteins might be lost in the process of aqueous two-phase partitioning. These should be covered by additional approaches specifically aimed, e.g., to secreted proteins [25].

Combination of the so far published large-scale glycosylation site datasets yields a unique collection of glycopeptides for use in cardiovascular studies. Two hundred fifty-seven distinct glycosylation sites cover a selection of the most important glycoproteins responsible for platelet function. They comprise a high number of adhesion receptors and signal transducers mandatory for platelet activation and aggregation. Some, such as glycoprotein VI, were not confirmed as their glycosylation sites are not accessible by proteolytic cleavage with trypsin as conducted in all three studies. We therefore estimate the number of potential *N*-glycosylation sites in platelets to at least 400–500, possibly more. However, using the so far identified sites, the adjacent sequence stretches were compared and visualized using WebLogo [26] (Fig. 5). A slight preference of threonine over serine could be observed at the third position of the consensus sequence NXS/T. Apart from this known consensus sequence, little features could be deduced even from a magnified version of the WebLogo visualization. It might be argued that a slight increase in the frequency for leucine, isoleucine, valine, alanine as well as serine—all rather small and uncharged amino acids—can be observed adjacent to asparagine. These differences are nevertheless not sufficient to deduce

Table 1 Proteins and glycosylation sites identified from human platelet samples using the ERLIC approach

Accession	Protein Name	Peptide	Position in protein sequence
P08195	4F2 cell-surface antigen heavy chain	K.DASSFLAEWQ#NITK.G	264
		K.SLVTQYL#NATGNR.W	323
		R.LLIAGT#NSSDLQQILSLLESNK.D	280
O14672	ADAM 10	R.#NISQVLEK.K	439
		R.I#NTTADEKDPTNPFR.F	278
P01009	Alpha-1-antitrypsin	K.YLG#NATAIFFLPDEGK.L	271
P35613	Basigin	K.ILLTCSL#NDSATEVTGHR.W	160
P20645	Cation-dependent mannose-6-phosphate receptor	R.EAG#NHTSGAGLVQINK.S	83
		R.L#NETHIF#NGSNWIMLIYK.G	107, 113
Q13740	CD166 antigen	K.IIISPEE#NVTLTCTAENQLER.T	480
P27701	CD82 antigen	R.DY#NSSREDSLQDAWDYVQAQVK.C	129
Q8IWA5	Choline transporter-like protein 2	K.TCNPETFPSS#NESR.Q	417
		R.K#NITDLVEGAK.K	200
P10909	Clusterin	R.LA#NLTQGEDQYYLR.V	374
P08174	Complement decay-accelerating factor	K.GSQWSDIEEFC#NR.S	95
P46977	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit STT3A	R.TILVDNNTW#NNTHISR.V	548
Q07108	Early activation antigen CD69	K.EFNNWF#NVTGSDK.C	166
Q9Y6C2	EMILIN-1	R.LGAL#NSSLQLLEDR.L	794
Q96AP7	Endothelial cell-selective adhesion molecule	R.LQGVPHVGA#NVTLSCQSPR.S	169
P42892	Endothelin-converting enzyme 1	K.EYLEQISTLI#NTTDR.C	383
		K.HLLE#NSTASVSEAER.K	166
		R.FF#NFSWR.V	539
		R.LGGW#NITGPWAK.D	210
P54852	Epithelial membrane protein 3	K.ESLNLWYDCTWN#NDTK.T	47
P02679	Fibrinogen gamma chain	K.VDKDLQSLEDILHQVE#NK.T	78
P02751	Fibronectin	K.LDAPTNLQFV#NETDSTVLVR.W	1007
		R.DQCIVDDITYNV#NDTFHK.R	528
Q9BX51	Gamma-glutamyltransferase-like protein 6	R.LHNQLLP#NVTTVER.N	167
Q68CP4	Heparan-alpha-glucosaminide N-acetyltransferase	R.FGEFG#NYSLLVK.N	162
P17301	Integrin alpha-2	K.GEYFV#NVTTR.I	1074
		K.QNNQVAIVY#NITLDADGFSSR.V	699
		K.T#NMSLGLILTR.N	112
		R.TASCS#NVTCWLK.D	1057
		R.YFF#NVSDEAALLEK.A	343
P08648	Integrin alpha-5	K.G#NLTYGYVTIL#NGSDIR.S	297, 307
		K.NAL#NLTFHAQNVGEGGAYEAELR.V	675
		R.TEKEPLSDPVGTCYLSTD#NFTR.I	182
P23229	Integrin alpha-6	K.EINSL#NLTESHNSR.K	930
		K.QLSCVANQ#NGSQADCELGNPFK.R	770
		K.YQTL#NCSVNVNCVNIR.C	966
		R.LW#NSTFLEEYSK.L	997
P06756	Integrin alpha-V	R.TAADTTGLQPILNQFTPA#NISR.Q	615
P05556	Integrin beta-1	K.DTCTQECSYF#NITK.V	669
		R.KCS#NISIGDEVQFEISITSNK.C	417
P05106	Integrin beta-3	K.DTGKDAV#NCTYK.N	680
		R.CNNG#NGTFECGVCR.C	478
		R.DLPEELSLSF#NATCLNNEVIPGLK.S	397
		R.LRPDDSK#NFSIQVR.Q	125
P13598	Intercellular adhesion molecule 2	K.AAPAPQEATATF#NSTADR.E	176
		K.GSLEV#NCSTTCNQPEVGGLETSLDK.I	47
		R.G#NETLHYETFGK.A	153
Q14392	Leucine-rich repeat-containing protein 32	R.LIYL#NLSNNLIR.L	271

Table 1 (continued)

Accession	Protein Name	Peptide	Position in protein sequence
Q9UIQ6	Leucyl-cystinyl aminopeptidase	K.SGVI#NLTEEVLWVK.V	682
P08575	Leukocyte common antigen	K.YA#NITVDYLYNK.E	232
		R.SFH#NFTLCYIK.E	419
Q08722	Leukocyte surface antigen CD47	K.SDAVSHTG#NYTCEVTELTR.E	111
		R.DIYTFDGAL#NK.S	73
P11279	Lysosome-associated membrane glycoprotein 1	K.E#NTSDPSLVIAFGR.G	83
		R.GHTLTL#NFTR.N	102
		R.YSVQLMSFVY#NLSDTHLFP#NASSK.E	120, 129
P13473	Lysosome-associated membrane glycoprotein 2	K.IAVQFGPGFSWIA#NFTK.A	101
P04156	Major prion protein	K.GE#NFTETDVK.M	197
Q6P4Q7	Metal transporter CNNM4	K.DLVVQQLV#NVSR.R	122
Q13201	Multimerin-1	K.SIHLSINFFSL#NK.T	921
		K.V#NESVVSIAAQQK.F	431
		R.L#NDSIQTLVNDNQR.Y	783
		R.SILYYESL#NK.T	507
Q92542	Nicastrin	K.A#NNSWFQSILR.Q	530
		K.IYIPL#NK.T	45
P51575	P2X purinoceptor 1	K.CVAF#NDTVK.T	153
		K.TLHPLCPVFQLGYVVQESGQ# NFSTLAEK.G	242
		R.EAE#NFTLFIK.N	184
P53801	Pituitary tumor-transforming gene 1 protein-interacting protein	K.TCEECLK#NVSCLWCNTNK.A	54
P16284	Platelet endothelial cell adhesion molecule	K.AVYSVMAMVEHSG#NYTCK.V	301
		K.DDVLFY#NISSMK.S	84
		K.EGKPFYQMTS#NATQAFWTK.Q	551
		R.L#NLSCSIPGAPPA#NFTIQK.E	344, 356
P16671	Platelet glycoprotein 4	K.#NCTSYGVLDISK.C	321
		R.NYIVPILWL#NETGTIGDEK.A	417
		R.QFWIFDVQNPQEVMM#NSSNIQVK.Q	79
P40197	Platelet glycoprotein V	K.LLDLSGN#NLTHLPK.G	181
		R.ISALGLPT#NLTHILLFGMGR.G	51
		R.LP#NLSSLTLSR.N	243
Q6UX71	Plexin domain-containing protein 2	R.V#NLSFDFPFYGHFLR.E	160
O15031	Plexin-B2	R.ALS#NISLR.L	127
		R.SI#NVTGQGFSLIQR.F	1002
		R.TEAGAFEYVPDPTFE#NFTGGVK.K	1099
Q9NZQ7	Programmed cell death 1 ligand 1	K.LF#NVTSTLR.I	192
P43119	Prostacyclin receptor	R.#NSSLLGLAR.G	78
O95866	Protein G6b	R.V#NLSCGGVSHPIR.W	32
Q9H0X4	Protein ITFG3	R.YKPDTLAVAVE#NGTGTDR.Q	389
P16109	P-selectin	K.AYSW#NISR.K	54
		R.AFQYDT#NCSFR.C	411
		R.V#NCSHPFGAFR.Y	460
Q12913	Receptor-type tyrosine-protein phosphatase eta	K.IHVAGETDSSNL#NVSEPR.A	413
		K.YEIDVG#NESTTLGYYNGK.L	910
		R.ACVAGFT#NITFHPQNK.G	937
		R.VEITT#NQSIIIGGLFPGTK.Y	501
		R.Y#NATVYSQAA#NGTEGQPQAIEFR.T	342, 351
Q9Y210	Short transient receptor potential channel 6	K.AQSIIDA#NDTLKDLTK.V	561
Q9NXL6	SID1 transmembrane family member 1	R.VYV#NSSSENLNYPVLVVVR.Q	83
Q8IVB4	Sodium/hydrogen exchanger 9	K.LTFSPSTLLV#NITDQVYEYK.Y	96
Q99523	Sortilin	K.DITDLI#NNTFIR.T	162
		R.HLYTTTGGETDFT#NVTSLR.G	406

Table 1 (continued)

Accession	Protein Name	Peptide	Position in protein sequence
Q6UWL2	Sushi domain-containing protein 1	R.EQVPVVCLDLYPTTDYTV#NVTLLR.S	470
		R.TPEVCLALYPGT#NYTV#NISTAPPR.R	367, 371
		R.WYLA#NFSHATSF#NFTTR.E	439, 447
P07996	Thrombospondin-1	K.GCSSSTSVLLTLDNNVV#NGSSPAIR.T	248
		K.VV#NSTTGPGEHLR.N	1067
Q9BX74	TM2 domain-containing protein 1	R.LSIT#NETFR.K	197
Q9NXH8	Torsin family protein C9orf167	R.FVLQ#NASR.A	314
Q8NFQ8	Torsin-1A-interacting protein 2	K.HL#NASNPTEPATIIFTAAR.E	286
P01137	Transforming growth factor beta-1	R.LASPPSQGEVPPGPLPEAVLALY#NSTR.D	82
O15321	Transmembrane 9 superfamily protein member 1	R.IIFA#NVSVR.D	178
Q9HD45	Transmembrane 9 superfamily protein member 3	R.IVDV#NLTSEGK.V	174
Q4KMQ2	Transmembrane protein 16F	K.L#NITCESSK.K	361
Q9Y274	Type 2 lactosamine alpha-2,3-sialyltransferase	K.NAYH#NVTAEQLFLK.D	308
Q15904	Vacuolar ATP synthase subunit S1	R.ILFWAQ#NFSVAYK.D	273

Glycosylation sites are marked by "#" within the peptide sequence

a pattern for enhanced prediction of *N*-glycosylation sites. Therefore, mass spectrometric sequencing remains the method of choice for elucidation of glycosylation sites.

Conclusion

Although sequencing of proteins on a proteome-wide scale is still necessary to obtain an overview of the cellular protein content, the analytical focus is continuously shifting toward quantification of differentially expressed proteins or even their posttranslationally modified isoforms. Modifications such as phosphorylations enable fast cellular responses to extracellular stimuli and are responsible for numerous changes in protein function during platelet activation and aggregation. In contrast, carbohydrates are responsible for enabling, e.g., cell–cell recognition and cellular adhesion caused by their presence on membrane glycoproteins. However, both phospho- and glycopeptides are less abundant compared to the overwhelming number of non-modified peptides hampering analysis by undersampling and suppression effects. For efficient analysis, both modifications require enrichment of modified peptides before mass spectrometric detection.

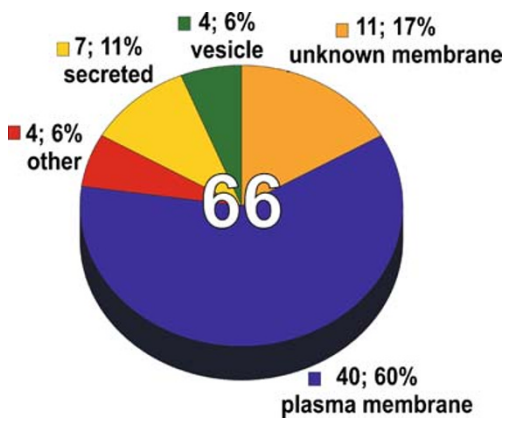

Fig. 4 Subcellular localization of identified glycoproteins. The annotation of subcellular localization information is based on the Swiss-Prot database (http://www.expasy.org)

The approach presented here, which uses electrostatic repulsion hydrophilic interaction chromatography, clearly enabled the accumulation of glycopeptides with an efficiency demonstrated also for previously used approaches like SCX-prefractionation. Application to platelet plasma membrane sample resulted in the annotation of 125 glycosylation sites, many thereof on proteins essential for platelet function. The knowledge of these sites has two major benefits. First, it enables the targeted functional characterization of influences of the elucidated sites on platelet function. As platelets are hardly accessible to standard genomic techniques because of their enucleate structure, platelet-specific knock-out models are one of the few possibilities to directly study influences of modification sites *in vivo*. However, these experiments require the exact site of the modification to be known. After the current study and in combination with two of our previous works over 250 *N*-glycosylation sites on platelet proteins have been mapped and can be used for further characterizations. Second, a major benefit of correct assignment of glycosylation sites—also with respect of clinical relevance—is their potential for the identification and quantification of biomarkers. Mass spectra of the respective peptides are the

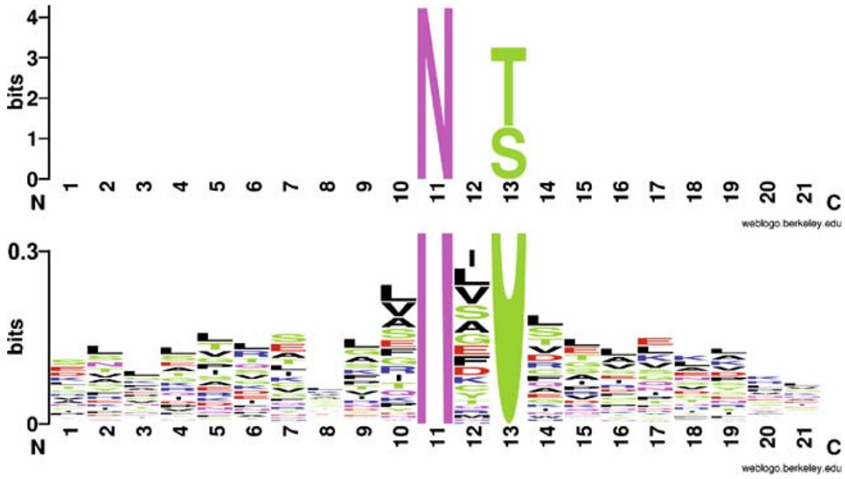

Fig. 5 Motif analysis of so far annotated platelet glycosylation sites using WebLogo [26]. Minor preferences of threonine over serine within the consensus sequence as well as a slight prevalence for small uncharged amino acids in adjacent positions (lower, magnified view) can be deduced

basis for targeted multireaction monitoring (MRM) approaches, which allow for simultaneous label-free quantification for scores of proteins, as demonstrated by Stahl-Zeng et al. [7]. The repository presented for the platelet system can therefore be used for the design of assays monitoring the presence of glycoproteins on the platelet surface or the complete cell. This enables fast and targeted screening of major protein components, e.g., during bleeding disorders. In this context, the high selectivity and sensitivity of MRM might even avoid the need for complicated sample preparation steps rendering the approach attractive for clinical use.

Apart from clinical applications using annotated glycosylation sites, the current approach also offers a new purification strategy for glycopeptides. Current results render ERLIC a complementary technique to established oxidative hydrazide coupling, lectin-based approaches, or SCX-enrichment of glycopeptides. Nevertheless, the presence of further peptide species bearing additional negative charges such as phosphorylations has to be anticipated as the ERLIC principle is not limited to glycopeptide isolation. With regard to phosphorylations, this obstacle can, e.g., easily be overcome using a dephosphorylation step before ERLIC. However, within the presented analysis of the platelet membrane samples, it is clear that spectra with glycopeptide marker ions made up for the dominant ion species when omitting the PNGaseF deglycosylation step. This argues for negligible effects of co-purified peptide species. The 125 glycosylation sites identified here show a major overlap with the previously employed techniques but also add new candidates to the number of known glycosylation sites. We therefore consider ERLIC to be advantageous for glycopeptide characterization on other glycosylation-related research fields as well. Preliminary results confirm distinct enrichment of glycopeptides from soluble fractions of *Arabidopsis thaliana* total lysates. This application to a complete organism renders ERLIC a feasible approach for complex body fluids as well. Moreover, ERLIC exhibits potential use for the separation of isoforms caused by carbohydrate microheterogeneities. This is of profound interest for basic research as it enables the direct access to intact glycopeptides. Furthermore, it reduces the complexity of individual fractions by spreading different glycopeptides over the entire gradient providing higher separation capacity than, e.g., bulk elution approaches. However, this feature might also limit the application of ERLIC for label-free quantitative profiling in turn. In contrast to, e.g., oxidative hydrazide coupling, ERLIC-separated peptides elute over a range of fractions. Intermediary but necessary clean-up steps of these fractions reduce the approach to a qualitative one as long as they are not severely standardized or the buffer conditions of ERLIC are adapted for direct mass spectrometric analysis. Nevertheless, this fact should not interfere with quantification approaches based on chemical derivatization, such as iTRAQ or ICAT. The exact retention mechanism of glycopeptides by ERLIC remains yet unclear. A major overlap (63 of 125 sites in the current study) with the previous SCX-based approach could hint to a possible enrichment of charged sialylated glycopeptides. In turn, the equally high number of additional glycopeptide identifications from the ERLIC approach argues for a more generalized retention mechanism apart from charge-bearing sialic acids.

Acknowledgements The authors thank Andrew Alpert (poly LC, Columbia, MD, USA) and Tim Tetaz (Chromatographic Technologies, Basel, Switzerland) for discussion and support regarding chromatographic separations, as well as the Institute for Transfusion Medicine, University Würzburg for supplying platelet aphaeresis samples This work was supported in part by the Deutsche Forschungsgemeinschaft, the Forschungszentrum FZT-82 and the Sonderforschungsbereich Grant 688 (to A.S.).

References

1. Kunicki TJ, Cheli Y, Moroi M, Furihata K. The influence of N-linked glycosylation on the function of platelet glycoprotein VI. Blood 2005;106(8):2744–9.
2. Hoffmeister KM, Felbinger TW, Falet H, et al. The clearance mechanism of chilled blood platelets. Cell 2003;112(1):87–97.
3. Josefsson EC, Gebhard HH, Stossel TP, Hartwig JH, Hoffmeister KM. The macrophage alphaMbeta2 integrin alphaM lectin domain mediates the phagocytosis of chilled platelets. J Biol Chem 2005;280(18):18025–32.
4. Nurden AT, Nurden P. Inherited disorders of platelets: an update. Curr Opin Hematol 2006;13(3):157–62.
5. Ju T, Cummings RD. Protein glycosylation: chaperone mutation in Tn syndrome. Nature 2005;437(7063):1252.
6. Ju T, Cummings RD. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. Proc Natl Acad Sci USA 2002;99(26):16613–8.
7. Stahl-Zeng J, Lange V, Ossola R, Aebersold R, Domon B. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics 2007;6 (10):1809–17.
8. Kaji H, Saito H, Yamauchi Y, et al. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. Nat Biotechnol 2003;21(6):667–72.
9. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol 2003;21(6):660–6.
10. Alvarez-Manilla G, Atwood J 3rd, Guo Y, Warren NL, Orlando R, Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. J Proteome Res 2006;5(3):701–8.
11. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P. A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. J Proteome Res 2004;3 (3):556–66.
12. Omaetxebarria MJ, Hagglund P, Elortza F, Hooper NM, Arizmendi JM, Jensen ON. Isolation and characterization of glycosylphosphatidylinositol-anchored peptides by hydrophilic interaction chromatography and MALDI tandem mass spectrometry. Anal Chem 2006;78(10):3335–41.
13. Lewandrowski U, Zahedi RP, Moebius J, Walter U, Sickmann A. Enhanced N-glycosylation site analysis of sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. Mol Cell Proteomics 2007;6(11):1933–41.
14. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat Methods 2007;4(3):231–7.
15. Larsen MR, Jensen SS, Jakobsen LA, Heegaard NH. Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. Mol Cell Proteomics 2007;6(10):1778–87.
16. Alpert AJ. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. Anal Chem 2008;80(1):62–76.
17. Tandon NN, Lipsky RH, Burgess WH, Jamieson GA. Isolation and characterization of platelet glycoprotein IV (CD36). J Biol Chem 1989;264(13):7570–5.
18. Lewandrowski U, Moebius J, Walter U, Sickmann A. Elucidation of N-glycosylation sites on human platelet proteins: a glycoproteomic approach. Mol Cell Proteomics 2006;5(2):226–33.
19. Schindler J, Lewandrowski U, Sickmann A, Friauf E, Gerd Nothwang H. Proteomic analysis of brain plasma membranes isolated by affinity two-phase partitioning. Mol Cell Proteomics 2006;5(2):390–400.
20. Fujiki Y, Hubbard AL, Fowler S, Lazarow PB. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. J Cell Biol 1982;93(1):97–102.
21. Geer LY, Markey SP, Kowalak JA, et al. Open mass spectrometry search algorithm. J Proteome Res 2004;3(5):958–64.
22. Beausoleil SA, Jedrychowski M, Schwartz D, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci USA 2004;101(33):12130–5.
23. Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Anal Chem 2004;76(14):3935–43.
24. Harvey DJ, Wing DR, Kuster B, Wilson IB. Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. J Am Soc Mass Spectrom 2000;11(6):564–71.
25. Hernandez-Ruiz L, Valverde F, Jimenez-Nunez MD, et al. Organellar proteomics of human platelet dense granules reveals that 14-3-3zeta is a granule protein related to atherosclerosis. J Proteome Res 2007;6(11):4449–57.
26. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004;14(6):1188–90.