

Quantitative Analysis of Human Red Blood Cell Proteome

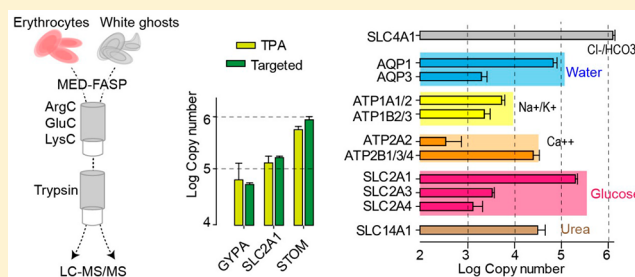
Agata H. Bryk[†] and Jacek R. Wiśniewski^{*ID}

Biochemical Proteomics Group, Department of Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

S Supporting Information

ABSTRACT: Red blood cells (RBCs) are the most abundant cell type in the human body. RBCs and, in particular, their plasma membrane composition have been extensively studied for many years. During the past decade proteomics studies have extended our knowledge on RBC composition; however, these studies did not provide quantitative insights. Here we report a large-scale proteomics investigation of RBCs and their “white ghost” membrane fraction. Samples were processed using the multienzyme digestion filter-aided sample preparation (MED-FASP) and analyzed using Q-Exactive mass spectrometer. Protein abundances were computed using the total protein approach (TPA). The validation of the data with stable isotope-labeled peptide-based protein quantification followed. Our in-depth analysis resulted in the identification of 2650 proteins, of which 1890 occurred at more than 100 copies per cell. We quantified 41 membrane transporter proteins spanning an abundance range of five orders of magnitude. Some of these, including the drug transporter ABCA7 and choline transporters SLC44A1 and SLC44A2, have not previously been identified in RBC membranes. Comparison of protein copy numbers assessed by proteomics showed a good correlation with literature data; however, abundances of several proteins were not consistent with the classical references. Because we validated our findings by a targeted analysis using labeled standards, our data suggest that some older reference data from a variety of biochemical approaches are inaccurate. Our study provides the first “in-depth” quantitative analysis of the RBC proteome and will promote future studies of erythrocyte structure, functions, and disease.

KEYWORDS: erythrocyte, red blood cell, white ghosts, membrane proteomics, FASP, MED FASP, copy number, total protein approach, TPA



INTRODUCTION

Red blood cells (RBCs) are the most abundant human blood cells. Apart from their pivotal role in oxygen and carbon dioxide transport, RBCs are involved in the transfer of glycoposphatidylinositol-linked protein immune complexes.¹ The plasma membrane of the RBCs is mechanically supported by spectrin tetramers connected by actin junctional complexes, forming a 2D six-fold triangular network anchored to the lipid bilayer.² To date, the membrane protein composition and structure has been investigated by chemical staining,^{3,4} immunoassay methods,^{5–7} crystallography,⁸ scanning electron microscopy, and atomic force microscopy.⁹

In contrast with the conventional biochemical approaches targeting single or few proteins, large-scale proteomics enables insights into thousands of proteins in a single experiment, including their abundances. However, all early proteomics analyses of RBCs failed to provide comprehensive data sets. The reasons were the presence of hemoglobin interfering with identification of lower abundant proteins and the limitations of proteomics technology to analyze membrane proteins. Further progress in analysis of cytosolic fraction of RBCs was possible using extensive fractionation of proteins^{10,11} and by reduction of hemoglobin content using peptide libraries.¹² Depletion of hemoglobin by metal affinity chromatography¹³ allowed

reproducible measuring and relative quantification of over 1000 proteins across cytosolic preparations from a series of RBCs donors.¹⁴ Using the N-terminomics procedure, Overall et al. identified 1369 human erythrocyte proteins.¹⁵ Analysis of the membrane preparation of RBCs from Diamond–Blackfan anemia patients, using extensive gel fractionation, allowed the identification of >1100 mainly nonmembrane proteins.¹⁶ Applying diverse depletion strategies to a triton insoluble membrane skeleton fraction enabled semiquantitative profiling of 77 proteins.¹⁷

Despite its potential, so far proteomics has not provided a comprehensive and quantitative picture of the erythrocyte. We aimed to (1) perform an in-depth proteomics analysis of whole RBCs and their “white ghost” (WG) preparations and to (2) quantify proteins in whole cells and their membrane fraction, including low abundance membrane transporters.

Received: January 13, 2017

Published: July 8, 2017



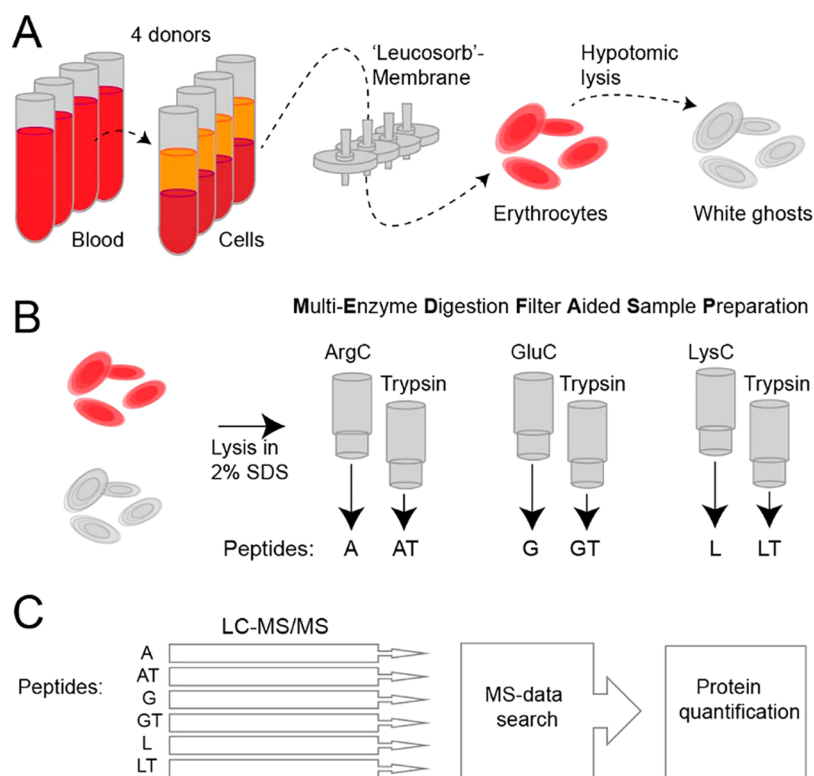


Figure 1. Overview of the analytical workflow. (A) Venous blood was collected from four healthy donors. Erythrocytes were collected by centrifugation and remaining leucocytes were depleted by filtration. We used purified erythrocytes for isolation of membrane fraction, so called the “White Ghosts”. (B) The purified cells and the membrane fractions were lysed in buffer containing 2% SDS, and the protein extracts were processed using the MED FASP method. Each sample was successively digested using endoproteinase ArgC and trypsin, endoproteinase GluC and trypsin, and endoproteinase LysC and trypsin. (C) The isolated peptides were chromatographed on a C_{18} column coupled to mass spectrometer. The mass spectra were searched using MaxQuant software. Proteins were quantified using the “Total Protein Approach”.

MATERIALS AND METHODS

Erythrocytes

Blood samples were provided by the authors and further laboratory employees who gave their informed consent. Nine mL of fasting venous blood was collected with minimal stasis into EDTA tubes. The procedure was modified from previously published methods¹⁶ (Figure 1). To minimize proteolysis, all samples and buffers were kept at 0–4 °C during all steps. RBCs were collected from the plasma by centrifugation for 5 min at 150g and were suspended in PBS containing 1 mM EDTA. Then, the slurry was passed through a leucocyte depletion filter (Acrodisc, PSF Syringe Filters, Pall, Ann Arbor, MI). RBCs were washed four times with PBS and stored overnight on ice at 4 °C in PBS containing 10 mM glucose and EDTA-free protease inhibitor cocktail tablets (Roche, Mannheim, Germany).

Preparation of “White Ghosts”

WGs were isolated according to Pesciotta et al.¹⁴ In brief, the purified RBCs were lysed using 10 volumes of hypotonic lysis buffer composed of 5 mM sodium phosphate, 1 mM EDTA, complete, mini EDTA-free protease inhibitor cocktail, pH 8.0; the WGs were then isolated by centrifugation at 21 000g for 40 min. WG were purified by four to five additional washes and centrifugations.

Lysis of the Cells and Membrane Preparations

Whole RBCs and WGs were lysed in 0.1 M Tris-HCl containing 2% SDS and 0.05 M DTT at 100 °C for 5 min.

Protein and Peptide Determination

Total protein concentration in lysates and the peptide content in the digests (see below) were assayed using a tryptophan fluorescence-based WF assay in microtiter plate format.¹⁸ For measurements, Corning Costar 96-well black flat-bottomed polystyrene plates (Sigma-Aldrich, Taufkirchen) were used.

Multienzyme Digestion Filter-Aided Sample Preparation (MED FASP)

Sample aliquots containing 100 μ g total protein were processed using the MED FASP method¹⁹ with modifications described recently.²⁰ Proteins were first cleaved overnight using endoproteinase ArgC, GluC, or LysC, and in the second step they were digested with trypsin for 2 h. The enzyme-to-protein ratio was 1:50. Digestions with LysC and trypsin were carried out in 50 mM Tris-HCl, pH 8.5. For the cleavage with ArgC the buffer was supplemented with 10 mM $MgCl_2$. The digestion with GluC was carried out in 25 mM ammonium bicarbonate. All reactions were conducted at 37 °C. Aliquots containing 10 μ g total peptide were concentrated to a volume of \sim 5 μ L and were stored frozen at –20 °C until mass spectrometric analysis. Ten μ g aliquots of the peptides were deglycosylated with 4 units PNGase F at room temperature for 2 h.

Liquid Chromatography–Tandem Mass Spectrometry

Analysis of peptide mixtures was performed using a QExactive HF mass spectrometer (Thermo-Fisher Scientific, Palo Alto). Aliquots containing 5 μ g total peptide were chromatographed on a 50 cm column with 75 μ m inner diameter packed C_{18} .

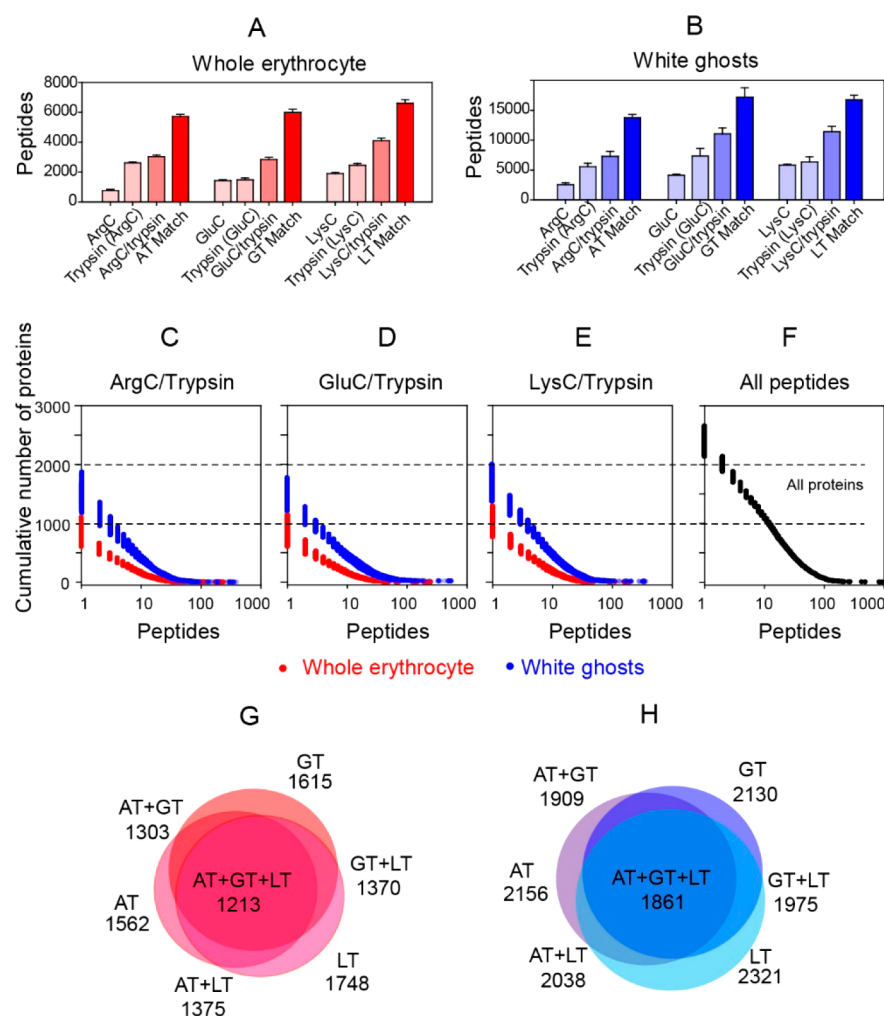


Figure 2. Peptide and protein identification and quantification in lysates of whole erythrocytes (red–pink colors) and their white ghosts (blue–purple colors). (A,B) Identification of peptides obtained by three different cleavage approaches using either ArgC followed by trypsin, GluC followed by trypsin, LysC followed by trypsin. In each group the first, second, third, and fourth bars indicate number of peptides identified from the digest with the first enzyme and the second enzyme (in all trypsin), the composite number of peptides from both digestions, and the total number of peptides from both digestions using the “matching between runs” algorithm of the MaxQuant software. (C–E) Cumulative number of proteins sorted by the number of matching peptides from individual digestion approaches. (F) Cumulative number of proteins sorted by the number of matching peptides from all approaches. (G,H) Venn plots with the number of identified proteins using various digestion approaches.

material (100 Å pore size; Dr. Maisch, Ammerbuch-Entringen, Germany). Peptide separation was carried out at 300 nL/min for 75 min using a two-step acetonitrile gradient 5–40% over the first 60 min and 40–95% for the following 15 min. The backpressure varied between 450 and 650 bar. The temperature of the column oven was 55 °C.

The mass spectrometer operated in data-dependent mode with survey scans acquired at a resolution of 50 000 at m/z 400 (transient time 256 ms). Up to the top 15 most abundant isotope patterns with charge $\geq +2$ from the survey scan (300–1650 m/z) were selected with an isolation window of 1.6 m/z and fragmented by HCD with normalized collision energies of 25. The maximum ion injection times for the survey scan and the MS/MS scans were 20 and 60 ms, respectively. The ion target value for MS1 and MS2 scan modes was set to 3×10^6 and 3×10^5 , respectively. The dynamic exclusion was 25 s and 10 ppm. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository²¹ with the data set identifier: PXD005200.

Stable Isotope-Labeled Peptide-Based Protein Quantification

Peptides with isotopically labeled lysine ($^{13}\text{C}_6$, $^{15}\text{N}_2$) and arginine ($^{13}\text{C}_6$, $^{15}\text{N}_4$) residues were purchased from (Sigma-Aldrich). The peptide matching sequences of erythrocyte band 7 integral membrane protein/stomatin (SLC4A1/STOM) (TISFDIPPQEILTK, YLQTLTIAAEK), glycophorin A (GYPA) (TVYPPEEETGER), and glucose transporter 1/solute carrier family 2 member A (GLUT1/SLC2A1) (TFDEIASGFR) were selected on the basis of the proteomics analysis using the LysC and trypsin digestion. The selected peptides were repeatedly identified in the analyses of the WG preparations. Ten μg aliquots of the digest of the WG lysate were spiked with 0.5, 1, 2, or 5 pmol of the labeled peptides and were LC–MS/MS analyzed as described above. Ratios of heavy to light MS1 intensities of “light” and “heavy” (standard) peptides were manually extracted from the spectra (Xcalibur Qual Browser) and used for calculation of concentrations of targeted proteins (Table S1).

Data Analysis

The spectra were searched using MaxQuant software. All raw files were searched together in single MaxQuant run using separate parameter options for MS files from analyses of Arg C, LysC, GluC, and tryptic peptides. For the tryptic peptides generated from the GluC-precleaved material a mixed GluC/tryptic specificity was set. A maximum of two missed cleavages was allowed. Carbamidomethylation was set as fixed modification. The “matching between the runs” option was used. The maximum false peptide and protein discovery rate was specified as 0.01. Protein abundances were calculated using the “total protein approach” (TPA) method.²² The calculations were performed in Microsoft Excel.

The copy number of protein was calculated using the relationship

$$\text{Copy number} = \text{RBC volume} \times N_A \times \text{total protein content per cell}$$

where

$$c(i) = \frac{MS - \text{signal}(i)}{\text{total MS} - \text{signal} \times MW(i)} \left[\frac{\text{mol}}{\text{g total protein}} \right]$$

and N_A is the Avogadro number.

We assumed the mean volume of RBCs of 90 fL²³ and the average total protein content per cell as 20% of RBCs mass.²⁴ The values obtained using different proteinase were averaged.

Calculation of copy numbers using data from the analysis of the WG fractions ($\text{Copy number}_{\text{WG}}$) was performed using the membrane protein enrichment (MPE) factor. This factor reflected the ratio of protein concentration in the whole RBCs sample and the respective protein concentration in the subsequent WG sample. MPE was computed for the 10 most abundant integral membrane proteins (Table S2).

$$\text{Copy number} = \text{Copy number}_{\text{WG}}(i) / \text{MPE}$$

Membrane proteins are often glycosylated with predominant N-glycosylation. To assess whether the N-glycosylation has an effect on the label free quantitation, we compared the proteins abundances assessed from analysis of the deglycosylated and untreated samples.

RESULTS

Proteomics Analysis

In this study we analyzed RBC samples originating from four donors (Figure 1). We analyzed whole cell lysates and membrane fractions prepared by hypotonic lysis of RBCs, resulting in WGs. Proteins were extracted with 2% SDS and subjected to the filter-aided sample preparation (FASP) approach, which allows unbiased analysis of soluble and membrane proteins.^{20,25} To cover the highest number of RBCs proteins and to achieve the height sequence coverage, each protein extract was consecutively digested using two enzymes. In the first cleavage step, endoproteinase Arg C, GluC, or LysC was used. After elution and collection of peptides, in the second step, the material remaining in the filter was treated with trypsin (Figure 1B). Protein digestions were carried out in the absence of any denaturant, a condition ensuring high digestions yields with very low rates of peptides with missed cleavage sites.²⁰ Each peptide fraction was analyzed separately by LC–MS/MS, and all recorded spectra were

searched together (Figure 1C). Each of the protein digestion protocols resulted in a similar number of identified peptides per analyzed sample of 2500–3000 in the analysis of the whole RBCs and 6000–7000 peptides in the WG analysis (Figure 2A,B). The cumulative numbers of identified peptides per sample were about 6000 and 15 000 for the whole cell and WG, respectively. Comparing the three digestion strategies used, the combination of LysC and trypsin enabled identification of the highest number proteins; however, the two other enzyme combinations significantly extended the total number of identified proteins (Figure 2C–E). The total number of proteins that were identified per donor sample matching RBCs and WG analyses was about 2500 (Figure 2F, Table S3). The Venn plots show that 30% of proteins were identified using only one or two cleavage approaches (Figure 2G,H). This result emphasizes the importance of multicleaveage strategies for comprehensive characterization of proteomes.²⁶

RBC Purity

Purity of the RBCs samples is prerequisite for their unbiased proteome analysis. In this study, we followed the efficient, reproducible RBC processing protocol by Pesciotta et al.¹⁶ RBCs were stored for 24 h to reduce the number of reticulocytes in 4 °C.

Next, RBCs were depleted of small cells, including platelets and microparticles, by repeated sample washing and low-speed centrifugation. Finally, leucocytes, neutrophils, and other cells larger than RBCs were removed by filtration (Figure 1). Because our analysis provided absolute abundance values of proteins, we could estimate the extent of contamination of the RBCs by platelets and leucocytes. In contrast with RBCs, leucocytes and neutrophils are nucleated cells. Each nucleus of a 2n human cell contains $\sim 2.7 \times 10^8$ copies of all core and linker histones. We identified $(1.5 \text{ to } 7.0) \times 10^4$ histone molecules across the samples (Table S3). This value accounts for one leucocyte or neutrophil per 10^4 RBCs. Out of the platelet-specific antigens such as CD36, CD41, CD42, and CD61 occurring at 20–80 000 copies per platelet,²⁷ we identified only CD36 at a copy number of 64 ± 17 . This suggests a maximal platelet contamination below 1 platelet per 1000 erythrocytes. In fact, the contamination of our preparations RBC and WG by platelets could be even lower. Notably, we did not identify some of the top ten most abundant platelet proteins such as platelet factor 4 or platelet basic protein, which occur in these cells at 530 000 and 355 000 copies, respectively.²⁷ Thus quantitative proteomics data confirmed the expected depletion of leucocytes and platelets from the RBC samples. Because a minute contamination of the erythrocyte preparations by other blood cells cannot be avoided we consider proteins above a threshold of 100 copies as those occurring in erythrocytes. However, we do not rule out the possibility that some proteins with abundances below this limit may also originate from RBCs. In addition we cannot exclude contamination of the RBC and their membrane preparations by plasma proteins. For example, the most abundant plasma protein, albumin, was present at several thousand copies.

To estimate contamination of the RBC preparation we compared the abundances of ribosomal subunits that were calculated in this study with abundances abundance of these proteins in reticulocytes.²⁸ In our preparation the median abundance of these subunits was 200 copies per cell, whereas in reticulocytes the corresponding value was 220 000. This

suggests that RBC preparations used in our analyses were contaminated by <0.1% reticulocytes

Protein Quantification

Across the preparations the MPE was in the range of 8–12 (Figure 3A). 1890 identified proteins occurred at >100 copies

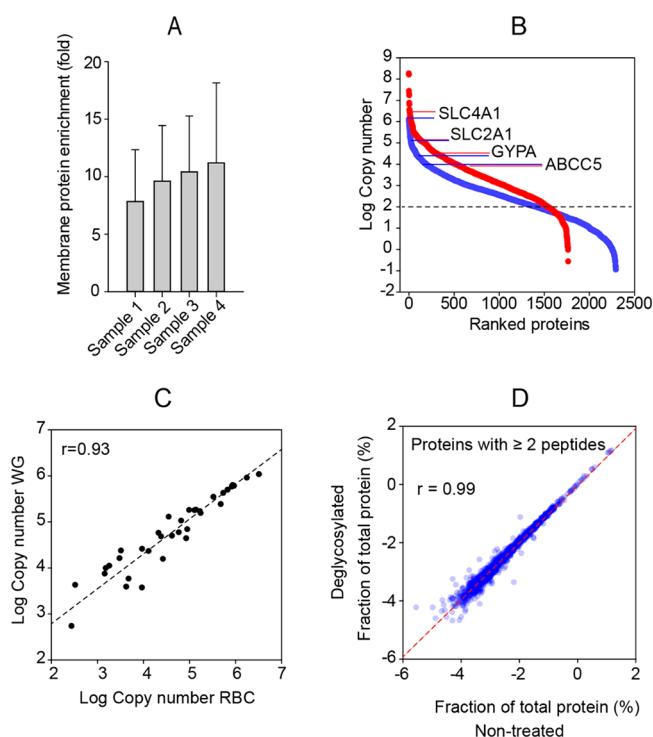


Figure 3. Quantitative analysis. (A) Average enrichment of top 10 most abundant membrane proteins in the WG fractions relative to whole RBCs. (B) Distribution of protein copy numbers across all proteins in whole cells (red circles) and WG fractions (blue circles). The dashed line shows the 100 copy per cell sample purity threshold. (C) Correlation of the copy numbers for the known plasma membrane proteins calculated using data from analysis of whole cells and WG fractions. (D) Correlation of assessed protein abundance with and without protein deglycosylation.

per cell; 1536 identified in the whole cell lysate and 1406 found in the WG fraction (Figure 3B). For the known RBC membrane proteins we observed an excellent correlation [$r(\text{Pearson}) = 0.93$] between copy numbers of proteins calculated using data from both whole cells and WGs data sets (Figure 3C). The analysis revealed that N-glycosylation has practically no effect on protein quantification (Figure 3D).

Analysis of the Whole RBCs

Analysis of whole RBCs lysates allowed the identification of 2090 proteins. Of those, 450 were quantified across all samples (Figure 3B). Hemoglobin is the dominant protein of the RBCs. Its subunits alpha (HbA) and beta (HbB) exist at about 200 million copies, whereas the less abundant subunits delta (HbD) and gamma (HbG) occur at about 30 and 4 million copies, respectively (Figure 4A). Surprisingly, we also found significant amounts of the subunits theta (HbQ) (400 000 copies), zeta (HbZ) (900 000 copies), and mu (HbM) (100 000 copies); these are generally considered to be embryonic forms of hemoglobin. Identification of these isoforms in the presence of >100-fold excess of HbA and HbB stresses the power of mass spectrometry compared with classical biochemical methods.

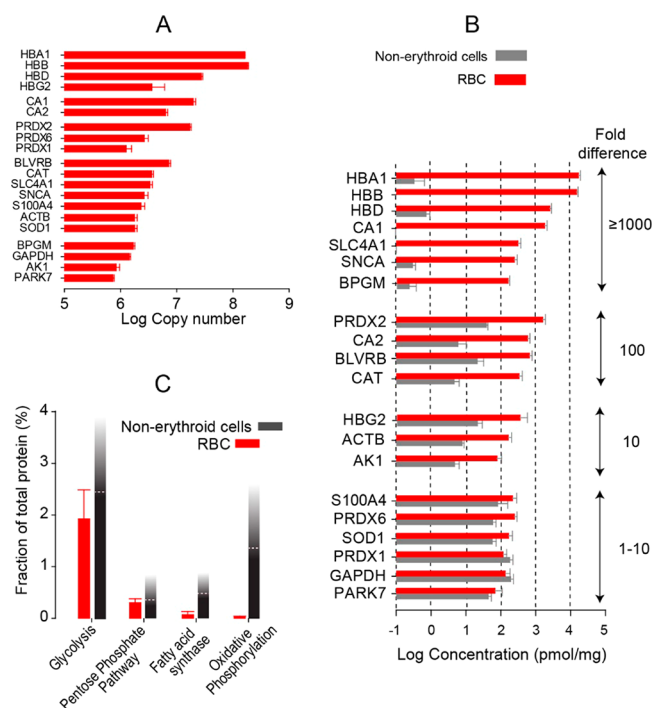


Figure 4. Proteomics analysis of the whole RBCs. (A) Copy numbers of the 20 most abundant proteins in RBCs. (B) Comparison of the concentrations of the most abundant proteins in RBCs compared with nonerythroid human cells. The values for the nonerythroid cells are based on proteomics analyses of human hepatocytes³¹ and cultured cells.^{31,32} (C) Total protein content of basic metabolic processes in RBCs compared with the nonerythroid cells. The black bars show the average value across different human cells. The gradient area of the bars shows standard deviation.

Carbonic anhydrase is the second most abundant protein. Its isoform 1 (CA1) is present at about 20 million copies. Peroxiredoxin-1, -2, and -6 (PRDX1, PRDX12, and PRDX16) occur with similar copy numbers. The next group of highly abundant proteins consists of biliverdin reductase B (BLVRB), which catalyzes reduction of methemoglobin, and catalase (CAT) and superoxide desmutase (SOD1), which protect from toxic effects of oxygen peroxide and radicals, respectively. Bisphosphoglycerate mutase (BPGM), the central enzyme of the Rapoport–Luebering cycle, is present at 1.6 million copies. BPGM converts 1,3-bisphosphoglycerate, which is formed by glycolysis, to 2,3-bisphosphoglycerate (2,3-BPG). The main role of 2,3-BPG is to shift the equilibrium of hemoglobin toward the deoxy state. Housekeeping enzymes glyceraldehyde-P dehydrogenase (GAPDH) and deglycase (PARK7)²⁹ are present at 1.4 and 0.7 million copies, respectively. Calcium binder S100A4 is present at about 2 million copies. Another enzyme found to be present at 2–3 million copies was α -synuclein. Notably, it has recently been reported that RBCs as the major source of α -synuclein in blood.³⁰

To compare abundance of proteins between RBCs and nonerythroid cells, such as hepatocytes³¹ and cultured cells,^{31,32} we used protein concentrations, which are not affected by cell size (Figure 4B). Hemoglobin subunits, CA1, SLC4A1, α -synuclein, and BPGM exist in RBCs at concentration higher by a least 1000 times compared with common cells. The concentrations of proteins protecting from exposure to oxygen, PRDX2, BLVRB, and CAT, are 100 times higher compared with nonerythroid cells. Proteins with “housekeeping” functions

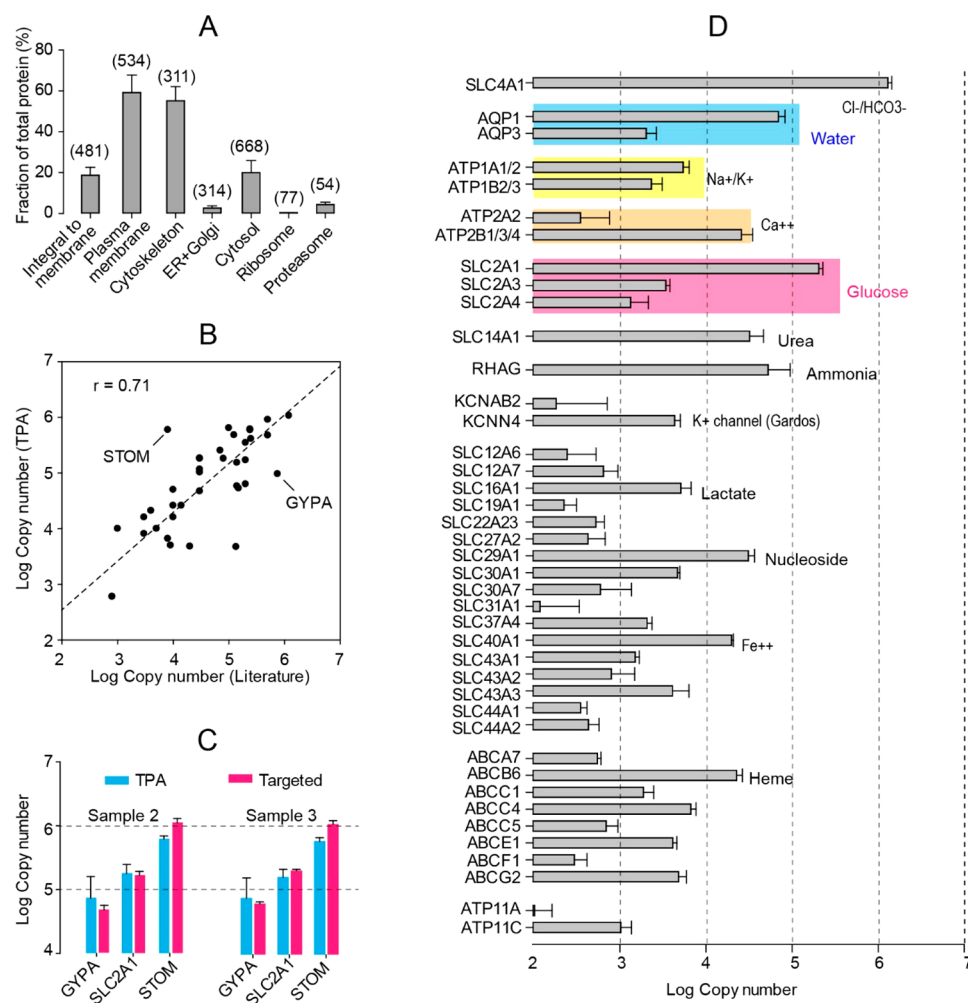


Figure 5. Composition of the WG preparation. (A) Global composition of WGs based on Gene Ontology (GO) annotations. In parentheses is the total number of proteins per GO term. (B) Comparison of the protein copy numbers determined by total protein approach (TPA) with literature values (Table 1). r , Spearman's correlation coefficient. (C) Comparison of the copy numbers of Glycophorin A (GYPA), Glut1 (SLC2A1), and Band 7-Stomatins (STOM) determined by TPA and targeted analysis. Two donor samples were analyzed. (D) Copy numbers membrane transporter proteins identified with >100 copies.

have similar concentrations in RBCs and nonerythroid cells. Enzymes involved in the glycolytic and pentose phosphate pathways constitute a similar portion of total protein of RBCs compared with nonerythroid cells. In contrast fatty acid synthase and enzymes of oxidative phosphorylation are present at significantly lower amounts in RBCs. Notably, not all key enzymes of the oxidative phosphorylation were identified, indicating interruption of this process. In addition, we were unsuccessful in identification of some of these proteins by Western blotting (Figure S1). This suggests their occurrence in a partially degraded state and most probably only in “young” RBCs. Proteins originating from nucleus, mitochondria, and other organelles identified in our and previous proteomics RBCs studies^{10,15} were suggested by us and the previous authors to represent nonfunctional remnants of erythrocyte maturation.

Analysis of the WG Fraction: Identification of Novel Membrane Proteins

Analysis of WG fractions resulted in the identification of 2400 proteins. The plasma membrane and cytoskeletal proteins constitute the most abundant portion of total protein (Figure 5A). Approximately 20% of the total WG protein is predicted

to be integral to membranes. The WG is a crude membrane preparation and hence contains a substantial proportion of nonplasma membrane proteins. Cytosolic proteins contribute in 20% of total protein to WG preparation, and more than 300 proteins have ER or Golgi “Gene Ontology” annotation. An almost complete set of ribosomal and proteosomal proteins was identified in this fraction.

For about 50 proteins of the WG fraction the copy numbers have been reported in the past. We found that copy number values calculated by the TPA approach were similar to the literature data for most proteins (Table 1, Figure 5B). A comparison of the data sets provides a Spearman correlation coefficient r of 0.71. However, obvious discrepancies were found for some proteins. The TPA-based copy number values for STOM were about 2 orders of magnitude higher compared with the literature data. In contrast, GYPA appeared to be 5–10 times less abundant compared with previous studies. Across references the copy numbers of SLC2A1 vary between 200 000 and 700 000.³³ To clarify these differences we have quantified these proteins by means of a so-called “targeted approach”. Two donor WG samples were spiked with stable isotope-labeled peptide standards and were analyzed by LC–MS/MS. The results confirmed the accuracy of the TPA quantification of

Table 1. Comparison of Protein Copy Numbers Obtained by the Proteomics Analysis and the Literature Data^a

protein	gene	copies ($\times 10^{-3}$) literature	copies \pm st. dev ($\times 10^{-3}$) proteomics
AE1 (band 3)	SLC4A1	1200	1074 \pm 114
b-Actin	ACTB	500	908 \pm 65
Stomatin (band 7)	STOM	10	643 \pm 79 (1100) ^b
b-Spectrin	SPTB	242	612 \pm 69
a-Spectrin	SPTA1	242	590 \pm 67
Glyceraldehyde-P DH	GAPDH	500	475 \pm 120
Ankyrin	ANK1	124	479 \pm 32
Protein 4.2	EPB42	250	411 \pm 20
Protein 4.1	EPB41	200	350 \pm 35
Tropomyosin	TPM3	70	253 \pm 66
b-Adducin	ADD2	30	181 \pm 14
p55	MPP1	80	181 \pm 9.6
Tropomodulin	TMOD1	30	184 \pm 26
Glut1	SLC2A1	200–700	170 \pm 18 (200)
Dematin (band 4.9)	EPB49	140	153 \pm 20
CD59	CD59	20–40	116 \pm 39
a-Adducin	ADD1	30	103 \pm 5.1
Rh polypeptide	RHCE, RHD	200 (D and CE)	63 \pm 17
Glycophorin A	GYPA	500–1000	96 \pm 67 (52)
Aquaporin 1	AQP1	120–160	58 \pm 10
Kell glycoprotein	KEL	3–18	50 \pm 5.4
CD47	CD47	10–50	47 \pm 5.4
RhAG	RHAG	100–200	53 \pm 25
Nucleoside transporter	SLC29A1	10	26 \pm 4.2
LW glycoprotein	ICAM4	3–5	21 \pm 10
Urea transporter	SLC14A1	14	26 \pm 8.2
ACHE	ACHE	10	16 \pm 2.9
Lutheran glycoprotein	BCAM	1.5–4	16 \pm 5.7
CD99	CD99	1	10 \pm 2.6
CD58 (LFA-3)	CD58	3–7	10 \pm 0.9
KX	XK	1.5–4	8.1 \pm 3.2
CD44	CD44	6–10	6.6 \pm 2.2
DAF (CD55)	CD55	20	4.8 \pm 2.6
Duffy	FY	6–13	5.0 \pm 2.5
CR1	CR1	0.02–1.5	0.6 \pm 0.3

^aFrom data compilation by Burton and Bruce.³³ ^bIn parentheses are values that were assessed by targeted proteomics analysis.

STOM, GYPA, and SLC2A1 (Figure 5C), suggesting that some older quantification data based on biochemical approaches needs to be revised.

Membrane channels and transporters are essential for the exchange of metabolites and ions between the cytoplasm and the environment of the cell. Our analysis provides a comprehensive picture of the occurrence of these proteins in the RBCs membrane. 41 proteins of this class were identified at more than 100 copies per cell (Figure 5D). The most abundant one is the solute carrier family 4, member 1/band 3 anion exchange protein (SLC4A1/AE1) present in >1 000 000 copies per RBCs, whereas the glucose transporter SLC2A1 occurs at about 200 000 copies. Proteins involved in transport of water, urea, ammonia, nucleosides, calcium, and iron exist at 20 000–60 000 copies. The Na⁺/K⁺ ATPase and the K⁺ channel *Gardos* are present at 3000–5000 copies. Other transporter proteins, often without known specificity, exist at fewer than 1000 copies. Some of these transporters, including the drug transporter, ATP-binding-cassette (ABC) subfamily A, member 7 (ABCA7) and the choline transporters, solute carrier family 44, members 1 and 2 (SLC44A1 and SLC44A2), have not previously been reported for RBCs.

In addition to the well known structural and transport proteins of the plasma membrane, our analyses mapped many other proteins with predicted transmembrane domains (Table S3). For example, 16 proteins with names “Transmembrane protein *number*” (TMEM_{number}) and 7 “Transmembrane emp24 domain-containing” proteins were found at 100–8000 copies. Most of these have not been functionally characterized.

DISCUSSION

Proteomics Analysis

To our knowledge, this study provides the most comprehensive proteome analysis of the human erythrocyte so far. We provide absolute protein abundance estimations in RBCs and their WG fractions. To increase the depth of analysis and quantitation accuracy we applied three different protein cleavage strategies using the MED FASP format. We identified 2650 proteins, among them 1890 proteins occurring at more than 100 copies per cell. Notably, this study allowed identification of low abundant membrane transporters and other membrane proteins with unknown functions. Because our data are quantitative, the created database can serve to the community

as a useful tool for further studies on RBCs diseases and erythrocyte aging during storage in blood banks.³⁴

Protein Copy Numbers

The composition of the plasma membrane of human RBCs has been studied over the past several decades. Previously, copy numbers of membrane proteins were determined by classical and biochemical methods, such as radioimmunoassay^{6,7} and densitometry of Coomassie-stained polyacrylamide gels.⁴ Because all of these methodologies allowed only indirect measurements, the accuracy of the data appears to be moderate. For instance, the copy numbers reported by Anstee et al. for CD44 varied from 1888 to 10,450 between individual assays and antibodies used.⁶ Relatively high copy numbers for GYPA were derived from Scatchard plots,⁷ which are known to be prone to generating large errors.³⁵ Copy numbers assessed by Coomassie staining⁴ are biased by uneven binding of the stain to proteins varying in their amino acid composition. This phenomenon was not taken into consideration. In addition, many of these calculations assumed a standard protein content of 0.6 pg per single WG, an estimate provided by Dodge et al.,³⁶ but overlooked the fact that the total membrane protein content can vary between methods used for WG preparation.

Because quantitative analyses using only single enzymes for protein digestion may be heavily biased against proteins lacking specific cleavage sites, in this study we applied three different protein digestion strategies. The protein abundance data reported here were calculated by averaging the TPA values obtained from the three different protein digests. In the present study copy number values were computed using the TPA approach. The copy numbers calculated by TPA showed excellent correlation with the literature data.²² Accuracy of the approach was validated by comparing with data from targeted approaches³⁷ and demonstrated that on average, the deviation between values obtained by TPA and targeted proteomics was 1.5 fold. We provided a global proteome quantification; for many proteins the abundance values reported here are only estimates. Nonetheless the copy number values attained for the RBCs membrane correlate well with the literature (Figure S5). Although for a few proteins, such as GYPA and STOM, the TPA and previous reports were quite different, targeted analysis showed that the values were close to the TPA copy number estimates. Similarly, the TPA values for GLUT1 copy number agreed well with the labeled standard-based quantification.

Membrane Transporters

Membrane transporters are of pivotal importance for cell homeostasis and play a crucial role in diseases. This study is the first to provide comprehensive and quantitative data set of membrane transporter proteins in the RBCs membrane.

We identified eight members of the ABC transporters family (ABCA7, ABCB6, ABCC1, ABCC4, ABCC5, ABCE1, ABCF1, ABCG2). Some of them had already been identified in previous nonproteomics studies.³⁸ ABCB6 and ABCB10 were assigned to mitochondria, although it has been recently shown that ABCB6 is secreted in exosomes, and resides on the erythrocyte cell surface.³⁸ ABCB6 is functionally relevant in the plasma membrane, where its expression prevents the accumulation of specific porphyrins in the cell.³⁹ We found ABCB6 at 10 000–20 000 copies in RBCs, whereas in human hepatocytes and cultured hepatoma cells this protein occurs at 50 000–75 000 copies.³¹ This three- to five-fold difference may be attributed to the fact that due to their function the RBCs require an efficient porphyrin clearance system.

The presence of human multidrug resistance protein 1 (MRP1/ABCC1) in the human erythrocyte membrane is well established.⁴⁰ Recently, the indirect evidence suggesting that ABCC4 (MRP4) and ABCC5 (MRP5) contribute to cGMP transport in RBCs was confirmed with a new inhibitor of ABCC4 transport.⁴¹ In another study, immunoprecipitation of ABCC5 from CHAPS-solubilized extracts reduced active transport by ~45%, showing that MRP5 is an important cGMP-transporting protein in human erythrocytes.⁴² We provided the first evidence of the presence of the ABCA7 transporter on the red blood cell membrane. The potential roles of this protein, such as binding of APOA1 and phospholipid efflux from cells,⁴³ have to be further investigated in the context of the red blood cell membrane.

Loss of some membrane surface results in hereditary spherocytosis syndrome manifesting as hemolytic anemia,⁴⁴ which may be attributed not only to the qualitative but also quantitative changes of the RBC membrane proteins. For example, absence of band 3 causes severe hereditary spherocytosis,⁴⁵ whereas a decrease in human erythroid protein 4.1R may result in hereditary elliptocytosis.⁴⁶ SDS–polyacrylamide gel electrophoresis (PAGE) is the technique recommended by the International Council for Standardization in Haematology for the quantitation of relative membrane protein contents in diseases such as hereditary elliptocytosis.⁴⁷ Currently, there is an increasing need for a precise tool for the detection of copy numbers of red blood cell membrane proteins, also in relation to such diseases as Alzheimer's disease or cancer. Data obtained by flow cytometry indicate that in late onset Alzheimer disease, the GLUT1, ABCA1, and the ABCG2 transporters are significantly increased in comparison with age-matched control subjects.⁴⁸ Because ABCG2 polymorphisms are known to modify essential pharmacokinetic parameters, uric acid metabolism, and cancer drug resistance, a direct determination of the erythrocyte membrane ABCG2 protein expression may provide valuable information for assessing these conditions or for devising drug treatments.⁴⁹

■ ASSOCIATED CONTENT

§ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.7b00025.

Figure S1. Western blot analysis of selected proteins. (PDF)

Table S1. Stable isotope labeled peptide based protein quantification. Table S2. Enrichment of membrane proteins in the white ghost fraction. Table S3. List of identified proteins and their concentrations in the whole red cells and white ghosts. (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jwisniew@biochem.mpg.de. Tel: +49 89 8578 2205. Fax: +49 89 8578 2219.

ORCID

Jacek R. Wiśniewski: 0000-0002-8452-5095

Present Address

†A.H.B.: Institute of Cardiology, Jagiellonian University Medical College; John Paul II Hospital, 80 Pradnicka Str., 31-202 Cracow, Poland.

Author Contributions

A.B. and J.R.W. designed the research, performed experiments, analyzed results, and wrote the paper.

Funding

The work was funded by Max-Planck Society for the Advancement of Science by the German Research Foundation (DFG/Gottfried Wilhelm Leibniz Prize). A.B. was supported by a short-term fellowship of the Deutscher Akademischer Austauschdienst (DAAD).

Notes

The authors declare no competing financial interest.

The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository²¹ with the data set identifier: PXD005200.

ACKNOWLEDGMENTS

We are grateful to Prof. Matthias Mann for continuous support. We thank Katharina Zettl and Korbinian Mayr for technical support. We thank Dr. G. Borner for critical reading of the manuscript.

REFERENCES

- (1) De Rosa, M. C.; Carelli Alinovi, C.; Galtieri, A.; Scatena, R.; Giardina, B. The plasma membrane of erythrocytes plays a fundamental role in the transport of oxygen, carbon dioxide and nitric oxide and in the maintenance of the reduced state of the heme iron. *Gene* **2007**, 398 (1–2), 162–71.
- (2) Li, H.; Lykotraftis, G. Erythrocyte membrane model with explicit description of the lipid bilayer and the spectrin network. *Biophys. J.* **2014**, 107 (3), 642–53.
- (3) Steck, T. L. The organization of proteins in the human red blood cell membrane. A review. *J. Cell Biol.* **1974**, 62 (1), 1–19.
- (4) Jones, M. N.; Nickson, J. K. Monosaccharide transport proteins of the human erythrocyte membrane. *Biochim. Biophys. Acta, Rev. Biomembr.* **1981**, 650 (1), 1–20.
- (5) Petty, A. C.; Green, C. A.; Daniels, G. L. The monoclonal antibody-specific immobilization of erythrocyte antigens assay (MAIEA) in the investigation of human red-cell antigens and their associated membrane proteins. *Transfus. Med.* **1997**, 7 (3), 179–88.
- (6) Anstee, D. J.; Gardner, B.; Spring, F. A.; Holmes, C. H.; Simpson, K. L.; Parsons, S. F.; Mallinson, G.; Yousaf, S. M.; Judson, P. A. New monoclonal antibodies in CD44 and CD58: their use to quantify CD44 and CD58 on normal human erythrocytes and to compare the distribution of CD44 and CD58 in human tissues. *Immunology* **1991**, 74 (2), 197–205.
- (7) Gardner, B.; Parsons, S. F.; Merry, A. H.; Anstee, D. J. Epitopes on sialoglycoprotein alpha: evidence for heterogeneity in the molecule. *Immunology* **1989**, 68 (2), 283–9.
- (8) Arakawa, T.; Kobayashi-Yurugi, T.; Alguet, Y.; Iwanari, H.; Hatae, H.; Iwata, M.; Abe, Y.; Hino, T.; Ikeda-Suno, C.; Kuma, H.; Kang, D.; Murata, T.; Hamakubo, T.; Cameron, A. D.; Kobayashi, T.; Hamasaki, N.; Iwata, S. Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science* **2015**, 350 (6261), 680–4.
- (9) Buys, A. V.; Van Rooy, M. J.; Soma, P.; Van Papendorp, D.; Lipinski, B.; Pretorius, E. Changes in red blood cell membrane structure in type 2 diabetes: a scanning electron and atomic force microscopy study. *Cardiovasc. Diabetol.* **2013**, 12, 25.
- (10) Pasini, E. M.; Kirkegaard, M.; Mortensen, P.; Lutz, H. U.; Thomas, A. W.; Mann, M. In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood* **2006**, 108 (3), 791–801.
- (11) Kakhniashvili, D. G.; Bulla, L. A., Jr.; Goodman, S. R. The human erythrocyte proteome: analysis by ion trap mass spectrometry. *Mol. Cell. Proteomics* **2004**, 3 (5), 501–9.
- (12) Roux-Dalvai, F.; Gonzalez de Peredo, A.; Simo, C.; Guerrier, L.; Bouyssie, D.; Zanella, A.; Citterio, A.; Burlet-Schiltz, O.; Boschetti, E.; Righetti, P. G.; Monsarrat, B. Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced mass spectrometry. *Mol. Cell. Proteomics* **2008**, 7 (11), 2254–69.
- (13) Ringrose, J. H.; van Solinge, W. W.; Mohammed, S.; O'Flaherty, M. C.; van Wijk, R.; Heck, A. J.; Slijper, M. Highly efficient depletion strategy for the two most abundant erythrocyte soluble proteins improves proteome coverage dramatically. *J. Proteome Res.* **2008**, 7 (7), 3060–3.
- (14) Pesciotta, E. N.; Lam, H. S.; Kossenkova, A.; Ge, J.; Showe, L. C.; Mason, P. J.; Bessler, M.; Speicher, D. W. In-Depth, Label-Free Analysis of the Erythrocyte Cytoplasmic Proteome in Diamond Blackfan Anemia Identifies a Unique Inflammatory Signature. *PLoS One* **2015**, 10 (10), e0140036.
- (15) Lange, P. F.; Huesgen, P. F.; Nguyen, K.; Overall, C. M. Annotating N termini for the human proteome project: N termini and Nalpha-acetylation status differentiate stable cleaved protein species from degradation remnants in the human erythrocyte proteome. *J. Proteome Res.* **2014**, 13 (4), 2028–44.
- (16) Pesciotta, E. N.; Sriswasdi, S.; Tang, H. Y.; Speicher, D. W.; Mason, P. J.; Bessler, M. Dysferlin and other non-red cell proteins accumulate in the red cell membrane of Diamond-Blackfan Anemia patients. *PLoS One* **2014**, 9 (1), e85504.
- (17) Basu, A.; Harper, S.; Pesciotta, E. N.; Speicher, K. D.; Chakrabarti, A.; Speicher, D. W. Proteome analysis of the triton-insoluble erythrocyte membrane skeleton. *J. Proteomics* **2015**, 128, 298–305.
- (18) Wisniewski, J. R.; Gaugaz, F. Z. Fast and sensitive total protein and Peptide assays for proteomic analysis. *Anal. Chem.* **2015**, 87 (8), 4110–6.
- (19) Wisniewski, J. R.; Mann, M. Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. *Anal. Chem.* **2012**, 84 (6), 2631–7.
- (20) Wisniewski, J. R. Quantitative Evaluation of Filter Aided Sample Preparation (FASP) and Multienzyme Digestion FASP Protocols. *Anal. Chem.* **2016**, 88 (10), 5438–43.
- (21) Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolome, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, 32 (3), 223–6.
- (22) Wisniewski, J. R.; Rakus, D. Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the *Escherichia coli* proteome. *J. Proteomics* **2014**, 109, 322–31.
- (23) McLaren, C. E.; Brittenham, G. M.; Hasselblad, V. Statistical and graphical evaluation of erythrocyte volume distributions. *Am. J. Physiol.* **1987**, 252 (4 Pt2), H857–66.
- (24) Brown, G. C. Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. *J. Theor. Biol.* **1991**, 153 (2), 195–203.
- (25) Wisniewski, J. R. Tools for phospho- and glycoproteomics of plasma membranes. *Amino Acids* **2011**, 41 (2), 223–33.
- (26) Meyer, J. G. In Silico Proteome Cleavage Reveals Iterative Digestion Strategy for High Sequence Coverage. *ISRN Comput. Biol.* **2014**, 2014, 1.
- (27) Burkhardt, J. M.; Vaudel, M.; Gambaryan, S.; Radau, S.; Walter, U.; Martens, L.; Geiger, J.; Sickmann, A.; Zahedi, R. P. The first comprehensive and quantitative analysis of human platelet protein composition allows the comparative analysis of structural and functional pathways. *Blood* **2012**, 120 (15), e73–82.
- (28) Gautier, E. F.; Ducamp, S.; Leduc, M.; Salnot, V.; Guillonnet, F.; Dussiot, M.; Hale, J.; Giarratana, M. C.; Raimbault, A.; Douay, L.; Lacombe, C.; Mohandas, N.; Verdier, F.; Zermati, Y.; Mayeux, P. Comprehensive Proteomic Analysis of Human Erythropoiesis. *Cell Rep.* **2016**, 16 (5), 1470–84.
- (29) Wisniewski, J. R.; Mann, M. A Proteomics Approach to the Protein Normalization Problem: Selection of Unvarying Proteins for

MS-Based Proteomics and Western Blotting. *J. Proteome Res.* **2016**, *15* (7), 2321–6.

(30) Barbour, R.; Kling, K.; Anderson, J. P.; Banducci, K.; Cole, T.; Diep, L.; Fox, M.; Goldstein, J. M.; Soriano, F.; Seubert, P.; Chilcote, T. J. Red blood cells are the major source of alpha-synuclein in blood. *Neurodegener. Dis.* **2008**, *5* (2), 55–9.

(31) Wisniewski, J. R.; Vildhede, A.; Noren, A.; Artursson, P. In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes. *J. Proteomics* **2016**, *136*, 234–47.

(32) Wisniewski, J. R.; Koepsell, H.; Gizak, A.; Rakus, D. Absolute protein quantification allows differentiation of cell-specific metabolic routes and functions. *Proteomics* **2015**, *15* (7), 1316–25.

(33) Burton, N. M.; Bruce, L. J. Modelling the structure of the red cell membrane. *Biochem. Cell Biol.* **2011**, *89* (2), 200–15.

(34) D'Alessandro, A.; Kriebardis, A. G.; Rinalducci, S.; Antonelou, M. H.; Hansen, K. C.; Papassideri, I. S.; Zolla, L. An update on red blood cell storage lesions, as gleaned through biochemistry and omics technologies. *Transfusion* **2015**, *55* (1), 205–19.

(35) Keightley, D. D.; Cressie, N. A. The Woolf plot is more reliable than the Scatchard plot in analysing data from hormone receptor assays. *J. Steroid Biochem.* **1980**, *13* (11), 1317–23.

(36) Dodge, J. T.; Mitchell, C.; Hanahan, D. J. The preparation and chemical characteristics of hemoglobin-free ghosts of human erythrocytes. *Arch. Biochem. Biophys.* **1963**, *100*, 119–30.

(37) Wisniewski, J. R.; Hein, M. Y.; Cox, J.; Mann, M. A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics* **2014**, *13* (12), 3497–506.

(38) Prenni, J. E.; Vidal, M.; Olver, C. S. Preliminary characterization of the murine membrane reticulocyte proteome. *Blood Cells, Mol. Dis.* **2012**, *49* (2), 74–82.

(39) Paterson, J. K.; Shukla, S.; Black, C. M.; Tachiwada, T.; Garfield, S.; Wincovitch, S.; Ernst, D. N.; Agadir, A.; Li, X.; Ambudkar, S. V.; Szakacs, G.; Akiyama, S.; Gottesman, M. M. Human ABCB6 localizes to both the outer mitochondrial membrane and the plasma membrane. *Biochemistry* **2007**, *46* (33), 9443–52.

(40) Bobrowska-Hagerstrand, M.; Wrobel, A.; Rychlik, B.; Ohman, I.; Hagerstrand, H. Flow cytometric monitoring of multidrug drug resistance protein 1 (MRP1/ABCC1) -mediated transport of 2',7'-bis-(3-carboxypropyl)-5-(and-6)- carboxyfluorescein (BCPCF) into human erythrocyte membrane inside-out vesicles. *Mol. Membr. Biol.* **2007**, *24* (5–6), 485–95.

(41) de Wolf, C. J.; Yamaguchi, H.; van der Heijden, I.; Wielinga, P. R.; Hundscheid, S. L.; Ono, N.; Scheffer, G. L.; de Haas, M.; Schuetz, J. D.; Wijnholds, J.; Borst, P. cGMP transport by vesicles from human and mouse erythrocytes. *FEBS J.* **2007**, *274* (2), 439–50.

(42) Boadu, E.; Sager, G. Reconstitution of ATP-dependent cGMP transport into proteoliposomes by membrane proteins from human erythrocytes. *Scand. J. Clin. Lab. Invest.* **2004**, *64* (1), 41–8.

(43) Wang, N.; Lan, D.; Gerbod-Giannone, M.; Linsel-Nitschke, P.; Jehle, A. W.; Chen, W.; Martinez, L. O.; Tall, A. R. ATP-binding cassette transporter A7 (ABCA7) binds apolipoprotein A-I and mediates cellular phospholipid but not cholesterol efflux. *J. Biol. Chem.* **2003**, *278* (44), 42906–12.

(44) Gallagher, P. G. Abnormalities of the erythrocyte membrane. *Pediatr. Clin. North Am.* **2013**, *60* (6), 1349–62.

(45) Ribeiro, M. L.; Alloisio, N.; Almeida, H.; Gomes, C.; Texier, P.; Lemos, C.; Mimoso, G.; Morle, L.; Bey-Cabet, F.; Rudigoz, R. C.; Delaunay, J.; Tamagnini, G. Severe hereditary spherocytosis and distal renal tubular acidosis associated with the total absence of band 3. *Blood* **2000**, *96* (4), 1602–4.

(46) Moriniere, M.; Ribeiro, L.; Dalla Venezia, N.; Deguillien, M.; Maillet, P.; Cynober, T.; Delhommeau, F.; Almeida, H.; Tamagnini, G.; Delaunay, J.; Baklouti, F. Elliptocytosis in patients with C-terminal domain mutations of protein 4.1 correlates with encoded messenger RNA levels rather than with alterations in primary protein structure. *Blood* **2000**, *95* (5), 1834–41.

(47) King, M. J.; Garcon, L.; Hoyer, J. D.; Iolascon, A.; Picard, V.; Stewart, G.; Bianchi, P.; Lee, S. H.; Zanella, A. International Council

for Standardization in, H., ICSH guidelines for the laboratory diagnosis of nonimmune hereditary red cell membrane disorders. *Int. J. Lab Hematol* **2015**, *37* (3), 304–25.

(48) Varady, G.; Szabo, E.; Feher, A.; Nemeth, A.; Zambo, B.; Pakaski, M.; Janka, Z.; Sarkadi, B. Alterations of membrane protein expression in red blood cells of Alzheimer's disease patients. *Alzheimers Dement (Amst)* **2015**, *1* (3), 334–8.

(49) Kasza, I.; Varady, G.; Andrikovics, H.; Koszarska, M.; Tordai, A.; Scheffer, G. L.; Nemeth, A.; Szakacs, G.; Sarkadi, B. Expression levels of the ABCG2 multidrug transporter in human erythrocytes correspond to pharmacologically relevant genetic variations. *PLoS One* **2012**, *7* (11), e48423.