CHAPTER 3

Red blood cells: proteomics, physiology and metabolism

Erica M. Pasini, Alan W. Thomas, Matthias Mann
1. Introduction
The discovery of the pulmonary circulation by Ibn al-Nafis and Servetus in the 13th century went largely unnoticed due to systematic destruction of Servetus' manuscripts, but the subsequent description of the systemic circulation by William Harvey was followed by the first successful blood transfusions between animals and attempted transfusions from animal to man, which were repeatedly fatal and which were ultimately banned at the end of the 17th century. Around the time that such transfusions were being banned the red blood cell (RBC) became a focus of scientific interest after being described almost simultaneously by the Dutch biologists Jan Swammerdam and Anton van Leeuwenhoek, made possible by the invention of the microscope (1).

Since those early days, developments in the RBC field, such as the establishment of safe transfusion technology after the discovery of A, B, 0 and Rh blood group antigens on the RBC membrane in the early 20th century, have been closely associated with the development of new technologies that have changed science and had a huge impact on the quality of life.

During the 20th century in-depth biochemical, immunological and molecular characterisation of single cellular proteins led to significantly increased understanding of RBC structure and function. At the end of the century new technologies emerged that allowed determination of the entire m-RNA (micro-array, transcriptomics) and protein (mass-spectrometry, proteomics) make-up of cells. Although in principle both transcriptomics and proteomics provide information about protein expression, micro-array measurements may be skewed where cellular m-RNA species are not translated to proteins. Furthermore, mature RBCs lack nuclei, organelles and transcription/translation apparatus so that transcriptomics cannot be applied, leaving proteomics as the method of choice for global RBC protein determinations (Figure 1).

2. Proteomics and the RBC
Proteomics, a general term describing mass-spectrometric protein expression studies (2), is best defined as a pipeline involving sample isolation/purification, sample preparation and optional fractionation, mass spectrometric analyses, database query and validation of the resulting protein list (Figure 2).

Isolation/purification. As mass-spectrometry (MS) has become increasingly sensitive, the initial steps of sample isolation and purification have become ever more critical to the reliability of data generated. This is because components of minor contaminating populations may be detected, leading to protein misattribution

DISORDERS OF ERYTHROPOIESIS, ERYTHROCYTES AND IRON METABOLISM
CHAPTER 3 - RBC proteomics
**Figure 1: ‘Omics approaches**

DNA  
Genomics  
Static  
Transcriptomics  
Dynamic  
m-RNA  
protein  
rRNA  
rRNA  
Ribosom  
mRNA  
tRNA  
Amino acid  
Proteomics

**Figure 2: Typical ESI-MS proteomics pipeline**

Blood sample (isolation)  
Sample fractionation  
Plasma (+EDTA)  
White blood cells, the “buffy coat”  
Red blood cells  
Sample digestion  
Purification steps  
Protein identification  
IPI00022361  
Database query  
Protein list validation  
Sample analysis  
200nL/min
(e.g. attribution of platelet proteins to the RBC). False positives can also arise from database misidentification. However, the use of high resolution, high accuracy data in combination with proper statistical techniques (e.g. decoy databases) has very effectively reduced the chances of misidentification.

For RBCs isolated for proteomic analysis, low level contaminants are most likely to derive from other blood cells and plasma. To date, isolation/purification has involved some or all of the following: repeated washes with isotonic medium (removal of plasma and platelets), elimination of the buffy coat (upper layer) at each washing step, the use of filters (leucocyte removal), density gradients e.g. percol, renografin-percol (granulocyte, reticulocyte removal) and customised nets (granulocyte removal) (Figure 3). Each method has advantages and disadvantages and the choice of isolation/purification procedures should be guided by the aims of the proteomics study.

**Preparation/fractionation.** Typically, to generate peptides suitable for MS, a purified sample is then either enzymatically cleaved directly in solution or is fractionated prior to enzymatic cleavage. Fractionation step(s) help reduce sample complexity, thereby facilitating MS, and can help in subsequent protein allocation as they provide

---

**Figure 3: Purification of the RBC sample**

RBC Purification step

- Whole blood collected from mouse passed on filters (Plasmodipur®)
- Lympholite centrifugation
- Renografin-Percol\(^1\) centrifugation
- RBCs after nylon nets\(^1\) passage
- Washed RBCs
- Aliquoted blood
- Plasma/serum
- WBC
- WBC & PMN
- Optional in human, necessary in mice
- Retics & Platelets
- PMN & Platelets

**Purity check by manual and automatic counting:**

Normally: 1 WBC every 1000 RBC
After purification: ~ 1WBC :1,000,000 RBC
~ 0-2 Reticulocytes :1,000,000 RBC
~ 0-5 Platelets :1,000,000 RBC

further information regarding the proteins from which peptides were derived. e.g. molecular weight (MW) in the case of SDS-PAGE, or location when using membrane specific regimes. For SDS/PAGE fractionation, gels are typically sliced at known MW boundaries prior to in-gel proteolytic digestion and subsequent peptide elution. In some cases fractionation steps may involve procedures isolating specific protein types such as membrane proteins, and as such may contribute both to purification and fractionation. RBC membrane and soluble fractions are both characterised by large numbers of proteins with a wide range of expression e.g. Band 3 is 2000 times more abundant than the complement receptor 1. To obtain the most sensitive outcomes, such complex cells require a reduction in complexity to levels compatible with the dynamic range of the MS instrument being used (Table 1).

The use of different extraction methodologies (e.g. using various solvents, detergents, ion-chelators and ionic solutions to extract proteins from RBC membranes) allows the detection of proteins with different biochemical characteristics. These methodologies can be applied independently to aliquots of the same sample prior to MS analysis, furnishing a very broad picture of the sample at hand. For the RBC membrane samples have been prepared by hypotonic buffer lysis - a procedure ensuring the preservation of RBC membrane characteristics - followed by extraction procedures including delipidation (e.g. EtOH), differential disruption of ionic bonds (using calcium carbonate) and disanchoring of the cytoskeleton (e.g, EDTA) (Table 2). Strategies to improve the detection of low abundance RBC membrane proteins by selective solubilisation of Band 3, a very abundant, highly hydrophobic membrane component, using fluorinated solvents have been recently proposed.

The MS ionisation method has often guided the choice between one and two dimensional (1D- and 2D) -SDS/PAGE for sample fractionation. 1D has traditionally been associated with electrospray ionisation (ESI) and 2D with matrix-assisted laser desorption ionisation (MALDI) as this requires a more extensive sample deconvolution due to the low specificity of peptide-mass fingerprinting.

<table>
<thead>
<tr>
<th>Table 1: Comparison of the characteristics of MS-analysers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mass analyser</strong></td>
</tr>
<tr>
<td>Quantity measured</td>
</tr>
<tr>
<td>Mass accuracy</td>
</tr>
<tr>
<td>Resolution at m/z 1000</td>
</tr>
<tr>
<td>Sensitivity</td>
</tr>
<tr>
<td>Dynamic range</td>
</tr>
</tbody>
</table>
**MS approaches.** MALDI has generally been coupled to time of flight (TOF)-analysers and ESI with ion trap, quadrupole and Fourier transform ion cyclotron (FT-MS) analysers. More recently, different peptide ionisation methods have been coupled with different detectors. New MALDI instruments coupled with TOF/TOF analysers or quadrupole ion traps have made protein identification of complex samples significantly easier and may in some cases have supplanted binomial 2D-MALDI-MS, but automation of MALDI-coupled liquid chromatography (LC) remains a challenge (2D-gel electrophoresis has limitations in reproducibility, dynamic range and depth of analysis).

In ESI-LC (Figure 4), peptides are analysed via a high-pressure liquid chromatography
HPLC) coupled mass-spectrometer, where the HPLC column provides peptide separation prior to injection into the MS. Each peptide analysed by the MS has a typical mass-to-charge ratio (m/z) which is registered by a detector. The output of the MS are m/z ratio (or peaks) spectra for each peptide.

**Database query and validation.** These spectra are used to query a sample-related database (e.g. the human protein database) generated by software that performs an in-silico (computer-generated) digest of all database proteins and generates the related peptide spectra. These spectra are compared with the experimental spectra by searching for matches, generating a score representing the quality of the match. The cut off point for scores can be set to ensure that the peptides/proteins are likely to be real. A manual validation step with appropriate restrictive parameters enhances the accuracy of the final protein lists.

**Recent developments.** In a pioneering study of the human RBC membrane using both 2-D and 1-D in gel-digestion prior to MS analysis, Low et al. (3) applied MALDI-TOF technology to identify 121 proteins. Using ESI-LC, Goodman et al. (2004) identified around 100 membrane and 100 soluble human RBC components using well-characterised membrane fractions such as inside-out vesicles and low salt extracts (4). A year later (5), a study aimed at developing new immobilised trypsin monolayers
for digestion and 2D nano HPLC separation of peptides to profile RBC proteins identified a total of 272 RBC soluble and membrane proteins. However, in this study, the RBC was a test-cell for this innovative methodology and most assignments were low confidence. In 2006 we combined biochemical RBC sub-fractionation with recent advances in MS-instrumentation and state of the art protein identification technology, used statistically significant numbers of MS-runs and undertook in-depth post-proteomic validation and analysis of the emerging protein lists to identify 340 membrane and 252 soluble human RBC proteins (6). RBC proteomics data have recently been summarised by Goodman and colleagues (7), who began to sketch an RBC protein interactome (the whole set of molecular interactions in the cell). Here we put RBC proteomics findings in the context of RBC protein literature, choosing examples to highlight discrepancies, analogies and unexpected findings.

3. The RBC membrane

3.1 Physiology and proteomics
The RBC membrane is critical to maintenance of the characteristic biconcave shape of the RBC; it ensures appropriate pH and cation concentration differentials between the RBC and the plasma (low potassium, high sodium and calcium), an adequate area to volume ratio, fluidity and osmolarity (8). Deformability, a key attribute requiring rapid modification of the RBC membrane in response to environmental variables such as blood vessel diameter, depends on all the above parameters. These are in turn a function of interaction between RBC membrane lipids, proteins and the underlying cytoskeletal network. Fluidity and deformability are also a function of the differential charge between the two membrane leaflets (the outer is neutral, the inner negatively charged), which is created and maintained by different phospholipid compartmentalisation in the two leaflets (9). In contrast to proteomics, lipidomics is still in its infancy and faces a number of technological challenges. A first attempt to characterise phospholipid exchange by lipidomics used embryonic fibroblasts (10). In future this approach may help unravel poorly abundant membrane lipids with important physiological functions, disclose the secrets of raft formation, answer questions on preferential protein-lipid interactions and promote a more in-depth understanding of phospholipid exchange and repair.
The mouse RBC membrane contains fewer integral and membrane-associated proteins than the human RBC membrane (11). However, as would be expected, prevailing functions annotated for both membranes include binding; transport; signal transduction; catalytic, structural and antioxidant activity.
In terms of the cytoskeletal network, the presence of all major proteins known to
have a role in its formation and plasticity have been confirmed by proteomics. In addition, proteomics pointed to the presence of a number of low-abundance cytoskeletal proteins (myosin, moesin, ezrin, radixin and F-actin capping protein), the physiological function of which remains to be determined (6). The plasticity of the cytoskeletal network requires ATP as the primary energy source, calcium extrusion and the formation of calcium-calmodulin complexes. A variety of membrane-associated enzymes, including several kinases highlighted by proteomics (protein kinase A, protein kinase C, cdc kinase, casein kinase 1) are thought to regulate interactions within the network through induction of specific post-translational modifications (PTMs) (e.g. phosphorylation, methylation, myristylation, palmitylation, or farnesylation) (6). Future proteomic approaches may contribute to better understanding of the regulation of the cytoskeletal network by identifying and characterising novel PTM's.

The RBC membrane is to some extent permeable to water and anions but extremely impermeable to cations, nucleotides (ATP, GTP) and glycolytic intermediates. RBC transporters have important functions such as:

- nutrient import (e.g. glucose, the most important RBC energy source enters through three independent transporters: GLUT1, GLUT3, GLUT4);
- RBC response to sudden osmolarity changes (aquaporin-1 and 3, probably the urea transporter UT-B),
- regulation of RBC volume (pumps, channels) and
- maintenance of the ion balance (chiefly Band 3, the Na⁺-K⁺-ATPase and the calmodulin activated Mg²⁺-dependent Ca²⁺-ATPase) (6, 7).

Band 3 is the most important RBC anion transporter able to rapidly exchange bicarbonate and chloride anions and slowly exchange a number of larger anions (e.g. sulfate, phosphate, and superoxide).

Proteomics evidence underpins the presence of all these transporters with the exception of the highly lipophylic Aquaporin-3 pore. Two cation transporters were first identified by proteomics (splice isoform 4 of the Q04656 Copper-transporting ATPase 1 and the zinc transporter 1), while two different isoforms of calcium transporting ATPase, a ubiquitous form (isoform 1 or B) and a rare one (isoform 4 or XD) were found (6, 7). The presence of two transporters previously ascribed to RBCs (Na⁺-H⁺-exchanger and L-type amino acid transporter 1) was not confirmed in the RBC proteome.

However, in addition to the cation transporters mentioned above, proteomics has suggested that other novel transporters may be present in the RBC membrane, including:

- the low abundant solute carrier family 19 member 1 (likely to be a reduced-folate carrier but occasionally also an ion and sugar transporter);
• the hydrogen-lactate co-transporter (monocarboxylate transporter 1);
• the hypothetical ion transporter protein FLJ23931;
• a fatty acid transporter (solute carrier family 27 member 4);
• the low abundance voltage insensitive, pH-dependent potassium channel, quinine sensitive potassium channel subfamily K member 5, which may be involved in expelling potassium ions, a role traditionally ascribed exclusively to the Gardos channel;
• two large neutral amino acid carriers (solute carrier family 43, member 1 and 2) and
• a glutamate specific transporter (Solute carrier family 1, member 7), previously described only in dog erythrocytes.
These findings warrant further investigation as in several cases preliminary evidence (e.g. with monoclonal antibodies) suggests that some may represent genuine RBC membrane components. Other well known passive transporters (e.g. the Na⁺-K⁺-Cl⁻ and the K⁺-Cl⁻-co-transporters), the nucleoside transporter and the XK protein, believed to have a role in amino acid and oligo-peptide transport were also found (6, 7).

3.2 The RBC membrane and its environment
RBCs are highly dynamic blood components, evident in the multi-faceted interplay with other blood cells (e.g. rosetting with white blood cells, platelets, P. falciparum-infected RBCs), with endothelial cells (transfer of GPI-anchored proteins) and with plasma proteins (e.g. lipoproteins, albumin, immunoglobulins, complement factors). While the two GPI-anchored proteins preventing intravascular haemolysis by homologous complement attack (decay accelerating factor or DAF) and assembly of the membrane attack complex (protectin or CD59) are well-known and characterised, the transfer of GPI-anchored proteins from and to RBC membranes (12) is less well understood. Proteomic analysis of the human RBC has provided independent supporting evidence for the existence and association with the RBC membrane of high molecular covalent C3b2-IgG complexes. To date, such complexes have not been found at the surface of animal RBC proteomes, possibly indicative of a prevalence of the non-classical versus classical complement pathway in humans (13). In future, the combination of interactomics and bioinformatics with open-access protein databases (e.g. http://www.plasmaproteomedatabase.org) has the potential to increase understanding of the complex network of RBC interactions with other cells.

4. The RBC soluble fraction
Proteins in both human (6) (Figure 5) and mouse (11) soluble proteomes are predominantly implicated in cellular metabolism and/or transport; prevalent
functions include proteolysis, ATPase and dehydrogenase activity. Transport processes mainly involve ion and protein transport or the regulation thereof. Haemoglobin remains the sole gas transporter identified, and there are 3 different hydrogen transporters. The RBC is able to metabolise a diverse range of compounds e.g. amines, amino acids and their derivatives, aromatic compounds, lipids, co-factors, nucleic and organic acids, phosphorous, sulfur and vitamins. Most proteins identified by proteomics as being involved in metabolic processes are catabolic, perhaps not surprising considering that RBCs are dynamic blood components that continue a process of predominantly destructive development initiated well before they enter the blood-stream. A number of proteins and enzymes are involved in protection from oxidative stress (e.g. GSH, superoxide dismutase, glutathione reductase, catalase), household (e.g. nucleotide (ATP/ADP), glucose) and macromolecule metabolism. The complex with the greatest number of members was the proteasome (17 proteins). Other complexes identified included a number of RBC household complexes (ubiquitin ligase complex, 6-fructokinase complex, Hb complex, and phosphopyruvate hydratase complex) (6, 7).

5. Comparative RBC proteomics
Due to the importance of animal models, particularly the mouse, we recently undertook an innovative comparative proteomics approach in which samples were run in rapid succession under controlled, highly comparable conditions (11). Whilst the human RBC proteome comprised 340 RBC membrane proteins and 252 soluble RBC proteins, the mouse RBC proteome comprised 247 membrane only, 232 soluble only and 165 proteins that were both membrane bound and soluble. The need for a novel category of membrane bound/soluble proteins in the mouse reflected a
comparatively large proportion of proteins present in both fractions. In human RBC proteins present in both compartments were limited to the most abundant proteins and to proteins of the cytoskeletal network. Categorisation of both proteomes in terms of sub-cellular localisation, protein family, splice isoforms and function demonstrated the overall similarity between human and mouse RBCs, but highlighted specific differences at the level of cytoskeletal proteins (Table 3) and glycolytic enzymes (Table 4). Thus, although similar profiles emerged for cytoskeletal proteins and their isoforms there were important exceptions: for example, tropomyosin 2, actin cytoplasmic 2 (γ-actin) and adducin-γ appear to

<table>
<thead>
<tr>
<th>Protein family</th>
<th>Unique to mouse</th>
<th>Unique to human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropomyosin</td>
<td>Tropomyosin 2</td>
<td></td>
</tr>
<tr>
<td>Actin</td>
<td>γ-actin</td>
<td></td>
</tr>
<tr>
<td>Adducins</td>
<td>γ-adducin</td>
<td>Tubulin α-3</td>
</tr>
<tr>
<td>Tubulins</td>
<td></td>
<td>Common to human and mouse</td>
</tr>
<tr>
<td>Spectrins</td>
<td>BetaIV-spectrin sigma1</td>
<td>Spectrin β chain, brain (betaIV-spectrin)</td>
</tr>
<tr>
<td></td>
<td>Spectrin α chain, RBC</td>
<td>Spectrin α chain, RBC</td>
</tr>
<tr>
<td></td>
<td>Spectrin β chain, RBC</td>
<td>Spectrin β chain, RBC</td>
</tr>
<tr>
<td></td>
<td>Spectrin brain 1</td>
<td>Spectrin β chain, brain 1</td>
</tr>
<tr>
<td>Ankyrins</td>
<td>Ankyrin 1 (isoform 1)</td>
<td>Ankyrin 1 isoform 1</td>
</tr>
<tr>
<td></td>
<td>Ank1 isoform 4, erythrocyte</td>
<td>Ankyrin 1 isoform 4</td>
</tr>
<tr>
<td></td>
<td>Ank variant 2.2, erythrocyte</td>
<td>Alt ankyrin (Var. 2.2)</td>
</tr>
<tr>
<td></td>
<td>Ankyrin 3</td>
<td>Ankyrin 3</td>
</tr>
<tr>
<td>Tropomyosins</td>
<td>Tropomyosin α 4 chain</td>
<td>Isoform 1 Tropomyosin α 4 chain</td>
</tr>
<tr>
<td></td>
<td>Isoform 1 Tropomyosin 1 α chain</td>
<td>Isoform 1 Tropomyosin 1 α chain</td>
</tr>
<tr>
<td></td>
<td>Isoform 2 Tropomyosin α 3 chain</td>
<td>Tropomyosin α 3 chain</td>
</tr>
<tr>
<td>Actins</td>
<td>Actin, α skeletal muscle</td>
<td>Actin, skeletal muscle (α-actin1)</td>
</tr>
<tr>
<td></td>
<td>Actin, cytoplasmic 1</td>
<td>Actin, cytoplasmic 1 (β-actin1)</td>
</tr>
<tr>
<td>Adducins</td>
<td>α-adducin</td>
<td>α-adducin</td>
</tr>
<tr>
<td></td>
<td>β-adducin</td>
<td>β-adducin</td>
</tr>
<tr>
<td>Tubulins</td>
<td>Tubulin α-1</td>
<td>Tubulin α-6/α-1</td>
</tr>
<tr>
<td></td>
<td>Tubulin β-1</td>
<td>Tubulin β-5/β-1</td>
</tr>
</tbody>
</table>
be unique to mouse RBCs, whereas tubulin α-3 was only detected in the human proteome. Review of the literature indicates that expression by human RBCs of those cytoskeletal proteins unique to the mouse proteome has been linked to abnormalities such as shape alteration (tropomyosin 2), reduction of RBC deformability (γ-actin) and spherocytic hereditary elliptocytosis (adducin-γ).

In terms of glycolytic enzymes, mouse and human RBC differ in localisation. While all such enzymes appear to be either membrane or cytoskeletally bound in human RBCs, most are found in the soluble fraction of mouse RBCs. In this context, it has been suggested that fewer inward facing membrane-associated enzymes are needed
in the mouse to carry the membrane structure to the cell interior than in human (reference can be found in (11)).

There were a number of unexpected proteins present in both the mouse and human RBC proteomes. These were mainly common between the species, giving an extra degree of confidence in the findings. Intriguingly, these proteins generally showed an anomalous but consistent in-gel migration pattern (Figure 6). Proteins were considered to migrate anomalously when the expected and the experimental molecular weights were at variance. This opened two scenarios for the metabolic status of a protein: 1) a protein migrated with lower apparent MW than expected and was classified as degraded; 2) proteins co-migrated with ubiquitins at higher apparent MW than expected and were considered incompletely proteolysed by the proteasome. In both cases these proteins are likely to be non-functional reticulocyte legacies (e.g. 78-kDa glucose regulatory protein, 71-kDa heat shock cognate). In some cases unexpected proteins (e.g. 40S ribosomal proteins S3 and S6) migrated at their expected molecular weight in both mammals, opening the possibility that they may have roles in the mature RBC (Table 5).

The number of reticulocyte legacy proteins found in the cytoplasm of the mouse mature RBCs was higher than in humans. This is likely to be due to the different life spans of RBCs in the mouse (60 days) and human (120 days), leading to a RBC higher turnover in mice, thereby increasing the proportion of younger mature RBCs, which contain higher levels of reticulocyte legacy proteins which have not been fully metabolised. Taken together, these aspects provide evidence for the hypothesis that reticulocyte/RBC maturation is an on-going processes that covers the whole RBC life span and which may have a schedule shared between mammals for metabolic degradation of legacy-proteins.

Of functional interest, even when direct human/mouse protein orthologs were missing, an in-depth data analysis often uncovered functional orthologs belonging to either related or unrelated protein families.

6. RBC proteomics: limitations and potential

Proteomic analyses are complicated by factors that include large ranges in expression levels; post-translational modifications that affect peptide molecular weight; occurrence of closely related isoforms, absence of appropriate database descriptors for a specific protein and the ubiquitous presence of haemoglobin. In terms of large ranges in expression it is noteworthy that protein band 3 occurs at one million copies per cell, comprising 30% of the membrane proteome, and spectrin tetramer occurs at 100,000 copies per cell, comprising 75% of the cytoskeleton. Our recent experience suggests that use of the latest ESI-MS technology coupled to membrane
Figure 6: Most human and mouse soluble orthologs show the same in-gel behaviour
extraction methodologies and repeated runs allows the confident detection of proteins at levels down to 500 copies per cell (e.g. CR1). The high concentration of haemoglobin (97% of soluble RBC protein dry mass and 35% of the total mass) and its amphipatic nature (possessing both hydrophilic and lipophilic properties) present challenges to protein detection both at the level of the membrane and of the soluble fraction. During RBC lysis haemoglobin tends to associate strongly with the membrane, hampering the detection of low abundance proteins. To reduce this effect manipulations can be performed at low temperature (≤ 4°C). Currently chromatographic methods that allow elimination of most of the haemoglobin are favoured, allowing better detection of soluble, low abundance proteins. It remains unclear how selective these methods are, and whether a number of other proteins which strongly associate with haemoglobin are lost by their application.

### Table 5: Comparison of the in-gel behaviour of human and mouse membrane orthologs

<table>
<thead>
<tr>
<th>Accession N.</th>
<th>Orthologs human/mouse soluble fraction</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPI00134135.1</td>
<td>Acid phosphatase 1, soluble</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00462072.2</td>
<td>Alpha enolase</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI0010257.1</td>
<td>Alpha-haemoglobin stabilising protein</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI0011359.1</td>
<td>Glutathione reductase</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00283531.4</td>
<td>Glutathione S-transferase</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00127691.1</td>
<td>Glutathione synthetase</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00315554.2</td>
<td>E2 ubiquitin conjugating enzyme</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00110042.1</td>
<td>Selenium-binding protein 1</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00008380</td>
<td>Serine/threonine phosphatase 2A, 65 kDa regulatory subunit A, alpha isoform</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI0039953.1</td>
<td>Serine/threonine-protein kinase WNK1</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00113408.1</td>
<td>Ribose 5-phosphate isomerase</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00113655.1</td>
<td>40S ribosomal protein S6</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00134599.1</td>
<td>40S ribosomal protein S3</td>
<td>at MW</td>
</tr>
<tr>
<td>IPI00121427.1</td>
<td>Calcyclin</td>
<td>NOT at MW ~43kDa</td>
</tr>
<tr>
<td>IPI00110049.1</td>
<td>Exportin 7</td>
<td>NOT at MW ~80kDa</td>
</tr>
<tr>
<td>IPI00129928.2</td>
<td>Fumarate hydratase, mitochondrial precursor</td>
<td>NOT at MW &gt;80kDa</td>
</tr>
<tr>
<td>IPI0031444.4</td>
<td>Importin 7</td>
<td>NOT at MW &gt;188kDa</td>
</tr>
<tr>
<td>IPI00230429.3</td>
<td>Importin beta 3 subunit</td>
<td>NOT at MW &gt;188kDa</td>
</tr>
<tr>
<td>IPI00316740.2</td>
<td>damage specific DNA binding protein 1</td>
<td>NOT at MW &gt;188kDa</td>
</tr>
</tbody>
</table>
Distinguishing between protein isoforms and highly homologous family members is complex (Figure 7). They may be characterised by long identical and short highly specific stretches (e.g. different glucose transporters); differ only by single point mutations (MN antigens displayed by GPA and GPB; the Gerbich blood antigen on GPC and GPD) or be short/long versions of the same protein (Glycophorins C and D). While it is possible to apply targeted MS-software based approaches (e.g. inclusion and exclusion lists) when stretches highly specific to an isoform or family member are present, traditional MS analysis has limited potential when proteins differ only by single point mutations and offers no possibility of distinguishing short protein variants. Glycomics studies (14) that differentiate between proteins based on their characteristic sugars (e.g. GPD carries only O-linked sugar chains, while GPC also carries one N-linked sugar chains in position 8) may become extremely useful adjuncts to proteomics studies as well as being informative in their own right. Thus, glycomics will contribute to study of the complex RBC carbohydrate heterogeneity that gives rise to a much blood group antigen diversity and immunology.

Interactomics and bioinformatics-based domain analysis will, on the other hand, help move traditional, identification-based proteomics into the domain of providing functional insight, for example by providing maps of the interactions between adjacent membrane proteins that create specific micro-domains and epitopes with important functional implications (e.g. interaction between domains on Band 3 and GPA resulting in the Wright (Wr\(^b\)) blood type).

Perhaps even more exciting and informative will be studies that help unravel the dynamic relationship between the RBC and its environment. Information continues to emerge suggesting that the RBC may mirror systemic diseases such as schizophrenia (15), potentially allowing for early diagnosis and serving as surrogate for treatment investigations. Following a similar line of reasoning proteomic studies of RBC at the epidemiological level have recently been proposed, using large-scale population-based cohorts (16). By monitoring RBC membrane changes over time, these studies offer a chance of uncovering risk factors and RBC markers for cancer, metabolic syndromes, cardiovascular and other diseases.

Thus, while identification-based proteomics has limitations, and while it is not clear whether the levels of sensitivity already achieved can usefully be improved upon, related technologies such as interactomics, lipidomics, glycomics and population proteomics are rapidly coming on line. Coupled with bioinformatics and systems biology approaches, these have the potential to shed light on the systemic influence of the RBC and influences upon the RBC as it travels through the body and interacts with an ever changing environment.
Figure 7: Different types of isoforms

A. Identical isoforms: Short versus long form of the same protein. B. Isoforms differing only by a short stretch of amino acids. C. Isoforms differing only by single or multiple point mutation(s)
Acknowledgements
The authors wish to thank Dr. Hans Lutz from the ETH Zürich for helpful discussion. This research was supported by the Nederlands Organisation for Scientific Research (NWO Genomics, grant number 050-10-053).

References
Multiple Choice Questionnaire

To find the correct answer, go to http://www.esh.org/iron-handbook2009answers.htm

1. The m/z ratio is the ratio of:
   a) Mass/charge
   b) Methionine/other amino acids
   c) Membrane/cytoplasmic proteins
   d) Mass spectrometry/HPLC

2. Why are sample isolation and purification critical to proteomics?
   a) Because it is the first step
   b) Because components of minor contaminating populations may be detected
   c) Because mass spectrometry has become less accurate
   d) Because they are very complex

3. Which cytoskeletal protein is absent in human normocytes but present in mouse normocytes?
   a) Tropomodulin
   b) Ankyrin-1-isoform 4
   c) γ-adducin
   d) Tubulin-3

4. Band 3 represents a problem for the detection of other RBC membrane proteins, because:
   a) It forms complexes
   b) It is very abundant compared to other proteins
   c) It migrates in different gel bands
   d) It is highly amphipathic

5. What are reticulocyte legacy proteins?
   a) Proteins that are unique to reticulocytes
   b) Proteins that are unique to mature RBCs
   c) Inactive proteins residual of reticulocytes still present in mature RBCs
   d) Contaminants