Wilson, M., Trakarnsanga, K., Heesom, K., Toye, A., & Frayne, J. (2016). Comparison of the proteome of adult and cord erythroid cells, and changes in the proteome following reticulocyte maturation. Molecular and Cellular Proteomics, 15(6), 1938-1946. DOI: 10.1074/mcp.M115.057315

Comparison of the proteome of adult and cord erythroid cells, and changes in the proteome following reticulocyte maturation.

Marieangela C. Wilson[1], Kongtana Trakarnsanga[1,2], Kate J. Heesom[1], Nicola Cogan[3,4], Carole Green[3,4], Ashley M. Toye[1,3,4], Steve F. Parsons[3], David J. Anstee[3,4] and Jan Frayne[1,4].

[1]School of Biochemistry, University of Bristol, United Kingdom; [2]Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand; [3]Bristol Institute for Transfusion Sciences, National Health Service Blood and Transplant (NHSBT), Filton, Bristol, United Kingdom. [4]NIHR Blood and Transplant Research Unit, University of Bristol.

Running Title: Human RBC comparative proteomics

Corresponding author:

Jan Frayne, School of Biochemistry, University of Bristol, Bristol, BS8 1TD, United Kingdom.

E-mail: jan.frayne@bristol.ac.uk

Telephone +44 117 3312127

Fax: +44 117 331216

Summary

Cord blood stem cells are an attractive starting source for the production of red blood cells *in vitro* for therapy because of additional expansion potential compared to adult peripheral blood progenitors, and cord blood banks usually being more representative of national populations than blood donors. Consequently it is important to establish how similar cord RBCs are to adult cells. In this study we used Multiplex Tandem Mass Tag labeling combined with nanoLC-MS/MS to compare the proteome of adult and cord RBCs and reticulocytes. 2838 unique proteins were identified, providing the most comprehensive compendium of RBC proteins to date. Using stringent criteria 1674 proteins were quantified, and only a small number differed in amount between adult and cord RBC. We focussed on proteins critical for RBC function. Of these only the expected differences in globin subunits, along with higher levels of carbonic anhydrase 1 & 2 and aquaporin-1 in adult RBCs would be expected to have a phenotypic effect since they are associated with the differences in gaseous exchange between adults and neonates. Since the RBC and reticulocyte samples used were autologous, we catalog the change in proteome following reticulocyte maturation. The majority of proteins (>60% of the 1671 quantified) reduced in abundance between 2 and 100-fold following maturation. However, ~5% were at a higher level in RBCs, localised almost exclusively to cell membranes, in keeping with the known clearance of intracellular recycling pools during reticulocyte maturation. Overall, these data suggest that with respect to the proteome there is no barrier to the use of cord progenitors for the *in vitro* generation of RBCs for transfusion to adults other than the expression of fetal not adult haemoglobin.

Introduction

The generation of human red blood cells (RBCs) *in vitro* for transfusion purposes is a major goal of health services globally. In recent years advances in the development of systems for the generation of erythrocytes *in vitro* have progressed rapidly using progenitor cells isolated from adult peripheral blood (PB)(1) umbilical cord blood(2,3) and human pluripotent stem cells.(4-6) Cells from adult PB and cord can be expanded and induced to differentiate efficiently down the erythroid pathway to generate significant numbers of enucleated reticulocytes as an end point.(1) However, adult PB progenitors have a more limited proliferative capacity than cord, which restricts the number of red cells that can be obtained by *in vitro* culture methods and greatly impacts the economic viability of producing therapeutic quantities of red cells from this source. Therefore, progenitors isolated from cord are attractive as a starting material for *in vitro* blood production because of their potential for greater expansion capacity.(3,7) In addition cord stem cell banks are generally more representative of blood group diversity in the population compared to adult donor blood banks.

Although cord progenitors offer a realistic potential for generating therapeutic quantities of erythroid cells, these cells appear to maintain a fetal, rather than adult phenotype. The most obvious difference between erythroblasts generated from cord compared to adult PB is their expression of predominately γ- (fetal) rather than β- (adult) globin. Other differences have been reported,(8,9) including expression of i rather than I antigen, and weak expression of ABH antigens on cord cells. However, to date a comprehensive, comparative analysis of the proteome of cord cells compared with adult cells has not been undertaken, although such information is essential before such cells can be considered for therapeutic use.

Another poorly defined feature of erythropoiesis is the change in proteome as reticulocytes undergo extensive transformation to create mature erythrocytes. Alterations in membrane proteins have been observed to occur during enucleation,(10) and the role of autophagic vesicles in reticulocyte maturation has recently been described.(1,11) However, a more comprehensive understanding of the final stages of reticulocyte maturation is now required in order to facilitate the study of inherited and acquired anaemias exhibiting reticulocytosis,(12) and aid in extending *in vitro* culture systems to create mature erythrocytes.

We have undertaken multiplex Tandem Mass Tag (TMT) labelling and nanoLC-MS/MS to compare the proteome of adult and cord endogenous RBCs, and also of reticulocytes generated *in vitro* from adult PB and cord progenitors. In addition, as our culture system generates functional, mature reticulocytes(1) we were able to compare the proteome of these cells with that of the original donors own mature RBCs, mimicking as close as possible the *in vivo* maturation process.

Materials and Methods

Isolation of adult and Cord Blood RBCs and $CD34^+$ cells

Leukocyte Reduction System (LRS) cones and CB units were obtained from healthy donors with written informed consent for research use in accordance with the Declaration of Helsinki and approved by local Research Ethics Committees (Southmead Research Ethics Committee reference 08/H0102/26 and Bristol Research Ethics Committee reference 12/SW/0199). Adult RBCs and $CD34^+$ cells, and CB RBCs and $CD34^+$ cells were isolated from the same Leukocyte Reduction System cones and cord blood units respectively. $CD34^+$ cells were separated as described by Griffiths et al(1).

Erythroid differentiation of $CD34^+$ cells

$CD34^+$ cells were differentiated in our three-stage culture system,(1) harvested on day 19 and passed through a PALL WBF leukocyte filter(1) to isolate mature reticulocytes.

Preparation of membrane and cytosol fractions

The RBCs and reticulocytes were separated into membrane and cytoplasmic fractions to reduce protein complexity for Mass Spectrometry (MS). Cells were washed twice in cold PBS, followed by 2 washes with iso-osmotic buffer (Na_2HPO_4 103 mM, NaH_2PO_4 155 mM pH 7.4). The cells were then lysed in lysis buffer (1 in 20 dilution of iso-osmotic buffer with water, 1x complete protease inhibitor [Roche] and 0.5 mM PMSF [Sigma-Aldrich]) at 50x packed cell volume, followed by centrifugation at 15,000 rpm for 10 minutes at 4°C. The first supernatant was collected and kept as a cytoplasmic fraction. The lysis step was repeated an identical number of times for the adult and cord cells until the membrane pellet of both become white. The pellets were solubilized in sodium phosphate buffer with 1x complete protease inhibitor and 0.5mM PMSF.

SDS-PAGE and Western blot

Washed cell pellets were re-suspended in solubilisation buffer (20 mM Tris HCL pH 7.5, 150 mM NaCl, 10% glycerol, 1% Triton, 0.1% SDS) containing 1x complete protease inhibitor and 2 mM PMSF at 200 µl of buffer to $1x10^7$ cells. After 1 hour incubation on ice the protein samples were treated with 25 Uml^{-1} Bensonase (Sigma-Aldrich) for 1 hour on ice and then centrifuged at 17,000 g for 5 minutes, at 4°C. Cytoplasmic and nuclear fractions were prepared in the same way as samples for mass spectrometry (above). Proteins were resolved by SDS-PAGE and transferred to PVDF by Western blot. Membranes were blocked with 10% milk powder for 1 hour, followed by incubation with primary antibodies overnight at

4°C. Primary antibodies used were Aquaporin I (CHIP28) and Glycophorin A (CVDP) 1:1000 dilution, Ankyrin 1 (BRIC274) 1:100 dilution and Band 3 (BRIC170) 1:2000 dilution all validated and supplied by IBGRL (http://ibgrl.blood.co.uk/ResearchProducts/ResProdHome.htm). BCL11A[14B5] (ab19487) and CAII (ab6621) 1:1000 dilution, CAI (ab34978) and Myosin IIB (ab684) 1:2000 dilution, Tropomodulin 4 (ab67776) 1:500 dilution, all from Abcam; β–globin (37-8; sc21757) 1:5000 dilution, γ-globin (51-7; sc21756) and α-globin (D-16; sc31110) 1:1000 dilution all from Santa Cruz Biotechnology; β–actin (A1978;Sigma) 1:2000 dilution. Specificity of antibodies has been previously demonstrated(13-15); Abcam plc.

Lentiviral constructs and transduction of erythroblasts

pLKO.1-TRC BCL11A short hairpin (sh) RNA plasmids (B1 and B5) were designed by the Broad Institute and purchased from Open Biosystems. The sequence of hairpin B1 is 5'-CCGG-CGCACAGAACACTCATGGATT-CTCGAG-AATCCATGAGTGTTCTGTGCG-TTTTTG-3', targeting nucleotides 792-812 and the sequence of B5 is 5'-CCGG-GCATAGACGATGGCACTGTTA-CTCGAG-TAACAGTGCCATCGTCTATGC-TTTTTG-3', targeting nucleotides 2015-35 of BCL11A-XL (accession number NM_022893.3). Transduction of erythroblasts was performed as described by Satchwell *et al* and Trakarnsanga *et al*.(15,16)

Polymerase Chain Reaction (PCR)

RNA (400 ng) was reverse transcribed into cDNA using SuperScript II reverse transcriptase (Invitrogen). Globin expression was analysed using primers to β-globin 5' CTTTAGTGATGGCCTGGCTC and 5' GGCAGAATCCAGATGCTCAA, and γ-globin 5'

GGGCAAGGTGAATGTGGAAGAT and 5' GGGTCCATGGGTAGACAACCA as described previously.(15)

Tandem Mass Tag labelling, preparation of samples for mass spectrometry, database search parameters and acceptance criteria for identifications

For multiplexed comparative proteomics 100 μg of each cell lysate was digested with trypsin and labeled with Tandem Mass Tag (TMT) reagents according to the manufacturer's protocol (Thermo Fisher Scientific). After labeling, samples were combined in equal amounts, and 50 μg of pooled sample fractionated by strong cation exchange using an Ettan LC system (GE Healthcare) prior to analysis by nanoLC-MS/MS. The raw data files were processed and quantified using Proteome Discoverer software v1.2 and v1.4 (Thermo Scientific) and searched against the UniProt/SwissProt Human database release version 57.3 (20326 entries) using the SEQUEST (Ver. 28 Rev. 13) algorithm. Peptide precursor mass tolerance was set at 10 ppm, and MS/MS tolerance was set at 0.8 Da. Search criteria included oxidation of methionine (+15.9949) as a variable modification and carbamidomethylation of cysteine (+57.0214) and the addition of isobaric mass tags (+229.163) to peptide N-termini and lysine as fixed modifications. Searches were performed with full tryptic digestion and a maximum of 1 missed cleavage was allowed. The reverse database search option was enabled and all peptide data was filtered to satisfy false discovery rate (FDR) of 5%. The Proteome Discoverer software generates a reverse "decoy" database from the same protein database and any peptides passing the initial filtering parameters that were derived from this decoy database are defined as false positive identifications. The minimum cross-correlation factor (Xcorr) filter was readjusted for each individual charge state separately to optimally meet the predetermined target FDR of 5% based on the number of random false positive matches from the reverse decoy database. Thus each data set has its own passing parameters. Quantitation

was performed using a peak integration window tolerance of 0.0075 Da with the integration method set as the most confident centroid. Protein ratios represent the median of the raw measured peptide ratios for each protein. These methods are described in Trakarnsanga *et al*.(5) The mass spectrometry (MS) proteomics data have been deposited to the ProteomeXchange Consortium(17) via the PRIDE partner repository with the dataset identifier PXD003276.

For analysis only rank 1 peptides were used, and only quantifications obtained using 2 or more unique peptides with high/medium confidence per protein were considered (although for additional information proteins quantified from 1 unique peptide are sometimes included). Proteins recorded as uncharacterized by the software were returned with a gene ID. We selected a comparative protein threshold of 2, such that proteins that differed in level by 2.00 fold or more between adult and cord samples were considered differentially expressed. For classification by cellular component and molecular function proteins were analysed using WebGestalt GSAT V2.

Experimental design and Rationale

The human samples for this study were provided as a pool of autologous $CD34^+$ cells and RBCs from 16 adults, and a pool of autologous $CD34^+$ cells and RBCs from 4 cord samples, minimising intrinsic variability between individuals. To ensure robustness of our comparative data the level of proteins in the pooled adult sample were further compared to an additional cord RBC and reticulocyte sample, and only proteins that *also* differed between the adult and these second cord samples are reported as differentially expressed. For comparative analysis between adult and cord erythroid cells internal controls included proteins known to be consistent in level, eg. α-globin, and to differ in level, eg. β- and γ-globin, between these cells, as described. For comparative analysis between reticulocytes and mature RBCs, internal

controls included proteins known to change in level following reticulocyte maturation, as described.

Results and discussion

Membrane and cytosol fractions of the adult and cord RBC and reticulocytes were each labelled with a different TMT, pooled and analysed by MS. Using this methodology, 1880 proteins were identified in the membrane fractions, and 1602 in the cytosolic fractions (Supplemental Tables 1A & B, 2A & B). Certain proteins were detected in both fraction hence overall we identified 2838 unique proteins, which represents the most comprehensive identification of RBC proteins to date. No membrane or cytosol protein detected was unique to just adult or cord, or just RBCs or reticulocytes.

We detected 26 RBC blood group protein antigens in these data (Table 1). As stated previously(5), Duffy is a large glycosylated protein with few trypsin sites resulting in a large peptide that is not compatible with MS identification. We have previously shown that our culture system does not alter the blood group protein expression profile of cultured adult erythroid cells.(1) In the present study we further compared the abundance of the RBC blood group proteins between the endogenous RBCs and autologous cultured cells. The levels were either consistent or lower in the cultured cells (Table 1), reassuring that the cultured cells are not more immunogenic than endogenous cells. We performed the same analysis for RBCs and autologous cultured cells from a single cord blood donor, confirming abundance of the RBC blood group proteins was either consistent or lower in the cultured cells (Supplemental Table 3). We also performed serological analysis of cord blood cells, again comparing the expression of blood group antigens between endogenous and erythroid cells differentiated from $CD34^+$ cells from the same donor (Supplemental Fig 1). As expected, Ch and Rg were

not detected on the cultured cells as these antigens are expressed on Complement component C4, absorbed onto red cells from plasma. When blood group antigens defined by carbohydrate structures were examined (ABH, Ii, P1, P, Pk, LKE), cultured cord reticulocytes showed stronger expression of P1 and I antigens and weaker expression of H and LKE antigens. Weaker expression of ABH antigens and stronger expression of i antigen on cord RBC's when compared with adult RBC's is well known.(8,18,19) Such variations in the degree of N-glycosylation are unlikely to impact on the use of cultured cord reticulocytes for therapy although the possibility of exposure of novel cryptantigens cannot be entirely ruled out.

Previous reports of the RBC proteome include studies by Kakhniashvili *et al* who identified 181 proteins,(20) Pasini *et al,* (21) who reported 534 proteins with unique IDs, and Roux-Dalvai *et al*(22) who used hexapeptide libraries to reduce the signal from haemoglobin during MS and thus maximise identification of RBC cytosolic proteins, detecting 1578 proteins. Recently, Heqedus *et al*(23) performed MS for RBC membranes, identifying 419 proteins, pooling these data with that available in the literature to generate a database for the RBC membrane containing 846 proteins. D'Alessandro *et al*(24) also interrogated the literature compiling a list of 1989 RBC proteins, pathway and network analysis of which supported the concept that RBCs suffer exacerbated oxidative stress. Basu *et al*(25)recently catalogued proteins of the RBC membrane skeleton identifying unexpected components such as myosin-9, lipid raft proteins and multiple chaperone proteins. Such studies clearly reveal the power of proteomic technologies. Our data represents a significant advance in the number of erythroid proteins identified to date, and was achieved for both cytosolic and membrane proteins simultaneously from the same cells. The depth of our coverage is confirmed by identification

of the nine haemoglobin chains, including those low expression chains highlighted by Roux-Dalvai *et al*(22) as normally hidden by the α- and β-globin chains.

Proteomic comparison of protein levels between adult and cord erythroid cells

We compared the level of proteins between the adult and cord RBCs and adult and cord reticulocytes.

Of the identified proteins 943 membrane proteins and 731 cytosolic proteins were quantified (Supplemental Tables 4A & B). Of these only 2.6% varied in level between adult and cord reticulocytes, and 6% between adult and cord RBCs.

As expected the level of β- and δ-globin was higher in adult than cord RBCs (5 and 12 fold respectively) and reticulocytes (2.2 and 3.4 fold), whereas the level of γ-globin 1 and 2 was higher in cord than adult RBCs (12 and 66 fold respectively) and reticulocytes (3 and 9 fold). The level of ε- and ξ-globin was also higher in cord than adult RBCs (23 and 56 fold respectively) and reticulocytes (both 4 fold). The level of α-globin was consistent between adult and cord RBCs and between adult and cord reticulocytes.

We interrogated the data to ensure inclusion of proteins critical for RBC function. All proteins of the band 3-ankyrin macro-complex and 4.1R junctional complex(26-28) were quantified, except the Rh antigens and RhAG which were identified, but as from only 1 unique peptide their quantification was not considered. Protein levels were consistent between the membrane fraction of adult and cord RBCs, and adult and cord reticulocytes (Supplemental Table 5).

We next searched for proteins with known adhesion or transport function, metabolic enzymes and cytoskeleton proteins essential for RBC structure, identifying 137 and quantifying 92 (Supplemental Table 5).

The overall abundance of all was equivalent between adult and cord reticulocytes.

In RBCs, five proteins (carbonic anhydrase 1 & 2, aquaporin 1, BCAM [Lutheran] and semaphorin-7A) were at a higher level in the adult cells (Table 2). The higher level of carbonic anhydrase 1 and 2 and aquaporin-1, along with the expression of β-globin in adult RBCs can be attributed to known differences in gaseous exchange requirement between adults and neonates. We previously showed that transduction of cord derived erythroid cells with transcription factors KLF1 and BCL11A-XL induces the switch from γ- to β-globin.(15) Conversely, knockdown of BCL11A in adult erythroblasts reverses the globin switch, increasing the expression of γ-globin.(29) We knocked down the level of BCL11A in adult erythroid cells using two different shRNA, B1 and B5 (Supplemental Fig 2). There was a greater increase in γ-globin expression following transduction with shRNA B5, we therefore compared the protein expression profile of these cells with that of control cells using TMT labeling and MS. The level of β-globin decreased as expected, but interestingly the level of carbonic anhydrase 1 & 2 and aquaporin-1 (Table 3) was also reduced, suggesting co-regulatory mechanisms for these proteins. Variation in abundance of Semaphorin 7A which carries the JMH blood group antigens and BCAM with carries the Lutheran blood group antigens would not be expected to have any clinical sequalae.

Three proteins, myosin heavy chain 9 and 10 and myosin regulatory light chain 12A, all subunits of the myosin cellular motor proteins that interact with actin, were at a higher level

in cord than adult RBCs (Table 2). A higher level of myosin 9 has previously been reported in neonatal RBCs.(30) Differences in the level of these proteins could contribute to differences in membrane deformability between cord and adult RBCs, although data pertaining to such are conflicting.(31,32) However, 40 other cytoskeleton proteins, including those critical for RBC stability, were at equivalent abundance, suggesting an overall similarity between the cell's resistance to mechanical stress.

Quantification of western blots of adult and cord RBC lysates probed with antibodies to selected proteins corroborated the MS data (Supplemental Fig 3). Glycophorin A and α-globin, which were consistent in level between adult and cord RBCs in the MS data, served as controls.

The similarity between the proteome of adult and cord erythroid cells makes their distinct globin expression profiles even more striking, as clearly their differences in oxygen binding capacity are not associated with significant differences in the expression of other proteins.

Analysis of the change in proteome following reticulocyte maturation to RBC

Detailed analysis of the change in proteome during maturation of human reticulocytes to mature RBCs has been hampered by the difficulty isolating a suitable population and amount of reticulocytes. As we were able to generate a large number of mature functional reticulocytes by *in vitro* culture, we were able to compare the proteome of these cells with that of the autologous endogenous RBCs using our MS data.

We first confirmed that specific proteins known to decrease in level during maturation of reticulocytes to RBCs(1,33) did so in our study. Transferrin receptor 1, integrin β_1, CD98, 65

ribosome subunits, all subunits of the sodium potassium ATPase detected and all isoforms of tubulin detected were at a lower level in both adult and cord RBCs compared to reticulocytes.

We next looked at the change in proteome between the adult autologous reticulocytes and RBCs. A total of 943 proteins were quantified in the membrane and 728 in the cytosol fraction (Supplemental Tables 6A and B).

In the membrane fraction 500 proteins were at a lower level in RBCs compared to reticulocytes, of these 370 were reduced by 2-10 fold and 130 by >10 fold. There were just 86 proteins at a higher level in RBCs, and the number of proteins with a large magnitude of change was far less with only 3 proteins (Glycophorin A, Sorbitol dehydrogenase and Cathepsin E) >10 fold higher. In the cytosol fraction 588 proteins were at a lower level in RBCs, comprising 464 proteins reduced by 2-10 fold and 124 by >10 fold. Only 4 proteins were at a higher level in RBCs compared to reticulocytes, with just Carbonic anhydrase 1 >10 fold higher. All proteins that differed in level by 10 fold or more in the membrane and cytosol fraction are shown in Supplemental Tables 7A & B respectively.

We classified the proteins by cellular component and molecular function (Figure 1). Examples from these analyses show the majority of ribosomal and mitochondrial proteins are lost from reticulocytes as expected. The majority of proteins involved in translation regulation and chromatin binding were also lost from reticulocytes, whereas proteins involved in lipid and oxygen binding were at a higher level in RBCs than reticulocytes

As there was a distinct difference in the number of proteins that increased in level in the membrane compared to cytosol fraction of RBCs, we questioned whether some proteins

underwent differential partitioning between the cytosol to the membrane during maturation. To address this we interrogated our MS data for proteins at a higher level in the RBC than reticulocyte membrane fraction, with a corresponding decrease in level in the RBC cytosol (Supplemental Table 8), notably this included a disproportionate number of proteins belonging to the band 3-ankyrin and 4.1R junctional membrane complexes. To determine whether other proteins in these complexes displayed the same trend we employed reduced stringency of 1 unique peptide, which although below our normal robust stringency of 2 unique peptides per protein for comparative analysis, still ensures that the peptide is specific to that protein and gives an indication of abundance. Using 2 or more unique peptides reassures consistency between comparisons. Using this approach we found that all proteins detected displayed such partitioning (Supplemental Table 9) in both adult and cord cells. This may be due to proportional loss of other membrane proteins, but more likely re-localisation of proteins from the cytosol to the membrane occurs during reticulocyte maturation reflecting the known loss of intracellular endocytic pools of proteins during maturation.(1,34) Data for selected proteins was confirmed by quantification of western blots (Supplemental Fig 4).

Using the reduced stringency we also found all 14 α and β proteasome subunits at a higher level in RBCs than reticulocytes, which along with the identification of other proteasome subunits support a function for proteasome in RBCs.

Finally, we examined the profile of all globin subunits in the cytosol fraction (Table 4). In adult cells the level of β- and δ-globin was consistent between reticulocytes and RBCs, however the level of gamma 1 & 2, ε- and ζ-globin was lower in RBCs than reticulocytes. In cord cells the level of both γ subunits, ε- and ζ-globin was consistent between reticulocytes and RBCs, but the level of β- and δ-globin was lower in RBCs than reticulocytes. The level

of α-globin was consistent. These data suggest selective loss of globin subunits during reticulocyte maturation, or differences in the stability of different mRNAs resulting in the continued translation of some but not other transcripts. Similar mechanisms may function for other proteins, as overall levels were more consistent between adult and cord reticulocytes than RBCs.

Conclusion

Our data provide the most comprehensive identification to date of the RBC proteome, and insight into the change in proteome during maturation, which will serve as a useful resource. Apart from the known differences in globin expression and key proteins known to be involved in gaseous exchange (aquaporin 1 and carbonic anhydrase 1 & 2), these data indicate the proteome of adult and cord erythroid cells are reassuringly similar. Thus, with respect to the proteome there is no detectable barrier to using cord progenitors for the *in vitro* generation of RBCs for adult therapeutics other than the known differences in oxygen uptake and release of HbF and HbA.

Acknowledgements

This research was funded by the National Institute for Health Research Blood and Transplant Unit (NIHR BTRU) in Red Blood Cell Products at the University of Bristol in Partnership with NHS Blood and Transplant (NHSBT). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Footnote

The accession number F5H250 for RhAG has been updated to Q02094.

Authorship

Contribution: J.F. and D.J.A. conceived and supervised the study; J.F., D.J.A., M.C.W. and K.T. designed experiments; M.C.W., K.T. and CG performed experiments; K.J.H., S.F.P., A.M.T. and N.C. gave practical assistance; M.C.W., K.T., A.M.T., D.J.A. and J.F. analysed data; J.F. wrote the paper and D.J.A., K.T., M.C.W. and A.M.T. read and edited the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

References

1. Griffiths, R. E., Kupzig, S., Cogan, N., Mankelow, T. J., Betin, V. M., Trakarnsanga, K., Massey, E. J., Lane, J. D., Parsons, S. F., and Anstee, D. J. (2012) Maturing reticulocytes internalize plasma membrane in glycophorin A-containing vesicles that fuse with autophagosomes before exocytosis. *Blood* **119**, 6296-6306
2. Fujimi, A., Matsunaga, T., Kobune, M., Kawano, Y., Nagaya, T., Tanaka, I., Iyama, S., Hayashi, T., Sato, T., Miyanishi, K., Sagawa, T., Sato, Y., Takimoto, R., Takayama, T., Kato, J., Gasa, S., Sakai, H., Tsuchida, E., Ikebuchi, K., Hamada, H., and Niitsu, Y. (2008) Ex vivo large-scale generation of human red blood cells from cord blood CD34+ cells by co-culturing with macrophages. *International journal of hematology* **87**, 339-350
3. Giarratana, M. C., Kobari, L., Lapillonne, H., Chalmers, D., Kiger, L., Cynober, T., Marden, M. C., Wajcman, H., and Douay, L. (2005) Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. *Nature biotechnology* **23**, 69-74

4. Dias, J., Gumenyuk, M., Kang, H., Vodyanik, M., Yu, J., Thomson, J. A., and Slukvin, II. (2011) Generation of red blood cells from human induced pluripotent stem cells. *Stem cells and development* **20**, 1639-1647

5. Trakarnsanga, K., Wilson, M. C., Griffiths, R. E., Toye, A. M., Carpenter, L., Heesom, K. J., Parsons, S. F., Anstee, D. J., and Frayne, J. (2014) Qualitative and Quantitative Comparison of the Proteome of Erythroid Cells Differentiated from Human iPSCs and Adult Erythroid Cells by Multiplex TMT Labelling and NanoLC-MS/MS. *PloS one* **9**, e100874

6. Lu, S. J., Feng, Q., Park, J. S., Vida, L., Lee, B. S., Strausbauch, M., Wettstein, P. J., Honig, G. R., and Lanza, R. (2008) Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood* **112**, 4475-4484

7. Tanavde, V. M., Malehorn, M. T., Lumkul, R., Gao, Z., Wingard, J., Garrett, E. S., and Civin, C. I. (2002) Human stem-progenitor cells from neonatal cord blood have greater hematopoietic expansion capacity than those from mobilized adult blood. *Experimental hematology* **30**, 816-823

8. Bhattacharya, N., and Stubble, P. (2009) Frontiers of Cord Blood Science. *Springer-Verlag London Ltd*

9. Greer, J. P., Arber, D. A., Glader, B., List, A. F., Means, R. T. J., Paraskevas, F., Rodgers, G. M., and Foerster, J. (2014) Fetal and neonatal erythrocyte membrane and metabolism. *In Wintrobes Clinical Hematology. Publisher Walters Kluwer, Lippincott Williams &Wilkins* **13th Edition**

10. Bell, A. J., Satchwell, T. J., Heesom, K. J., Hawley, B. R., Kupzig, S., Hazell, M., Mushens, R., Herman, A., and Toye, A. M. (2013) Protein distribution during human erythroblast enucleation in vitro. *PloS one* **8**, e60300

11. Mankelow, T. J., Griffiths, R. E., Trompeter, S., Flatt, J. F., Cogan, N. M., Massey, E. J., and Anstee, D. J. (2015) Autophagic vesicles on mature human reticulocytes explain phosphatidylserine-positive red cells in sickle cell disease. *Blood* **126**, 1831-1834

12. Liu, J., Guo, X., Mohandas, N., Chasis, J. A., and An, X. (2010) Membrane remodeling during reticulocyte maturation. *Blood* **115**, 2021-2027

13. Richardson, B. M., Heesom, K. J., Parsons, S. F., Anstee, D. J., and Frayne, J. (2009) Analysis of the differential proteome of human erythroblasts during in vitro erythropoiesis by 2-D DIGE. *Proteomics Clin Appl* **3**, 1123-1134

14. Senol, O., Schaaij-Visser, T. B., Erkan, E. P., Dorfer, C., Lewandrowski, G., Pham, T. V., Piersma, S. R., Peerdeman, S. M., Strobel, T., Tannous, B., Saydam, N., Slavc, I., Knosp, E., Jimenez, C. R., and Saydam, O. (2015) miR-200a-mediated suppression of non-muscle heavy chain IIb inhibits meningioma cell migration and tumor growth in vivo. *Oncogene* **34**, 1790-1798

15. Trakarnsanga, K., Wilson, M. C., Lau, W., Singleton, B. K., Parsons, S. F., Sakuntanaga, P., Kurita, R., Nakamura, Y., Anstee, D. J., and Frayne, J. (2014) Induction of adult levels of beta-globin in human erythroid cells that intrinsically express embryonic or fetal globin by transduction with KLF1 and BCL11A-XL. *Haematologica* **99**, 1677-1685

16. Satchwell, T. J., Hawley, B. R., Bell, A. J., Ribeiro, M. L., and Toye, A. M. (2015) The cytoskeletal binding domain of band 3 is required for multiprotein complex formation and retention during erythropoiesis. *Haematologica* **100**, 133-142

17. 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., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature biotechnology* **32**, 223-226

18. Marsh, W. L. (1961) Anti-i: a cold antibody defining the Ii relationship in human red cells. *Br J Haematol* **7**, 200-209
19. Marsh, W. L., Nichols, M. E., and Reid, M. E. (1971) The definition of two I antigen components. *Vox sanguinis* **20**, 209-217
20. Kakhniashvili, D. G., Bulla, L. A., Jr., and Goodman, S. R. (2004) The human erythrocyte proteome: analysis by ion trap mass spectrometry. *Mol Cell Proteomics* **3**, 501-509
21. Pasini, E. M., Kirkegaard, M., Mortensen, P., Lutz, H. U., Thomas, A. W., and Mann, M. (2006) In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood* **108**, 791-801
22. 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., and Monsarrat, B. (2008) Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced mass spectrometry. *Mol Cell Proteomics* **7**, 2254-2269
23. Hegedus, T., Chaubey, P. M., Varady, G., Szabo, E., Saranko, H., Hofstetter, L., Roschitzki, B., Stieger, B., and Sarkadi, B. (2015) Inconsistencies in the red blood cell membrane proteome analysis: generation of a database for research and diagnostic applications. *Database (Oxford)* **2015**, bav056
24. D'Alessandro, A., Righetti, P. G., and Zolla, L. (2010) The red blood cell proteome and interactome: an update. *J Proteome Res* **9**, 144-163

25. Basu, A., Harper, S., Pesciotta, E. N., Speicher, K. D., Chakrabarti, A., and Speicher, D. W. (2015) Proteome analysis of the triton-insoluble erythrocyte membrane skeleton. *Journal of proteomics* **128**, 298-305

26. Bruce, L. J., Beckmann, R., Ribeiro, M. L., Peters, L. L., Chasis, J. A., Delaunay, J., Mohandas, N., Anstee, D. J., and Tanner, M. J. (2003) A band 3-based macrocomplex of integral and peripheral proteins in the RBC membrane. *Blood* **101**, 4180-4188

27. Mohandas, N., and Gallagher, P. G. (2008) Red cell membrane: past, present, and future. *Blood* **112**, 3939-3948

28. van den Akker, E., Satchwell, T. J., Williamson, R. C., and Toye, A. M. (2010) Band 3 multiprotein complexes in the red cell membrane; of mice and men. *Blood cells, molecules & diseases* **45**, 1-8

29. Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., Mikkola, H. K., Hirschhorn, J. N., Cantor, A. B., and Orkin, S. H. (2008) Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* **322**, 1839-1842

30. Matovcik, L. M., Groschel-Stewart, U., and Schrier, S. L. (1986) Myosin in adult and neonatal human erythrocyte membranes. *Blood* **67**, 1668-1674

31. Gross, G. P., and Hathaway, W. E. (1972) Fetal erythrocyte deformability. *Pediatric research* **6**, 593-599

32. Johnson, R. M., Panchoosingh, H., Goyette, G., Jr., and Ravindranath, Y. (1999) Increased erythrocyte deformability in fetal erythropoiesis and in erythrocytes deficient in glucose-6-phosphate dehydrogenase and other glycolytic enzymes. *Pediatric research* **45**, 106-113

33. Ney, P. A. (2011) Normal and disordered reticulocyte maturation. *Current opinion in hematology* **18**, 152-157

34. Satchwell, T. J., Bell, A. J., Pellegrin, S., Kupzig, S., Ridgwell, K., Daniels, G., Anstee, D. J., van den Akker, E., and Toye, A. M. (2011) Critical band 3 multiprotein complex interactions establish early during human erythropoiesis. *Blood* **118**, 182-191

Table 1. RBC blood group proteins in pooled erythroid cell samples.

Endogenous adult and cord RBCs and reticulocytes (Retic) differentiated *in vitro* from adult peripheral blood and cord $CD34^+$ cells (RBCs and $CD34^+$ cells pooled from 16 and 4 individuals respectively) were fractionated and proteins subjected to trypsin digest, with resultant peptides labelled with TMTs for nanoLC-MS/MS based quantitation. Values show the ratio of protein levels between adult RBCs and reticulocytes (Retic) and cord RBCs and reticulocytes. Proteins were quantified from at least 2 unique peptides. Peptides and unique peptides; the total number of peptide sequences and number of unique peptides identified for that protein. Proteome Discoverer software v1.4 was used for analysis.

Accession	Description	System	Unique Peptides	Peptides	Adult RBC/Retic	Cord RBC/Retic
C9JGQ9	ACHE	Yt	5	5	2.742	1.406
E9PC21	AQP1	Colton	2	2	5.165	3.412
Q9NP58	ATP-binding cassette sub-family B member 6	LAN	14	15	0.816	1.166
Q9UNQ0	ATP-binding cassette sub-family G member 2	JR	3	3	2.204	2.023
P02730	Band 3 anion transport	Diego	27	27	3.826	3.604

	protein					
P50895	Basal cell adhesion molecule	Lutheran	10	10	2.373	1.155
Q54A51	Basigin	Ok	11	11	0.954	1.260
Q02161	Blood group Rh(D) polypeptide	Rh	1	2	2.872	8.032
A6H8M8	C4A protein	Chido/ Rodgers	1	1	6.252	1.575
B6EAT9	CD44	Indian	6	6	1.946	1.188
E9PNW4	CD59	CD59	3	3	1.259	2.748
A6NIW1	CD99 antigen	Xg	1	1	2.014	0.741
E9PDY4	CR1	knops	2	2	1.885	1.083
Q14UF5	Decay-accelerating factor	Cromer	11	11	2.113	2.026
Q93070	Ecto-ADP-ribosyltransferase 4	Dombrock	3	3	3.086	3.262
Q96PL5	Erythroid membrane-associated protein	Scianna	10	10	3.127	2.423
B8Q185	Glycophorin A MNS blood group	MNS	3	3	12.251	11.878
P04921	Glycophorin-C	Gerbich	4	4	2.685	2.892
Q14773	Intercellular adhesion molecule 4	Landsteiner-Wiener	5	5	2.414	4.418
P23276	Kell blood group glycoprotein	Kell	8	8	3.245	2.189
P51811	Membrane transport protein XK	Kx	6	6	2.411	2.520
F5H250	RHAG	Rh associated	1	1	1.431	5.235

		glycoprotein				
E7EWZ5	RHCE	Rh	1	2	1.431	5.235
O75326	Semaphorin-7A	John Milton Hagen	8	8	5.814	3.986
E9PR61	SLC14A1	Kidd	1	5	2.675	1.397
Q13336	Urea transporter 1 SLC14A1	Kidd	3	7	2.844	2.518

Table 2. Proteins that differ in amount between adult and cord RBCs

See legend for Table 1 for experimental details. Values show the ratio of protein levels between adult and cord RBCs and adult and cord reticulocytes.

Accession	Description	Unique Peptides	Peptides	RBC Adult/Cord	Retic Adult/Cord
E9PC21	Aquaporin 1	2	2	2.104	1.565
P50895	Basal cell adhesion molecule	10	10	3.236	1.560
P00915	Carbonic anhydrase 1	11	11	8.706	1.075
P00918	Carbonic anhydrase 2	11	11	4.976	1.470
O75326	Semaphorin-7A	8	8	2.152	1.399
P19105	Myosin regulatory light chain 12A	4	4	0.339	0.918
P35579	Myosin-9	83	96	0.375	0.796
P35580	Myosin-10	74	88	0.175	0.737

Table 3. Alteration in protein levels following knockdown of BCL11A in adult erythroid cells

Adult erythroid cells differentiated from peripheral blood $CD34^+$ cells were transduced with a BCL11A shRNA at day 3 in culture, with cells collected on day 8. Cells were lysed and protein levels analysed as described in legend for Table 1. Values show the ratio of protein levels between control erythroid cells and those transduced with BCL11A shRNA (BCL11A KD). Glycophorin A and Band 3 were included to show that changes in protein expression following BCL11A knockdown were not generic or due to delayed differentiation. Globin subunits were quantified using high stringency, all other proteins were quantified using medium stringency peptide ID.

Description	**Unique Peptides**	**Peptides**	**Control/ BCL11A KD**
β-globin	2	13	7
γ-globin	7	8	0.3
ε-globin	2	3	0.09
Aquaporin 1	3	3	3
Carbonic anhydrase 1	10	10	11
Carbonic anhydrase 2	14	14	3
Band 3	28	30	1
Glycophorin A	2	3	1

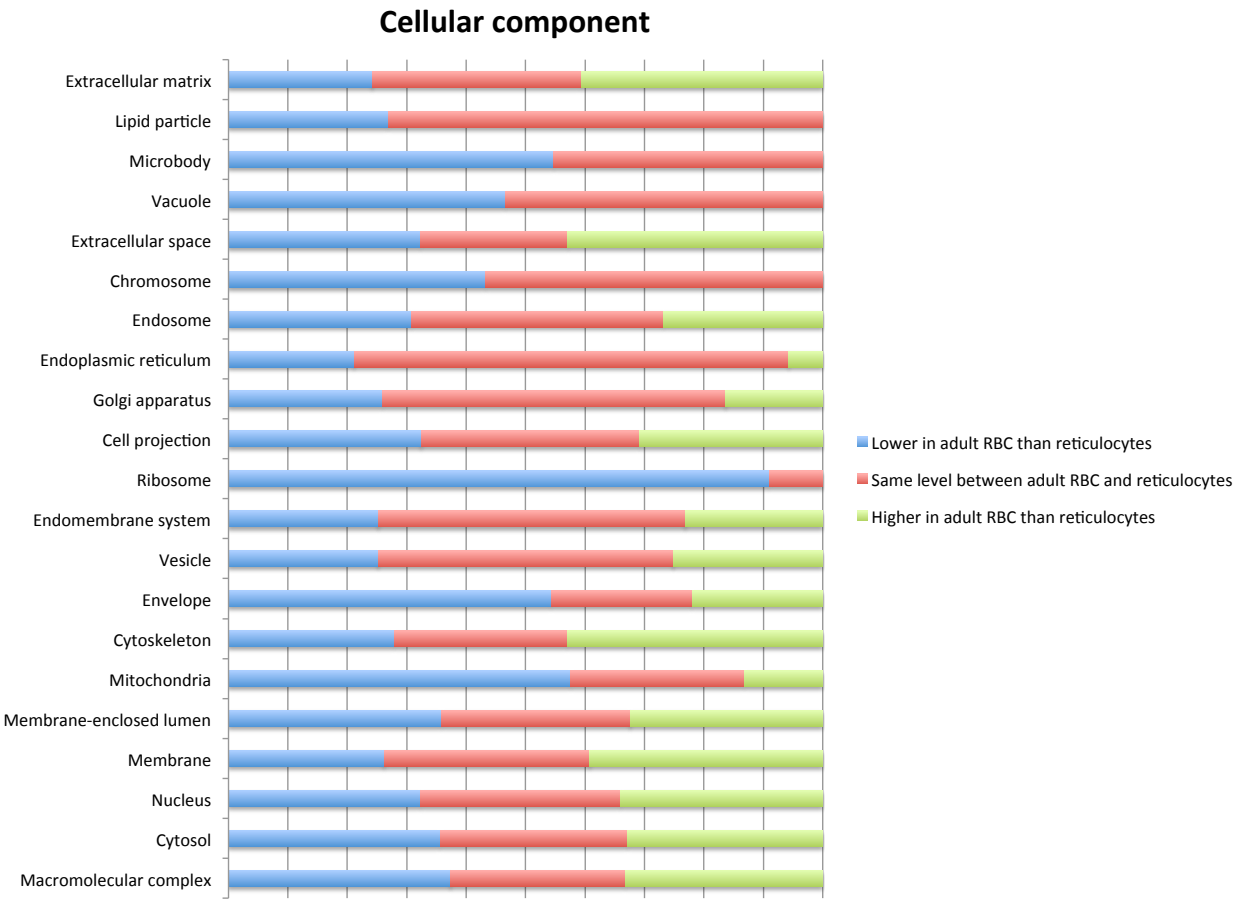
Table 4. Comparing the level of globin subunits between adult and cord erythroid cells, and between RBCs and reticulocytes

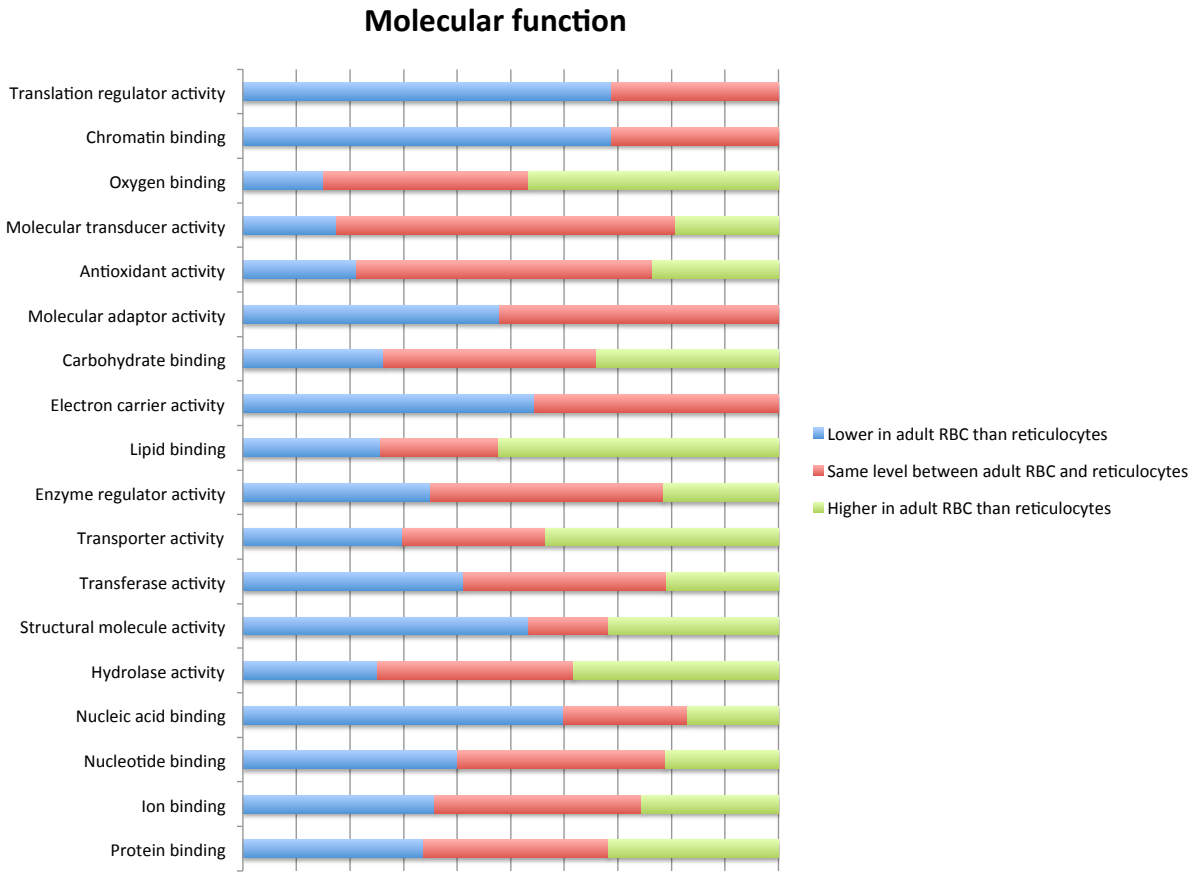
See legend for Table 1 for experimental details. Proteins were analysed in the cytosol fraction of cells. Values show the ratio of protein levels between RBCs and reticulocytes (Retic), and between adult and cord cells.

Accession	Globin subunit	Unique peptides	Peptides	RBC Adult/Cord	Retic Adult/Cord	Adult RBC/Retic	Cord RBC/Retic
P69905	Alpha	13	14	1.115	0.951	0.894	0.725
P68871	Beta	8	16	4.842	2.203	0.970	0.450
P02042	Delta	6	14	11.660	3.431	0.901	0.268
D9YZU8	Gamma-1	4	15	0.085	0.328	0.253	0.933
P69892	Gamma-2	3	14	0.015	0.106	0.174	1.171
P02100	Epsilon	3	6	0.044	0.281	0.123	0.968
P02008	Zeta	8	9	0.018	0.231	0.119	1.291
Q6B0K9	Mu	5	5	0.206	0.362	0.161	0.496
P09105	Theta-1	3	3	0.627	0.690	0.720	0.652

Figure 1. Cellular component and molecular function of proteins that decrease, increase and remain at the same level during maturation of adult reticulocytes to RBCs.

Proteins from endogenous adult RBCs and reticulocytes differentiated *in vitro* from adult peripheral blood $CD34^+$ cells were subjected to trypsin digest and resultant peptides labelled with TMTs for nanoLC-MS/MS based quantitation. Proteins were quantified using at least 2 unique peptides and analysed using WebGestalt GSAT V2. The proportion of proteins as a percentage of the total in each of the groups, cellular component and molecular function, was calculated.


Cellular component
Extracellular matrix
Lipid particle
Microbody
Vacuole
Extracellular space
Chromosome
Endosome
Endoplasmic reticulum
Golgi apparatus
Cell projection
Ribosome
Endomembrane system
Vesicle
Envelope
Cytoskeleton
Mitochondria
Membrane-enclosed lumen
Membrane
Nucleus
Cytosol
Macromolecular complex
Lower in adult RBC than reticulocytes
Same level between adult RBC and reticulocytes
Higher in adult RBC than reticulocytes


Molecular function
Translation regulator activity
Chromatin binding
Oxygen binding
Molecular transducer activity
Antioxidant activity
Molecular adaptor activity
Carbohydrate binding
Electron carrier activity
Lipid binding
Enzyme regulator activity
Transporter activity
Transferase activity
Structural molecule activity
Hydrolase activity
Nucleic acid binding
Nucleotide binding
Ion binding
Protein binding
Lower in adult RBC than reticulocytes
Same level between adult RBC and reticulocytes
Higher in adult RBC than reticulocytes