MINIREVIEW

The Human Red Blood Cell Proteome and Interactome

STEVEN R. GOODMAN,*,†,[1] ANASTASIA KURDIA,‡ LARRY AMMANN,§ DAVID KAKHNIASHVILI,* AND OVIDIU DAESCU‡

*Department of Molecular and Cell Biology, University of Texas at Dallas, Richardson, Texas 75080; †Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390; ‡Jonsson School of Engineering and Computer Science, University of Texas at Dallas, Richardson, Texas 75080; and §Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75080

The red blood cell or erythrocyte is easily purified, readily available, and has a relatively simple structure. Therefore, it has become a very well studied cell in terms of protein composition and function. RBC proteomic studies performed over the last five years, by several laboratories, have identified 751 proteins within the human erythrocyte. As RBCs contain few internal structures, the proteome will contain far fewer proteins than nucleated cells. In this minireview, we summarize the current knowledge of the RBC proteome, discuss alterations in this partial proteome in varied human disease states, and demonstrate how *in silico* studies of the RBC interactome can lead to considerable insight into disease diagnosis, severity, and drug or gene therapy response. To make these latter points we focus on what is known concerning changes in the RBC proteome in Sickle Cell Disease. Exp Biol Med 232:1391–1408, 2007

Key words: red blood cell; erythrocyte; proteomics; interactome; systems biology; Sickle Cell Disease

The red blood cell (RBC) or erythrocyte travels through our circulatory system for 120 days, during which time it must constantly change its shape from a biconcave disc of 8 μm diameter to a cigar shape able to traverse passage ways that narrow to 1 μm in diameter. A two dimensional meshwork of proteins called the spectrin membrane skeleton, found on the cytoplasmic surface of the plasma membrane, gives the RBC its properties of elasticity and flexibility that allows for the success of this journey (1–5). Defects in this membrane skeleton lead to misshapen and osmotically fragile RBCs (6–10).

This tortuous journey traveled by the erythrocyte is for the purpose of carrying oxygen from our lungs to cells, tissues and organs throughout the body; and returning carbon dioxide to our lungs. This essential RBC function is conducted by hemoglobin, the major protein constituent of the RBC cytosol. The erythrocyte is the simplest of human cells as it lacks internal organelles; lost during the process of erythropoiesis. The ease of obtaining blood, lack of internal organelles, and important physiologic function of the RBC has made it a major focus of biochemical study during the 20th and 21st century. As a result, we know the functions of erythrocyte proteins in greater detail than any other human cell type.

The simplicity of the human erythrocyte cell structure has also made it an optimal cell for proteomic study. While nucleated cells contain 20,000 to 30,000 proteins (11–14), RBCs which lack nuclei and other organelles contain far fewer. It is, therefore, the cell type where we are most likely to approach a complete proteome in the foreseeable future. This fact, along with the comprehensive current literature on RBC function, will allow us to soon understand the relationship of the erythrocyte proteome and interactome to RBC function. Further, we are beginning to understand how changes in the RBC proteome and interactome relate to

Work described in this minireview was supported by a Sickle Cell Center grant from the National Institutes of Health HL070588, Project 11, awarded to Steven R. Goodman.

[1] To whom correspondence should be addressed at 2601 North Floyd Road, P.O. Box 830688, Richardson, TX 75083-0688. E-mail: sgoodmn@utdallas.edu

DOI: 10.3181/0706-MR-156
1535-3702/07/23211-1391$15.00

erythrocyte disorders. The proteome and interactome of the normal and abnormal human erythrocyte is the subject of this mini review.

A Status Report of the Human Erythrocyte Proteome

To understand the human erythrocyte proteome we would need to know the identity of every protein, their amino acid sequence and posttranslational modifications (PTMs), the number of copies of each protein per cell, and all of their interactions (the interactome). The use of modern mass spectrometry, with appropriate search algorithms and comprehensive human protein databases allows one to study the complete RBC proteome; by the process called proteomics.

Here we will discuss the identification of RBC proteins by modern mass spectrometry and database searching. The story behind this portion of RBC proteomic data has unfolded over the past five years and closely follows the technical advances in the proteomics field over this period of time. The first serious study of the RBC membrane's proteome was performed by Low et al and published in 2002 (15). Their focus was on plasma membrane proteins and they used the classical approach of two dimensional IEF–SDS PAGE (2D electrophoresis), followed by silver staining, in gel trypsin digestion and MALDI TOF mass spectrometry. To avoid some of the problems related to using 2D gels for proteomic identification (problems with high and low MW proteins, proteins with high and low isoelectric points, and hydrophobic proteins), they also performed one dimensional SDS PAGE. By a combination of these techniques, Low et al (15) were able to identify 84 distinct RBC membrane proteins. While the one dimensional gels did allow them to identify some hydrophobic proteins (ex. sorbitol dehydrogenase and the glucose transporter) other major RBC transmembrane hydrophobic proteins were not detected (ex glycophorin A and C).

Our laboratory utilized classical techniques for RBC fractionation, in combination with a shotgun proteomic approach and tandem mass spectrometry, to identify 181 unique proteins in the erythrocyte membrane and cytosol (16).

In this study by Kakhniashvili et al, tryptic digests were prepared individually from intact cells (to identify surface exposed proteins), inside out spectrin depleted membrane vesicles (to identify proteins exposed on the cytoplasmic membrane surface) and a low ionic strength membrane extract (to identify spectrin and other components of the membrane skeleton) (16). Each tryptic digest was individually separated by reverse phase HPLC and analyzed by tandem mass spectrometry utilizing a ThermoFinnigen LCQ DECA XP Ion Trap mass spectrometer. We identified 91 unique membrane proteins. In addition we separated the cytosolic proteins, by gel filtration chromatography, into 21 fractions. Each fraction was digested with trypsin and analyzed by LC – MS/MS as described above. Again 91 unique cytosolic proteins were identified. Glyceraldhyde–3-P–dehydrogenase was placed in both the membrane and cytosolic lists, because of the high number of hits in both fractions. Therefore, the total number of proteins identified was 181 which was the most complete erythrocyte proteome list available in 2004 (16). In this study by Kakhiashvili et al not only were the glycophorin A and C identified in the membrane fraction, but also proteins that are present in just a few hundred copies per RBC (ex B-CAM) (16).

In both the Low et al (15) and Kakhniashvili et al (16) studies many proteasomal subunits were identified. As previous studies had claimed no proteasomal degradation of proteins in RBCs (17), it was possible that these proteins were coming from a small number of contaminating reticulocytes. However, we have recently demonstrated by confocal immuno-fluorescence and western blot analysis that proteasomal subunits and intact proteasomes do exist in mature RBCs (18). Therefore, proteomic studies have now led to an exciting question for future exploration: What is the physiologic function of RBC proteasomes?

Collectively Low et al (15) and Kakhniashvili et al (16) had identified over 200 unique proteins. Studies that followed over the next year included the use of trypsin associated with self assembled monolayers on gold to create an enzyme chip for digestion of RBC protein prior to MudPit (SCX coupled to RPHPLC) and tandem mass spectrometry (19); and the use of soft Immobiline gels instead of IPG strips for 2D Electrophoresis prior to MALDI TOF (20). The former study by Tyon et al identified 272 proteins, but only 30 by 2 or more unique peptides (19). The latter study by Bruschi et al allowed greater visualization of high molecular weight protein spots on 2D Gels (20). While both studies contributed valuable new technologies, for specialized situations, neither added significantly to the number of RBC proteins that had been identified by proteome technology by 2004.

The next substantive increase in the number of RBC proteins identified by proteomic technology came at the end of 2006 (21). Pasini et al, utilizing both an Applied Biosystem Quadropole TOF Q–STAR mass spectrometer and a Thermo Electron hybrid linear ion trap Fourier transform MS (LTQ–FT MS), were able to identify 314 RBC membrane proteins and 252 soluble proteins (21). While Pasini et al (21) were careful to optimize the extraction of membrane proteins, with combinations of ethanol and Na Carbonate, the major contributor to the substantial increase in protein identification was technological. Proteomic advances are closely tied to technologic advances in mass spectrometers and Pasini et al (21) were using instruments with extremely high sensitivity (LTQ–FTMS has 1 ppm accuracy) and very high mass accuracy (attamoles). The result was their identification of 566 unique RBC proteins.

By combining all of the proteomic data discussed in this section, and ignoring PTMs including proteolysis, we have created a comprehensive list of RBC proteins identified, by proteomic technology, to date (Table 1). As

Table 1. Erythrocyte Proteins That Have Been Identified By Proteomic Methods

1. 14-3-3 protein epsilon
2. 14-3-3 protein gamma
3. 14-3-3 protein tau
4. 14-3-3 protein zeta/delta (Protein kinase C inhibitor protein-1, KCIP-1)
5. 17 kDa cyclophilin A
6. 2,3-bisphosphoglycerate mutase
7. 2', 3'-cyclic-nucleotide 3'-phosphodiesterase
8. 26S protease regulatory subunit 4
9. 26S protease regulatory subunit 7
10. 26S protease regulatory subunit S10
11. 26S proteasome non-ATPase regulatory subunit 12
12. 26S proteasome non-ATPase regulatory subunit 13
13. 26S proteasome non-ATPase regulatory subunit 14
14. 26S proteasome non-ATPase regulatory subunit 3
15. 26S proteasome non-ATPase regulatory subunit 5
16. 26S proteasome non-ATPase regulatory subunit 6
17. 26S proteasome non-ATPase regulatory subunit 7
18. 26S proteasome regulatory subunit S14
19. 3-mercaptopyruvate sulfurtransferase
20. 3-oxo-5-beta-steroid-4-dehydrogenase
21. 40S ribosomal protein S3
22. 40S ribosomal protein S6
23. 6-phosphofructokinase
24. 6-phosphogluconate dehydrogenase, decarboxylating
25. 78 kDa glucose-regulated protein precursor
26. Acid phosphatase isoenzyme Af
27. Actin 2, alpha aortic smooth muscle
28. Actin binding protein anillin
29. Actin, alpha skeletal muscle
30. Actin, cytoplasmic 1 (Beta-actin)
31. Actin, cytoplasmic 2 (Gamma-actin)
32. Actin-like protein 2
33. Actin-like protein 3 (Actin-2)
34. Acylamino-acid-releasing enzyme
35. ADD1 protein
36. Adducin 1 (alpha) isoform c
37. Adducin alpha subunit, erythrocyte
38. Adenylate kinase isoenzyme 1
39. ADP-ribosylation factor 3
40. Aldehyde dehydrogenase 1A1
41. Aldehyde dehydrogenase 3B1
42. ALDOC protein
43. Aldolase A
44. Alpha enolase
45. Alpha-2,8-sialyltransferase 8C
46. Alpha-actinin 4
47. Alpha-hemoglobin stabilizing protein
48. Alpha-soluble NSF attachment protein (SNAP-alpha)
49. Aminolevulinate, delta -, dehydratase
50. Amyloid beta A4 protein (fragment)
51. Amyloid protein-binding protein 1
52. Ankyrin 1
53. ankyrin 1 isoform 1
54. Ankyrin 1, isoform 2, erythrocytic
55. Ankyrin 1, isoform 4, erythrocytic
56. Ankyrin 1, splice form 2
57. Annexin A11 protein
58. annexin I
59. annexin VII isoform 2
60. Antioxidant protein 2
61. ANXA4 protein
62. Apolipoprotein E precursor

Table 1. (Continued)

63. Aquaporin 1
64. Arginase type 1
65. Arginase type 1 erythroid variant
66. ARPC5 protein
67. Aspartate Aminotransferase 1
68. Aspartyl-tRNA synthetase
69. ATP citrate lyase
70. ATP synthase, alpha subunit
71. ATP synthase beta chain, mitochondrial precursor
72. ATP-binding cassette half-transporter
73. ATP-binding cassette, sub-family B, member 6, mitochondrial precursor
74. ATP-binding cassette, subfamily C, member 6
75. ATP-binding cassette, sub-family G, member 2
76. ATP-citrate synthase
77. bA421H8.2 (novel protein)
78. BAG-family molecular chaperone regulator-2
79. B-CAM protein
80. Bcl-2-related protein A1
81. Beta adducin
82. beta-Actin
83. Beta-soluble NSF attachment protein
84. Bifunctional coenzyme A synthase
85. Biliverdin reductase A precursor
86. Biliverdin reductase B
87. Biotin- -protein ligase
88. Block of proliferation 1
89. C-1-tetrahydrofolate synthase, cytoplasmic
90. C4B1
91. Calcium and integrin-binding protein 1
92. Calcium binding protein 39
93. Calcium transporting ATPase 4
94. Calcyclin
95. Calmodulin 2
96. Calnexin precursor
97. Calpain 1, large [catalytic] subunit
98. Calpain 5
99. Calpain inhibitor (Calpastatin)
100. Calreticulin precursor
101. cAMP-dependent protein kinase type I- alpha regulatory chain
102. cAMP-dependent protein kinase, alpha-catalytic subunit
103. Carbonic anhydrase I
104. carbonic anhydrase II
105. Carbonic anhydrase III
106. Carbonyl reductase [NADPH] 1
107. Carnitine O-palmitoyltransferase I, mitochondrial liver isoform
108. Casein kinase I, alpha isoform
109. Catalase
110. Catalase 5.
111. Cathepsin G [precursor]
112. Cationic trypsinogen
113. CD 59 antigen p18-20
114. CD44 antigen
115. CD59 glycoprotein [precursor]
116. CGI-150 protein
117. CGI-26 protein
118. Channel-like integral membrane protein
119. Chaperonin Containing TCP1, subunit 2
120. Chaperonin Containing TCP1, subunit 6A isoform a
121. Chaperonin containing TCP1, subunit 7
122. Chloride intraCellular Channel 1

Table 1. (Continued)

123. Chloride intracellular channel protein 3
124. Chromosome 2 open reading frame 32
125. chromosome 22 open reading frame 13
126. Chromosome 9 open reading frame 78
127. chromosome 9 open reading frame 19
128. Clathrin coat assembly protein AP180
129. Clathrin coat assembly protein AP50
130. Clathrin heavy Chain 1
131. Clusterin precursor
132. Coactivator-associated arginine methyltransferase 1
133. Coagulation factor IX [precursor]
134. Cofilin 1 (nonmuscle)
135. Cofilin 2
136. Complement C3b
137. Complement receptor type 1 precursor
138. Conserved hypothetical protein
139. COP9 subunit 4
140. Copine II
141. Copine III
142. Copper chaperone for superoxide dismutase
143. Creatine kinase, muscle
144. Cullin homolog 2
145. Cullin homolog 5
146. Cypa complexed with Hagpia, chain A
147. cytochrome b5 reductase membrane-bound isoform
148. Cytosolic acetoacetyl-coenzyme A thiolase
149. DC 38
150. DC-TM4F2 protein
151. D-dopachrome tautomerase
152. delta globin
153. Delta-aminolevulinic acid dehydratase isoform b
154. Dematin (Erythrocyte membrane protein band 4.9)
155. Dematin 52 kDa subunit
156. Diacylglycerol kinase,
157. Diacylglycerol kinase, Delta 130kDa isoform 1
158. Diaphanous 1
159. Dimeric dihydrodiol dehydrogenase
160. Dipeptidyl-peptidase III
161. DKFZP434C212 protein
162. DNA damage binding protein 1
163. DNA-damage inducible protein 2
164. Down syndrome cell adhesion molecule 2
165. Duodenal cytochrome b
166. Early endosome antigen 1
167. Ecto-ADP-ribosyltransferase 4 precursor
168. Egl nine homolog 2
169. EH-domain containing protein 1
170. Elongation factor 1-alpha 1
171. Elongation factor 1-alpha 2
172. Embrionic Gower Ii carbonmonoxy hemoglobin F chain
173. Endoplasmic reticulum protein ERp29 [precursor]
174. Enhancer protein
175. Enolase 2
176. ENSEMBL:ENSP00000296811 Tax_Id=9606
177. ENSEMBL:ENSP00000309219 Tax_Id=9606
178. ENSEMBL:ENSP00000311603 Tax_Id=9606
179. Eosinophil granule major basic protein precursor
180. Epsilon globin
181. Epsilon polypeptide, mono-oxygenase activation protein
182. Erythroblast membrane-associated protein
183. Erythrocyte 55 kDa membrane protein
184. Erythrocyte band 3 membrane protein, anion transporter

Table 1. (Continued)

185. Erythrocyte band 4.1 membrane protein
186. Erythrocyte band 7 integral membrane protein (Stomatin) (Protein 7.2B)
187. Erythrocyte membrane protein band 4.2 (Pallidin)
188. Erythroid membrane-associated protein
189. Erythroid protein 4.1isoform B
190. Esterase D
191. Eukaryotic initiation factor 4A-I or Eukaryotic initiation factor 4A-II
192. Eukaryotic initiation factor 5A
193. Eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa
194. Eukaryotic translation initiation factor 2C 2
195. Eukaryotic translation initiation factor 4B
196. Exportin 7
197. Ezrin
198. F-actin capping protein alpha-1 subunit (CapZ alpha-1)
199. F-actin capping protein beta subunit (CapZ beta)
200. Fascin
201. Fatty acid synthase
202. F-box only protein 7
203. FK506-binding protein 1A
204. FK506-binding protein 2 precursor
205. Flavin reductase
206. FLJ00257 protein
207. Flotillin 2
208. Flotillin-1
209. FtsJ homolog 3
210. Fumarate hydratase, mitochondrial precursor
211. Fumarylacetoacetase
212. G protein pathway suppressor 1 isoform 2
213. Galactosylgalactosylxylosylprotein-3-beta-glucurono-syltransferase 2
214. Galectin-3 (Galactose-specific lectin 3)
215. GDP-L-fucose synthetase
216. gene rich cluster, C3f gene
217. Glucose transporter
218. Glucosidase II beta subunit precursor
219. Glutamate- -cysteine ligase regulatory subunit
220. Glutamate-cysteine ligase, catalytic subunit
221. Glutaredoxin (thioltransferase)
222. Glutathione reductase, mitochondrial precursor
223. Glutathione S-transferase P
224. Glutathione synthetase
225. Glutathione transferase
226. Glyceraldehyde 3-phosphate dehydrogenase
227. Glycophorin A
228. Glycophorin A precursor
229. Glycophorin B
230. Glycophorin C
231. Glycophorin C, isoform 1
232. Glycophorin Erik (STA) precursor
233. GMP reductase
234. GOLGA7 protein
235. gp25L2 protein
236. GTP-binding nuclear protein RAN
237. Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3
238. Guanine nucleotide binding protein (G protein), q polypeptide
239. Guanine nucleotide-binding protein G(i), alpha-2 subunit
240. Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-5 (like) subunit

Table 1. (Continued)

241. Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 2
242. Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 1
243. Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 1 (Transducin beta chain 1)
244. Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 2 (Transducin beta chain 2)
245. Guanine nucleotide-binding protein G(S), alpha subunit
246. Guanine nucleotide-binding protein, alpha-13 subunit
247. Guanine nucleotide-binding protein, alpha-14 subunit
248. Heat shock 70 kDa protein 1
249. Heat shock 70 kDa protein 4
250. Heat shock 70 kDa protein 8
251. heat shock 90kDa protein 1, alpha
252. Heat shock cognate 71 kDa protein
253. Heme binding protein 1
254. Hemoglobin alpha chain
255. Hemoglobin gamma-A and gamma-G chains
256. Hemoglobin gamma-G
257. Hemoglobin mu chain
258. Hemoglobin, ß chain
259. HGTD-P
260. High mobility group protein 1 - AMIGO2 protein (Amphoterin)
261. Histone H2A.F/Z variant, isoform 1
262. Histone H4
263. Horf6, a novel human peroxydase enzyme, chain A
264. Hsc70-interacting protein
265. Hydroxyacylglutathione hydrolase
266. Hypothetical protein CGI-109 precursor
267. Hypothetical protein DKFZp564D0478
268. Hypothetical protein DKFZp564E227
269. Hypothetical protein DKFZp564J0863
270. Hypothetical protein DKFZp686H13163
271. Hypothetical protein DKFZp686L1653
272. Hypothetical protein DKFZp761K0511
273. Hypothetical protein DKFZp762A227
274. Hypothetical protein FLJ12878
275. Hypothetical protein FLJ14347
276. Hypothetical protein FLJ14877
277. Hypothetical protein FLJ16766
278. Hypothetical protein FLJ20187
279. Hypothetical protein FLJ25678
280. Hypothetical protein FLJ31842
281. Hypothetical protein FLJ32597
282. Hypothetical protein FLJ39623
283. Hypothetical protein FLJ40269
284. Hypothetical protein FLJ42537
285. Hypothetical protein FLJ45139
286. Hypothetical protein FLJ45640
287. Hypothetical protein KIAA0153
288. Hypothetical protein MGC10204
289. Hypothetical protein MGC34680
290. Hypothetical protein ORF9 precursor
291. Hypothetical protein PSEC0098
292. Hypothetical protein XP_061743
293. Hypothetical protein XP_089225
294. Hypothetical protein XP_091430
295. Hypothetical protein XP_091724
296. Hypothetical protein XP_092517
297. Hypothetical protein XP_093639
298. Hypothetical protein XP_095686

Table 1. (Continued)

299. Hypothetical protein XP_095819 84 Far upstream element binding protein
300. Hypothetical protein XP_100510
301. Hypothetical protein XP_100619
302. Hypothetical protein XP_100665
303. Hypothetical protein XP_100925 1
304. Hypothetical protein XP_103707
305. Hypothetical protein XP_106269
306. Hypoxanthine-guanine phosphoribosyltransferase
307. Ig gamma-1 chain C region
308. Ig heavy chain V-V region
309. Ig kappa chain C region
310. Importin 7
311. Importin 9
312. Importin beta-1 subunit
313. Importin beta-3
314. Inosine-5′-monophosphate dehydrogenase 2
315. Insulin gene enhancer protein ISL-1
316. Insulin-like growth factor binding protein 5 [precursor]
317. Integrin beta-2 precursor
318. Intermediate conductance calcium-activated potassium channel protein 4
319. ion transporter protein
320. Junctional adhesion molecule 1 precursor
321. JWA protein regulates intracellular concentrations of taurine and glutamate
322. Kell blood group glycoprotein
323. KIAA0340
324. KIAA0462 protein
325. KIAA0747 protein
326. KIAA0830 protein
327. KIAA0851 protein
328. KIAA1228 protein
329. KIAA1258 protein
330. KIAA1363 protein
331. KIAA1617 protein
332. KIAA1741 protein
333. KIHUA adenylate kinase 63 Bleomycin hydrolase, chain A 64 Complexed carbonic anhydrase I
334. KRT8 protein
335. Lactate dehydrogenase A chain
336. Lactate dehydrogenase B
337. Lactate dehydrogenase H chain
338. Lactotransferrin [precursor]
339. Lactoylglutathione lyase
340. LanC-like protein 1
341. Latexin
342. Leucine-rich repeat, typical subtype containing protein
343. Leukemia inhibitory factor receptor [precursor]
344. Leukotriene A4 hydroLase
345. LFA-3
346. LGALS3 protein
347. Liver phosphofructokinase, isoform a
348. L-lactate dehydrogenase A chain
349. Long-chain-fatty-acid- -CoA ligase 3
350. Low molecular weight phosphotyrosine protein phosphatase
351. L-plas
352. Lutheran blood group glycoprotein precursor
353. Lyn B protein
354. Lysozyme C precursor
355. Malate dehydrogenase, cytoplasmic
356. Malate dehydrogenase, mitochondrial precursor
357. Membrane alanine aminopeptidase precursor

Table 1. (Continued)

358. Membrane associated progesterone receptor component 2
359. Membrane protein p55, erythrocytic, (palmitoylated)
360. Membrane transport protein XK
361. Mesenchymal stem cell protein DSCD75
362. methylenetetrahydrofolate dehydrogenase 1
363. Methylosome subunit pICln
364. MGC16733 protein
365. moesin
366. Monocarboxylate transporter 1
367. Multidrug resistance-associated protein 4
368. Multifunctional protein ADE2
369. Multiple inositol polyphosphate phosphatase-like protein
370. Myosin light chain 1, embryonic muscle/atrial isoform
371. Myosin regulatory light chain 2, nonsarcomeric
372. Myotrophin
373. NADH-ubiquinone oxidoreductase AGGG subunit
374. Neuropathy target esterase
375. Neutrophil gelatinase associated lipocalin [precursor] (NGAL)
376. Nicotinate phosphoribosyltransferase-like protein
377. NipSnap2 protein
378. Nm23 protein
379. Nucleophosmin
380. Nucleoside diphosphate kinase B
381. Nucleosome assembly protein 1-like 1
382. Nucleosome assembly protein 1-like 4
383. NYD-SP11 protein
384. OCP2
385. OTU-like cysteine protease family protein
386. Oxidative-stress responsive 1
387. P47
388. PB39
389. PDCD6IP protein
390. Peflin
391. Peptidyl-prolyl cis-trans isomerase
392. Peptidylprolyl isomerase A
393. peptidylprolyl isomerase A isoform 1
394. Peptidylprolyl isomerase B
395. Perodoxiredoxin 3 isoform b
396. Peroxiredoxin 1
397. Peroxiredoxin 2
398. PHD finger protein 3
399. Phosphatidylethanolamine-binding protein
400. Phosphatidylinositol 4-kinase type II
401. Phosphatidylinositol-4-phosphate 5 kinase, type III
402. Phosphatidylinositol-4-phosphate 5-kinase type II alpha
403. Phosphoglucose isomerase chain A
404. Phosphoglycerete kinase 1
405. Phospholipid scramblase 1
406. Phospholipid scramblase 4
407. Phosphoribosyl pyrophosphate synthetase
408. Phosphoribosyl pyrophosphate synthetase 1-like
409. Phosphoribosyl pyrophosphate synthetase-associated protein 2
410. Phosphoribosylformylglycinamidine synthase
411. PICALM protein
412. Placental ribonuclease inhibitor
413. Platelet-activating factor acetylhydrolase IB beta subunit
414. Poly (A)-specific ribonuclease
415. Polyposis locus protein 1

Table 1. (Continued)

416. Polyubiquitin
417. Potassium channel subfamily K member 5
418. PREDICTED: KIAA1755 protein
419. PREDICTED: similar to D-dopachrome tautomerase (Phenylpyruvate tautomerase II)
420. PREDICTED: similar to dual specificity phosphatase 5
421. PREDICTED: similar to KIAA1693 protein
422. PREDICTED: similar to RAB1B, member RAS oncogene family subset match by rab-10 and 8A
423. PREDICTED: similar to RIKEN cDNA C730027E14
424. PREDICTED: similar to ring finger protein 129
425. PREDICTED: upstream regulatory element binding protein 1
426. Presenilin-associated protein
427. PRKAR2A protein
428. PRO0974
429. Programmed cell death protein 6
430. Prohibitin
431. Prolactin
432. Proliferation-associated factor
433. Prolyl endopeptidase
434. Prostatic binding protein (neuropolypeptide h3
435. Proteasome 26S ATPase subunit 2
436. Proteasome 26S non-ATPase subunit 11
437. Proteasome 26S non-ATPase subunit 5
438. Proteasome 26S subunit, ATPase, 6
439. Proteasome activator complex subunit 1
440. Proteasome activator complex subunit 2
441. Proteasome alpha 2 subunit
442. Proteasome delta chain
443. Proteasome inhibitor PI31 subunit
444. Proteasome subunit alpha type 1 (Proteasome component C2)
445. Proteasome subunit alpha type 3 (Proteasome component C8)
446. Proteasome subunit alpha type 4 (Proteasome component C9)
447. Proteasome subunit alpha type 5
448. Proteasome subunit alpha type 6 (Proteasome iota chain)
449. Proteasome subunit alpha type 7 (Proteasome subunit RC6-1)
450. Proteasome subunit beta type 1
451. Proteasome subunit beta type 2 (Proteasome component C7-I)
452. Proteasome subunit beta type 3 (Proteasome theta chain)
453. Proteasome subunit beta type 4
454. Proteasome subunit beta type 5 [precursor] (Proteasome epsilon chain)
455. Proteasome subunit beta type 7 precursor
456. protein (peptidyl-prolyl cis/trans isomerase) NIMA-interacting, 4 (parvulin)
457. Protein arginine N-methyltransferase 5
458. Protein BAP28
459. Protein C14orf102
460. Protein C2orf4
461. Protein disulfide-isomerase A6 precursor
462. Protein disulfide isomerase A3
463. Protein disulfide isomerase A3 [precursor]
464. Protein disulfide-isomerase A2 precursor
465. Protein FAM38A
466. Protein phosphatase 2A, regulatory subunit B'

Table 1. (Continued)

467. Protein-L-isoaspartate(D-aspartate) O-methyltransferase (EC 2.1.1.77) (Protein-beta-aspartate methyltransferase) (PIMT)
468. PSMC3 protein
469. Purine nucleoside phosphorylase
470. Puromycin-sensitive aminopeptidase
471. Putative RNA-binding protein 3
472. RAB 35, RAS oncogene family
473. Rab GDP dissociation inhibitor alpha
474. Rab GDP dissociation inhibitor beta
475. Rabphilin-3 A-integrating protein
476. RAD23 homolog A
477. Radixin
478. Ral A binding protein
479. RAP1A, member of RAS oncogene family or
480. RAP1B
481. RAP2B, member of RAS oncogene family
482. Ras-related C3 botulinum toxin substrate 1
483. Ras-related C3 botulinum toxin substrate 4
484. Ras-related protein Rab-10
485. Ras-related protein Rab-14
486. Ras-related protein Rab-2 (A or B)
487. Ras-related protein Rab-21
488. Ras-related protein Rab-33B
489. Ras-related protein Rab-35
490. Ras-related protein Rab-5B
491. Ras-related protein Rab-5B TAMNVNDLFLAIAK unique
492. Ras-related protein Rab-5C
493. Ras-related protein Rab-6B
494. Ras-related protein Rab-8A
495. Ras-related protein Rab-8B
496. Ras-related protein Ral-A
497. Ras-related protein RAP-1A
498. Ras-related protein Rap-1b
499. Ras-related protein Rap-2b
500. Ras-related protein RAP-2B
501. RECS1 protein homolog
502. Red cell acid phosphatase 1, isozyme S
503. Reticulon protein 3
504. Rh blood CE group antigen polypeptide
505. Rh blood D group antigen polypeptide
506. RhD protein
507. Rhesus blood group-associated glycoprotein
508. Rhesus C/E antigens
509. Rhesus D category VI type III protein
510. Rho-GTPase-activating protein 1
511. Ribonuclease/angiogenin inhibitor or placental ribonuclease inhibitor
512. Ribonucleoside-diphosphate reductase large subunit
513. Ribose-phosphate pyrophosphokinase I (PRS-I)
514. RP42 protein
515. RuvB-like 2
516. S100 calcium-binding protein A3 2
517. S100 calcium-binding protein A4
518. Sec1 family domain containing protein 1
519. Secretory carrier-associated membrane protein 4
520. Semaphorin 7A precursor
521. Serine/threonine protein phosphatase 2A,
522. Serine/threonine protein phosphatase 5
523. Serine/threonine-protein kinase VRK1
524. Serum albumin precursor
525. SH3 domain-binding glutamic acid-rich-like protein 3
526. Similar to 26S protease regulatory subunit 8

Table 1. (Continued)

527. Similar to Adenosylhomocysteinase
528. Similar to adhesive plaque matrix protein precursor
529. Similar to ankyrin 1
530. Similar to Bifunctional purine biosynthesis protein PURH
531. Similar to Calpain inhibitor
532. Similar to discs, large (Drosophila) homolog 3 (neuroendocrine-dlg)
533. Similar to DJ776F14.1
534. Similar to DJ776F14.1
535. Similar to Erythrocyte cytosolic protein of 51 kDa, EC
536. Similar to Erythrocyte cytosolic protein of 51 kDa, EC
537. Similar to fibronectin type 3 and SPRY domain-containing protein
538. similar to FKSG30
539. Similar to flotillin
540. Similar to flotillin 2
541. Similar to glyceraldehyde-3-phosphate dehydrogenase
542. Similar to glycophorin A
543. Similar to Lutheran blood group
544. Similar to olfactory receptor-like protein
545. Similar to oxidized protein hydrolase
546. Similar to PMMLP
547. Similar to prostatic binding protein
548. Similar to proteasome subunit type 7
549. Similar to proteasome subunit, type, 3
550. Similar to RAS-related protein RAB-15
551. Similar to RAS-related protein RAL-A
552. Similar to Ribose 5-phosphate isomerase
553. Similar to RIKEN cDNA 1500009M05 gene
554. similar to RIKEN cDNA 4732495G21 gene HQGVMVGMGQKDCYVGDEAQSK unique
555. Similar to Selenium-binding protein 1
556. Similar to Ser/Thr-rich protein T10 in DGCR region
557. Similar to suppression of tumorigenicity 13
558. Similar to SWAP-70
559. Similar to Transaldolase
560. Similar to tropomyosin
561. Similar to tropomyosin 4
562. Similar to ubiquitin conjugating enzyme E2L
563. Similar to valosin-containing protein
564. Site specific mutant of carbonic anhydrase Ii, chain A 28 Hemoglobin -G chain
565. Small nuclear ribonucleoprotein D3 polypeptide
566. S-methyl-5-thioadenosine phosphorylase
567. Sodium/potassium-transporting-ATPase-alpha-2 precursor
568. Solute carrier family 1 (Glutamate transporter), member 7
569. solute carrier family 19 member 1 isoform b
570. Solute carrier family 2 (facilitated glucose transporter), member 1
571. Solute carrier family 2, facilitated glucose transporter, member 3
572. Solute carrier family 2, facilitated glucose transporter, member 4 ERPLSLLQLLGSR unique
573. Solute carrier family 27 (Fatty acid transporter), member 4
574. Solute carrier family 29 (nucleoside transporter), member 1
575. Solute carrier family 40, member 1
576. Sorbitol dehydrogenase
577. Spectrin alpha chain, erythrocyte
578. Spectrin beta chain, erythrocyte

Table 1. (Continued)

579. S-phase kinase-associated protein 1A isoform a
580. SPla/RYanodine receptor SPRY domain containing protein
581. Splice isoform Short of Q9UKU0 Long-chain-fatty-acid--CoA ligase 6
582. Splice isoform 1 2 of P11277 Spectrin beta chain, erythrocyte
583. Splice isoform 1 of P08174 Complement decay-accelerating factor precursor
584. Splice isoform 1 of P59190 Ras-related protein Rab-15
585. Splice Isoform 1 of 26S protease regulatory subunit 6B
586. Splice Isoform 1 of 26S proteasome non-ATPase regulatory subunit 1
587. Splice Isoform 1 of Adenylosuccinate lyase
588. Splice Isoform 1 of Alpha-synuclein
589. Splice Isoform 1 of Clathrin light chain A
590. Splice Isoform 1 of COP9 signalosome complex subunit 2
591. Splice Isoform 1 of Cullin homolog 1
592. Splice Isoform 1 of Cullin homolog 3
593. Splice Isoform 1 of Dynamin 2
594. Splice Isoform 1 of Fibronectin precursor
595. Splice Isoform 1 of Hypothetical protein C14orf9
596. Splice Isoform 1 of Importin-alpha re-exporter
597. Splice Isoform 1 of Long-chain-fatty-acid--CoA ligase 1
598. Splice isoform 1 of O14818 Proteasome subunit alpha type 7
599. Splice isoform 1 of O95292 Vesicle-associated membrane protein-associated protein B/C
600. Splice isoform 1 of P07226 Tropomyosin alpha 4 chain
601. Splice isoform 1 of P08237 6-phosphofructokinase, muscle type
602. Splice isoform 1 of P09493 Tropomyosin 1 alpha chain
603. Splice isoform 1 of P11142 Heat shock cognate 71 kDa protein
604. Splice isoform 1 of P11171 Protein 4.1
605. Splice isoform 1 of P17612 cAMP-dependent protein kinase, alpha-catalytic subunit
606. Splice isoform 1 of P22303 Acetylcholinesterase precursor
607. Splice isoform 1 of P26599 Polypyrimidine tract-binding protein 1
608. Splice isoform 1 of P35611 Alpha adducin
609. Splice isoform 1 of P35612 Beta adducin
610. Splice isoform 1 of P47756 F-actin capping protein beta subunit
611. Splice isoform 1 of P62820 Ras-related protein Rab-1A
612. Splice Isoform 1 of Phosphofurin acidic cluster sorting protein 1
613. Splice Isoform 1 of Potential phospholipid-transporting ATPase IA
614. Splice Isoform 1 of Probable ATP-dependent RNA helicase
615. Splice Isoform 1 of Probable ubiquitin carboxyl-terminal hydrolase FAF-X
616. Splice Isoform 1 Of Proteasome subunit
617. Splice Isoform 1 Of Protein-glutamine gamma-glutamyltransferase

Table 1. (Continued)

618. Splice Isoform 1 Of Pyruvate kinase, isozymes R/L
619. Splice isoform 1 of Q12955 Ankyrin 3 ILGNKATFSPIVTVEPR unique
620. Splice isoform 1 of Q16563 Pantophysin
621. Splice isoform 1 of Q8NB25 Hypothetical protein C6orf60
622. Splice isoform 1 of Q92542 Nicastrin precursor
623. Splice isoform 1 of Q9H254 Spectrin beta chain, brain 3
624. Splice isoform 1 of Q9Y4D1 Disheveled associated activator of morphogenesis 1
625. Splice isoform 1 of Q9Y666 Solute carrier family 12 member 7
626. Splice Isoform 1 Of Serine/threonine-protein kinase WNK1
627. Splice Isoform 1 Of Spectrin alpha chain, brain
628. Splice Isoform 1 Of Thyrotropin receptor precursor
629. Splice Isoform 1 Of Ubiquitin carboxyl-terminal hydrolase 15
630. Splice Isoform 1 Of Ubiquitin carboxyl-terminal hydrolase 5
631. Splice Isoform 1 Of Ubiquitin-conjugating enzyme E2 variant 1
632. Splice isoform 1 or 2 of P05089 Arginase 1
633. Splice Isoform 2 Of Adapter-related protein complex 2 alpha 1 subunit
634. Splice Isoform 2 Of Adapter-related protein complex 2 beta 1 subunit
635. Splice Isoform 2 Of Glucose-6-phosphate 1-dehydrogenase short
636. Splice isoform 2 of P06753 Tropomyosin alpha 3 chain
637. Splice isoform 2 of P11171 Protein 4.1
638. Splice isoform 2 of P60953 Cell division control protein 42 homolog
639. Splice Isoform 2 Of Porphobilinogen deaminase
640. Splice isoform 2 of Q14697 Neutral alpha-glucosidase AB precursor
641. Splice isoform 2 of Q16570 Duffy antigen/chemokine receptor
642. Splice Isoform 2 Of Unc-112 related protein 2
643. Splice isoform 2 or 1 of Q9H3Z4 DnaJ homolog subfamily C member 5
644. Splice isoform 2 or 1 of P35613 Basigin precursor
645. Splice isoform 2B of P01116 Transforming protein p21
646. Splice isoform 4 of P11171 Protein 4.1
647. Splice isoform 4 of Q04656 Copper-transporting ATPase 1
648. Splice isoform A of P15154 Ras-related C3 botulinum toxin substrate 1
649. Splice isoform Alpha-S2 of P63092 Guanine nucleotide-binding protein G(s), alpha subunit
650. Splice isoform APP770 of P05067 Amyloid beta A4 protein precursor
651. Splice isoform B of P20020 Plasma membrane calcium-transporting ATPase 1
652. Splice isoform B or A of Q96RL7 Vacuolar protein sorting 13A
653. Splice isoform CNPI of P09543 2′,3′-cyclic-nucleotide 3’-phosphodiesterase
654. Splice isoform Delexon-17 of P33527 Multidrug resistance-associated protein 1
655. Splice isoform Glycophorin C of P04921 Glycophorin C

Table 1. (Continued)

656. Splice isoform I of P14209 T-cell surface glycoprotein E2 precursor
657. Splice isoform Long of O00182 Galectin-9
658. Splice isoform Long of P05023 Sodium/potassium-transporting ATPase alpha-1 chain precu
659. Splice isoform Long of Q01082 Spectrin beta chain, brain 1
660. Splice isoform Long of Q14773 Intercellular adhesion molecule-4 precursor
661. Splice isoform M1 or M2 of P14618 Pyruvate kinase, isozymes M1/M2
662. Splice isoform OA3-293 of Q08722 Leukocyte surface antigen CD47 precursor
663. Splice isoform RHVIII of P18577 Blood group Rh(CE) polypeptide
664. Splice isoform Short of Q08495 Dematin
665. Splice isoform SNAP-23a of O00161 Synaptosomal-associated protein 23
666. Splice isoform XD of P23634 Plasma membrane calcium-transporting ATPase 4
667. SRPK1a protein kinase
668. Steroid dehydrogenase homolog, Aldo-keto reductase family 1 member C3
669. Stomatin isoform a
670. Stress-induced-phosphoprotein 1
671. Stromal cell-derived receptor-1 alpha
672. Superoxide dismutase
673. Syntaxin 4
674. Syntaxin 7
675. T-complex protein 1, alpha subunit
676. T-complex protein 1, beta subunit
677. T-complex protein 1, delta subunit
678. T-complex protein 1, epsilon subunit
679. T-complex protein 1, eta subunit
680. T-complex protein 1, gamma subunit
681. T-complex protein 1, theta subunit
682. TGF-beta receptor type I [precursor]
683. Thimet oligopeptidase
684. Thioredoxin
685. Thioredoxin domain containing protein 1 precursor
686. Thioredoxin mutant with Cys-73 replaced by Ser 11,
687. Thioredoxin peroxidase 1 (Thioredoxin-dependent peroxide reductase 1)
688. Thioredoxin reductase 1, cytoplasmic
689. Thioredoxin-like protein KIAA1162 precursor
690. Thrombospondin 1 precursor, glycoprotein IV, also in mature RBCs
691. THYRO1000951 protein
692. TIP120 protein
693. Titin
694. titin isoform novex-1
695. Toll-interacting protein
696. TPM1 protein
697. TPR repeat containing protein
698. Transducin chain 1
699. Transducin chain type 2
700. Transforming protein N-Ras
701. Transgelin 2
702. Transglutaminase 2
703. Transitional endoplasmic reticulum ATPase
704. Transketolase
705. Translation initiation factor 2C
706. Translation initiation factor 5A
707. Translocon-associated protein, delta subunit precursor

Table 1. (Continued)

708. transmembrane protein 24
709. Transmembrane protein Tmp21 precursor
710. Triosephosphate isomerase
711. Tripeptidyl peptidase II
712. Tropomodulin 1
713. Tropomyosin 3, cytoskeletal
714. Tubulin alpha-6 chain
715. Tubulin beta-1 chain
716. tubulin, alpha 3
717. Tyrosine-protein kinase JAK2 (Janus kinase 2)
718. UBE3B variant 1
719. Ubiquitin and ribosomal protein S27a,
720. Ubiquitin carboxyl-terminal hydrolase 14
721. Ubiquitin conjugating enzyme E2 variant 1
722. Ubiquitin isopeptidase T
723. Ubiquitin-activating enzyme E1
724. Ubiquitin-conjugating enzyme E2 N
725. Ubiquitin-like protein SUMO-1 conjugating enzyme
726. Ubiquitin-specific protease 7 isoform
727. UDP-glucose:glycoprotein glucosyltransferase 1 precursor
728. UEV1Bs
729. Uncharacterized hematopoietic stem/progenitor cells protein MDS032
730. UniProt/Swiss-Prot:P62805
731. Unknown protein 46,884.2
732. UPF0198 protein CGI-141
733. Urea transporter, erythrocyte
734. Uroporphyrinogen decarboxylase
735. UTP--glucose-1-phosphate uridylyltransferase 1
736. Vacuolar ATP synthase 16 kDa proteolipid subunit
737. Vacuolar ATP synthase catalytic subunit A, ubiquitous isoform
738. Vacuolar ATP synthase subunit d
739. Valosin-containing protein
740. Vesicle trafficking protein SEC22b
741. Vesicle-associated membrane protein 2 (synaptobrevin 2)
742. Vesicle-associated membrane protein-associated protein A
743. Vesicular integral-membrane protein VIP36 precursor
744. Vesicular-fusion protein NSF
745. WNT-3 proto-oncogene protein [precursor]
746. Xaa-Pro dipeptidase
747. XRP2 protein
748. Zinc finger protein
749. Zinc transporter 1
750. Zn-finger, RING domain containing protein
751. Zona pellucida binding protein

shown in Table 1, we now know the identity of 751 proteins through proteomic technology. We have a long way to go to identify all RBC proteins and to know their complete sequence, PTMs, and number of copies per RBC. But as will be discussed in the following section, we can begin to assemble a preliminary interactome from the currently known RBC protein composition.

The Human RBC Interactome

To better understand how disease processes, such as sickle cell anemia, affect the functions of erythrocyte protein

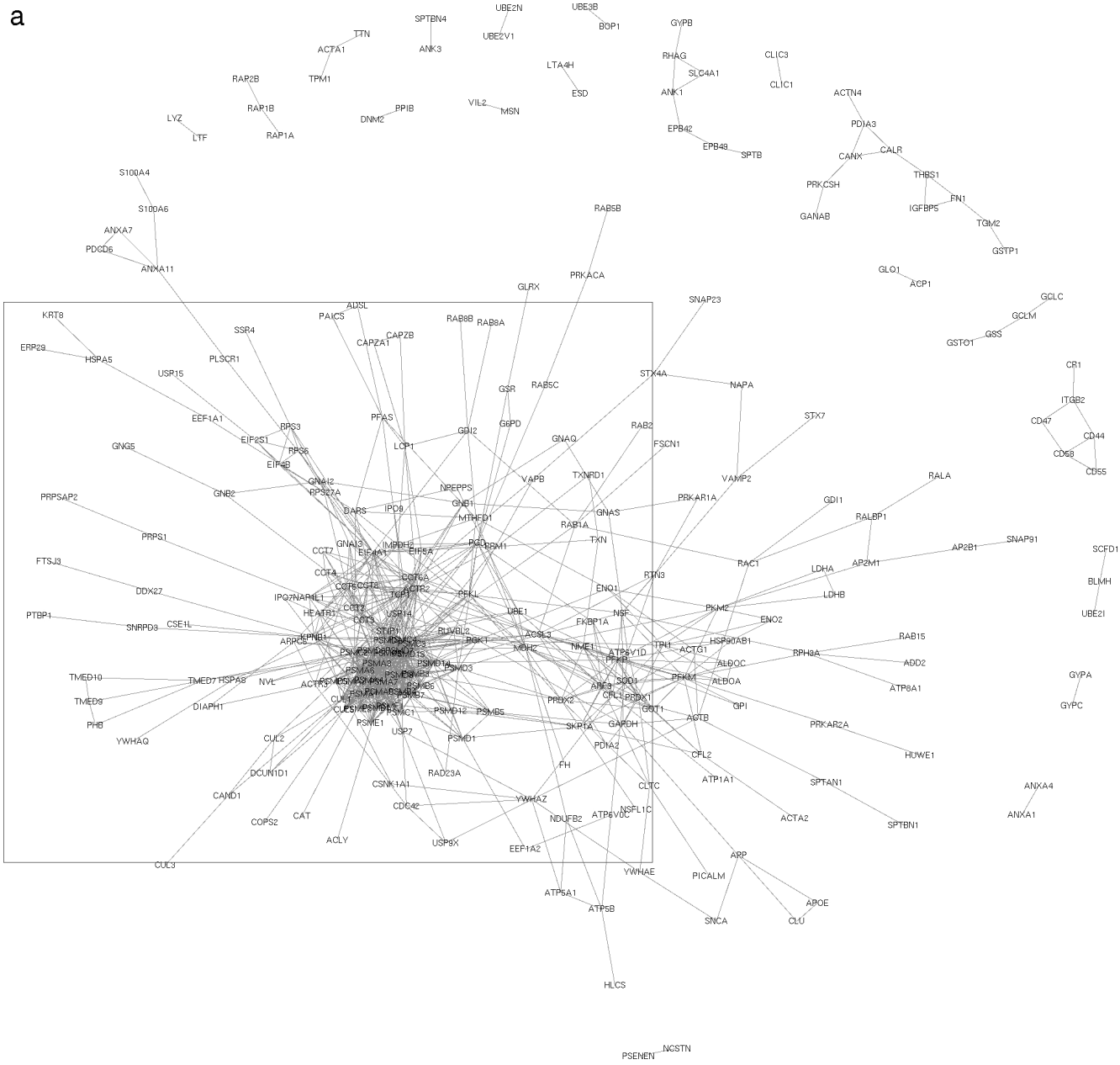

Figure 1a. Preliminary RBC Interactome Network
Nodes give gene symbols of erythrocyte proteins and lines show interactions between proteins. The ROD box shows the region of the network that contains chaperonin, heat shock, and proteasome proteins along with their interacting partners and is expanded in 1b.

networks, it is important to examine the interrelationships that exist among these proteins. A first step in this approach is to extract protein interaction information from databases that are currently available, and to express this information as a network graph. Compilation of such information is far from complete, and the information that is available contains false positives (listed interactions that don't really exist in a particular organism) and false negatives (existing interactions that have not been discovered). Nevertheless, significant progress has been made to extend and improve protein interaction databases, and to include in these databases measures of confidence for the interactions.

The preliminary human erythrocyte interactome network presented here was derived from protein interaction data obtained from the Unified Human Interactome (UniHI) (22). To reduce the likelihood of including false positives in these networks, we only included interactions among erythrocyte proteins that had a Spearman correlation of at least 0.3. The Spearman correlation reported by UniHI is derived from gene expression experiments and represents a measure of confidence for the interaction (23). We selected a lower bound of 0.3 to include an interaction in the network since that threshold represents the 60th percentile of all returned interaction correlations and hence gives moderately

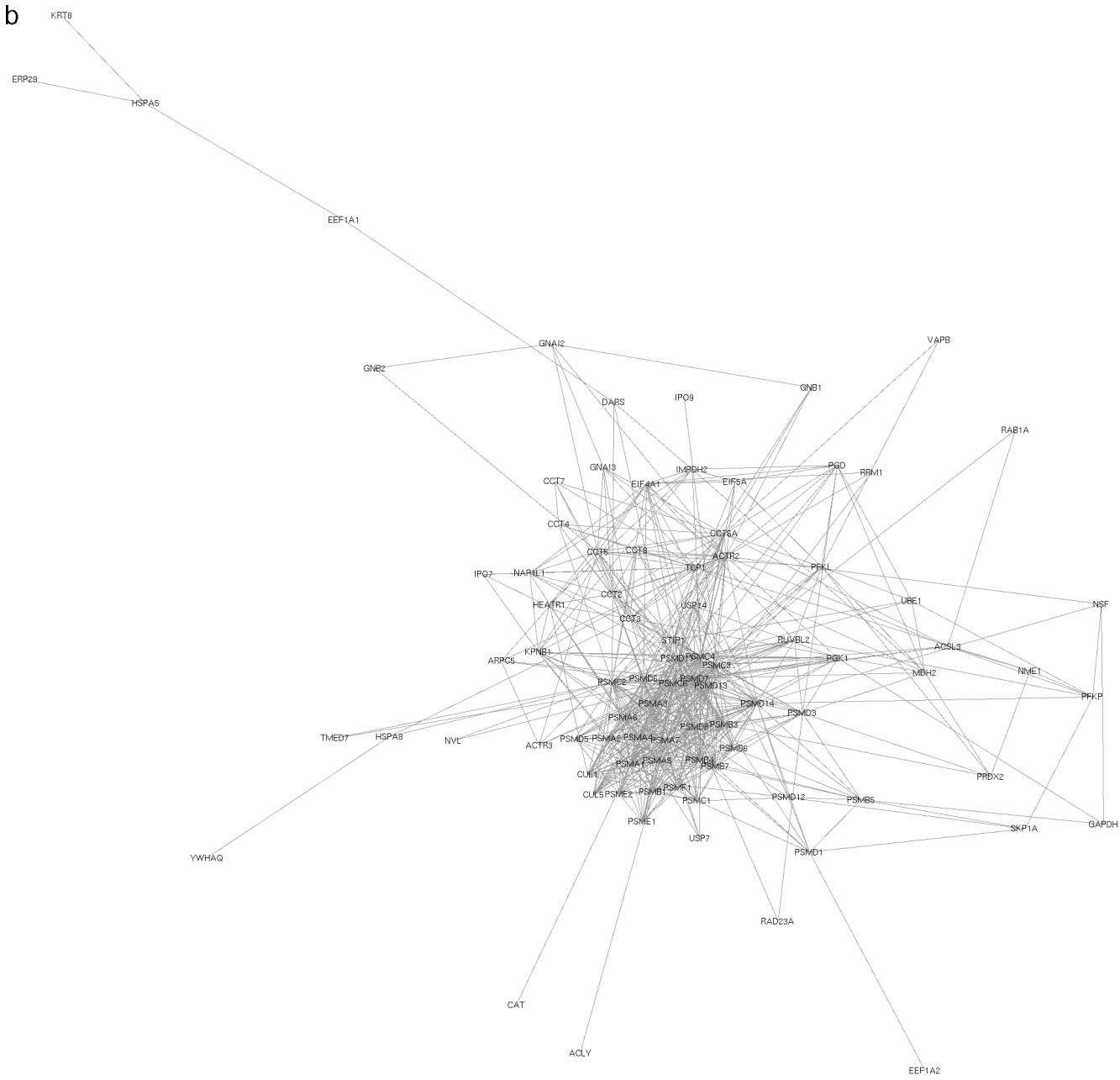

Figure 1b.

strong confidence that the interaction is not a false positive. The graphics were generated with R (24) and the igraph package (25).

The resulting protein-protein interaction (PPI) network is depicted in Figure 1a. The nodes of the network are gene ID's of the erythrocyte proteins, and the connecting line segments represent interactions between the corresponding proteins. Some of the erythrocyte proteins have no interactions between them and any other erythrocyte proteins that satisfied the threshold for retention. These proteins are omitted on all but the last figure. The structure of the network is visibly irregular with sparsely connected subgraphs, densely connected subgraphs, several clusters containing only a few nodes, and one large cluster containing the majority of proteins. The boxed region in Figure 1a is expanded and displayed in Figure 1b. This region is defined by the proteasomal, chaperonin, and heat shock proteins along with their immediate neighbors in the network. We refer to this box as the **R**epair **O**r **D**estroy (**ROD**) Box.

The ROD box contains proteins that utilize the energy of ATP hydrolysis to fold nascent proteins or refold damaged proteins (heat shock proteins and chaperonins). As mature Red Blood Cells are thought not to synthesize nascent proteins only the latter function is relevant to this discussion. The ROD box also contains proteins involved in

the proteasomal degradation of ubiquitinated proteins (ex proteasomal subunits). The recent demonstration, by our laboratory (18), that proteasomes are present in mature RBCs raises the important question of whether ubiquitin dependent proteolytic degradation exists in RBCs. As can be seen in the Figure 1B, the proteins involved in repair of damaged proteins or destruction of proteins beyond repair are interacting within the ROD Box.

RBC Proteomics as it Relates to Human Disease

Protein profiling, where the protein content is compared in control and diseased tissue, has become a powerful tool for diagnosis and the identification of new therapeutic targets. Early proteomic protein profiling studies utilized image analysis software to compare protein stained two dimensional gels of disease versus control cells and tissues. An example of this approach can be found in the comparison of RBC membrane proteins where control subjects are compared to participants with type 2 diabetes (26). In this study by Jiang et al (26), they found 27 spots upregulated and 15 protein spots down regulated in RBC membranes from type 2 diabetes patients. Those proteins that demonstrated significant change were identified by in-gel trypsin digestion followed by MALDI TOF Mass spectrometry. The lipid raft protein, flotilin 1, was increased 4.8 fold; syntaxin 1 C was decreased 8.3 fold; and arginase was upregulated by 37.2 fold. Jiang et al speculated that decreases in the target membrane fusion protein, syntaxin 1C, could lead to reduction in translocation of the Glut 4 glucose transporter to the RBC plasma membrane during erythropoesis in type 2 diabetes (26). Further they suggested that the large increase in arginase could down regulate nitric oxide (NO) production, by creating a competition for the L-arginine substrate with NO synthase.

While this approach of image analysis of protein stained two dimensional gel has been historically valuable, as exemplified above, it is limited by the fact that protein migration from gel to gel is not identical. As a result, the technique of two dimensional difference gel electrophoresis (2D DIGE) was developed (27). In this approach the proteins in two samples (diseased versus control) are first labeled with Cy3 and Cy5 fluorescent dyes (sometimes using a Cy 2 labeled reference standard), then combined prior to two dimensional electrophoresis. The two different dyes can then be detected and fluorescent intensities determined for Cy 3 and Cy 5 for every protein spot. This approach was used by our laboratory for performing protein profiling on RBC membranes derived from homozygous (SS) Sickle Cell Disease (SCD) patient versus control (AA) (28). From over 500 fluorescent protein spots we found 38 that demonstrated $\geq$ 2.5 fold increase in sickle cell RBC membranes and 11 protein spots that demonstrated $\geq$ 2.5 fold decrease in SS RBC membranes (28). We chose a 2.5 fold threshold because when comparing RBC membrane proteins between control subjects we found that 99.8% of all determined ratios were within a 2.5 fold difference from the expected value of 1.0. We identified 44 protein spots by trypsin in–gel digestion followed by nanoLC-MS/MS tandem mass spectrometry. These 44 identifications included 22 unique proteins with various PTMs (28). As shown in Table 2, the proteins that were changing fell into five protein groups. Three of these groups were related to the extreme oxidative stress observed in SS RBCs: Protein repair participants (heat shock 70 kDa protein isoforms, chaperonin containing TCP1 subunits and T-complex protein 1 delta); proteasome components (proteasome 26S ATPase subunit 6, proteasome alpha subunit 1 isoform 1 and proteasome beta 1 subunit); and scavengers of oxygen radicals (catalase, peroxiredoxin 1 and peroxiredoxin 3 isoform a) (28). All of these proteins, except one (chaperonin subunit zeta 1), were increased greater than or equal to 2.5 fold in sickle cell RBC membranes. This demonstrates that the RBC and its erythropoetic precursors are launching an adaptive response against the elevated levels of oxygen radicals and diminished reduced glutathione found within SS RBCs (28). The two major lipid raft components in RBCs, flotilin 1 and stomatin, had four different protein spots all diminished in SS RBC membranes (28). This suggested that when unstable sickle cell RBC membranes bleb off vesicles, these vesicles are enriched in cholesterol and sphingolipid rich lipid rafts. Finally various post-translationally modified forms of spectrin membrane skeleton components changed by $\geq$ 2.5 fold in sickle cell RBC membranes (28). Two known proteolytic fragments of ankyrin (29–32) were present in greater quantity in SS RBC membranes; six PTM forms of proteins 4.1 were increased while one was decreased; and one protein spot each from protein 4.9 and tropomyosin 3 were decreased (Table 2). The value of 2D DIGE is that it allows us to detect multiple post translationally modified forms of the same protein, while the limitation of using two dimensional gels to separate and quantitate proteins has already been discussed above. We followed this initial study on sickle cell RBC membranes, with the use of cleavable isotope coded affinity tag (cICAT) technology coupled to tandem mass spectrometry to study the sickle cell versus control RBC membrane skeletal components (33). In this study, we found that the cICAT technology and changes in control populations led to only 20% variance in membrane skeleton protein ratios. We also found that these ratios from SS versus AA samples were not significantly different from 1 for spectrin α subunit, spectrin β subunit, β-actin, and protein 4.1 (33). Therefore, while various PTM forms of proteins 4.1 change in SS versus AA RBC membranes as measured by 2D DIGE (28), the total content of protein 4.1 does not vary in the SS versus AA membrane skeleton as measured by cICAT technology (33). This is an example of why these two technologies are complementary and it is useful to utilize both. cICAT protein profiling is able to measure ratios of total protein 4.1 in SS versus AA samples with great precision ($\sim$20% variation in the technique); while 2D

Table 2. The altered proteins in SS RBC membranes placed in functional categories

Protein Group	#	Identified Protein	Spot #[a] Decrease	Spot #[a] Increase
Spectrin & Actin Accessory Proteins	1	Ankyrin 1, erythrocyte splice form 2 or Ankyrin 1 isoform 1; ankyrin R		1, 2
	2	Protein 4.1 (band 4.1)	3	4–9
	3	Dematin (band 4.9)	30	
	4	Tropomyosin 3	33	
Protein Repair Participants	5	Heat shock 70kDa protein 8 isoform 1		10
	6	Heat shock 70kDa protein 1		11, 12
	7	Chaperonin containing TCP1, subunit 2		13,16,29
	8	Chaperonin containing TCP1, subunit 6A (zeta 1)	15	14
	9	Chaperonin containing TCP1, subunit 7 (eta)		19–22
	10	T-complex protein 1, delta		19–21
Lipid Rafts	11	Flotilin 1	23,24	
	12	Stomatin isoform a (band 7.2)	35,36	34
Proteasome Components	13	Proteasome 26S ATPase subunit 6		31
	14	Proteasome alpha 1 subunit isoform 1		32
	15	Proteasome beta 1 subunit		38
Scavengers of Oxygen Radicals	16	Catalase		17
	17	Peroxiredoxin 3 isoform a precursor		37
	18	Peroxiredoxin 1		38
Other Categories	19	EH-domain containing 1; testilin	15	
	20	Tubulin alpha 6		18
	21	ATP-synthase		25–28
	22	RAB-8b protein		38

[a] We indicate spot numbers from Kakhniashvili et al (28), figure 2 and table 2, where the protein decreased or increased by at least 2.5 fold when comparing sickle cell versus control RBC membrane protein content.

DIGE is able to determine ratios of specific post-translationally modified forms of protein 4.1 in SS versus AA samples, but with lower precision (2.5 fold changes). As the cICAT technique labels cysteine containing peptides only (34), it will miss many post-translational changes (33) that will not be missed by 2D DIGE (28). More recent use of iTRAQ (isobaric technique for relative and absolute quantitation) reagents overcome this cICAT related problem; as it labels terminal amines and ε–amino groups and therefore labels all tryptic peptides (35). It also allows comparisons of up to four distinct samples and overcomes the possibility of cysteine oxidation affecting ICAT results (34, 35). Because of the separate strengths of 2D DIGE, cICAT and iTRAQ technologies, we now apply all three in protein profiling studies.

Recently Prabakaran et al (36) have performed 2D DIGE analyses on RBC proteins derived from 20 schizophrenic subjects versus 20 controls. As they were studying whole RBC protein content, they utilized IEF fractionation to limit the hemoglobin content by discarding the protein fraction in the isoelectric point (pI) range of 6.2 – 10.0. This, of course, causes the loss of many other proteins that fell into this broad range of pI. The elimination of hemoglobin is an important and difficult problem in RBC proteomics. The use of hemoglobin antibodies is prohibitively costly, and not successful, because of the very high intracellular hemoglobin content in the RBC. We have found that gel filtration chromatography is an acceptable compromise between diminishing the hemoglobin content while limiting the loss of other proteins (16). Despite discarding this broad pI range hemoglobin containing fraction, Prabakaran et al observed $\sim$ 1200 fluorescent RBC protein spots, 49 of which were significantly changed in the samples from schizophrenia patients ($p \leq 0.05$). After in–gel trypsin digestion and tandem mass spectrometry, 8 unique proteins were found to demonstrate significant change. Two proteins were increased (selenium binding protein and glutathione reductase), while the other six decreased in schizophrenic patients (thioredoxin peroxidase, heat shock 70 kD protein, serum albumin, apolipoprotein A1, erythroid α spectrin and β-actin). The changes in erythroid α spectrin and β-actin probably reflect post-translational changes, and not total content, as described above. Selenium binding protein 1 (SBP1), glutathione reductase and thioredoxin are known to quench reactive oxygen species (ROS), oxidative stress is occurring in brain and RBCs in schizophrenia (36), and SBP1 has previously been suggested as a potential biomarker for schizophrenia. This proteomic study (36) allows the possibility of utilizing RBC proteomics for schizophrenia disease detection.

Although a description of malaria is beyond the scope of this review, we will mention one proteomic study that may supply new targets for antimalarial drugs and vaccines. Flovens et al biotinylated the surface proteins of Plasmo-

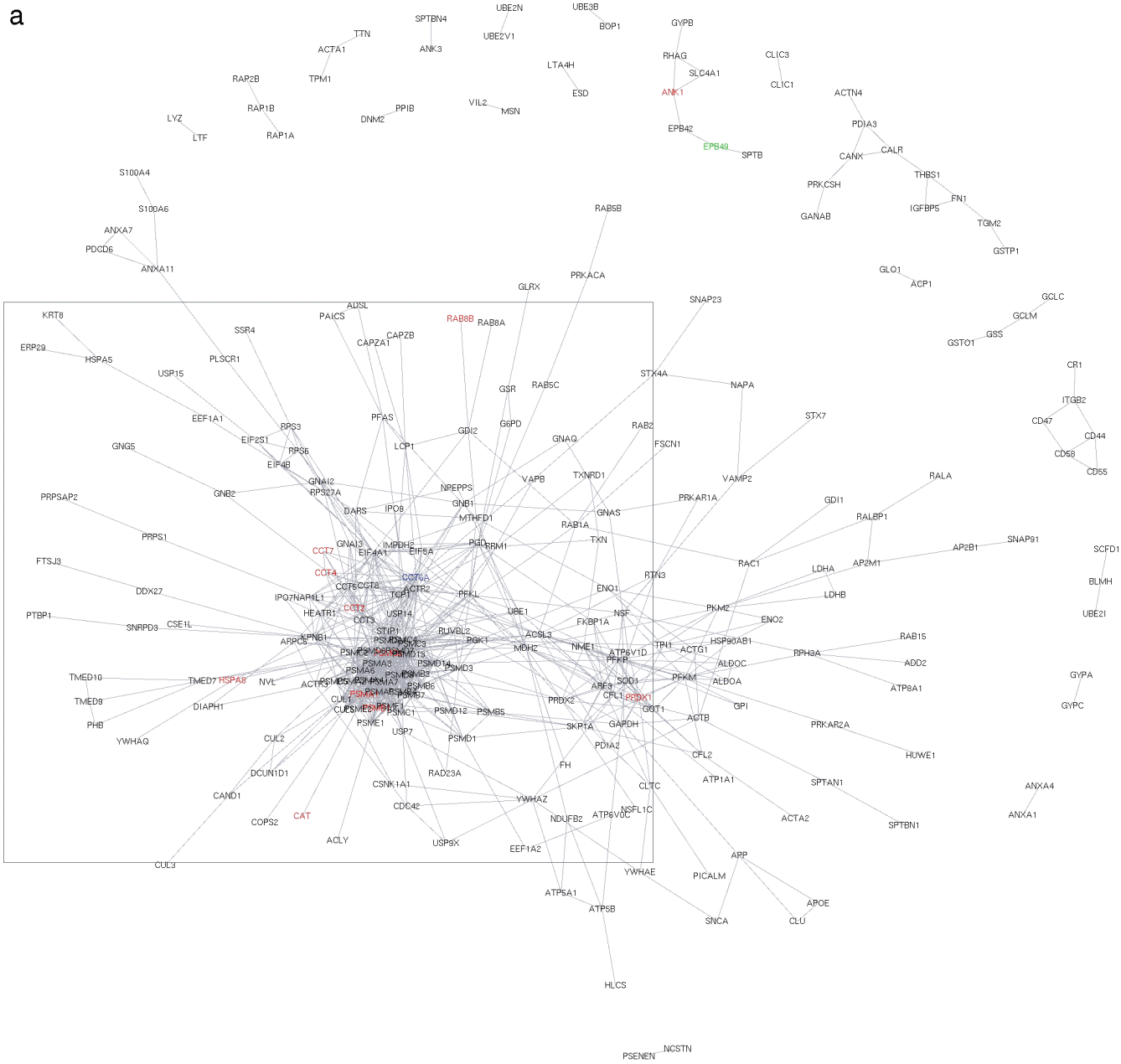

Figure 2a. Preliminary SCD RBC Interactome Network
Same as Figures 1a and 1b with SCD-altered protein symbols colored red: SCD-altered proteins that increased > 2.5-fold; green: SCD-altered proteins that decreased > 2.5-fold; and blue: SCD-altered proteins that had some modified forms that increased > 2.5-fold and others that decreased > 2.5-fold.

dium parasite infected erythrocytes (37). They then used streptavidin affinity chromatography on solubilized infected RBC membranes to isolate parasite encoded proteins on the surface of the infected erythrocyte (PIESPs). Trypsin digestion, MuDPiT separation of peptides, tandem mass spectrometry, and bioinformatic analysis led to identification of 36 candidate PIESPs. The authors characterized two: PIESP1 (154 kD) and PIESP2 (49 kD) and demonstrated both to be associated with knob – like protrusions on the surface of parasite infected RBCs (37). These two proteins, both encoded by single copy genes, may prove to be excellent future therapeutic targets for malaria.

The Altered Interactome in Sickle Cell RBCs

The powerful combination of proteomic and *in silico* techniques, described in this minireview, gives researchers the ability to identify protein content and PTM changes related to a specific disease, or disease severity, or drug and gene therapy treatment. These linked approaches represent the future of clinical identification of specific disease, its

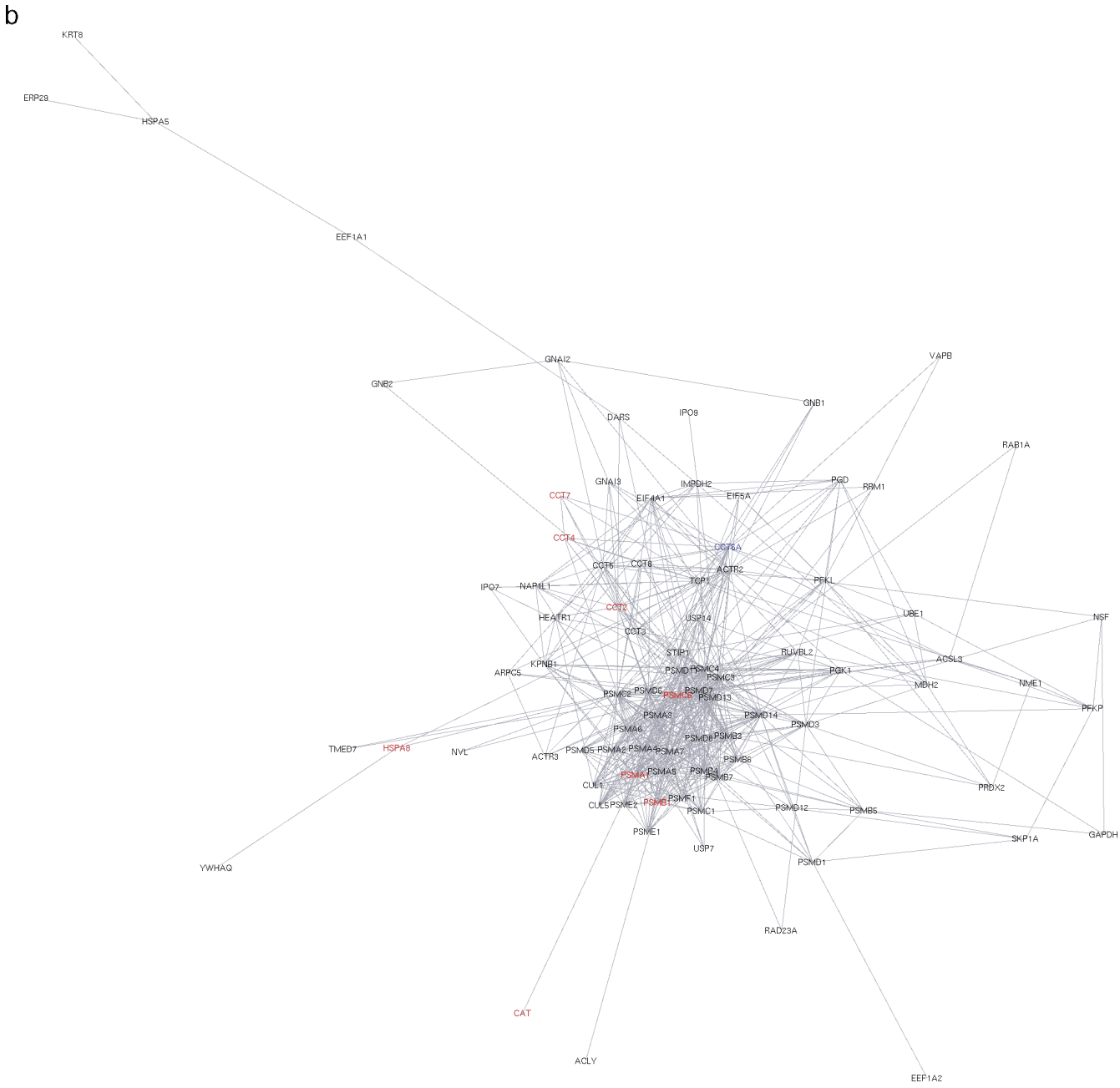

Figure 2b.

current stage or severity, and the effects and side effects of various treatments. We are utilizing protein profiling and interactome maps for all of these purposes in Sickle Cell Disease.

To visualize the impact of SCD on the erythrocyte interactome PPI network, proteins that were significantly altered in SCD patients are marked on the network graphs shown in Figures 2a and 2b. Figure 2a is the same as Figure 1a with SCD-altered protein gene symbols that are colored red, green, or blue. Red symbols represent proteins that increased at least 2.5-fold among SCD subjects compared to non-SCD subjects, green symbols represent proteins that decreased at least 2.5-fold among SCD subjects compared to non-SCD subjects, and blue symbols represent proteins with some isoforms or PTM forms that showed a decrease of at least 2.5-fold, and with other isoforms or PTM forms that showed an increase of at least 2.5-fold. Figure 2b is the ROD Box proteins with SCD-altered protein symbols colored as in Figure 2a.

Figure 2a illustrates how the altered proteins interact with other proteins in the complete network. Some of the former have a large number of connections with other proteins in the network (such as proteasomal subunits PSMA1, PSMB1, PSMC6 and chaperonins CCT 2, 4, 6A,

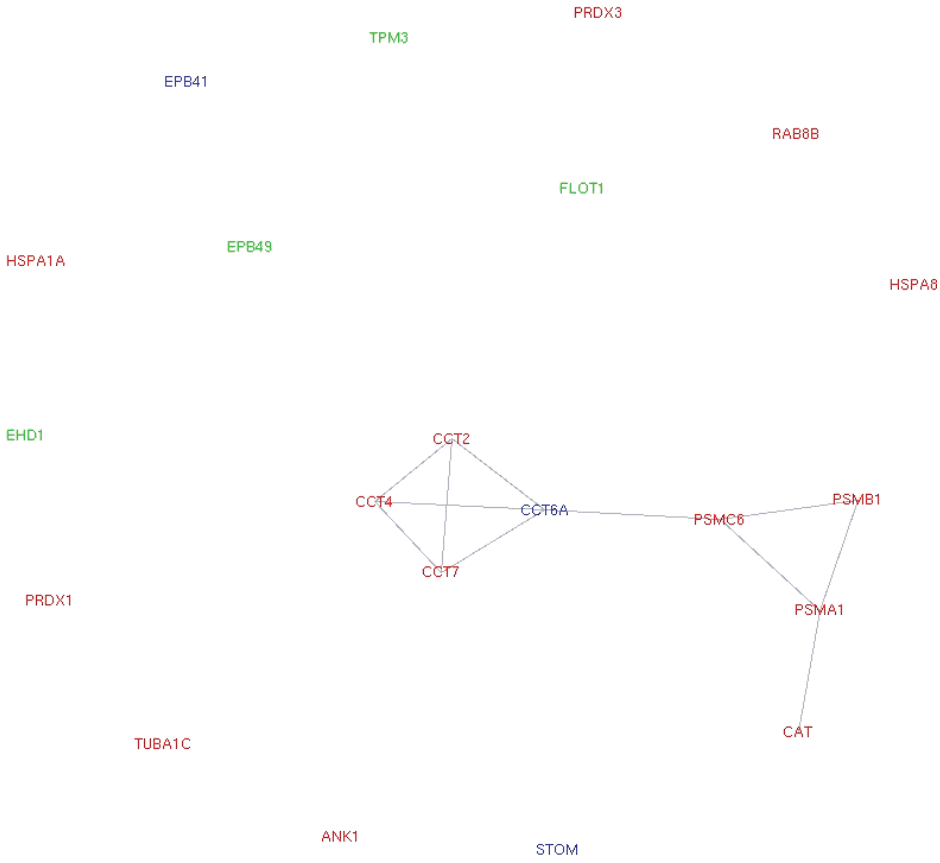

Figure 3. PPI network of SCD-altered proteins
Lines represent interactions among SCD-altered proteins only.

7), whereas some others are relatively isolated with very few links to the rest of the network (ex. catalase CAT). It can be seen from Figure 2b that 9 of the SCD altered proteins are closely connected to the subset of proteasome, chaperonin, and heat shock proteins (including catalase). That is, these 9 altered proteins are interacting partners with at least one protein in that subset. Therefore, many of the changes in the RBC membrane in SCD are found in the ROD Box.

Figure 3 shows that, with respect to their interactions with each other, SCD-altered proteins form two distinct groups: one small, connected group and the other containing pair-wise disconnected proteins. The connected component of this SCD interaction graph has an interesting structure from a graph theoretic standpoint as it is composed by three disjoint cliques of sizes four, three, and one, with the three-clique connecting the one-clique with the four-clique, where a clique is a complete subgraph (a subset of vertices of the graph that are all connected to each other).

Thus, the connected component is a linear sequence of cliques. We also observe from Figure 2 that certain SCD-altered proteins, such as ankyrin 1 (ANK1), dematin erythrocyte membrane protein band 4.9 (EPB49), heat shock protein 8 (HSPA8), proteasome subunit alpha type 1 (PSMA1) and proteasome subunit beta type 1 (PSMB1), are articulation points of the entire PPI network. If we remove such a point, then a previously connected component of the network gets partitioned into two or more disjointed subgraphs. Similarly, it can be seen in Figure 3 that chaperonin containing TCP1, subunit 6A isoform a (CCT6A), proteasome subunit alpha type 1 (PSMA1) and proteasome 26S subunit, and ATPase 6 (PSMC6) are the articulation points of the SCD network. For example, removing CCT6A from this network would separate the interconnected chaperonin proteins from the interconnected proteasomal proteins.

Final Thoughts

The combination of proteomic and *in silico* approaches allows one to not only identify disease and/or drug related changes in the proteome, but to predict the changes in the protein interactome network that will cause disease related change of function. As an example, we are currently utilizing this unified approach to determine how proteomic changes in the protein profile of RBCs, WBCs and plasma are linked to SCD severity. We are also utilizing this approach to determine hydroxy urea dependent changes in the SCD RBC membrane protein profile (38). The result of this study has already demonstrated increases in catalase, chaperonin containing TCP1, aldehyde dehydrogenase and p55. Catalase (CAT) and chaperonin containing TCP1

(CCT6A) are in the ROD box. Aldehyde dehydrogenase and p55 are not in the ROD box and the confidence scores for their interactions were not sufficiently high.

The great sensitivity of the modern mass spectrometers indicates that contamination of an RBC preparation by even low levels of reticulocytes or WBCs can be problematic. All previous studies on the RBC proteome have made an attempt to limit this contamination. Reticulocytes can represent 0.1 to 1.0 percent of control RBC preparations, but increases greatly in the case of hematologic disorders such as SCD. We therefore have recently established a rapid, high volume and throughput Percoll density approach which allows the reduction of reticulocytes within control and SCD RBC preparations to less than 3 ppm (39). As all other technologies such as FACS, MACS, and long incubations at 4°C are far less efficacious, our new technology should become the standard within the RBC proteome field (39).

The combination of starting with highly purified reticulocyte free RBC preparations, constantly improving mass spectrometers and data bases, and the use of interactome *in silico* approaches have the field of RBC proteomics and disease ready to explode over the next few years. We hope that this review will help illuminate the path and approach for investigators within the field.

1. Goodman SR, Shiffer K. The spectrin membrane skeleton of normal and abnormal human erythrocytes: A review. The American Journal of Physiology 244(3):C121–141, 1983.
2. Goodman SR, Krebs KE, Whitfield CF, Riederer BM, Zagon IS. Spectrin and related molecules. CRC Critical Reviews in Biochemistry 23(2):171–234, 1988.
3. Bennett V, Lambert S. The spectrin skeleton: From red cells to brain. Journal of Clinical Investigation 87(5):1483–1489, 1991.
4. Hsu YJ, Goodman SR. Spectrin and ubiquitination: A review. Cellular and Molecular Biology (Noisy-Le-Grand, France) Suppl 51:OL801–807, 2005.
5. Dhermy D, Schreìvel J, Lecomte M-. Spectrin-based skeleton in red blood cells and malaria. Current Opinion in Hematology 14(3):198–202, 2007.
6. Palek J, Sahr KE. Mutations of the red blood cell membrane proteins: From clinical evaluation to detection of the underlying genetic defect. Blood 80(2):308–330, 1992.
7. Delaunay J. Genetic disorders of the red cell membrane. Critical Reviews in Oncology/Hematology 19(2):79–110, 1995.
8. Iolascon A, Perrotta S, Stewart GW. Red blood cell membrane defects. Reviews in Clinical and Experimental Hematology 7(1):22–56, 2003.
9. Goodman SR. The irreversibly sickled cell: A perspective. Cellular and Molecular Biology (Noisy-Le-Grand, France) 50(1):53–58, 2004.
10. Gallagher PG. Update on the clinical spectrum and genetics of red blood cell membrane disorders. Current Hematology Reports 3(2):85–91, 2004.
11. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 422(6928):198–207, 2003.
12. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nature Biotechnology 21(3):255–261, 2003.
13. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nature Reviews Molecular Cell Biology 5(9):699–711, 2004.
14. Yates III JR, Gilchrist A, Howell KE, Bergeron JJM. Proteomics of organelles and large cellular structures. Nature Reviews Molecular Cell Biology 6(9):702–714, 2005.
15. Low TY, Seow TK, Chung MCM. Separation of human erythrocyte membrane associated proteins with one-dimensional and two-dimensional gel electrophoresis followed by identification with matrix-assisted laser desorption/ionization-time of flight mass spectrometry. Proteomics 2(9):1229–1239, 2002.
16. Kakhniashvili DG, Bulla Jr. LA, Goodman SR. The human erythrocyte proteome: Analysis by ion trap mass spectrometry. Molecular and Cellular Proteomics 3(5):501–509, 2004.
17. Pickart CM, Vella AT. Levels of active ubiquitin carrier proteins decline during erythroid maturation. J Biol Chem 263:12028–12035, 1988.
18. Neelam, S.N., Kakhniashvili, D.G. and Goodman S.R., The Discovery of Proteasomal Subunits in Human Erythrocytes: Relationship to Sickle Cell Disease. National Sickle Cell Disease Program. Abstract 42, 2006.
19. Tyan Y-, Jong S-, Liao J-, Liao P-, Yang M-, Liu C-, Klauser R, Himmelhaus M, Grunze M. Proteomic profiling of erythrocyte proteins by proteolytic digestion chip and identification using two-dimensional electrospray ionization tandem mass spectrometry. Journal of Proteome Research 4(3):748–757, 2005.
20. Bruschi M, Seppi C, Arena S, Musante L, Santucci L, Balduini C, Scaloni A, Lanciotti M, Righetti PG, Candiano G. Proteomic analysis of erythrocyte membranes by soft immobiline gels combined with differential protein extraction. Journal of Proteome Research 4(4): 1304–1309, 2005.
21. Pasini EM, Kirkegaard M, Mortensen P, Lutz HU, Thomas AW, Mann M. In-depth analysis of the membrane and cytosolic proteome of red blood cells. Blood 108(3):791–801, 2006.
22. Chaurasia G., Iqbal, Y., Hänig, C., Herzel, H., Wanker, E. and Futschik, M. UniHI: an entry gate to the human protein interactome, Nucleic Acids Res. 35 Database issue:D590–604, 2007.
23. Balasubramanian, R., LaFramboise, T., Scholtens, D., and Gentleman, R. A graph theoretic approach to testing associations between disparate sources of functional genomics data. Bioinformatics, 20, 3353–3362, 2004.
24. R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL http://www.R-project.org, 2007.
25. Csardi, G. igraph: Routines for simple graphs, network analysis. R package version 0.3.3. URL: http://cneurocvs.rmki.kfki.hu/igraph, 2006.
26. Jiang M, Jia L, Jiang W, Hu X, Zhou H, Gao X, Lu Z, Zhang Z. Protein disregulation in red blood cell membranes of type 2 diabetic patients. Biochemical and Biophysical Research Communications 309(1):196–200, 2003.
27. Tonge RP, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1(3):377–396, 2001.
28. Kakhniashvili DG, Griko NB, Bulla Jr. LA, Goodman SR. The proteomics of sickle cell disease: Profiling of erythrocyte membrane proteins by 2D-DIGE and tandem mass spectrometry. Experimental Biology and Medicine 2005;230(11):787–92.
29. Goodman SR, Branton D. Spectrin binding and the control of membrane protein mobility. Journal of Supramolecular and Cellular Biochemistry 8(4):455–463, 1978.
30. Yu J, Goodman SR. Syndeins: The spectrin-binding protein(s) of the human erythrocyte membrane. Proceedings of the National Academy of Sciences of the United States of America 76(5):2340–2344, 1979.
31. Siegel DL, Goodman SR, Branton D. The effect of endogenous proteases on the spectrin binding proteins of human erythrocytes. Biochimica Et Biophysica Acta 598(3):517–527, 1980.
32. Wallin R, Culp EN, Coleman DB, Goodman SR. A structural model of human erythrocyte band 2.1: Alignment of chemical and functional

domains. Proceedings of the National Academy of Sciences of the United States of America 81(13 I):4095–4099, 1984.

33. Chou J, Choudhary PK, Goodman SR. Protein profiling of sickle cell versus control rbc core membrane skeletons by icat technology and tandem mass spectrometry. Cellular and Molecular Biology Letters 11(3):326–337, 2006.
34. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology 17(10):994–999, 1999.
35. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular and Cellular Proteomics 3(12): 1154–1169, 2004.
36. Prabakaran S, Wengenroth M, Lockstone HE, Lilley K, Leweke FM, Bahn S. 2-D DIGE analysis of liver and red blood cells provides further evidence for oxidative stress in schizophrenia. Journal of Proteome Research 6(1):141–149, 2007.
37. Florens L, Liu X, Wang Y, Yang S, Schwartz O, Peglar M, Carucci DJ, Yates III JR, Wu Y. Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. Molecular and Biochemical Parasitology 135(1):1–11, 2004.
38. Ghatpande S, Goodman SR. Protein Modifications of the Sickle Cell RBC Membrane Caused by Hydroxyurea Treatment. Sickle Cell Disease Program in press, 2007.
39. Kakhniashvili DG, Hughes KMH, Neelam S, Goodman SR, The Isolation of Reticulocyte Free Human Red Blood Cells For The Proteomic Study of Sickle Cell Disease. Sickle Cell Disease Program in press, 2007.