Proc. Nat. Acad. Sci. USA
Vol. 69, No. 6, pp. 1445–1449, June 1972

Chemical Characterization and Surface Orientation of the Major Glycoprotein of the Human Erythrocyte Membrane

(tryptic peptides/cyanogen bromide peptides/intramembranous particles/receptors/blood-group antigens)

V. T. MARCHESI, T. W. TILLACK*, R. L. JACKSON†, J. P. SEGREST, AND R. E. SCOTT‡

Laboratory of Experimental Pathology, National Institute of Arthritis and Metabolic Diseases, National Institute of Health, Bethesda, Maryland 20014

Communicated by George E. Palade, March 15, 1972

ABSTRACT **The major glycoprotein of the human erythrocyte membrane has been isolated by treatment with lithium di-iodosalicylate and found to be a single polypeptide chain with a molecular weight of about 50,000. This molecule, which is 60% carbohydrate and 40% protein, carries multiple blood-group antigens, the receptors for influenza viruses, and various plant agglutinins. Four unique carbohydrate-containing peptides (α-1, α-2, α-3, and β) are produced by tryptic digestion of the isolated glycoprotein; their order in the molecule has been determined by sequential tryptic digestion of intact erythrocyte membranes and partially digested glycoprotein fragments. Cleavage of the native protein with cyanogen bromide produces five fragments; two of these (C-5 and C-1) contain most of the carbohydrate in the molecule and are derived from the N-terminal half of the polypeptide chain. The nonpolar amino acids of this glycoprotein are located predominantly in the C-terminal fragment (C-2).**

Phytohemagglutinin conjugated to ferritin has been used to map the distribution of glycoprotein receptors over the surfaces of intact erythrocytes by freeze-etching and electron microscopy. This label localizes to sites on the membrane that overlie the intramembranous particles. These findings suggest that the glycoprotein is oriented at the cell surface with its oligosaccharide-rich N-terminal end exposed to the exterior, while its C-terminal segment interacts with other components in the interior of the membrane to form intramembranous particles.

Glycoproteins comprise about 10% of the total protein of the human erythrocyte membrane (1). Their carbohydrate moieties are antigenic determinants (2, 3) and receptors for viruses and plant agglutinins (4), and their sialic acid residues are responsible for most of the negative charge at the cell surface (5).

Glycoproteins have been isolated from erythrocyte membranes with various solvents, such as phenol (6), butanol (7), pyridine (8), sodium dodecyl sulfate (SDS) (9), or formic acid (10). Most of the glycoproteins isolated in these ways are about 60% carbohydrate by weight and have MN blood group activity and influenza virus receptors, but differ in that their reported molecular weights vary from 31,000 (6) to 160,000 (8).

This report summarizes studies on the properties of the major glycoprotein of the human erythrocyte membrane extracted from membranes by a new procedure with lithium di-iodosalicylate as a dissociating agent. A water-soluble glycoprotein has been isolated in high yield that has the chemical properties and biological activities characteristic of the native membrane-bound molecules. This glycoprotein has been partially characterized after tryptic digestion and cyanogen bromide cleavage, and its location in the membrane has been determined by electron microscopy.

MATERIALS AND METHODS

Human blood was obtained fresh from donors, and the erythrocytes were washed three times in phosphate-buffered saline (pH 7.4) before preparation of ghost membranes by the procedure of Dodge *et al.* (11). Glycoprotein was extracted from freeze-dried ghost membranes with 0.3 M lithium di-iodosalicylate and purified by phosphocellulose chromatography (12). Samples were electrophoresed in acrylamide gels containing Tris·glycine (pH 8.4), according to the procedure of Davis (13) and in the presence of sodium dodecyl sulfate (14). Purified glycoprotein, extracted with chloroform–methanol (2:1) three times at room temperature, was assayed for AB and MN blood-group activities by hemagglutinin-inhibition assays with appropriate antisera. Receptors for influenza viruses, phytohemagglutinin (PHA), and wheat germ agglutinin were also assayed by hemagglutination-inhibition reactions by use of purified virus (obtained from Eli Lilly) and purified lectins.

Glycoprotein preparations were assayed for hexose by the phenolsulfuric acid method (15), and for sialic acid by the thiobarbituric acid method (16). Amino sugars and amino acids were determined on a Hitachi KLA-3B amino-acid analyzer after hydrolysis in 3 N HCl at 100° for 4 hr and 6 N HCl at 110° for 22 hr, respectively.

Purified glycoproteins were cleaved by treatment with cyanogen bromide in 70% formic acid at room temperature for 48 hr. The resulting peptides were separated and purified by gel filtration in 25% formic acid.

Soluble glycoproteins were digested with trypsin that was treated to remove chymotrypsin activity (Worthington, 2.5% by weight) in Tris·HCl buffer (pH 8.2) at 37° for 24 hr; acid-

Abbreviations: PHA, phytohemagglutinin; SDS, sodium dodecyl sulfate; PAS, periodic acid-Schiff.

Present Addresses: * Department of Surgical Pathology, Barnes Hospital Plaza, St. Louis, Mo. 63110; † Department of Medicine, The Methodist Hospital, Houston, Texas 77025; and ‡ Department of Pathology, Jackson Hall, University of Minnesota, Minneapolis, Minn. 55455.

soluble glycopeptides were fractionated on Sephadex G-150. Intact erythrocytes were treated with TPCK-trypsin (250 mg of trypsin was added to 1500 ml of packed erythrocytes) in 0.1 M phosphate-buffered saline (pH 8.0)–0.1 M NaCl for 90 min at 37°. Glycopeptides released from intact erythrocytes by trypsin were fractionated on phosphocellulose and Sephadex G-150. Glycoprotein fragments that remain bound to the erythrocyte membranes after this tryptic digestion were extracted from ghosts with lithium di-iodosalicylate. The soluble glycoprotein fragments were treated with trypsin a second time (2% trypsin at 37° for 20 hr), and the products were fractionated on diethylaminoethyl (DEAE)-cellulose.

Phytohemagglutinin (*Phaseolus vulgaris*) was obtained from Difco and purified on sulfoethyl Sephadex (17) followed by gel filtration on Sephadex G-200. Purified PHA was conjugated to ferritin (6 times crystalized, Pentex) with glutaraldehyde (Polysciences) by the procedure described (18). PHA–ferritin complexes were purified on Agarose A-1.5 columns and concentrated with a Diaflo apparatus (Amicon Corp.). Erythrocytes and ghosts were labeled with PHA–ferritin complexes by addition of small aliquots of washed erythrocytes to concentrated solutions of lectins in order to label the cells completely and minimize agglutination. Labeled erythrocytes were washed and lysed as described above.

Erythrocyte ghosts were freeze-etched in a Balzers apparatus (19), and platinum–carbon replicas were examined in a Phillips 200 electron microscope.

RESULTS AND DISCUSSION

Properties of the isolated glycoprotein

The isolated glycoprotein is soluble in water or in neutral salt solutions and migrates as a single component on polyacrylamide gels both in Tris·glycine and when electrophoresed in the presence of SDS. The mobility of the glycoprotein in SDS–polyacrylamide gels relative to other standard proteins suggested a molecular weight of about 100,000 (20), but more recent studies indicate that this value is incorrect due to the anomalous migration of glycoproteins under these conditions. Glycoproteins bind proportionally less SDS than "standard" proteins (21), and as a result they have a slower mobility than polypeptides of comparable size. This anomalous mobility is less evident when glycoproteins are run in gels of low porosity (greater crosslinking), and a molecular weight of 50,000–55,000 is obtained when the erythrocyte-membrane glycoprotein is electrophoresed in 12.5% acrylamide gels (21). This value is comparable to that calculated on the basis of the peptides obtained after cleavage with cyanogen bromide, as described below.

The electrophoretic mobility of the isolated glycoprotein is identical to that of the major periodic acid-Schiff (PAS)-positive band obtained when erythrocyte membranes are dissolved in SDS and electrophoresed without prior extraction. This suggests that the glycoprotein isolated by the lithium di-iodosalicylate method is similar in size to the native molecule and is not altered to any detectable extent during the purification steps.

60% of the dry weight of the purified glycoprotein is carbohydrate, consisting of sialic acid (25%), galactose (10–12%), *N*-acetylglucosamine (6%), *N*-acetylgalactosamine (12%), and small amounts of mannose and fucose. Glycoproteins isolated from appropriate erythrocytes have AB and MN blood-group activities, and all preparations have receptors for influenza viruses, phytohemagglutinin (prepared from kidney beans), and wheat-germ agglutinin. Since AB blood-group activity has been reported on erythrocyte-membrane glycolipids (22), glycoprotein preparations were treated with chloroform–methanol to insure that contaminating glycolipids were not responsible for the inhibiting activity. The presence of AB activity on this glycoprotein confirms the previous findings of others (23, 24).

TABLE 1. *Amino-acid analysis of cyanogen-bromide cleavage fragments*

	C-1	C-2	C-3	C-4	C-5	Sum	Complete glycoprotein
Asp	2	7	5	1	0	15	12
Thr	8	6	4	2	2	22	22
Ser	8	12	6	1	1	28	28
Glu	8	7	4	1	1	21	20
Pro	3	9	4	0	0	16	15
Gly	2	5	4	1	0	12	13
Ala	3	2	4	2	1	12	11
Val	6	6	2	1	1	16	17
Met*	1	0	1	1	1	4	4
Ile	3	8	3	0	0	14	14
Leu	2	8	4	0	1	15	13
Tyr	3	2	1	0	0	6	6
Phe	2	0	1	0	0	3	3
Lys	2	4	3	0	0	9	8
His	4	0	1	0	0	5	8
Arg	3	3	2	0	0	8	9
Total	60	79	49	10	8	206	203

* Determined as homoserine.

Cleavage of the glycoprotein with cyanogen bromide

Five unique fragments are obtained when the glycoprotein is treated with cyanogen bromide. These fragments have been purified by gel filtration and analyzed for their amino-acid and carbohydrate compositions (Table 1 and Segrest, Jackson, and Marchesi, in preparation). Four of the fragments (C-1, C-3, C-4, and C-5 designated by order of chromatographic elution) contain homoserine, and their absolute compositions have been calculated on the basis of a single residue of homoserine. The absolute composition of the C-terminal fragment (C-2), which lacks homoserine, was calculated on the basis of its molecular weight as determined by gel filtration.

The sum of the residues of the five fragments is in good agreement with the amino-acid composition of the intact glycoprotein containing four methionyls (Table 1); the intact glycoprotein has a single N-terminal residue of leucine in 50% yield and the C-terminal fragment (C-2) has a single N-terminal residue of alanine in 60% yield. These results suggest that the glycoprotein is one polypeptide chain with a molecular weight of about 50,000.

The fragments C-5 and C-1, containing most of the carbohydrate, have been shown by overlap with glycopeptides produced by treatment with trypsin to be the N-terminal fragments in the order given. The C-terminal peptide (C-2) does not contain carbohydrate but is rich in hydrophobic amino-acid residues. Fragments C-4 and C-3 have been only tentatively aligned. This general molecular topography is illustrated in Fig. 1.

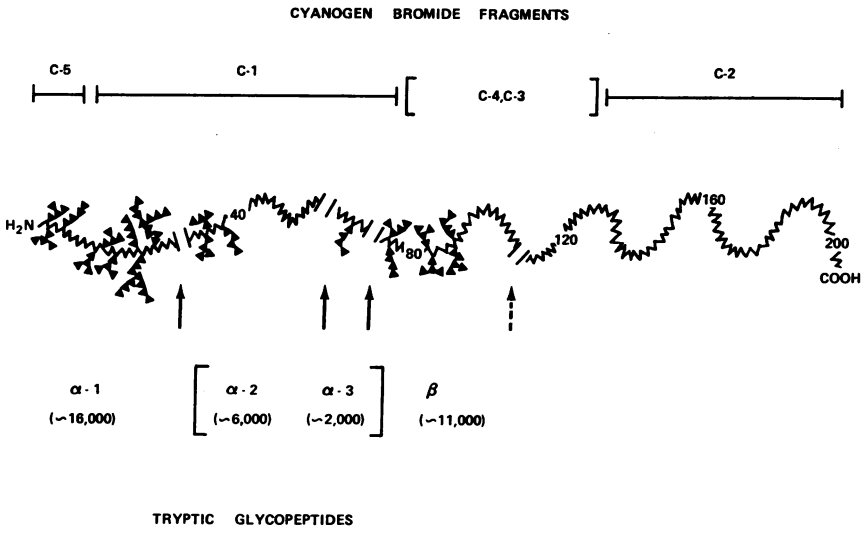

FIG. 1. *Arrows* mark the approximate locations of trypsin-sensitive sites on the polypeptide chain of the glycoprotein that result in the production of four unique glycopeptides. The *dashed arrow* indicates the linkage between the β-peptide and the rest of the molecule, which is not susceptible to trypsin while the molecule is in the intact membrane. The approximate molecular weights of the glycopeptides produced after trypsinization (*tryptic glycopeptides*) are given in parentheses. The overlap between the five peptides produced after cleavage with cyanogen bromide (*cyanogen bromide fragments*) with the tryptic glycopeptides is based on the carbohydrate and amino-acid compositions of C-5 and C-1 and on the absence of homoserine in C-2. The tentative alignment of the cyanogen-bromide fragments is supported by preliminary data on the amino-acid sequence of these fragments. Peptides produced by tryptic digestion of the C-terminal half of the glycoprotein that do not contain carbohydrate have not yet been characterized. The *solid triangles* represent saccharide units, which are added only to illustrate the approximate locations of sugar residues. The number and types of oligosaccharides on each peptide have not been determined.

Arrangement of glycopeptides that were produced by treatment with trypsin

Four unique, oligosaccharide-containing peptides are produced by tryptic digestion of the isolated glycoprotein that account for over 90% of the total carbohydrate content of the molecule (Fig. 2*A*). These glycopeptides, labeled α-1, α-2, α-3, and β, have been purified and characterized in terms of their composition, and antigen and receptor activities (Jackson, Segrest, and Marchesi, in preparation).

Treatment of intact erythrocytes with trypsin releases the α-glycopeptides (α-1, α-2, and probably α-3) from the glycoprotein (Fig. 2*B*, and *C*). The β-glycopeptide is released only after the partially digested glycoprotein is extracted from the trypsinized membranes with lithium di-iodosalicylate and then subjected to a second trypsin digestion (Fig. 2*D*). These results indicate that the oligosaccharide-containing peptides derived from cyanogen-bromide cleavage and trypsin digestion are concentrated at the N-terminal end of the molecule as depicted in Fig. 2.

Location of glycoprotein receptors on the erythrocyte surface

PHA and PHA–ferritin have been used to map the location of individual glycoproteins on the cell surface by freeze-etching electron microscopy (25). The freeze-etching technique is well suited for study of the location of labels on the cell surface since large areas of the surface membrane can be examined at reasonably high resolution [3–5 nm (30–50 Å)]. Internal regions of the membrane can also be examined by freeze-etching as a result of the tendency of cell membranes to cleave along a plane within the hydrophobic region of the membrane (26, 27). Thus, the position of specific labels on the external surface can be compared to structures within the membrane itself.

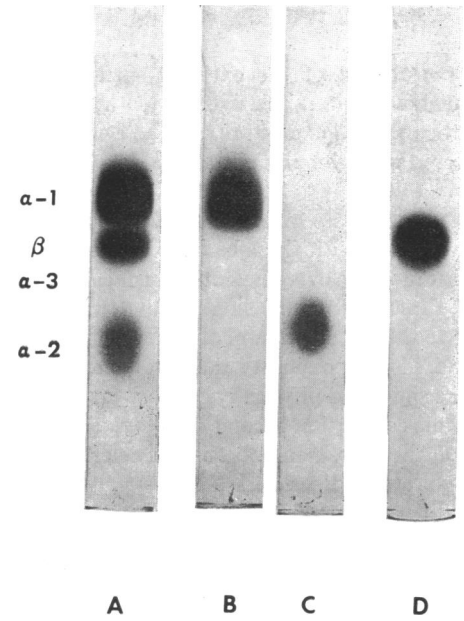

FIG. 2. Gel electrophoresis (12.5% acrylamide–0.1% SDS) of glycopeptides stained for carbohydrate with PAS. (*A*) Tryptic digestion of the isolated glycoprotein produces four unique glycopeptides: α-1, α-2, α-3, and β. α-3 is small and does not fix or stain permanently, but it runs in the position labeled. (*B,C*) α-1 and α-2 (and probably α-3) are released when intact erythrocytes are trypsinized. α-1 and α-2 have been separated by gel filtration before electrophoresis. None of the β-peptide appears in the medium. (*D*) Purified β-glycopeptide obtained when partially digested glycoproteins are extracted from trypsinized erythrocytes and subjected to a second trypsin digestion.

FIG. 3. Platinum replica of a freeze-etched human erythrocyte ghost after labeling with PHA–ferritin conjugate. The globular PHA–ferritin complexes are attached to the external surface of the membrane (*ES*) and appear to be distributed in the same pattern as that of the intramembranous particles. ×110,000.

A platinum replica of the external surface of a normal erythrocyte ghost appears relatively smooth; in contrast, the "faces" of the membrane that are exposed as a result of cleavage are covered with globular particles about 7.5 nm (75 Å) in diameter.

The external surfaces of erythrocytes labeled with PHA–ferritin are covered with these complexes (Fig. 3), which are distributed uniformly over the entire surface of the cell. The pattern of PHA–ferritin labeling on the surface seemed to correspond to the distribution of the intramembranous particles, and raised the possibility that the receptors on the external surface might be physically connected to the underlying particles. This correlation became more evident when the distribution of the intramembranous particles was artificially modified by protease digestion.

Trypsin digestion of erythrocyte ghosts causes the particles to clump into reticular patterns, and when trypsinized ghosts are treated with PHA–ferritin the label also appears in a reticular pattern on the cell surface that corresponds to the pattern of the underlying particles. Similar experiments with influenza virus as the labeling agent have produced an identical result (Tillack, Scott, and Marchesi, manuscript in preparation), and a recent report suggests that blood-group A antigens may also be localized to sites over the intramembranous particles (28).

Orientation of the glycoprotein *in situ*

On the basis of these results we propose that the major glycoprotein of the human erythrocyte membrane is oriented at the cell surface as depicted in Fig. 4. The orientation of the glycoprotein at the cell surface and the arrangement of its oligosaccharide moieties proposed here is consistent with the results of earlier studies on the effects of protease digestion of erythrocytes (29) and with recent attempts to label exposed proteins on cell surfaces with radioactive reagents (30, 31) and with lectin–ferritin conjugates (18). The carbohydrate-rich portion of the molecule extends outside the lipid barrier, while the C-terminal region interacts in some unknown manner with components in the interior of the membrane at the sites of the intramembranous particles. Since peptides isolated from the C-terminal region of the molecule have a high content of nonpolar amino acids it is reasonable to speculate that this end of the glycoprotein interacts through hydrophobic association with membrane lipids or with other polypeptides, as has been suggested (32, 33).

The apparent connection between the major glycoprotein and the intramembranous particles demonstrated here raises the possibility that these particles may be protein aggregates or lipid–protein complexes in the membrane. It is not possible to determine whether there is a single glycoprotein molecule per particle on the basis of these labeling experiments. However, the number of particles per erythrocyte (5×10^5) is close enough to the number of PHA receptors (4.5×10^5) and to the number of A_1 blood group sites (8×10^5) (35) to suggest that there may be a one-to-one correlation between glycoproteins and particles.

This glycoprotein appears to be a single, unique polypeptide chain rather than a family of closely related molecules as suggested earlier (33). This conclusion is based primarily on the finding of a single N-terminal amino acid (leucine) for the

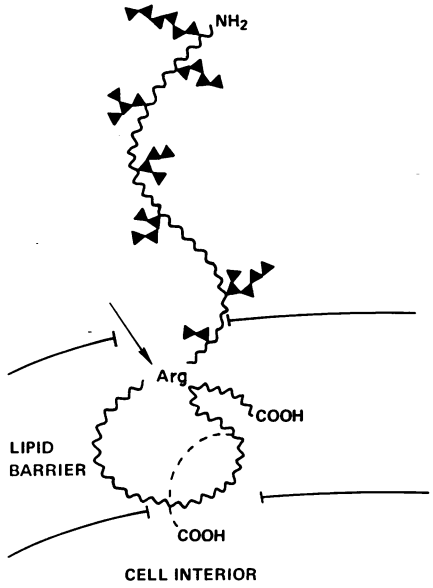

FIG. 4. The major glycoprotein of the human erythrocyte seems to be oriented as shown in this highly schematic figure. The *arrow* marks a trypsin-sensitive arginine linkage, which is inaccessible to proteolytic digestion when the glycoprotein is *in situ*, but is readily cleaved after the molecule is isolated. It has been suggested recently (31) that the C-terminal end of this molecule may extend completely across the lipid barrier as shown by the *dashed line*. This idea is based on results obtained from labeling membrane ghosts with isotopes specific for polypeptides that are exposed to the surrounding aqueous media. However, this interpretation depends on the still unproven assumption that the lipid barrier of the ghost membrane has the same permeability properties as that of the intact cell.

intact glycoprotein and on the results of cleavage with cyanogen bromide.

Preliminary results (Segrest, Terry, and Marchesi, unpublished observations) are consistent with the idea that the glycoprotein is a single polypeptide chain, but minor variations in sequence cannot be ruled out. This molecule contains most of the sialic acid of the human erythrocyte membrane and is clearly the major glycoprotein of the membrane. We suggest that this glycoprotein be called a glycophorin [Greek φορέω, to bear] with the idea that glycophorins be considered a special class of glycoproteins whose principal biological role is to bear oligosaccharides that serve as receptors or surface recognition sites on membranes. The molecule characterized here might be called erythrocyte glycophorin, and when other membrane glycoproteins are characterized they could be distinguished by cell type or alphabetically.

1. Winzler, R. J. (1970) *Int. Rev. Cytol.* **29**, 77–125.
2. Morton, J. A. (1962) *Brit. J. Haematol.* **8**, 134–148.
3. Cook, G. M. W. & Eylar, E. H. (1965) *Biochim. Biophys. Acta* **101**, 57–66.
4. Springer, G. F. Nagai, Y. & Tegtmeyer, H. (1966) *Biochemistry* **5**, 3254–3272.
5. Eylar, E. H., Madoff, M. A., Brody, O. V. & Oncley, J. L. (1962) *J. Biol. Chem.* **237**, 1992–2000.
6. Kathan, R. H., Winzler, R. J. & Johnson, C. A. (1961) *J. Exp. Med.* **113**, 37–45.
7. Maddy, A. H. (1966) *Biochim. Biophys. Acta* **117**, 193–200.
8. Blumenfeld, O. O., Gallop, P. M., Howe, C. & Lee, L. T. (1970) *Biochim. Biophys. Acta* **211**, 109–123.
9. Lenard, J. (1970) *Biochemistry* **5**, 1129–1132.
10. Laico, M. T., Ruoslahti, E. I., Papermaster, D. S. & Dreyer, W. J. (1970) *Proc. Nat. Acad. Sci. USA* **67**, 120–127.
11. Dodge, J. T., Mitchell, C. & Hanahan, D. J. (1963) *Arch. Biochem. Biophys.* **100**, 119–130.
12. Marchesi, V. T. & Andrews, E. P. (1971) *Science* **174**, 1247–1248.
13. Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404–427.
14. Maizel, J. V. (1966) *Science* **151**, 988–990.
15. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356.
16. Warren, L. (1959) *J. Biol. Chem.* **234**, 1971–1975.
17. Weber, T., Nordman, C. T. & Grasbeck, R. (1967) *Scand. J. Haematol.* **4**, 77–80.
18. Nicolson, G. L. & Singer, S. J. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 942–945.
19. Moor, H. & Muehlethaler, K. (1963) *J. Cell Biol.* **17**, 609–628.
20. Marchesi, V. T. (1970) *Fed. Proc.* **29**, 600.
21. Segrest, J. P., Jackson, R. L., Andrews, E. P. & Marchesi, V. T. (1971) *Biochem. Biophys. Res. Commun.* **44**, 390–395.
22. Hakomori, S. & Strycharz, G. D. (1968) *Biochemistry* **7**, 1279–1286.
23. Poulik, M. D. & Lauf, P. K. (1965) *Nature* **208**, 874–876.
24. Whittemore, N. B., Trabold, N. C., Reed, C. F. & Weed, R. I. (1969) *Vox Sang.* **17**, 289–299.
25. Tillack, T. W., Scott, R. E. & Marchesi, V. T., *J. Exp. Med.*, in press.
26. Branton, D. (1966) *Proc. Nat. Acad. Sci. USA* **55**, 1048–1056.
27. Tillack, T. W. & Marchesi, V. T. (1970) *J. Cell Biol.* **45**, 649–653.
28. Pinto da Silva, P., Douglas, S. D. & Branton, D. (1971) *Nature* **232**, 194–195.
29. Cook, G. M. W. (1968) *Biol. Rev.* **43**, 363–391.
30. Phillips, D. R. & Morrison, M. (1971) *Biochemistry* **10**, 1766–1771.
31. Bretscher, M. S. (1971) *Nature New Biol.* **231**, 229–232.
32. Morawiecki, A. (1964) *Biochim. Biophys. Acta* **83**, 339–347.
33. Winzler, R. J. (1969) in *Red Cell Membrane, Structure and Function*, eds Jamieson, G. A. & Greenwalt, T. J. (J. B. Lippincott Co., Philadelphia, Pa.), pp. 157–171.
34. Kornfeld, S. & Kornfeld, R. (1969) *Proc. Nat. Acad. Sci. USA* **63**, 1439–1446.
35. Economidou, J., Hughes-Jones, N. C. & Gardner, B. (1967) *Vox Sang.* **12**, 321–328.