J. Membrane Biol. 64, 205–215 (1982)

The Journal of
Membrane Biology

Primary Structure of Horse Erythrocyte Glycophorin HA. Its Amino Acid Sequence has a Unique Homology with those of Human and Porcine Erythrocyte Glycophorins

Jun-Ichiro Murayama, Motowo Tomita, and Akira Hamada
School of Pharmaceutical Sciences, Showa University, Hatanodai, Shinagawa-ku, Tokyo 142, Japan

Summary. The complete amino acid sequence of the major sialoglycoproteins of horse erythrocyte membranes, glycophorin HA, was determined by manual sequencing methods, using tryptic, chymotryptic, and cyanogen bromide fragments. Glycophorin HA is a polypeptide chain of 120 amino acid residues and contains 10 oligosaccharide units attached to the amino-terminal side of the molecule. Its amino terminus is pyroglutamic acid. All of the oligosaccharides are linked O-glycosidically to threonine or serine residues. The amino acid sequence is consistent with the transmembrane orientation of glycophorins.

There is no significant homology between the glycosylated domains of horse, human, and porcine glycophorins, but there is a considerable homology between the hydrophobic domains of the three glycophorins, which interact with the lipid bilayer of the erythrocyte membrane.

Key words erythrocyte membrane · horse glycophorin · amino acid sequence · glycopeptide · sequence homology

Introduction

Glycophorins are the major sialoglycoproteins of erythrocyte membranes. They are among the few membrane glycoproteins which can be readily isolated in preparative amounts, and their primary structures can be elucidated with presently available techniques. Indeed, the complete amino acid sequence of two glycophorins, human glycophorin A (Tomita, Furthmayr & Marchesi, 1978) and porcine glycophorin (Honma, Tomita & Hamada, 1980), has been determined. The comparison of the amino acid sequence of the two glycophorins shows that there is no homology in the sequence of the amino-terminal glycosylated domain, while striking homology is observed in the sequence of the hydrophobic domain.

These findings led us to propose that the glycosylated domain, which is entirely exposed at the outside of the cell and which does not show homology among animal glycophorins, may play an important role as a species-specific antigen. But, since this proposal was made from the comparison of only glycophorins from two species, further confirmatory evidence has to be obtained. The present paper reports the complete amino acid sequence of horse erythrocyte glycophorin HA. Horse erythrocyte membranes contain two glycophorins, and the major component was designated glycophorin HA (Murayama, Takeshita, Tomita & Hamada, 1981). The sequence data allowed a precise molecular description of glycophorin and were consistent with our proposal described above.

Materials and Methods

Materials

Trypsin (TPCK[1]-treated) and α-chymotrypsin were purchased from Worthington Biochemical, thermolysin from Seikagaku Kogyo, and subtilisin and carboxypeptidase A from Boehringer Mannheim Biochemicals. Sephadex and Sephacryl were obtained from Pharmacia, DEAE-cellulose DE-52 from Whatman, and Aminex 50W-X4 (20–30 μm) from Bio-rad Laboratories. Thin-layer plates were purchased from Cheng Chin Trading Co., and fluorescamine from Roche Diagnostics. The reagents for Edman degradation were specific grade materials for amino acid sequence analysis obtained from Wako Pure Chemical Industries. All other chemicals were analytical grade.

Preparation of Glycophorin

Erythrocyte membranes were prepared from fresh horse blood as described (Dodge, Mitchell & Hanahan, 1963). The glycophorin fraction was prepared by the method of Marchesi and Andrews (1971); this procedure was initially developed for the isolation of human glycophorins. Although the glycophorin fraction from horse erythrocyte membranes contains a minor component in addition to a single major component, the preparation was pure enough to isolate the peptides derived from the major glycophorin. For some studies the major component, designated glycophorin HA, was isolated from the glycophorin fraction by DEAE-cellulose chromatography in the presence of Ammonyx LO (Murayama et al., 1981). Glycophorin HA represented about 80% by mass of the glycophorin fraction.

[1] *Abbreviations used:* TPCK – L-(1-tosylamide-2-phenyl) ethyl chloromethyl ketone; PTH – 3-phenyl-2-thiohydantoin; DNS – 1-dimethylamino naphthalene-5-sulfonyl; DEAE – diethylaminoethyl.

0022-2631/82/0064-0205 $02.20

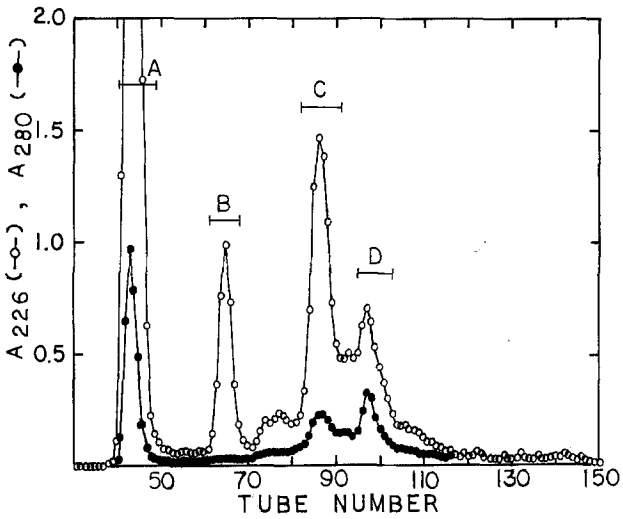

Fig. 1. Gel chromatography of chymotryptic peptides from horse glycophorin. The peptides obtained from 100 mg of glycophorin HA were applied to a Sephacryl S-200 column (2.5 × 100 cm), and eluted with 0.1 M NH_4HCO_3 at a flow rate of 10 ml/hr. Fractions (4 ml) were collected. The peptides were analyzed by absorbance at 226 and 280 nm

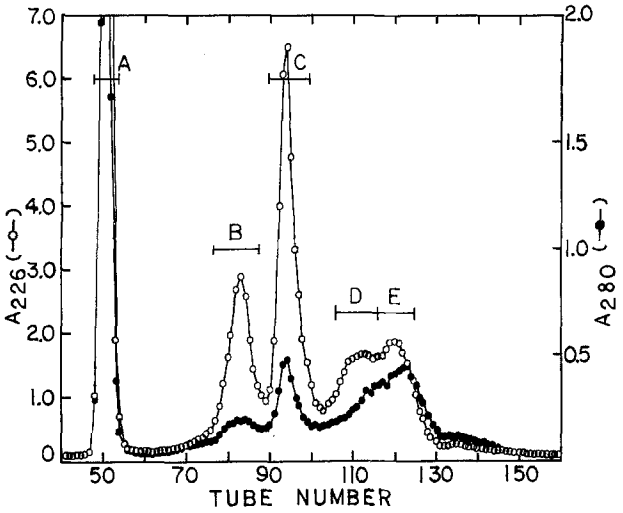

Fig. 2. Gel chromatography of tryptic peptides from horse glycophorin. The peptides obtained from 300 mg of the glycophorin fraction were applied to a Sephacryl S-200 column (2.5 × 140 cm), and eluted with 0.1 M NH_4HCO_3 at a flow rate of 12 ml/hr. Fractions (5 ml) were collected. The peptides were analyzed by absorbance at 226 and 280 nm

Enzymatic Hydrolysis

Glycophorin HA (100 mg) was dissolved in 10 ml of 0.1 M ammonium bicarbonate, pH 8.3, and digested with chymotrypsin (2 mg) at 37 °C for 16 hr. The incubation mixture was directly lyophilized. Tryptic digestion of the glycophorin fraction (300 mg) was performed with 5 mg of trypsin in 0.1 M ammonium bicarbonate (30 ml), pH 8.3, at 37 °C for 22 hr. The incubation mixture was lyophilized. Additional enzymatic digestions were performed on some of the isolated primary peptides. Thermolysin, subtilisin, trypsin, or chymotrypsin was used for the preparation of the secondary peptides at an enzyme/substrate ratio of 1:50 at 37 °C for 24 hr. The digestion was stopped by lyophilization.

Gel Chromatography of Enzymatic Digests

The chymotryptic and tryptic peptides were separated by gel chromatography on a column of Sephacryl S-200 which was previously equilibrated with 0.1 M ammonium bicarbonate. The elution profiles are shown in Figs. 1 and 2. The fractions were combined into several pools. Secondary peptides were separated by gel chromatography on a column of Sephadex G-50 (superfine, 2.5 × 150 cm) equilibrated with 0.05 M acetic acid at a flow rate of 15 ml/hr. Fractions (5 ml) were collected and analyzed by absorbance at 226 and 280 nm, and by the fluorescamine reaction (Udenfriend et al., 1972).

Ion-Exchange Chromatography of Enzymatic Digests

Hydrophilic peptides were further purified on a DEAE-cellulose column (1.0 × 30 cm) which was previously equilibrated with 0.001 M ammonium bicarbonate. The peptides were eluted with a linear gradient of 0.001–0.4 M ammonium bicarbonate (300 ml each) at a flow rate of 8 ml/hr. Fractions (4 ml) were collected and analyzed by absorbance at 226 nm. Some of the small peptides were further purified on a column of cation-exchange resin (Aminex 50W-X4, 20–30 μm, 0.5 × 50 cm). The peptides were eluted with a gradient of pH and ionic strength in pyridine-acetate buffer as described by Schroeder (1972).

Purification of Hydrophobic Peptides

Chymotryptic hydrophobic peptides were eluted in the void fraction (pool *A*) of the Sephacryl S-200 column (Fig. 1), but were contaminated with incompletely digested material. The hydrophobic peptides were isolated on a column of Sephadex LH-60 by the modified method of Gerber et al. (1979). Pool *A* was dissolved in 2 ml of 88% formic acid. Ethanol (7 ml) was added to the solution and allowed to stand at −20 °C overnight. The precipitate produced was removed by centrifugation at 14,000 rpm for 30 min. This precipitate contained the incompletely digested material. The supernatant was applied to a column of Sephadex LH-60 (2.5 × 90 cm). A major fraction was designated CH1 as shown in Fig. 3. Tryptic hydrophobic peptides were eluted in the void fraction of the Sephacryl S-200 column (Fig. 2), but were also contaminated with incompletely digested material. When pool *A* was dissolved in 9 ml of 88% formic acid-ethanol (2:7, vol/vol), the contaminating peptides were precipitated and a hydrophobic peptide designated T1 was recovered in the supernatant.

Cyanogen Bromide Cleavage

The glycophorin fraction (200 mg) was dissolved in 8 ml of 70% formic acid and cleaved with cyanogen bromide (100 mg) at room temperature for 6 hr. The reaction mixture was directly applied to a column of Sephadex G-75 (2.5 × 95 cm) which was previously equilibrated with 25% formic acid. Fractions (4 ml) were collected and analyzed by the ninhydrin method after alkaline hydrolysis (Hirs, 1967).

Amino Acid Analysis

Dry samples containing 10–50 μg of amino acids were hydrolyzed with 100 μl of 6 N HCl at 110 °C for 22 hr. The amino acid analyses were performed on a Hitachi 835 amino acid analyzer.

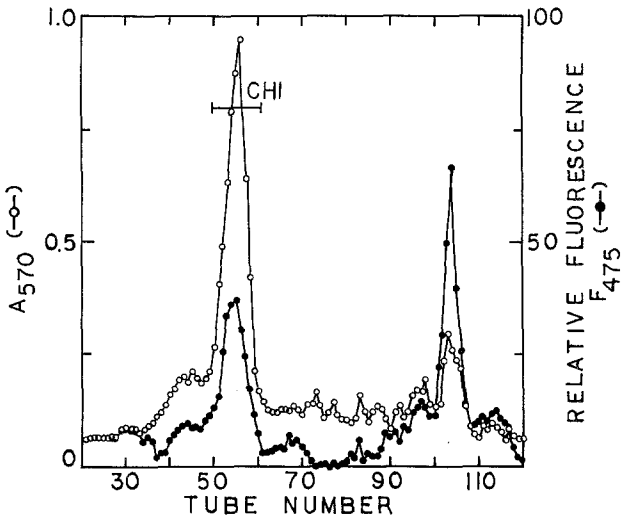

Fig. 3. Purification of hydrophobic peptide CH1 by gel chromatography on Sephadex LH-60. The sample (8 ml) was applied to a column (2.5 × 90 cm) of Sephadex LH-60, and eluted with 88% formic acid-ethanol (2:7, vol/vol) at a flow rate of 10 ml/hr. Fractions (3.5 ml) were collected. Each fraction was analyzed by the ninhydrin reaction at 570 nm after alkaline hydrolysis (Hirs, 1967) and by the fluorescamine reaction after alkaline hydrolysis (Lai, 1977)

Carbohydrate Analysis

Sugar composition was determined by gas-liquid chromatography of trimethylsilyl derivatives of the methyl glycosides released by methanolysis according to Reinhold (1972). Samples containing 20–100 μg of carbohydrate were used with 10 μg of inositol as an internal standard. Amino sugars were determined also on the Hitachi 835 analyzer after hydrolysis with 4 N HCl at 100 °C for 4 hr.

Sequence Analysis

Manual Edman degradation was performed by the method as described (Peterson, Nerrlich, Oyer & Steiner, 1972). Phenylthiohydantoin (PTH)-amino acids were identified by gas-liquid chromatography on a 10% DC-560 column (Pisano, Bronzert & Brewe, 1972), thin-layer chromatography on a polyamide sheet (Summers, Symthers & Oroszlan, 1973) and back-hydrolysis with 6 N HCl containing 0.1% $SnCl_2$ at 150 °C for 4 hr followed by amino acid analysis (Mendez & Li, 1975). Dansyl-Edman degradation was done by the method of Gray (1972) with a modification for sequencing of multiple glycosylated peptides (Tomita et al., 1978). For carboxyl-terminal sequence determination, peptides (50 nmol) were digested with carboxypeptidase A in 0.1 M ammonium bicarbonate, pH 8.0 at 37 °C. An aliquot of the digest for various periods was taken and lyophilized. The released amino acids were analyzed on the amino acid analyzer. Pyroglutamic acid residue at the amino-terminus of peptide was changed to glutamic acid residue by methanolysis (Kawasaki & Itano, 1972). Peptide (100–200 nmol) was treated with 0.5 N HCl in methanol (0.5 ml) at 50 °C for 4 hr. The reaction mixture was evaporated to dryness under a stream of nitrogen. The methanolyzed peptide could be directly sequenced by Edman degradation after one cycle of Edman procedure was performed without phenylisothiocyanate to hydrolyze the methyl ester group which produced during the methanolysis.

Results

Chymotryptic Peptides

Chymotryptic peptides of glycophorin HA were separated into four pools *A–D* by gel chromatography on Sephacryl S-200 (Fig. 1). Pool *A*, which was recovered in the excluded volume of the column, contained the hydrophobic peptides derived from the intramembranous segment of glycophorin HA. Isolation of the hydrophobic peptide CH1 from the incomplete digest was achieved by preferential solubilization in 88% formic acid-ethanol (2:7, vol/vol) and by gel chromatography on Sephadex LH-60 equilibrated with the same solvent (Fig. 3). Although pool *A* contains incompletely digested material, probably representing residues 1–108, the incompletely digested peptides were almost insoluble in the solvent and could be removed from the hydrophobic peptides by the solubilization procedure. The amino acid composition of CH1 is shown in Table 1. It should be noted that the hydrophobic peptides which were obtained from the crude glycophorin fraction instead of glycophorin HA were contaminated with a considerable amount of other peptides and were not pure enough for sequence analysis. Pool *B* was further fractionated by ion-exchange chromatography on DEAE-cellulose (Fig. 4*A*). The amino acid composition of CH2 is shown in Table 2. Pool *C* contains the peptide labeled CH3, which was further purified by ion-exchange chromatography on DEAE-cellulose (Fig. 5*A*). Pool *D* appears to contain several peptides, but the isolation of the peptides was not attempted, because these peptides were not essential for sequence determination of glycophorin HA. The results described above indicated that glycophorin HA was digested with chymotrypsin into three major peptides (CH1, CH2 and CH3) in addition to several small peptides. We previously reported the chymotryptic digestion of the total glycophorin fraction (Fukuda, Tomita & Hamada, 1980). These data suggested that the additional small peptides identified above are derived from the minor component in the horse glycophorin mixture.

Tryptic Peptides

Tryptic peptides of the glycophorin fraction were separated into five pools *A–E* by gel chromatography on Sephacryl S-200 (Fig. 2). The elution profile was similar to that of the chymotryptic peptides. Pool *A* contained the hydrophobic peptides originating from the intramembranous segment of glycophorin HA. The hydrophobic peptide T1 was purified by solubilization in 88% formic acid-ethanol (2:7, vol/vol). The glyco-

Table 1. Amino acid compositions of the peptides derived from the hydrophobic and carboxyl-terminal segment[a]

	T1[b] (35–75)[c]	T4 (76–77)	T3A (78–120)	T3B (79–120)	CH1[d] (41–108) (45–108)	CH3 (109–120)	T3ACH1 (78–108)
Asp	2.7 (2)	–	6.1 (6)	5.8 (6)	4.8 (4–5)	2.1 (2)	4.2 (4)
Thr	2.8 (1)	–	4.2 (4)	4.0 (4)	2.0 (3)	1.6 (2)	2.0 (2)
Ser	2.6 (2)	–	5.7 (6)	5.4 (6)	5.6 (6)	1.7 (2)	4.0 (4)
Glu	4.2 (4)	–	4.3 (4)	4.0 (4)	3.3 (3)	2.2 (2)	2.0 (2)
Pro	1.6 (1)	–	9.9 (10)	10.3 (10)	10.4 (10)	1.0 (1)	9.0 (9)
Gly	3.9 (5)	–	1.3 (1)	1.1 (1)	3.3 (3)	1.0 (1)	–
Ala	3.9 (3)	–	4.3 (4)	4.1 (4)	6.6 (6–7)	–	4.0 (4)
Cys	–	–	–	–	–	–	–
Val	2.7 (4)	–	4.0 (4)	4.1 (4)	7.2 (7)	1.0 (1)	2.8 (3)
Met	0.6 (1)	–	–	–	1.0 (1)	–	–
Ile	5.0 (7)	–	–	–	6.8 (7)	–	–
Leu	4.7 (6)	0.9 (1)	1.2 (1)	1.0 (1)	6.3 (6)	–	0.8 (1)
Tyr	0.8 (1)	–	0.7 (1)	0.9 (1)	1.0 (1)	0.8 (1)	–
Phe	1.4 (1)	–	–	–	0.8 (0–1)	–	–
Lys	–	–	0.9 (1)	0.3 (0)	1.0 (1)	–	0.8 (1)
His	0.9 (1)	–	–	–	0.8 (0–1)	–	–
Arg	1.9 (2)	1.1 (1)	1.1 (1)	0.9 (9)	4.0 (4)	–	1.0 (1)
Total residues	41	2	43	42	62–66	12	31
Yield	32%	25%	47%	20%	33%	52%	59%

[a] Values are expressed in mol/mol peptide. Numbers in parentheses indicate mol/mol peptide assumed from sequence. No entry denotes 0.1 residue or less.
[b] Amino acid composition of T1 does not correspond to that of the assumed sequence because of contamination of incomplete digest (residues 1–75).
[c] Numbers in parentheses indicate position of peptide within sequence.
[d] The peptide CH1 was a mixture of two peptides corresponding to residues CH1A (41–108) and CH1B (45–108).

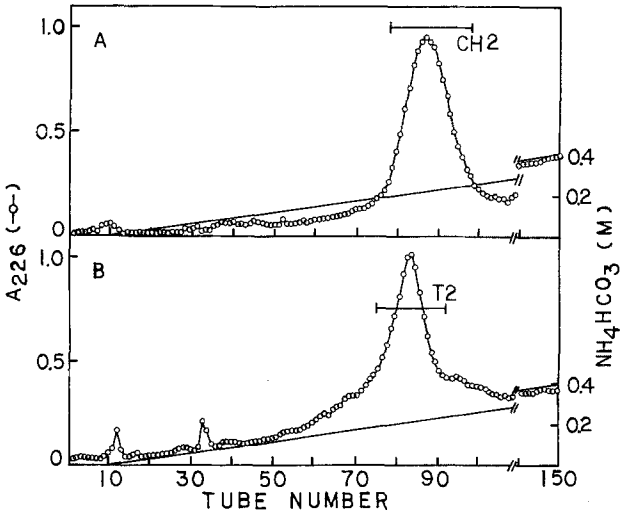

Fig. 4. Purification of major glycopeptides by DEAE-cellulose chromatography. Procedures are given under Methods. (*A*): The pool *B* (21 mg) of chymotryptic digests (Fig. 1) was loaded. (*B*): The pool *B* (27.9 mg) of tryptic digests (Fig. 2) was loaded

peptide T2 was purified from pool *B* by ion-exchange chromatography (Fig. 4*B*). The amino acid composition of T2 is similar to that of CH2 with the exception of Glu, Gly, Asp and Leu (Table 2), indicating that T2 is somewhat smaller than CH2. Pool *C* contains several peptides as shown in Fig. 5*B*. All of the peptides are similar in amino acid composition, indicating that these peptides are derived from the same segment of glycophorin HA. The amino acid composition of the major peptides, T3A and T3B, is shown in Table 1. The peptides T3A and T3B appear to differ only by the presence of an additional lysine in T3A. Pool *D* does not contain major peptides when fractionated by ion-exchange chromatography on DEAE-cellulose. It is likely that the minor component of horse glycophorin yields several peptides which elute with pool *D*. Pool *E* contains the peptide labeled T4 and Arg and Lys as the major components. It was purified by ion-exchange chromatography on Aminex 50W-X4. The amino acid composition of T4 is shown in Table 1. The results described above suggested that trypsin cleavage produces the four major fragments T1, T2, T3 and T4.

Cyanogen Bromide Peptides

The cyanogen bromide peptides of the glycophorin fraction were separated into five pools Vo, CB1A, CB1B, CB2 and Vi by gel chromatography on Sephadex G-75 (Fig. 6). Pool Vo contains uncleaved glycophorin in addition to aggregated forms of fragments,

Table 2. Amino acid compositions of the peptides derived from the glycosylated segment[a]

	T2 (1–34)[b]	AT2TH1 (1–9)	AT2TH2 (10–34)	CH2 (1–40)	CH2T2 (35–40)	AT2TH2-S3 (12–26)	AT2TH2-S5 (27–34)
Asp	2.3 (2)	–	1.8 (2)	2.2 (3)	1.1 (1)	1.0 (1)	0.8 (1)
Thr	8.5 (9)	2.0 (2)	5.9 (7)	9.2 (9)	–	4.6 (5)	2.0 (2)
Ser	4.2 (4)	1.0 (1)	2.9 (3)	4.4 (4)	–	2.6 (3)	0.4 (0)
Glu	3.6 (3)	1.0 (1)	2.0 (2)	4.5 (5)	1.9 (2)	–	2.1 (2)
Pro	2.7 (3)	2.2 (2)	1.0 (1)	3.1 (3)	–	1.0 (1)	–
Gly	3.6 (3)	1.1 (1)	2.5 (2)	4.4 (5)	1.7 (2)	1.0 (1)	1.0 (1)
Ala	3.9 (4)	1.1 (1)	2.9 (3)	4.1 (4)	–	2.3 (2)	0.2 (0)
Cys	–	–	–	–	–	–	–
Val	–	–	–	–	–	–	–
Met	–	–	–	–	–	–	0.2 (0)
Ile	2.8 (3)	1.1 (1)	2.0 (2)	3.2 (3)	–	1.3 (1)	–
Leu	0.9 (1)	–	1.1 (1)	2.6 (2)	1.0 (1)	0.9 (1)	–
Tyr	–	–	–	–	–	–	–
Phe	1.1 (1)	–	1.0 (1)	1.1 (1)	–	–	1.4 (1)
Lys	–	–	–	–	–	–	–
His	–	–	–	–	–	–	–
Arg	0.9 (1)	–	0.9 (1)	1.1 (1)	–	–	1.5 (1)
Total residues	34	9	25	40	6	15	8
Yield	40%	54%	63%	41%	40%	58%	25%

[a] Values are expressed in mol/mol peptide. Numbers in parentheses indicate mol/mol peptide assumed from sequence. No entry denotes 0.1 residue or less.
[b] Numbers in parentheses indicate position of peptide within sequence.

as suggested from data obtained by gel electrophoresis in the presence of 0.1% sodium dodecyl sulfate (data not shown). Pools CB1A and CB1B are essentially identical in amino acid composition and differences in carbohydrate content are probably responsible for their separation by gel chromatography. Peptide CB1B was further purified by rechromatography on the same column. The amino acid composition of CB1B and CB2 is shown in Table 3. Pool Vi does not contain a significant amount of peptide material, but contains sialic acid presumably released with 70% formic acid. These results indicate that horse glycophorin HA was cleaved into two fragments CB1B (or CB1A) and CB2 with cyanogen bromide. Indeed the amino acid composition of glycophorin HA was very similar to the sum of the amino acid composition of CB1B and CB2 (Table 3).

Amino Acid Sequence and Oligosaccharide Attachment Sites of Amino-Terminal Segment (Residues 1–40)

No PTH-amino acid was detected as the amino-terminal residue of glycophorin HA by Edman degradation; the amino-terminus was blocked. Since both glycopeptides T2 and CH2 had also the blocked amino-terminus, both glycopeptides presumably are derived from the amino-terminal segment of glycophorin HA. Although intact T2 was resistant to thermolysin, the desialated fragment T2 produced by acid treatment became sensitive to this protease: two fragments labeled AT2TH1 and AT2TH2 were purified from the digests by gel chromatography on Sephadex G-50 (Fig. 7*A*). The smaller peptide AT2TH1 did not react with fluorescamine (Fig. 7*A*), and therefore presumably represents the amino-terminal segment of T2. After methanolysis of AT2TH1, the amino acid sequence of AT2TH1 could be determined as shown in Table 4. Since methanolysis has been used to change pyroglutamic acid to glutamic acid (Kawasaki & Itano, 1972), the amino-terminal residue of AT2TH1 should be pyroglutamic acid. The yield of PTH-glutamic acid at the amino-terminus is about 30% (Table 4). This value is comparable to the yield reported on the methanolysis of cytochrome b_5 by Ozols, Gerard & Nobrega (1976). When myoglobin was treated as a standard peptide under the same conditions, no additional amino-terminal residue was observed in a significant yield, indicating that peptide bond cleavage is negligible during methanolysis. Table 4 also shows that no PTH-amino acid was identified at the 2nd, 5th and 7th positions of the methanolyzed AT2TH1. With the dansyl-Edman procedure, DNS-Thr was found at the 2nd and 5th positions, and DNS-Ser at the 7th position. This result indicated that these three residues are glycosylated. The amino acid sequence of the first 24 residues of the other thermolytic peptide AT2TH2 was determined by Edman degradation (Table 5). No PTH-amino acid was detected at the 4th, 8th, 9th, 11th, 12th, 15th and

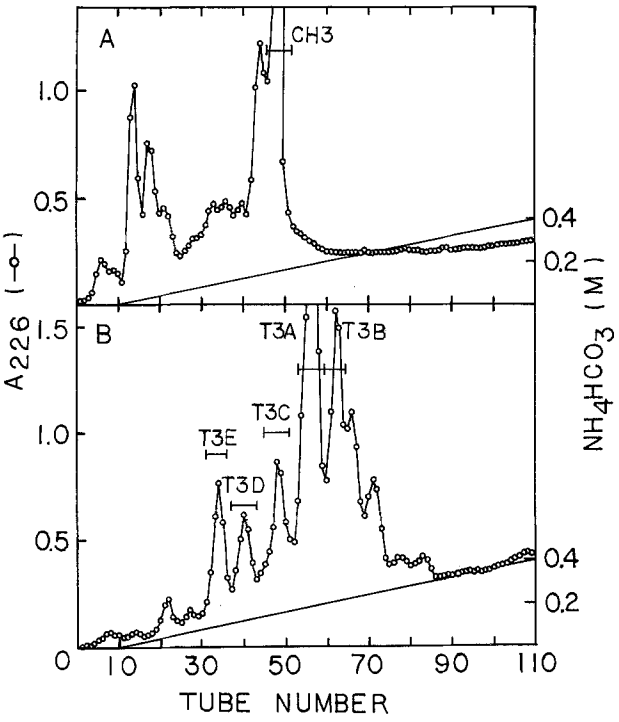

Fig. 5. Purification of tryptic and chymotryptic peptides by DEAE-cellulose chromatography. Procedures are given under Methods. (*A*): The pool *C* (19.9 mg) of chymotryptic digests (Fig. 1) was loaded. (*B*): The pool *C* (32.6 mg) of tryptic digests (Fig. 2) was loaded

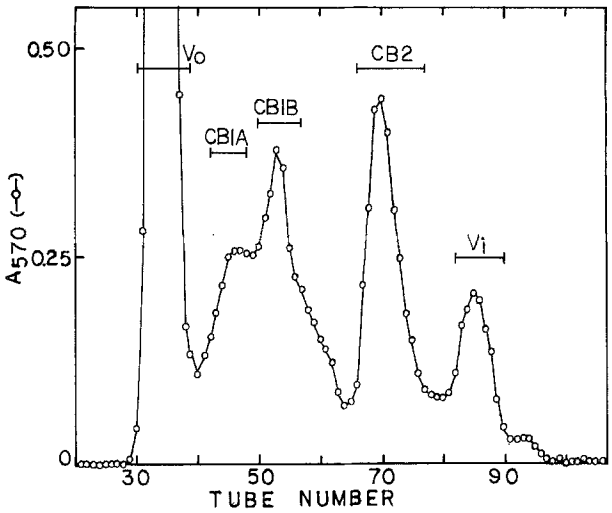

Fig. 6. Gel chromatography of cyanogen bromide fragments of horse glycophorin. The peptides were applied to a Sephadex G-75 column (2.5 × 100 cm) and eluted with 25% formic acid at a flow rate of 7 ml/hr. Fractions (4 ml) were collected. An aliquot of each fraction was analyzed by the ninhydrin assay after alkaline hydrolysis (Hirs, 1967)

19th positions of AT2TH2. As shown in Table 5, DNS-Thr was identified at the 4th, 9th, 11th and 15th positions and DNS-Ser at 8th and 12th positions. The 19th position seemed to be DNS-Thr but this was not conclusive. After T2 was methanolyzed, the amino acid sequence of the methanolyzed T2 was determined for the first 16 residues as shown in Table 5. Its first nine residues were identical to those of AT2TH1. This result also revealed that peptide AT2TH2 was derived from residues 10–34 of intact glycophorin HA. Since the tryptic peptide T2 should have Arg at the carboxyl-terminus, the entire sequence of T2 could be constructed by using the sequence of AT2TH1 and AT2TH2, except for the 28th position. The carboxyl-terminal sequence of T2 was confirmed using another secondary peptide, which was produced with subtilisin. The peptides labeled AT2TH2-S3 and AT2TH2-S5 were isolated from the subtilisin digest of AT2TH2 by gel chromatography on Sephadex G-50. The amino acid compositions of the two peptides are shown in Table 2. The amino acid sequence of AT2TH2-S5 containing one Arg residue revealed the sequence of residues 27–34. The second residue of AT2TH2-S5 was identified to be Thr only by dansyl-Edman procedure: the 28th position of T2, therefore, should be glycosylated Thr. Sequence analysis of AT2TH2-S3 also confirmed the sequence of residues 12–26 of T2.

When the chymotryptic peptide CH2 was treated with trypsin, a small peptide CH2T2 and a large peptide CH2T1 were produced with approximately 40% yield. The two peptides were purified by gel chromatography on Sephadex G-50. The amino acid composition of CH2T1 was identical to that of T2. The amino acid composition of CH2T2 (Table 2) indicated that this peptide is a hexapeptide, and the amino acid sequence of CH2T2 was determined as shown in Table 5. As determined with carboxypeptidase A treatment, the carboxyl-terminus of CH2 was Leu, which was also the carboxyl-terminus of CH2T2. Therefore, CH2T2 represents the carboxyl-terminal fragment of CH2. Although the overlapping from the position 34 to 35 was not completed, it is unlikely that we missed an additional sequence in this segment because the amino acid composition of CH2 is consistent with its proposed sequence.

Amino Acid Sequence of Hydrophobic Segment (Residues 41–75)

The amino acid sequence of T1 was determined for the first 19 residues (Table 5). The amino-terminus of T1 was identical with that of CH2T2 which represents residues 35–40 of intact glycophorin HA. The hydrophobic peptide CH1, purified by gel chromatography on Sephadex LH-60, still was a mixture of two peptides, because Edman degradation of CH1 showed two PTH-amino acids for each cycle. The

Table 3. Amino acid compositions of horse glycophorin HA and the cyanogen bromide peptides[a]

	Glycophirin HA (1–120)	CB1B (1–59)	CB2 (60–120)
Asp	9.6 (10)	3.6 (4)	6.1 (6)
Thr	12.2 (14)	9.7 (10)	4.8 (4)
Ser	10.9 (12)	5.5 (5)	5.5 (7)
Glu	10.8 (11)	6.2 (7)	4.5 (4)
Pro	13.9 (14)	4.5 (4)	9.4 (10)
Gly	8.3 (9)	5.3 (6)	3.5 (3)
Ala	12.2 (11)	5.4 (5)	5.6 (6)
Cys	–	–	–
Val	8.4 (8)	3.4 (3)	5.1 (5)
Met	1.4 (1)	–	–
Hse	–	+[c] (1)	–
Ile	9.8 (10)	6.1 (6)	3.1 (4)
Leu	9.4 (9)	4.3 (4)	5.0 (5)
Tyr	1.9 (2)	–	1.9 (2)
Phe	2.3 (2)	2.1 (2)	–
Lys	1.4 (1)	–	1.5 (1)
His	1.2 (1)	1.2 (1)	–
Arg	5.7 (5)	1.3 (1)	3.9 (4)
Total residues	120	59	61
Yield		36%	45%

[a] Values are expressed in mol/mol peptide. Numbers in parentheses indicate mol/mol peptide assumed from sequence. No entry denotes 0.1 residues or less.
[b] Numbers in parentheses indicate position of peptide within sequence.
[c] The amino acid was not determined quantitatively.

separation of the two peptides, tentatively designated CH1A and CH1B, was unsuccessful because of their similar characteristics. Quantitative determination of PTH-amino acids at each cycle of Edman degradation, however, revealed that CH1A had the additional residues Ala-His-Asp-Phe at the amino-terminus of CH1B. The amino acid sequence of CH1A was determined for the first 22 residues (Tables 5 and 6). The amino-terminal sequence of CH1A overlaps with the sequence from the position 9 of T1. Since the sequence of CH1A also contains Met in position 59, it was evident that the cyanogen bromide peptide CB1B represents residues 1–59 of intact glycophorin HA; the amino acid composition of CB1B was consistent with the sequence data of residues 1–59 (Table 3). The amino acid sequence of CH1A and CB2 shows that glycophorin HA has a sequence of 23 hydrophobic amino acids (residues 50–72), uninterrupted by charged amino acid residues. These sequence data indicate that T1 is derived from the residues 35–75, but the amino acid composition of T1 does not correspond to that of the assumed sequence (Table 1), while sequence analysis of T1 does not indicate contamination with other peptides (data not shown). These results can be explained by the fact that the preparation is contaminated with the glycopeptide

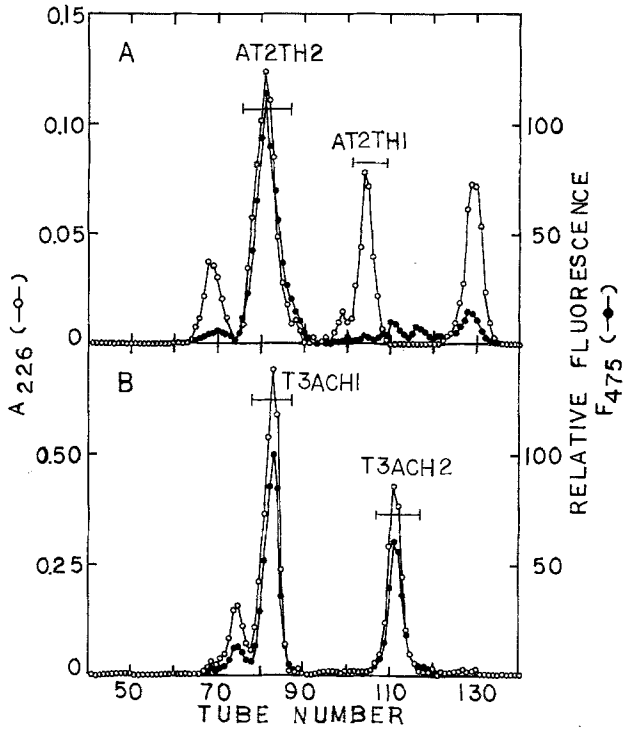

Fig. 7. Gel chromatography of the secondary peptides on Sephadex G-50. Procedures are given under Methods. (*A*): The thermolytic peptides of desialated T2 (7.0 mg). (*B*) The chymotryptic peptides of T3A (5.7 mg)

Table 4. Sequence analysis of peptide AT2TH1 after methanolysis

Cycle	PTH-amino acid identified	Yield[a] (nmol)	DNS-amino acid identified
1	Glu	42	Glu
2	–[b]	–	Thr
3	Ile	28	Ile
4	Ala	25	Ala
5	–	–	Thr
6	Gly	15	Gly
7	–	–	Ser
8	Pro	5	Pro

[a] Yield of PTH-amino acid was determined by gas-liquid chromatography.
[b] Dash denotes that no PTH-amino acid was obtained with a significant yield.

representing residues 1–75 (data not shown); since this glycopeptide has a blocked amino-terminus, it does not interfere with the sequence determination of T1. Nevertheless, at present the sequence obtained with T1 should be considered tentative.

Amino Acid Sequence of the Carboxyl-Terminal Segment (Residues 75–120)

The amino acid sequence of T3A was determined for the first 32 residues by Edman degradation

Table 5. Amino-terminal sequence of horse glycophorin and the peptides used for sequence studies

Proposed sequence		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
		<Glu-Thr-Ile-Ala-Thr-Gly-Ser-Pro-Pro-Ile-Ala-Gly-Thr-Ser-Asp-Leu-Ser-Thr-Ile-Thr-
		CHO CHO CHO CHO CHO CHO CHO
T2	(1–34)[a]	→ ( ) → → ( ) → ( ) → → → → → ( ) → → →
AT2TH1	(1–9)[a]	→[b] (→) → → (→) → (→) → →
AT2TH2	(10–34)	→ → → (→) → → → (→) (→) → (→)
AT2TH2-S3	(12–26)	→ (→) → → → (→) (→) → (→)
Proposed sequence		21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
		Ser-Ala-Ala-Thr-Pro-Thr-Phe-Thr-Thr-Glu-Gln-Asp-Gly-Arg-Glu-Gln-Gly-Asp-Gly-Leu-
		CHO CHO CHO
AT2TH2	(10–34)	(→) → → (→) → → → (→) → → → → →
AT2TH2-S3	(12–26)	(→) → → (→) → →
AT2TH2-S5	(27–34)	→ (→) → → → → → →
T1	(35–75)	→ → → → → →
CH2T2	(35–40)	→ → → → → →
Proposed sequence		41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
		Gln-Leu-Ala-His-Asp-Phe-Ser-Gln-Pro-Val-Ile-Thr-Val-Ile-Ile-Leu-Gly-Val-Met-Ala-
T1	(35–75)	→ → → → → → → → → → → → →
CH1A	(41–108)	→ →
CB2	(60–120)	→

[a] Edman degradation was performed after methanolysis with methanolic HCl.

[b] → denotes that PTH-amino acid was identified. (→) denotes that the residue was identified by the dansyl-Edman technique only. ( ) denotes that no PTH-amino acid was identified.

(Table 6). Its amino-terminal sequence overlapped with the sequence from the position 19 of CB2. Another peptide T3B was identical to T3A except for the lack of a Lys residue at the amino-terminus in T3A. The carboxyl-terminal sequence of T3A was determined by using another peptide. When T3A was digested with chymotrypsin, only the two peptides labeled T3ACH1 and T3ACH2 were obtained (Fig. 7*B*). The small peptide T3ACH2 was completely sequenced as shown in Table 6 and the carboxyl-terminal sequence of T3ACH1 was determined to be Thr-Ser-Leu by carboxypeptidase A treatment. Therefore, we could construct the entire sequence of T3A as shown in Table 6: T3A represented the residues 78–120 of intact glycophorin HA.

The peptide T4, determined to be Leu-Arg, should be derived from the residues 76–77. The chymotryptic peptide CH3 is identical with the secondary peptide T3ACH2 (Tables 1 and 6). These results indicate that T1 and CH1A represent residues 35–75 and 43–108, respectively. The amino acid composition of CH1 (the mixture of CH1A and CH1B) was consistent with the sequence data (Table 1).

The two peptides isolated from pool *C* of the tryptic digest (Fig. 2), T3C and T3D, showed the amino acid composition corresponding to residues 78–108 and 79–108, respectively. The two peptides could have been produced by residual chymotryptic activity in trypsin used. The carboxyl-terminal of glycophorin HA was determined to be Ser-Gln by carboxypeptidase A treatment. Thus, the peptides CH3 and T3A represent the carboxyl-terminal end of glycophorin HA. The complete amino acid sequence of glycophorin HA has been constructed as shown in Tables 5 and 6. The amino acid composition of glycophorin HA is consistent with the proposed amino acid sequence (Table 3).

Oligosaccharide Units of Glycophorin HA

Glycophorin HA contains 30% carbohydrate by weight, and the composition suggests the presence of galactose, N-acetylgalactosamine and N-glycolylneuraminic acid. The structure of the major oligosaccharide units has been determined (Fukuda et al., 1980). The carbohydrate composition of glycophorin HA is consistent with that determined from the proposed structure of the oligosaccharides. The ten residues of glycosylated amino acids are found within the amino-terminal domain of glycophorin HA at positions 2, 5, 7, 13, 17, 18, 20, 21, 24 and 28 (Table 5) when PTH-amino acid and DNS-amino acid are analyzed simultaneously for each cycle of Edman degradation. Glycophorin HA does not contain the N-glycosidic oligosaccharide which has been found in human glycophorin A and porcine glycophorin.

Table 6. Carboxyl-terminal sequence of horse glycophorin HA and the peptides used for sequence studies

Proposed sequence		61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 Gly-Ile-Ile-Gly-Ile-Ile-Leu-Leu-Leu-Ala-Tyr-Val-Ser-Arg-Arg-Leu-Arg-Lys-Arg-Pro-
CH1A	(41–108)	→ → → →
CB2	(60–120)	→ → → → → → → → → → → → → → → → → → → →
T4	(76–77)	→ →
T3A	(78–120)	→ → →
T3B	(79–120)	→ →

Proposed sequence		81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 Pro-Ala-Asp-Val-Pro-Pro-Pro-Ala-Ser-Thr-Val-Pro-Ser-Ala-Asp-Ala-Pro-Pro-Pro-Val-
CB2	(60–120)	→ → → →
T3A	(78–120)	→ → → → → → → → → → → → → → → → → → → →
T3B	(79–120)	→ → → → → → → → → → → → → → → → → → → →

Proposed sequence		101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 Ser-Glu-Asp-Asp-Glu-Thr-Ser-Leu-Thr-Ser-Val-Glu-Thr-Asp-Tyr-Pro-Gly-Asp-Ser-Gln
T3A	(78–120)	→ → → → → → → → →
T3B	(79–120)	→ → → → → → → → →
CH3	(109–120)	→ → → → → → → → → → → →
T3ACH1	(78–108)	← ← ←[a]
T3ACH2	(109–120)	→ → → → → → → → → → → →

[a] ← denotes that the sequence was determined with carboxypeptidase A.

Discussion

Human erythrocyte membranes contain at least three different glycophorins (Furthmayr, Tomita & Marchest, 1975). Similarly, horse erythrocyte membranes contain two different glycophorins, glycophorin HA as a major component and glycophorin HB as a minor component (Murayama et al., 1981). Therefore, the possibility must be considered that some of tryptic peptides as well as cyanogen bromide peptides were derived from the minor component, because those peptides were prepared from the glycophorin fraction which was a mixture of glycophorins HA and HB. This possibility was almost excluded by the fact that the yields of all the peptides used for sequence analysis are well over 20%; glycophorin HB accounts for at most 20% of the amino acids in the glycophorin fraction.

The major glycophorin of the horse erythrocyte membrane, glycophorin HA, is composed of 120 amino acids and 10 oligosaccharide chains and has a blocked amino-terminus. All of the oligosaccharides are linked to Thr or Ser residues by O-glycosidic bonds. A molecular weight of approximately 20,000 can be estimated for glycophorin HA from the amino acid sequence and the carbohydrate content.

Figure 8 shows the amino acid sequence of horse, human and porcine glycophorins. Gaps are included to optimize homology. In a previous report (Honma et al., 1980), a comparison was made between the amino acid sequence of porcine and human glycophorins. It was suggested that two general features are characteristics of the glycophorin structure. The first is that glycophorins are organized into three distinct domains: an amino-terminal glycopeptide domain, a hydrophobic domain and a carboxyl-terminal domain. The second is that glycophorins do not exhibit homology in amino acid sequence of the glycosylated domain which is exposed at the outside of the cell, but do exhibit striking homology in sequence of the hydrophobic domain which interacts with the lipid bilayer. The amino acid sequence of glycophorin HA is consistent with the two general features of glycophorin structure. Remarkable differences in amino acid sequence of the glycopeptide domains obtained from the various species are evident (Fig. 8). In addition there are differences in the number of the constituent amino acid residues, the distribution of the glycosylated residues, and the structure of the oligosaccharide chains. Since glycophorins contain the major portion of the carbohydrate of the erythrocyte membrane, and since the glycophorins of each species have markedly different structures, it is suggested that these proteins function as blood-group antigens, as virus receptors, or as markers of cell differentiation.

The complete amino acid sequence of hemagglutinins of influenza virus, another membrane glycoprotein, has recently been determined by DNA-sequencing techniques (Gething, Bye, Skehel & Waterfield,

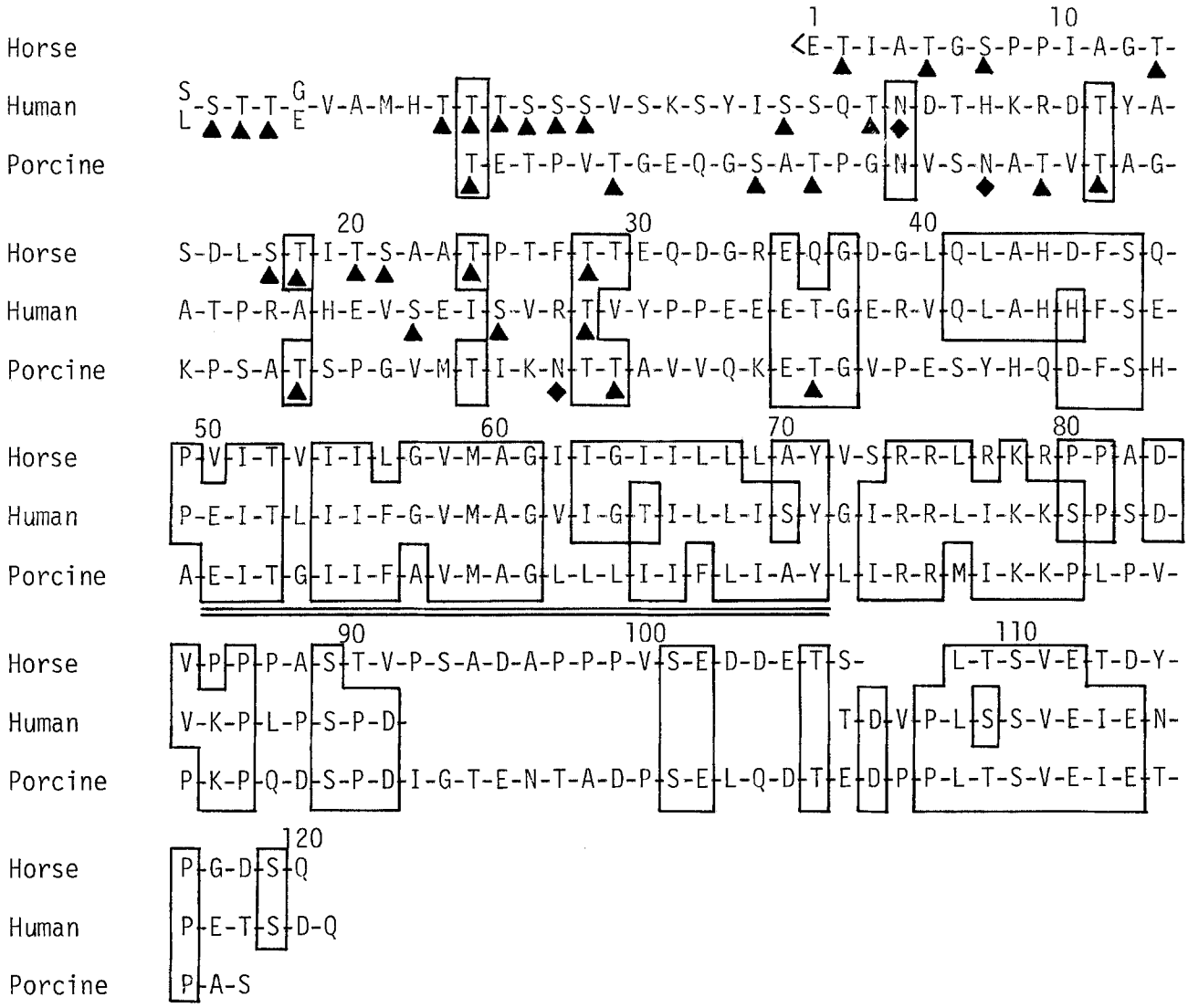

Fig. 8. Comparison of the amino acid sequence among horse, human and porcine glycophorins. Boxes enclose the identical residues in the sequences. Symbols (▲) indicate O-glycosidic oligosaccharide units and symbols (◆) indicate N-glycosidic oligosaccharide units

1980). The hydrophobic domain of the two types of the hemagglutinins shows only weak homology, indicating that a cluster of hydrophobic amino acids, regardless of the amino acid sequence, is indicative of a transmembrane structure. On the other hand, the glycophorins isolated from several species show striking homology in amino acid sequence of the hydrophobic domain. This could suggest that the structure of the hydrophobic domain is conserved to serve an important but yet unknown biological function.

The amino acid sequence of the three glycophorins reveals another interesting finding, namely that His residues are located at the amino-terminal side adjacent to the hydrophobic domain. The His residues possibly interact with the phosphate groups of phospholipids. In addition, there is a cluster of basic amino acids (residues 74–79) at the carboxyl-terminal side adjacent to the hydrophobic domain. This cluster also could interact with the phosphate groups of phospholipids.

References

Dodge, J.T., Mitchell, C., Hanahan, D.J. 1963. The preparation and chemical characteristics of hemoglobin-free ghosts of human erythrocytes. *Arch. Biochem. Biophys.* **100**:119–130

Fukuda, K., Tomita, M., Hamada, A. 1980. Isolation and characterization of alkaline-labile oligosaccharide units from horse glycophorin. *J. Biochem.* **87**:687–693

Furthmayr, H., Tomita, M., Marchesi, V.T. 1975. Fractionation of the major sialoglycopeptides of the human red blood cell membrane. *Biochem. Biophys. Res. Commun.* **65**:113–121

Gerber, G.E., Anderegg, R.J., Herlihy, W.C., Gray, C.P., Biemann, K., Khorana, H.G. 1979. Partial primary structure of bacterio-

rhodopsin: Sequencing methods for membrane proteins. *Proc. Natl. Acad. Sci. USA* **76**:227–231

Gething, M., Bye, J., Skehel, J., Waterfield, M. 1980. Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. *Nature (London)* **287**:301–306

Gray, W.R. 1972. Sequence analysis with dansyl chloride. *In:* Methods in Enzymology. C.H.W. Hirs and S.N. Timasheff, editors. Vol. 25B, pp. 333–334. Academic Press, New York

Hirs, C.H.W. 1967. Detection of peptides by chemical methods. *In:* Methods in Enzymology. C.H.W. Hirs, editor. Vol. 11, pp. 325–329. Academic Press, New York

Honma, K., Tomita, M., Hamada, A. 1980. Amino acid sequence and attachment sites of oligosaccharide units of porcine erythrocyte glycophorin. *J. Biochem.* **88**:1679–1691

Kawasaki, I., Itano, H.A. 1972. Methanolysis of the pyrolidone ring of amino-terminal pyroglutamic acid in model peptides. *Anal. Biochem.* **48**:546–556

Lai, C.Y. 1977. Detection of peptides by fluorescence methods. *In:* Methods in Enzymology. C.H.W. Hirs and S.N. Timasheff, editors. Vol. 47E, pp. 236–243. Academic Press, New York

Marchesi, V.T., Andrews, E.P. 1971. Glycoproteins: Isolation from cell membranes with lithium diiodosalicylate. *Science* **174**:1247–1248

Mendez, E., Li, C.Y. 1975. Regeneration of amino acids from thiazolinones formed in the Edman degradation. *Anal. Biochem.* **68**:47–53

Murayama, J., Takeshita, K., Tomita, M., Hamada, A. 1981. Isolation and characterization of two glycophorins from horse erythrocyte membranes. *J. Biochem.* **89**:1593–1598

Ozols, J., Gerard, C., Nobrega, F.G. 1976. Proteolytic cleavage of horse liver cytochrome b_5. Primary structure of the heme-containing moiety. *J. Biol. Chem.* **251**:6767–6774

Peterson, J.D., Nerrlich, S., Oyer, P.E., Steiner, D.F. 1972. Determination of the amino acid sequence of the monkey, sheep and dog proinsulin C-peptides by a semi-micro Edman degradation procedure. *J. Biol. Chem.* **247**:4866–4871

Pisano, J.J., Bronzert, Y.J., Brewe, H.B. 1972. Advances in the gas chromatographic analysis of amino acid phenyl- and methylthiohydantoins. *Anal. Biochem.* **45**:43–59

Reinhold, V.N. 1972. Gas liquid chromatographic analysis of constituent carbohydrates in glycoproteins. *In:* Methods in Enzymology. C.H.W. Hirs and S.N. Timasheff, editors. Vol. 25B, pp. 244–249. Academic Press, New York

Schroeder, W.A. 1972. Separation of peptides by chromatography on columns of Dowex 50 with volatile developers. *In:* Methods in Enzymology. C.H.W. Hirs and S.N. Timasheff, editors. Vol. 25B, pp. 203–213. Academic Press, New York

Summers, M.R., Symthers, G.W., Oroszlan, S. 1973. Thin-layer chromatography of sub-nanomole amounts of phenylthiohydantoin (PTH) amino acids on polyamide sheets. *Anal. Biochem.* **53**:624–628

Tomita, M., Furthmayr, H., Marchesi, V.T. 1978. Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry* **17**:4756–4770

Udenfriend, S., Stein, S., Bohler, P., Pairman, W., Leimgruber, W., Weigele, M. 1972. Fluorescamine: A reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science* **178**:871–872

Received 24 April 1981; revised 20 July 1981