The Structure of the Major Protein of the Human Erythrocyte Membrane

CHARACTERIZATION OF THE INTACT PROTEIN AND MAJOR FRAGMENTS

By ROBERT E. JENKINS and MICHAEL J. A. TANNER
Department of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.

(Received 20 May 1976)

Polypeptide 3, the major membrane-penetrating protein of the human erythrocyte membrane, was characterized, together with two major fragments derived by specific proteolysis of the native protein in the membrane. One fragment (fragment 3f) was obtained from thermolysin cleavage in the extracellular region of the protein, and the other (fragment T1) was derived from tryptic cleavage in the intracellular region of the protein. The results of N- and C-terminal group analysis suggest that fragment 3f contains the N-terminal region of polypeptide 3 and fragment T1 contains the C-terminal part of the molecule. The carbohydrate contents of the polypeptides suggest that carbohydrates are present in three regions of the molecule, much of this carbohydrate being present in the C-terminal part of the molecule. This region of the protein also contains the receptors for concanavalin and the lectins from Phaseolus vulgaris and Ricinis communis, and our results suggest that there is heterogeneity in the carbohydrate chains present in the C-terminal region of polypeptide 3. These data are related to the folding of polypeptide 3 in the erythrocyte membrane.

Materials and Methods

Carbohydrate analysis was done by a g.l.c. method (Clamp, 1974; Tanner et al., 1976). The methods used for sodium dodecyl sulphate/polyacrylamide-gel electrophoresis, labelling of polypeptides with 125I and the preparation of peptide ‘maps’ have been described (Boxer et al., 1974; Tanner et al., 1976). Amino acid analysis was carried out as described by Tanner & Gray (1971), with a Technicon TSM-1 amino acid analyser.

Preparation of polypeptide 3 and fragments 3f and T1

The method used for the preparation of intact polypeptide 3 was described previously (Tanner et al., 1976). The two fragments 3f and T1 were prepared by the addition of suitable proteolysis steps to the procedure used to prepare intact polypeptide 3. An outline of these preparations is shown in Scheme 1. Each preparation was carried out on 1 blood-bank unit of whole blood. Trypsin or thermolysin digestion of intact erythrocytes was carried out as described by Jenkins & Tanner (1975), except that the enzyme concentration was lowered from 0.5mg/ml to 0.1mg/ml, and the incubation was allowed to proceed for 90 min at 37°C. Trypsin digestion of erythrocyte ‘ghosts’ was done in the
R. E. JENKINS AND M. J. A. TANNER

Washed erythrocytes

Trypsin digestion, wash

Trypsin-treated erythrocytes

Hypo-osmotic lysis, wash

'Ghosts' from trypsin-treated erythrocytes

Trypsin digestion
(iso-osmotic conditions)

wash

Trypsin-treated 'ghosts' from trypsin-treated erythrocytes

Extraction with NaOH

Insoluble residue

Dissolution in sodium dodecyl sulphate, gel filtration

Fragment T1

Thermolysin digestion, wash

Thermolysin-treated erythrocytes

Hypo-osmotic lysis, wash

'Ghosts' from thermolysin-treated erythrocytes

Extraction with NaOH

Insoluble residue

Dissolution in sodium dodecyl sulphate, gel filtration

Fragment 3f

Scheme 1. Preparation of fragments 3f and T1

Experimental details are given in the Materials and Methods section.

presence of iso-osmotic KCl and 5mM-magnesium acetate as described in the preceding paper (Jenkins & Tanner, 1977). In each case the fragment produced as a result of proteolysis was purified by extraction of the 'ghosts' with cold 0.1M-NaOH, followed by gel filtration of the insoluble residue on a column (7.5cm x 80cm) of Sephadex G-150 equilibrated with 1% sodium dodecyl sulphate/1mM-EDTA, pH8.0, as previously described (Tanner et al., 1976).

The fragments were detected in the column fractions by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis, and the required fractions were pooled and concentrated by pressure dialysis under N2 pressure over an Amicon PM-10 membrane. The concentrate (approx. 20ml) was then dialysed against several changes each of 8 litres of water for 24h, 500ml of 40% (v/v) methanol for 12h, and for a further 24h against several changes of 8 litres of water. The material was then freeze-dried and stored at -18°C. Each preparation yielded 80-120mg of white fluffy solid material. With preparations of fragment 3f it was sometimes necessary to extract this material further by suspending it at a concentration of 2mg/ml in water for 2min, to remove traces of the erythrocyte sialoglycoprotein. The insoluble material remaining after two washes with water was collected and freeze-dried.

The purity of the protein preparations was assessed by cutting out and weighing the peaks obtained from densitometer scans of sodium dodecyl sulphate/polyacrylamide gels of the samples after staining with Coomassie Blue. The gels were immersed in 7.5% (v/v) acetic acid/5% (v/v) methanol in a glass spectrophotometer cell and scanned at 550nm in a Gilford linear transporter attached to a Gilford spectrophotometer. Radioautographs were scanned as described by Jenkins & Tanner (1977).
Estimation of molecular weight

The molecular weight of fragment T1 was estimated by gel filtration through a column (2 cm × 125 cm) of Sephadex G-100 equilibrated with 1% sodium dodecyl sulphate/1 mM-EDTA, pH 8.0. The column was calibrated with ovotransferrin (mol. wt. 80000, a gift from Dr. J. Williams, University of Bristol, Bristol, U.K.), bovine serum albumin [mol. wt. 67000, Sigma (London) Chemical Co., London S.W.6, U.K.], glutamate dehydrogenase (mol. wt. 56000, a gift from Dr. J. J. Holbrook, University of Bristol), lactate dehydrogenase (mol. wt. 36000, a gift from Mr. D. Parker, University of Bristol), soya-bean trypsin inhibitor (mol. wt. 22500, Sigma) and horse heart cytochrome c [mol. wt. 13400, Boehringer Corp. (London) Ltd., Ealing, London W.5, U.K.]. All the molecular weights given are of the polypeptide subunits.

N-Terminal and C-terminal analysis

The method of Stark (1967) was used for quantitative N-terminal analysis. Carbamoylation was carried out in the presence of 0.5% sodium dodecyl sulphate on 5 mg of each protein preparation. N-Terminal analysis was also done by the dansylation procedure (Gray, 1972) with the modification described by Langdon (1974). The dansylation was done in the presence of 0.3% sodium dodecyl sulphate.

C-Terminal amino acids were determined by digestion with a mixture of carboxypeptidase A (Worthington Biochemical Corp., Freehold, NJ, U.S.A.) and carboxypeptidase B (Sigma). Both enzymes were pretreated with di-isopropyl phosphorofluoridate. A sample (21 mg) of each peptide was digested with 0.75 mg of carboxypeptidase A and 0.37 mg of carboxypeptidase B. The reaction mixture contained 200 nm-norleucine, 0.5 M-N-ethylmorpholine/acetate buffer, pH 8.6, and 1.2% (w/v) sodium dodecyl sulphate (Guidotti, 1960). The mixture was incubated for 2 h at 37°C, samples being removed before incubation and at intervals of 10, 20, 60 and 120 min and then treated as described by TANNER & BOXER (1972).

Determination of γ-glutamyl-e-lysine in intact polypeptide 3

γ-Glutamyl-e-lysine was measured by isotope dilution. The authentic 14C-labelled dipeptide was synthesized by the incorporation of [14C]lysine (The Radiochemical Centre, Amersham, Bucks., U.K.) into casein, as described by Matacic & Loewy (1968). Human plasma Factor XIII was a gift from Dr. R. D. Cooke, University of Bristol. The γ-glutamyl-e-[14C]lysine formed was released from casein by an enzymic-digestion procedure similar to that of Pisano et al. (1969). The dry casein (20 mg) was dissolved in 5 ml of 0.25 M-ammonium acetate/12.5 mM-CaCl2 and the pH adjusted to 8.0 with NH3.

Pronase (Calbiochem, Hereford, U.K.; 4 mg) was added, and incubation carried out at 37°C for 4 days, after which an additional 4 mg of Pronase was added. After a further 4 days of incubation, 2 mg of imidodipeptidase (Miles-Seravac, Stoke Poges, Bucks., U.K.), 1 mg of leucine aminopeptidase (Sigma) and 3 mg of peptidase (grade III, Sigma) were added, and incubation continued for 5 more days. The mixture was then freeze-dried. Bacterial growth was inhibited by thymol.

The dried digest was dissolved in water, made 12% (w/v) in trichloroacetic acid and centrifuged at 1000 g for 5 min. The clear supernatant was diluted with 5 vol. of water and applied to a column (1 cm × 10 cm) of Bio-Rad AG50-X4 (H+ form; 200–400 mesh). The column was washed with 120 ml of water and the radioactivity eluted with 60 ml of 1 M-pyridine (Pisano et al., 1969). The pyridine eluate was freeze-dried, and the radioactive material further purified by paper chromatography in butan-1-ol/acetic acid/water (7:2:10, by vol.) as solvent (Tanner & Gray, 1971) and paper electrophoresis (at pH 1.65 in 6.7%, v/v, formic acid). After each step the dipeptide was located by radioautography, eluted from the paper with water, and freeze-dried. The chromatographically pure material was found on dansylation to contain the dipeptide glutamyl-lysine as a contaminant. To destroy this dipeptide the preparation was incubated for 24 h with 0.25 mg of carboxypeptidase B in 100 μl of 1 M-N-ethylmorpholine/acetate buffer, pH 8.6. The material was again purified by paper chromatography and paper electrophoresis as described above. On amino acid analysis the final preparation was found to contain equimolar amounts of glutamic acid and lysine. Only α-Dns*-lysine and Dns-glutamate were obtained after dansylation and acid hydrolysis. The specific radioactivity was determined by liquid-scintillation counting, and amino acid analysis, a suitable quench correction factor being derived by external standardization. Purified γ-glutamyl-e-[14C]lysine (100 nmol) was added to 10 mg of intact polypeptide 3, and the mixture was then subjected to the enzymic-digestion procedure described above. The radioactive γ-glutamyl-e-lysine was re-purified as described above, and the specific radioactivity again determined.

Results and Discussion

Characterization of polypeptide 3 and fragments 3f and T1

Fragment 3f (which we have previously referred to as 'ε'; Boxer et al., 1974; Jenkins & Tanner, 1975) is the 60000-mol. wt. fragment derived from polypeptide 3 which is retained in the membrane on treatment of erythrocytes with thermolysin, Pronase,

* Abbreviation: Dns-, 5-dimethylaminonaphthalene-1-sulphonyl-.
chymotrypsin or subtilisin (Jenkins & Tanner, 1975). Fragment T1 is a product resulting from the cleavage of polypeptide 3 by trypsin when erythrocyte 'ghosts' are treated with trypsin [see the preceding paper, Jenkins & Tanner (1977)]. Fragment 3f results from proteolytic cleavage in the extracellular region of polypeptide 3, whereas fragment T1 arises as a result of cleavage at the intracellular region of the protein. Scheme 1 summarizes the methods used to prepare the two fragments.

Figs. 1(a) and 1(b) show the sodium dodecyl sulphate/polyacrylamide-gel electrophoresis patterns obtained for purified fragments 3f and T1 on gels containing 8% (w/v) acrylamide. The band containing fragment 3f accounted for 75% of the total stain. Fragment T1 migrated as a diffuse band with a long trailing edge which contained a sharply defined impurity band. The diffuse band in the fragment-T1 preparation contained 91% of the total stain. The sodium dodecyl sulphate/polyacrylamide-gel electrophoresis pattern of the polypeptide-3 preparation used for the experiments described here has been described by Tanner et al. (1976).

Fragment T1 migrated as a diffuse band on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis in an 8% acrylamide gel (Jenkins & Tanner, 1977, and Fig. 1b). On electrophoresis in a 5% (w/v) acrylamide gel the diffuse band was further resolved into two peaks (T1a and T1b, Fig. 1c) and a broad trailing edge containing the sharply defined impurity band. A sample of fragment T1 was labelled with 125I and the labelled components were separated by electrophoresis in a 5% (w/v) acrylamide gel (Fig. 1d). The bands corresponding to T1a, T1b and the trailing edge were eluted from the gel. Peptide 'maps' were prepared from thermolysin digests of these labelled fractions, and the labelled peptides detected by radioautography. Fragments T1a and T1b gave the same pattern of labelled peptides, and the trailing edge contained some minor labelled peptides in addition to those found in the peptide 'maps' of fragments T1a and T1b. The additional minor labelled peptides found in the trailing-edge fraction probably arose from the sharply defined impurity band present in the fragment-T1 preparations (Figs. 1b and 1c). This suggests that the polypeptide chains of fragments T1a and T1b are the same on the basis of this criterion. Evidence that suggests that fragments T1a and T1b differ in carbohydrate content is discussed below. The T1 preparation migrated as a single band with mol. wt. 60000 on gel filtration on a column of Sephadex G-100 equilibrated with 1% sodium dodecyl sulphate/1mM-EDTA (disodium salt), pH 8.0, which had been calibrated with marker proteins of known molecular weights.

The amino acid compositions of the three polypeptides are shown in Table 1. The data for polypeptide 3 are in good agreement with those of Yu & Steck (1975) and Ho & Guidotti (1975), though they agree less well with our own earlier data (Tanner & Boxer, 1972). Fragment 3f contains noticeably more glutamic acid and/or glutamine than fragment T1, whereas the latter has higher
Table 1. Amino acid content of polypeptide 3 and fragments 3f and T1

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Polypeptide 3</th>
<th>Fragment 3f</th>
<th>Fragment T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>63 ± 4</td>
<td>71 ± 5</td>
<td>59 ± 2</td>
</tr>
<tr>
<td>Thr</td>
<td>53 ± 1</td>
<td>52 ± 3</td>
<td>53 ± 1</td>
</tr>
<tr>
<td>Ser</td>
<td>72 ± 3</td>
<td>70 ± 3</td>
<td>68 ± 2</td>
</tr>
<tr>
<td>Glu</td>
<td>116 ± 3</td>
<td>136 ± 9</td>
<td>77 ± 3</td>
</tr>
<tr>
<td>Pro</td>
<td>71 ± 4</td>
<td>72 ± 5</td>
<td>71 ± 1</td>
</tr>
<tr>
<td>Gly</td>
<td>81 ± 3</td>
<td>78 ± 4</td>
<td>79 ± 3</td>
</tr>
<tr>
<td>Ala</td>
<td>77 ± 6</td>
<td>75 ± 2</td>
<td>70 ± 3</td>
</tr>
<tr>
<td>Δ-Cys</td>
<td>8.5 ± 1.5</td>
<td>9.6 ± 3.3</td>
<td>8.7</td>
</tr>
<tr>
<td>Val</td>
<td>56 ± 2</td>
<td>51 ± 3</td>
<td>75 ± 2</td>
</tr>
<tr>
<td>Met</td>
<td>30 ± 1</td>
<td>20 ± 0.5</td>
<td>34</td>
</tr>
<tr>
<td>Ile</td>
<td>34 ± 1</td>
<td>35 ± 2</td>
<td>46 ± 2</td>
</tr>
<tr>
<td>Leu</td>
<td>134 ± 4</td>
<td>133 ± 1</td>
<td>127 ± 2</td>
</tr>
<tr>
<td>Tyr</td>
<td>26 ± 4</td>
<td>26 ± 1</td>
<td>30 ± 1</td>
</tr>
<tr>
<td>Phe</td>
<td>53 ± 1</td>
<td>47 ± 1</td>
<td>66 ± 1</td>
</tr>
<tr>
<td>Lys</td>
<td>38 ± 2</td>
<td>33 ± 5</td>
<td>56 ± 1</td>
</tr>
<tr>
<td>His</td>
<td>28 ± 2</td>
<td>33 ± 4</td>
<td>23 ± 1</td>
</tr>
<tr>
<td>Arg</td>
<td>63 ± 7</td>
<td>59 ± 4</td>
<td>61 ± 4</td>
</tr>
</tbody>
</table>

Cysteine and methionine were measured as cysteic acid and methionine sulphone respectively in hydrolysates of performic acid-oxidized samples (Hirs, 1967). The values given are the means ± s.d. for each polypeptide as follows: polypeptide 3, three preparations, all of which were performic acid-oxidized; fragment 3f, three preparations, two of which were performic acid-oxidized; fragment T1, one preparation, from which both untreated and performic acid-oxidized samples were analysed. The data for cysteine and methionine were derived only from the performic acid-oxidized samples.

Table 2. Terminal amino acid analysis of polypeptide 3 and fragments 3f and T1

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Polypeptide 3</th>
<th>Fragment 3f</th>
<th>Fragment T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>—</td>
<td>0.53</td>
<td>—</td>
</tr>
<tr>
<td>Val</td>
<td>—</td>
<td>0.55</td>
<td>—</td>
</tr>
<tr>
<td>Leu</td>
<td>0.23</td>
<td>0.28</td>
<td>0.43</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.30</td>
<td>1.59</td>
<td>—</td>
</tr>
<tr>
<td>Phe</td>
<td>0.24</td>
<td>0.39</td>
<td>0.13</td>
</tr>
<tr>
<td>His</td>
<td>—</td>
<td>—</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Quantitative determination of N-terminal amino acids was done on all three preparations. The major amino acid obtained from both polypeptide 3 and fragment 3f was serine, but the yields of this amino acid were only 0.12 mol/mol of polypeptide 3 and 0.15 mol/mol of fragment 3f after correction for losses during the procedure (Stark, 1967). No other amino acid was recovered in yields of more than 0.07 mol/mol of polypeptide 3 or 0.12 mol/mol of fragment 3f. No terminal amino acid could be detected by the dansylation procedure (Gray, 1972) in either of these polypeptides. When the cyanate procedure was applied to fragment T1 a significant yield of glycine was obtained (0.53 mol/mol of polypeptide), but no other amino acid accounted for more than 0.16 mol/mol of polypeptide. A clear indication of the presence of N-terminal glycine was also obtained when the dansylation procedure was done on fragment T1.

The low recoveries of amino acids on N-terminal analysis of polypeptide 3 and fragment 3f suggest that they probably both have blocked N-terminal residues. The result for polypeptide 3 is consistent with our earlier data (Tanner & Boxer, 1972), and more recent results obtained by the dansylation procedure alone (Ho & Guidotti, 1975; Yu & Steck, 1975). The similarity between the N-terminus of polypeptide 3 and fragment 3f could mean that fragment 3f contains the N-terminal region of polypeptide 3, whereas fragment T1, which has N-terminal glycine, does not.

Analysis of the C-terminal amino acids of the three preparations was done by using digestion with carboxypeptidases A and B in the presence of sodium dodecyl sulphate (Tanner & Boxer, 1972).

### Contents

...methionine, valine, phenylalanine and lysine. When the amino acids are grouped as apolar residues, charged residues and residues of intermediate polarity as described by Yu & Steck (1975), neither of the fragments shows any noticeable increase in hydrophobic or apolar character compared with the intact protein.

Birkbichler et al. (1973) have shown that certain membrane proteins contain covalent cross-links of the γ-glutamyl-e-lysyl type. We have been unable to detect any γ-glutamyl-e-lysyl in our polypeptide-3 preparation when using the isotope-dilution method. Langdon (1974) reported that the collagen and elastin type of lysine-derived cross-link is present in the proteins of the human erythrocyte 'ghost'. A more detailed investigation (Bailey et al., 1976) suggests that the compounds thought to be derived from lysine–aldehyde cross-links are in fact hexosyl-lysine derivatives. There is no evidence for the presence of any covalent inter-chain cross-links in polypeptide 3 which could give rise to any multi-chain subunit structure of the 93000-mol.wt. form of the protein.

Vol. 161
The time-dependence of the digestion was examined, and the values obtained after a 20 min digestion period are shown in Table 2; 1.6 mol of tyrosine, approx. 0.5 mol each of valine and aspartic acid/mol of polypeptide and smaller amounts of phenylalanine and leucine were obtained from fragment 3f at this stage. All of the tyrosine was released rapidly, followed by the slower release of valine and aspartic acid. Polypeptide 3 and fragment T1 yielded amino acids in amounts lower than expected. Phenylalanine and leucine were obtained in both cases, but histidine and tyrosine were also released on digestion of polypeptide 3, and arginine was obtained on treatment of fragment T1. In both cases leucine was the first amino acid to be released.

The results for fragment 3f suggest that it may have a C-terminal sequence of Tyr-Tyr. We cannot exclude the possibility that contaminating peptides fortuitously contribute to the large amount of tyrosine obtained, but even in this case the results strongly suggest that fragment 3f has a C-terminal tyrosine residue. It is not possible to come to a firm conclusion about the nature of the C-terminal amino acids of polypeptide 3 and fragment T1. The low recoveries of amino acids obtained in these cases could result from incomplete digestion by the enzyme, or may be due to heterogeneity of the C-termini in the preparations. The small amount of contaminants found on polyacrylamide-gel electrophoresis of these preparations suggests that the latter is unlikely to be the case unless the heterogeneity exists within the polypeptide-3 and fragment-T1 bands.

Fragment T1 is a product of trypsin digestion, whereas fragment 3f is derived from thermolysin digestion of polypeptide 3. The initial release of leucine from both polypeptide 3 and fragment T1 suggests that trypsin cleavage does not occur at the C-terminal side of fragment T1, and that both polypeptides may have C-terminal leucine residues. The smaller amount of more-slowly released arginine obtained from fragment T1 is probably due to contaminating peptides, as most of these would be expected to have a lysine or an arginine residue at their C-terminal. Fragment T1 probably contains the C-terminal region of polypeptide 3, whereas fragment 3f contains the N-terminal region of the molecule. This interpretation is consistent with the results of N- and C-terminal analysis, and with the 'S-shaped' structure that we have proposed for the protein and the sites of cleavage by thermolysin and trypsin (Jenkins & Tanner, 1975, 1977). Consideration of the molecular weights of fragments 3f and T1 suggests that a substantial region of polypeptide chain must be common to both.

**Carbohydrate contents of polypeptide 3, fragment 3f and fragment T1**

The carbohydrate contents of the three polypeptides are shown in Table 3. The values are based on the sum of the measured protein and carbohydrate contents, because of the presence of residual detergent in these preparations, all of which were derived from chromatography in sodium dodecyl sulphate (Tanner & Boxer, 1972). The distribution of carbohydrates found in polypeptide 3 is similar to that reported by other workers, but our value for the carbohydrate content of polypeptide 3 is higher than the values obtained by Ho & Guidotti (1975) and Yu

<table>
<thead>
<tr>
<th>Carbohydrate contents of polypeptide 3 and fragments 3f and T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The values given are means ± s.d. of analyses on two preparations of fragment 3f and two preparations of fragment T1. The carbohydrate content is expressed as the percentage of the sum of the protein and carbohydrate contents. Protein was determined by the method of Lowry et al. (1951). Mol. wts. of 93000, 60000 and 60000 were used for polypeptide 3, fragment 3f and fragment T1 respectively (see the text) for calculating the data, and the predicted carbohydrate contents of fragment T2, the region not found in fragment 3f, and the region common to both fragments 3f and T1 are based on calculated mol. wts. of 33000, 33000 and 27000 for each of these regions.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Content (mol/mol of polypeptide)</th>
<th>Analytical results</th>
<th>Calculated (predicted) values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Region not found in fragment 3f</td>
<td>Region common to fragments 3f and T1</td>
</tr>
<tr>
<td>Polypeptide 3*</td>
<td>Fragment 3f</td>
<td>Fragment T1</td>
</tr>
<tr>
<td>Fucose</td>
<td>4 ± 1</td>
<td>2 ± 0</td>
</tr>
<tr>
<td>Mannose</td>
<td>7 ± 1</td>
<td>5 ± 1</td>
</tr>
<tr>
<td>Galactose</td>
<td>24 ± 4</td>
<td>10 ± 3</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>25 ± 5</td>
<td>9 ± 2</td>
</tr>
<tr>
<td>GalNAc</td>
<td>4 ± 1</td>
<td>3 ± 1</td>
</tr>
<tr>
<td>Sialic acids</td>
<td>5 ± 3</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>Carbohydrate (%)</td>
<td>15.1</td>
<td>9.8</td>
</tr>
</tbody>
</table>

* Values taken from Tanner et al. (1976).
& Steck (1975). However, these workers quote values based on the weight, rather than the protein content, of their preparations, and their samples are likely to have contained residual amounts of the detergents used in their preparation.

Fragment T1 contains substantially more carbohydrate than does fragment 3f (Table 3), but both fragments have a similar distribution of sugars. From these data it is possible to predict the approximate carbohydrate composition of the regions of polypeptide 3 which are not found in fragment T1 (i.e. fragment T2; Jenkins & Tanner, 1977) and those which are not found in fragment 3f (i.e. the regions cleaved from the protein to give rise to fragment 3f on thermolysin treatment of intact erythrocytes; Jenkins & Tanner, 1977). In a similar way it is possible to estimate the carbohydrate content of the region of polypeptide common to both fragments 3f and T1. These calculated values are given in Table 3.

On the basis of mol.wts. of 93000, 60000 and 60000 for polypeptide 3 and fragments 3f and T1 respectively (Tanner & Boxer, 1972; Boxer et al., 1974), the mol.wts. of the polypeptide chains of the three polypeptides can be estimated to be 80000, 54000 and 30000 respectively, when allowance is made for the carbohydrate content of each molecule. The region of polypeptide common to both fragments 3f and T1 has a mol.wt. of 24000.

Lectin receptors on fragment T1 and fragment 3f
Studies of the lectin-binding components present in human erythrocyte membranes have shown that concanavalin A and the lectins from Ricinis communis and Phaseolus vulgaris bind to polypeptide 3 (Findlay, 1974; Adair & Kornfeld, 1974; Tanner & Anstee, 1976). The binding of these lectins to the preparations of fragments T1 and 3f separated by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis was studied by the method of Tanner & Anstee (1976). The results for fragment T1 are shown in Fig. 2. Concanavalin A and the lectin from Phaseolus vulgaris bound predominantly to the T1s band, whereas the lectin from Ricinis communis gave a broader pattern of binding which was primarily associated with the T1b band. This suggests that there are differences in the carbohydrate contents of the two bands. A similar heterogeneity has been observed in the binding of concanavalin A to the intact protein. This lectin only binds to the leading edge of the intact polypeptide-3 band (Tanner & Anstee, 1976). Findlay (1974) has shown that only a proportion of polypeptide 3 can bind to a concanavalin A-Sepharose column, and Yu & Steck (1975) have also demonstrated heterogeneity in the carbohydrate content of different regions of the diffuse band obtained for polypeptide 3 on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis. The diffuse nature of both the polypeptide-3 and fragment-T1 bands probably results from this carbohydrate heterogeneity. By contrast, experiments similar to those described in Fig. 2 showed that fragment 3f did not bind any of the three lectins. Thus the receptors for concanavalin A and the lectins from Ricinis communis and Phaseolus vulgaris are located in the C-terminal (T1) region of the protein.

The relationship between polypeptide 3 and fragments 3f and T1, and the distribution of oligosaccharide chains and lectin receptors on the polypeptide chain of the protein, are illustrated in Fig. 3. Fig. 4 relates this information to what is
Fig. 3. Alignment of fragments 3f and T1

Y indicates the presence of oligosaccharide chains. The number of oligosaccharide chains in each region is not known, and the diagram indicates only the relative proportions of carbohydrate present in each region.

Fig. 4. Structure of polypeptide 3

Fragments 3f and T1 and carbohydrates are related to the folding of polypeptide 3 in the erythrocyte membrane (Jenkins & Tanner, 1975, 1977). Y indicates the presence of oligosaccharide chains, and the diagram indicates the relative proportions of carbohydrate in each region, rather than the exact number of oligosaccharide chains.
known about the folding of the native protein in the erythrocyte membrane. It should be noted that, although the extracellular arms of the protein are shown separately, it is likely that extensive interactions might occur between them. In the preceding paper (Jenkins & Tanner, 1977) the results of Steck et al. (1976) are considered in relation to the model shown in Fig. 4.

The question of the homogeneity of polypeptide 3 and the fragments that we have studied merits further discussion. There is increasing evidence which suggests that the carbohydrates bound to polypeptide 3 are heterogeneous (Findlay, 1974; Yu & Steck, 1975; Tanner & Anstee, 1976), and our results suggest that this is due to heterogeneity in the carbohydrates in the C-terminal (T1) region of the polypeptide chain. The question of the homogeneity of the polypeptide chains in the preparations must also be considered. It is known that several distinct erythrocyte-‘ghost’ polypeptides migrate in the same region as polypeptide 3 on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis. These include the phosphorylated intermediate of the erythrocyte Na\(^+\)-K\(^+\)-activated adenosine triphosphatase (Avruch & Fairbanks, 1972), acetylcholinesterase (Bellhorn et al., 1970) and a component which is more resistant to the action of trypsin on erythrocyte ‘ghosts’ than is polypeptide 3 (Reichstein & Blostein, 1975; Jenkins & Tanner, 1977). However, all these components are much less abundant than polypeptide 3 in erythrocyte ‘ghosts’, and, if present, they are unlikely to cause significant contamination (in a protein-chemical sense) of our polypeptide-3 preparations. An alternative possibility is that the polypeptide-3 band may contain a family of closely related proteins (Jenkins & Tanner, 1975; Yu & Steck, 1975). On radioactive labelling of tyrosine residues and subsequent fragmentation, the protein behaves as a single species (Boxer et al., 1974; Jenkins & Tanner, 1975, 1977). The mode of generation of fragment 3f, its homogeneity on polyacrylamide-gel electrophoresis and the presence of the expected amounts of a single C-terminus in the fragment-3f preparation are consistent with homogeneity in a substantial portion which includes the N-terminal region of the protein. The presence of a single N-terminus in fragment T1 (which overlaps fragment 3f) is also consistent with this. However, the interpretation of the results of our C-terminal analyses on polypeptide 3 and fragment T1 is more ambiguous, and we cannot exclude the possibility that the results reflect heterogeneity in the C-termini present in the preparation. This would suggest that, if any heterogeneity exists in the polypeptide of band 3, it is likely to be restricted to the C-terminal region of the protein. It is clear that detailed amino acid-sequence studies on the protein will be necessary before this question can be fully resolved.

We are grateful to Professor J. R. Clamp for providing our carbohydrate analysis. R. E. J. was the recipient of a Science Research Council Studentship. The work was supported in part by grants from the Medical Research Council and Wellcome Trust.

**References**


