The Major Human Erythrocyte Membrane Protein

EVIDENCE FOR AN S-SHAPED STRUCTURE WHICH TRAVERSES THE MEMBRANE TWICE AND CONTAINS A DUPLICATED SET OF SITES

By ROBERT E. JENKINS and MICHAEL J. A. TANNER

Department of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.

(Received 14 October 1974)

The structure of the major human erythrocyte membrane protein (protein E) was investigated by studying the products of proteolysis of the native protein in the membrane. The distribution and location of the tyrosine residues labelled by radioiodination by lactoperoxidase was determined. Proteolysis of the extracellular region of the protein by thermolysin released four tyrosine-containing peptides, all of which were also found to remain in the major fragment that is retained in the membrane. The presence of these duplicated sites in the extracellular region of the protein was confirmed by limited trypsin digestion of the intracellular region of the protein. Two groups of fragments were obtained. Both groups contained a set of the extracellular labelled sites, but they differed in containing distinct groups of intracellular sites, showing that the two sets of extracellular sites are linked by an intracellular region of the protein. The polypeptide chain thus traverses the membrane twice. An S-shaped model which is consistent with these data is proposed.

The human erythrocyte 'ghost' contains two major proteins which have been shown to traverse the permeability barrier of the cell (Bretscher, 1971*a,b*; Segrest *et al.*, 1973; Boxer *et al.*, 1974). One of these proteins, the erythrocyte sialoglycoprotein, has been extensively investigated and sequence analysis carried on portions of it (Segrest *et al.*, 1972). The protein carries an unusually large amount of carbohydrate, and the composition and distribution of carbohydrate chains on the polypeptide have been studied (Thomas & Winzler, 1969). All the carbohydrate is bound to the extracellular region of the protein, whereas the polypeptide backbone crosses the membrane once and contains a hydrophobic region which is believed to be buried within the membrane (Segrest *et al.*, 1972). The other major protein, protein E (band III of Fairbanks *et al.*, 1971), has a subunit molecular weight of 93 000, and is the most abundant polypeptide in the membrane. Although this protein has been isolated, and has been shown to contain some carbohydrate (Tanner & Boxer, 1972; Steck & Dawson, 1974), little is known about its structure.

We report the results of an investigation into the structure of the large, potentially complex, protein E. We have studied the distribution of radioiodinated tyrosine residues along the polypeptide chain of protein E, by using the fragments produced by proteolytic digestion *in situ* of protein E in the erythrocyte membrane. Our data suggest that the protein contains a duplicated set of extracellular

sites. By using the known cellular location of the labelled sites we show that the protein traverses the membrane at least twice. An 'S-shaped' structure is proposed for the native protein in the erythrocyte membrane.

Methods

The preparation of labelled human erythrocytes, 'ghosts', and the procedures used for sodium dodecyl sulphate–polyacrylamide-gel electrophoresis, isolation of labelled proteins, and the preparation of peptide 'maps' have been described (Boxer *et al.*, 1974).

Bovine serum albumin [Sigma (London) Chemical Co., Kingston-upon-Thames, Surrey, U.K.; fraction V powder] was pretreated with di-isopropyl phosphor-fluoridate to inhibit contaminating proteinases, by treating 100 ml of 10% (w/v) bovine serum albumin in 0.13 M-NaCl–10 mM-sodium phosphate, pH 7.4, with 0.05 ml of di-isopropylphosphorfluoridate at 18°C for 60 min.

Proteinase treatment of intact erythrocytes

For trypsin digestion, labelled cells were washed with 0.1 M-sodium phosphate buffer, pH 8.0. A 20% suspension of cells in the same buffer was digested at 37°C for 30 min with 0.5 mg of trypsin (toluene-*p*-sulphonamidophenylalanine chloromethyl ketone-treated)/ml. The cells were washed four times with 0.15 M-NaCl before lysis or subsequent digestion

with thermolysin. Thermolysin digestion was done at pH 8.5 as described by Boxer *et al.* (1974). The supernatant from thermolysin digestion was made 10 mM in EDTA, and heated at 100°C for 10 min. The sample was freeze-dried before preparing peptide 'maps'.

When digestion was done as described by Cabantchik & Rothstein (1974b), trypsin digestion was carried out as described above, then the cells were washed three times with 0.15 M-NaCl containing 0.5% bovine serum albumin, and twice with 0.15 M-NaCl. Thermolysin digestion at pH 7.4 was done at 37°C for 30 min on a 20% suspension of cells equilibrated with iso-osmotic phosphate buffer, pH 7.4 (0.155 M-NaH₂PO₄ adjusted to pH 7.4 with 0.103 M-Na₂HPO₄), and at a thermolysin concentration of 0.5 mg/ml. Thermolysin digestion at pH 8.5 and the treatment of supernatants was as described above. The digested cells were washed three times with 0.15 M-NaCl containing 0.5% bovine serum albumin, and twice with 0.15 M-NaCl before preparation of 'ghosts'. The identity of corresponding peptides from the peptide 'maps' was confirmed by elution and co-electrophoresis at pH 1.7 and pH 6.5.

Trypsin digestion of labelled 'ghosts'

Labelled 'ghosts', derived from trypsin-treated erythrocytes, were digested in 5 mM-sodium phosphate buffer, pH 8, for 1 h at 37°C with 10 µg of trypsin (toluene-*p*-sulphonamidophenylalaninechloromethyl ketone-treated)/mg of 'ghost' protein, as described by Steck (1972a). The labelled fragments derived from protein E which remained in the membrane were separated by electrophoresis on 8% (w/v) polyacrylamide gels (Boxer *et al.*, 1974). Fragments with a molecular weight of less than 30 000 were concentrated by freeze-drying, and freed of sodium dodecyl sulphate by passage through a column (0.5 cm × 20 cm) of Sephadex G-25 (fine grade) equilibrated with 5 mM-(NH₄)₂CO₃-0.1 mM-EDTA. They were then digested with thermolysin, and the peptides separated on peptide 'maps' as described by Boxer *et al.* (1974).

Results and Discussion

Fragmentation of protein E at the extracellular surface

Two classes of proteolytic enzymes could be distinguished by their ability to cleave differentially protein E and the sialoglycoprotein in the intact erythrocyte. One class was represented by trypsin, and included papain and bromelain. Gel electrophoresis of 'ghosts' derived from trypsin-treated erythrocytes showed that the periodate-positive sialoglycoprotein band was lost (Steck *et al.*, 1971), whereas protein E remained unchanged (Plate 1, gels *a* and *b*).

A second class of proteolytic enzymes was

represented by thermolysin, and included Pronase, chymotrypsin and subtilisin. In these cases the carbohydrate-containing band corresponding to the sialoglycoprotein was again lost; however, protein E was converted into a fragment of molecular weight 60 000 (*e*) (Boxer *et al.*, 1974) and yielded a gel-electrophoresis pattern which was identical with that shown in Plate 1(*c*). The fragment 'e' produced by the different enzymes had the same mobility on polyacrylamide-gel electrophoresis. The sequential treatment of trypsin-digested cells with thermolysin yielded the same pattern of degradation of protein E as did direct thermolysin treatment (Plate 1*c*). This sequential treatment was designed to make it possible to obtain the fragments of protein E released from the membrane on thermolysin digestion with the minimum of contamination by digestion products from the sialoglycoprotein.

Intact erythrocytes were pre-labelled by iodination with lactoperoxidase, and then treated with trypsin. After subsequent digestion with thermolysin, which released 26% of the radioactivity from the cells, the fragment retained in the membrane (*e*) and the fragments released from the cell surface (derived from protein E) were isolated and the labelled peptides studied on peptide 'maps'. An outline of the experiment is shown in Fig. 1.

Five major mobile peptides were obtained from protein E isolated from the radioiodinated intact erythrocyte (Boxer *et al.*, 1974) as shown in Plate 2(*a*). Protein E isolated from trypsin-treated erythrocytes gave a peptide 'map' identical with that described above, showing that the labelled sites on this protein were indeed retained on trypsin treatment. The fragment 'e' remaining in the membrane after sequential digestion with trypsin and thermolysin of labelled erythrocytes also yielded the same peptide 'map' as unmodified protein E. The same peptide 'maps' (Plate 2*a*) were obtained from fragment 'e' derived from erythrocytes directly digested with thermolysin (Boxer *et al.*, 1974) or Pronase. Peptides 3 and 5, although always present in peptide maps of fragment 'e' preparations, irrespective of the particular proteolytic enzyme used, showed rather variable intensities.

The labelled peptides released from protein E into the supernatant from thermolysin digestion of trypsin-treated iodinated erythrocytes gave the peptide 'map' shown in Plate 2(*b*). Peptides 1, 3, 4 and 5 were again found, and two new peptides A and B were obtained. Redigestion of the supernatant with thermolysin (under the conditions used to prepare peptide 'maps') gave the same peptides.

The origin of peptides A and B is not clear. They are not likely to be derived from protein E, as they were not found on thermolysin digestion of either the intact protein E or protein E from trypsin-

EXPLANATION OF PLATE I

Sodium dodecyl sulphate-polyacrylamide-gel electrophoresis of radioiodinated samples

Radioautographs of dried slices of 1.1 cm gels are shown. SP, sialoglycoprotein. (a)–(c) Washed erythrocytes were labelled with ^{125}I by the direct addition of H_2O_2 in the presence of butylated hydroxytoluene, as described by Boxer *et al.* (1974). The cells were sequentially digested with trypsin and thermolysin as described in the Methods section. Bovine serum albumin was absent from the washing solution. Electrophoresis was done in a 7.5% acrylamide gel. (a) 'Ghosts' from iodinated untreated erythrocytes. (b) 'Ghosts' from trypsin-treated iodinated erythrocytes. (c) 'Ghosts' from iodinated erythrocytes that had been sequentially treated with trypsin and thermolysin. (d)–(g) Washed erythrocytes were labelled as described in (a)–(c) but with a mixture of 5 mCi of Na^{131}I and 0.5 mCi of Na^{125}I per 0.6 ml of packed erythrocytes. Sequential digestion with trypsin and thermolysin was done as described in the Methods section, with bovine serum albumin present in the washing solution. Electrophoresis was done in an 8% acrylamide gel. (d) 'Ghosts' from iodinated untreated erythrocytes. (e) 'Ghosts' from trypsin-treated iodinated erythrocytes. (f) 'Ghosts' from iodinated erythrocytes which had been treated with trypsin and then treated with thermolysin at pH 7.4. (g) 'Ghosts' from iodinated erythrocytes which had been treated with trypsin and then treated with thermolysin at pH 8.5. (h)–(i) Washed trypsin-treated erythrocytes, or 'ghosts' from trypsin-treated erythrocytes, were labelled with 2.5 mCi of Na^{131}I and 1 mCi of Na^{125}I per 0.2 ml of packed cells or 'ghosts', by using the glucose oxidase method as described (Boxer *et al.*, 1974). The labelled 'ghosts', and 'ghosts' prepared from the labelled trypsin-treated erythrocytes, were digested with trypsin as described in the Methods section. Electrophoresis was done in an 8% gel. (h) Trypsin digest of 'ghosts' from iodinated trypsin-treated erythrocytes. (i) Trypsin digest of iodinated 'ghosts' which were derived from trypsin-treated erythrocytes.

EXPLANATION OF PLATE 2

Radioautographs of peptide 'maps' of protein E and fragments derived from protein E from labelled erythrocytes

(a) Protein E from iodinated erythrocytes, prepared as in Plate 1(a), was digested with thermolysin and a peptide 'map' obtained as described in the Methods section. (b) Fragments released from trypsin-treated labelled erythrocytes by subsequent digestion with thermolysin at pH 8.5 were separated on a peptide 'map'. Preparation is as described in the Methods section.

EXPLANATION OF PLATE 3

Radioautographs of thermolysin peptide 'maps' of protein E and tryptic fragments of protein E

The intact protein and the tryptic fragments were prepared from labelled 'ghosts' that had been derived from trypsin-treated erythrocytes. (a) Intact protein E, prepared as described by Boxer *et al.* (1974). (b) Tryptic fragment TA, prepared as in Plate 1(i). (c) Tryptic fragment TG, prepared as in Plate 1(i).

Fig. 1. Outline of method for the isolation of labelled fragment 'e', and of labelled fragments released from protein E on digestion of labelled erythrocytes

Details of the procedure are given in the Methods section.

treated erythrocytes (Plate 2a). Jackson *et al.* (1973) showed that one extracellular tyrosine-containing glycopeptide of the sialoglycoprotein was not released from the erythrocyte by tryptic digestion. Peptides A and B probably resulted from subsequent thermolytic cleavage of this fragment. We were not able to detect this fragment in 'ghosts' from trypsin-treated erythrocytes. It may be that this fragment was particularly labile to traces of proteinase absorbed to the 'ghosts', as is the case with fragment e_2 , discussed below.

In the above experiments, proteinase-treated cells were washed with 0.15M-NaCl before the preparation of 'ghosts'. Cabantchik & Rothstein (1974b) reported that if proteinase-treated cells were washed with a solution containing bovine serum albumin before cell lysis, additional fragments derived from protein E were recovered in the 'ghosts'. To study these fragments, iodinated erythrocytes were digested with trypsin, washed with bovine serum albumin, and digested either under the conditions of Cabantchik & Rothstein (1974b) at pH 7.4, or under our usual conditions (pH 8.5), and again washed with bovine serum albumin. Polyacrylamide-gel patterns of the 'ghosts' from the cells obtained from these two experiments are shown in Plate 1 (gels *d-g*). In each case three additional fragments were found in the cells that had been digested with trypsin and thermolysin (e_1 , e_2 and e_3 , with apparent molecular weights of 58000, 35000

and 20000, Plate 1, gels *f* and *g*). The pair of sharp bands with a slightly lower mobility than fragment e_2 were not derived from protein E, as these were also found in the trypsin-treated cells (Plate 1, gel *e*). In contrast with fragment 'e', e_2 and e_3 yielded broad diffuse bands similar to protein E. Cabantchik & Rothstein (1974b) have also noted the diffuse appearance of fragment e_2 . Only small amounts of fragments e_1 , e_2 and e_3 were obtained compared with fragment 'e'. Digestion with thermolysin at pH 7.4 was less complete than at pH 8.5, as significant amounts of intact protein E remained in the former case. Thermolysin peptide 'maps' of fragments e_1 , e_2 and e_3 were the same as those of fragment 'e', but the peptides contained much less radioactivity. We conclude that fragments e_1 , e_2 and e_3 are minor digestion products, the preferred path of digestion giving rise to fragment 'e'.

As expected, the supernatant peptides from thermolysin digestion at pH 8.5 gave the same peptide 'map' whether the cells were washed with serum albumin or not (Plate 2b). Thermolysin digestion at pH 7.4 liberated the same peptides except that peptide 3 was not found. The radioactivity in the peptide-3 spot on peptide 'maps' from fragment 'e' produced at pH 7.4 and at pH 8.5 from the same labelled erythrocytes was also measured. The peptide-1 spot was used as an internal standard. On the basis of the same amount of peptide 1 in each case, peptide 3 from fragment 'e'

produced by thermolysin digestion at pH 7.4 contained 2.4 times more radioactivity than did peptide 3 from fragment 'e' produced by digestion at pH 8.5. Thus an extra site 3 occupies a terminal position in the fragment 'e' produced by thermolysin digestion of erythrocytes at pH 7.4. As the 'e' fragments obtained at the two pH values gave apparently identical molecular weights on sodium dodecyl sulphate-polyacrylamide-gel electrophoresis, one site 3 must be close to the point of cleavage that gives rise to fragment 'e' on digestion at pH 8.5. The comparative resistance to digestion of this site 3 accounts for the variability of intensity of peptide 3 in peptide 'maps' of fragment 'e' produced by different proteinases.

These data suggest that at least two sets of the labelled sites represented by peptides 1, 3, 4 and 5 are present in protein E.

The above experiments were carried out on cells labelled before proteinase treatment. Pretreatment of the cells with trypsin, thermolysin or Pronase, followed by radioiodination, resulted in a 4-5-fold increase in iodine incorporation compared with cells labelled under the same conditions and subsequently digested with proteinases. The increased degree of iodination is due to the enhanced accessibility of the sites to lactoperoxidase, and not to the appearance of new labelled sites, as the peptide 'maps' of protein E and fragment 'e' were identical whether labelling was carried out before or after digestion.

Bretscher (1971*a*) found that the peptide 'map' of protein E labelled at lysine residues with formyl-[³⁵S]methionyl sulphone methyl phosphate was the same as that of fragment 'e' obtained after Pronase treatment and subsequent labelling of erythrocytes. This observation is consistent with our suggestion that a duplicated set of sites exist in protein E, and implies that the duplication probably also extends to extracellular lysine residues in the protein.

Fragmentation of protein E at the intracellular surface

The arrangement of these sets of sites along the polypeptide sequence was further investigated by partial digestion of protein E at the intracellular surface of the erythrocyte 'ghost'. This protein has been shown to be susceptible to proteolysis by trypsin at the inner membrane surface but not from the outer surface (Steck, 1972*a*). To minimize interference by fragments from the sialoglycoprotein, 'ghosts' derived from trypsin-treated erythrocytes were used. Both the intracellular and extracellular sites of protein E were iodinated by labelling these unsealed erythrocyte 'ghosts' with lactoperoxidase (Boxer *et al.*, 1974) and the labelled 'ghosts' were then treated with low concentrations of trypsin. As trypsin does not cleave the extracellular region of protein E, all the cleavages will occur at intracellular sites. The fragments originating from protein E, and remaining in the membrane, were identified by comparison on polyacrylamide gels with the labelled fragments obtained by tryptic digestion, under the

Fig. 2. Outline of method for the isolation of labelled tryptic fragments of protein E

Details of the procedure are given in the Methods section.

Table 1. *Composition of tryptic fragments of protein E*

Protein E and tryptic fragments were prepared from labelled 'ghosts' as described in the legend to Plate 1(i). The labelled peptides were located from the radioautographs of the thermolysin peptide 'maps' of the tryptic fragments and the intact protein. The spots were cut out and counted for radioactivity in a scintillation counter in 4ml of a solution of 4g of 2,5-bis-(5-t-butylbenzoxazol-2-yl)thiophen scintillator/litre of toluene. To determine the content of labelled sites in each peptide, the radioactivity in peptide 1 was used as internal standard in each case. For each peptide 'map' the radioactivity in each peptide was calculated so that peptide 1 = 1.00. The radioactivity found in each labelled peptide on the peptide 'map' of the intact protein was taken to represent the number of each site present in the intact protein as shown in the Table and this information was used to calculate the content of 'sites' in the fragments. The calculation assumes that the two site-1 tyrosine residues in the intact protein carry an equal amount of label. A similar assumption is made for sites 3, 4 and 5. Although the data suggest that this approximation is not unreasonable, it is unlikely to be numerically accurate and may contribute to the apparently low recoveries of sites 2, 8, 9 and 10 in fragments TA and TB. Contaminating peptides (as they probably all contain sites 1, 4 and 5) also give an apparently decreased recovery of the sites other than sites 1, 4 and 5, as the data are calculated on the basis of 1.0 site 1 in each fragment preparation.

Site	Number of sites in				
	Intact protein	TA	TB	TF	TG
1	2	1.0	1.0	1.0	1.0
2	1	0.4	0.5	—	—
3	2	1.0	0.8	—	—
4	2	1.4	1.4	0.9	0.9
5	2	1.2	1.2	0.9	1.0
6	1	0.02	0.04	0.8	0.5
7	1	0.1	0.1	0.9	1.5
8	1	0.3	0.4	—	—
9	1	0.4	0.4	—	—
10	1	0.5	0.5	0.2	0.1
11	1	—	0.1	0.9	0.7
12	1	—	—	1.1	0.6
13	1	—	—	0.9	0.8
14	1	1.3	—	—	—
15	1	0.1	0.1	—	—

same conditions, of 'ghosts' derived from labelled trypsin-treated erythrocytes (in which protein E is essentially the only labelled protein; Plate 1, gel b). An outline of this experiment is shown in Fig. 2. The fragments were isolated by gel electrophoresis, and the patterns obtained are given in Plate 1 (gels h and i). The molecular weights of fragments TA, TB, TF and TG are 65000, 57000, 27500 and 24000 respectively. Thermolysin peptide 'maps' of fragments TA and TG are shown in Plates 3(b) and 3(c),

as is a thermolysin peptide 'map' of the unmodified protein from which the fragments were derived (Plate 3a). The radioactivity in individual peptides from the peptide 'maps' was also estimated, and the data for fragments TA, TB, TF and TG are shown in Table 1. The remaining fragments were impure. Two distinct sets of tryptic fragments were produced; fragments TA and TB contain sites 1, 2, 3, 4, 5, 8, 9 and 10, whereas fragments TF and TG lack sites 2 and 3, and contain sites 1, 4, 5, 6, 7, 11, 12 and 13. Fragments TA and TB differ by the presence of site 14 in TA. Together fragments TA and TF contain all the major labelled sites found in the intact protein except for site 15, and their combined molecular weight is the same as that of the intact protein. We did not recover site 15 in any of the major fragments. The small amount of peptides 7, 11 and 15 found in fragments TA and TB were probably derived from a contaminating peptide. The fragmentation pattern produced in different experiments was somewhat variable, probably because the extent of digestion is critically dependent on the trypsin and membrane concentrations. However, in three different experiments we were able to isolate products that were similar to fragments TA and TF.

From these data we conclude that the two sets of extracellular sites each containing sites 1, 4 and 5 are not arranged adjacent to each other in the polypeptide sequence. One set of sites 1, 4 and 5 are joined to sites 2, 3, 8, 9, 10 and 14, whereas the other set of sites 1, 4 and 5 are linked to sites 6, 7, 11, 12 and 13. As sites 1, 2, 3, 4 and 5 are located on the outside surface of the cell, and the remainder are intracellular (Boxer *et al.*, 1974), our results demonstrate that at least two lengths of the protein E sequence traverse the membrane. Shown in Fig. 3 is a model for the structure of protein E which is consistent with these findings. Also shown is the order of some of the labelled peptides on the protein sequence deduced from the composition of the fragments obtained. The diffuse appearance on gel electrophoresis of bands e_2 and e_3 suggests that they are not subfragments of fragment 'e'. They are found associated with the 'ghosts' even after extensive washing. This suggests that the fragments are either themselves bound to the membrane or remain tenaciously bound to fragment 'e' after proteolytic cleavage. As gel electrophoresis of fragment 'e' gives a sharply defined band, in contrast with the diffuse band given by the intact protein, the segment X-Y in Fig. 3 probably gives rise to fragments e_2 and e_3 and contains the region causing the diffuse banding characteristics of the intact protein E. As fragment e_2 is reported to have a high content of carbohydrate (Cabantchik & Rothstein, 1974b), some of the carbohydrate present in protein E (Tanner & Boxer, 1972) is located in this region between X and Y. If fragment e_2 is

Fig. 3. *S-shaped model for the structure of protein E in the erythrocyte membrane*

Details of the model are discussed in the text. Parentheses round the groups of labelled sites implies that the order of the sites within each group is not established. Although the order of the sites shown in the model is as established, their position along the sequence is not implied to be to scale.

indeed itself bound to the membrane, the arm containing the region X-Y may be reinserted into the membrane (shown as a broken line in the model, Fig. 3).

We were not able to obtain a reasonable recovery of peptide 15 in any of the tryptic fragments. Distinct but equally basic peptides, possibly derived from peptide 15, are found in the 'maps' of fragments TF and TG (marked C in Plate 3c). As the loss of peptide 15 results from the action of trypsin, it is probable that a basic residue in peptide 15 is a primary site of intracellular cleavage of protein E by trypsin.

Two possible schemes which account for the generation of fragment 'e' by thermolysin cleavage of protein E are shown in Figs. 4(a) and 4(b). If the region X-Y re-enters the membrane, the major route of cleavage in both schemes leaves the region Y embedded in the membrane. Because this segment contains no labelled sites, we would not be able to detect it on our radioautographs. Less-favoured cleavages in the region X gave rise to fragments e_2 and e_3 . In the above analysis we have assumed that the thermolysin peptides obtained on the peptide 'maps' ('sites') each represent distinct tyrosine-containing peptide sequences that are not related to each other (i.e. not being products of partial digestion). The high enzyme protein ratios and long digestion times (2:5, w/w; 18h; Boxer *et al.*, 1974) used in this work make it unlikely that digestion is incomplete. However, even if this is not the case our

Fig. 4. *Possible routes of cleavage by thermolysin at the extracellular region of protein E*

Details of the model are discussed in the text. The zig-zag regions represent tightly folded and rigid regions of the polypeptide chain. The solid arrows represent the major cleavage points by thermolysin. The broken arrows at X show the minor cleavage which gives rise to fragments e_2 and e_3 .

interpretation of the order of the sites and traverses of the polypeptide chain remain valid. The variability in the intensities of peptides 3 and 5 in the products of digestion of the intact cell (which use much milder conditions of enzyme digestion) suggest that peptide 5 may be related to peptide 3 (or, more likely, a peptide that co-migrates with peptide 3).

The proteinase-resistant regions containing the extracellular labelled tyrosine residues of fragment 'e' are probably protected from cleavage by the structure of the protein. Presumably although tyrosine residues in these regions of the protein are accessible to lactoperoxidase, the polypeptide backbone cannot be reached by the proteinases. The possible location of these regions is shown in Fig. 4. As fragment 'e' is produced by a range of different enzymes, it is probable that the sites where cleavage does occur are at particularly exposed and flexible regions of the polypeptide chain.

The intracellular portion of protein E is depicted in Fig. 3 as a continuous loop of polypeptide chain. However, we cannot exclude the possibility that the protein consists of two related polypeptide chains linked intracellularly by linkages other than disulphide bonds. Recently γ -glutamyl-lysyl linkages ('isopeptide linkages'; Birckbichler *et al.*, 1973) have been reported in membrane proteins.

The assumption that protein E is a homogeneous polypeptide is central to our interpretation of these results. Its behaviour on proteolytic cleavage (Bender *et al.*, 1971; Bretscher, 1971a; Steck, 1972a), and oxidation to form a dimeric species (Steck, 1972b) form strong evidence that this is the case. There is evidence that suggests some heterogeneity in the carbohydrate carried by this polypeptide (Steck & Dawson, 1974; Findlay, 1974), and this may account for the anomalous behaviour of protein E on proteolytic cleavage observed by Cabantchik & Rothstein (1974b). The available evidence suggests that it is correct to assume that protein E contains a homogeneous polypeptide chain. However, we cannot eliminate the possibility that band E contains a family of closely related proteins, in which case other models for the structure of the protein(s) would be possible.

The erythrocyte sialoglycoprotein penetrates the membrane only once. Protein E, which contains 4-5 times as many amino acid residues, traverses it at least twice. Models in which the polypeptide chain of protein E penetrates the membrane more often (for example with the intracellular loop reinserted into the membrane) would also be consistent with our data, as we have only monitored the distribution of tyrosine residues in the molecule.

Despite the fact that protein E is easily the most

abundant polypeptide in the membrane (1.2×10^6 copies per 'ghost'; Fairbanks *et al.*, 1971) its function has not been unequivocally established. Suggestions for the function of the protein include an involvement in erythrocyte anion transport (Bretscher, 1973; Cabantchik & Rothstein, 1974a), a cell-surface sugar carrier (Tanner & Boxer, 1972) and as a binding site for the intracellular glycolytic enzyme glyceraldehyde 3-phosphate dehydrogenase (Kant & Steck, 1973). The relatively complex structure which the protein possesses is compatible with all of these suggestions.

We acknowledge the skilful assistance of Mrs. Judith Jefferies and Mr. Alan Reading at stages in this work. R. E. J. holds a Science Research Council Studentship. The work was supported by the Medical Research Council.

References

- Bender, W. W., Garan, W. & Berg, H. C. (1971) *J. Mol. Biol.* **58**, 783-797
- Birckbichler, P. J., Dowben, R. M., Maticic, S. & Loewy, A. G. (1973) *Biochim. Biophys. Acta* **291**, 149-155
- Boxer, D. H., Jenkins, R. E. & Tanner, M. J. A. (1974) *Biochem. J.* **137**, 531-534
- Bretscher, M. S. (1971a) *J. Mol. Biol.* **59**, 351-357
- Bretscher, M. S. (1971b) *Nature (London) New Biol.* **231**, 229-232
- Bretscher, M. S. (1973) *Science* **181**, 622-629
- Cabantchik, Z. I. & Rothstein, A. (1974a) *J. Membrane Biol.* **15**, 207-226
- Cabantchik, Z. I. & Rothstein, A. (1974b) *J. Membrane Biol.* **15**, 227-248
- Fairbanks, G., Steck, T. L. & Wallach, D. F. M. (1971) *Biochemistry* **10**, 2606-2617
- Findlay, J. B. C. (1974) *J. Biol. Chem.* **249**, 4398-4403
- Jackson, R. L., Segrest, J. P., Kahane, I. & Marchesi, V. T. (1973) *Biochemistry* **12**, 3131-3138
- Kant, J. A. & Steck, T. L. (1973) *J. Biol. Chem.* **248**, 8457-8464
- Segrest, J. P., Jackson, R. L., Marchesi, V. T., Guyer, R. B. & Terry, W. (1972) *Biochem. Biophys. Res. Commun.* **49**, 964-969
- Segrest, J. P., Kahane, I., Jackson, R. L. & Marchesi, V. T. (1973) *Arch. Biochem. Biophys.* **155**, 167-183
- Steck, T. L. (1972a) in *Membrane Research* (Fox, C. F., ed.), pp. 71-93, Academic Press, New York
- Steck, T. L. (1972b) *J. Mol. Biol.* **66**, 295-305
- Steck, T. L. & Dawson, G. (1974) *J. Biol. Chem.* **249**, 2135-2142
- Steck, T. L., Fairbanks, G. & Wallach, D. F. H. (1971) *Biochemistry* **10**, 2617-2624
- Tanner, M. J. A. & Boxer, D. H. (1972) *Biochem. J.* **129**, 333-347
- Thomas, D. B. & Winzler, R. J. (1969) *J. Biol. Chem.* **244**, 5943-5946