The human erythrocyte anion-transport protein

Further amino acid sequence from the integral membrane domain homologous with the murine protein

Christopher J. BROCK* and Michael J. A. TANNER

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

Sequences from the human erythrocyte anion-transport protein homologous with residues 417–449 and 794–813 of the murine erythrocyte anion-transport protein have been determined. The former sequence includes the putative transmembrane helix closest to the *N*-terminus of the protein. The latter sequence traverses almost all of the lipid bilayer and is located towards the *C*-terminus of the protein. Sites have been identified by alignment with the murine sequence in the integral membrane domain that are accessible to proteolytic enzymes. Sequences from the integral membrane domain of the erythrocyte anion-transport protein are highly conserved.

INTRODUCTION

The anion-transport protein, Band 3, is the most abundant integral membrane glycoprotein of the human erythrocyte [for reviews, see Knauf (1979) and Wieth *et al.* (1982)]. The polypeptide chain comprises two distinct functional domains. The *N*-terminal portion of about 40 kDa is located intracellularly and binds the erythrocyte cytoskeleton, many glycolytic enzymes and haemoglobin (reviewed by Gillies, 1982). The *C*-terminal part of the protein is mainly buried in the lipid bilayer and mediates the strict one-for-one exchange of anions across the membrane (Grinstein *et al.*, 1978) by a sequential kinetic mechanism (Gunn & Frohlich, 1978, 1979).

In previous work (Brock *et al.*, 1983) the strategy we adopted to obtain amino-acid-sequence information from the protein was to cleave the protein into major peptide fragments while it was still in the lipid bilayer with other erythrocyte membrane proteins. Peptides were assigned as originating from the anion-transport protein if they contained, or at least were continuous with, known labelled sites in the anion-transport protein. Recently the complete sequence of 929 amino acids of the murine erythrocyte anion-transport protein was deduced from the nucleotide sequence of a cDNA clone (Kopito & Lodish, 1985). There is extensive homology between this sequence and those of fragments of the human protein which have been determined previously (Mawby & Findlay, 1982; Brock *et al.*, 1983). It is now possible to assign peptides which originate from the human protein by their homology with the murine sequence.

We present here evidence for further amino acid sequences from the human erythrocyte protein which are not only homologous with the murine protein, but could also form transmembrane α -helices. Sequences close to sites of proteolytic cleavage in the integral membrane domain are also described which may be useful in defining the topology of the polypeptide chain.

EXPERIMENTAL AND RESULTS

Materials

All chemicals were analytical or sequencer-grade reagents (Brock *et al.*, 1983). All solutions were prepared with deionized glass-distilled water.

Peptide nomenclature

GTh1, produced by thermolytic cleavage, is similar to, but not identical with, a chymotryptic peptide (Chb) studied by Mawby & Findlay (1982). The covalent binding site for the disulphonic stilbenes and some other anion-transport inhibitors is located near the *C*-termini of these peptides (Mawby & Findlay, 1982; Kopito & Lodish, 1985).

Peptides P(416), PCB(795) and Ch(672) are named according to whether they were produced by cleavage with pepsin (P), pepsin followed by CNBr (PCB) or chymotrypsin (Ch). The number which follows refers to the residue number in the murine sequence which aligns with the *N*-terminus of the peptide.

Peptide purification and sequencing

Haemoglobin-free erythrocyte ghosts were prepared by the method of Dodge *et al.* (1963). Digestion of erythrocyte ghosts with pepsin and thermolysin and purification of the membrane-integrated fragment GTh1 was as described by Kempf *et al.* (1981). The peptide P(416) is the main component of the shoulder on the trailing edge of the peak comprising peptide P5 when fractionated on Sephadex G-50 (Kempf *et al.*, 1981). It was obtained in pure form when the material from this region was rechromatographed on a column (20 mm $\times$ 1500 mm) of Sephacryl S-200 (superfine grade) equilibrated with 1% SDS/1 mM-sodium EDTA, pH 8.0.

Peptide P(416) was DITC-coupled and subjected to solid-phase Edman degradation (Laursen, 1971; Wachter *et al.*, 1975; Brock *et al.*, 1983). Amino acid PTH

Abbreviations used: PTH, phenylthiohydantoin; DITC, *p*-phenylene di-isothiocyanate.

* Present address: A.F.R.C. Food Research Institute – Bristol, Langford, Bristol BS18 7DY, U.K.

Fig. 1. Amino acid sequences of peptides P(416) and PCB(795)

Sequences were assigned to the human erythrocyte anion-transport protein by homology with the murine protein. Identical residues are boxed. Residues in the human sequence shown in **bold** typeface are based on quantitative evidence from h.p.l.c. analysis of amino acid PTH derivatives released after successive cycles of solid-phase Edman degradation. Yield was defined as the increase measured in the amino acid PTH derivative release over that in the previous cycle. All amino acid PTH derivatives decreased to background level in the cycle subsequent to that in which they were identified. (a) P(416): 4 nmol of peptide was successfully coupled and sequenced with a repetitive yield of 88% (Ala-4 → Ala-19). Alanine was identified in cycle 12 because it only decreased by 0.3 nmol compared with cycle 11. It returned to background level in cycle 13. Serine in cycles 6 and 16 was identified qualitatively as PTH-serine (a shoulder on PTH-glutamine) in poor yield and a peak just after PTH-glycine, which in our experience is characteristic of PTH-serine in a sequence. No other amino acid PTH derivatives were released after cycles 6, 12 or 16 in sufficient yield to justify assignment. (b) PCB(795): 4 nmol of peptide was successfully coupled and sequenced with a repetitive yield of 96% (Leu-4 → Leu-12). Serine was identified qualitatively in cycle 5 as the characteristic peak eluted after PTH-glycine. PTH-serine was co-eluted with PTH-glutamine in this experiment and so was not clearly detectable. No other amino acid PTH derivative was released in sufficient yield to justify its assignment to this position. PCB(745) was contaminated with P5CB2 and P5CB6 (Brock *et al.*, 1983) to the extent of about 10% and 30% of the level of the major sequence respectively. The effect of this was to distort the yields of some amino acid PTH derivatives. Consequently the yields after cycles 2, 7 and 19 were high, owing to a contribution from the contaminating peptides, and those after cycles 8, 9 and 17 were low, owing to the effect of the contaminating peptides on the preceding cycle in each case.

derivatives were identified by h.p.l.c. as described by Gates *et al.* (1979) and Brock *et al.* (1983). Quantitative evidence for residues 2–19 was obtained (except for residues 6, 12 and 16; see the legend to Fig. 1). This sequence is homologous (12/18) with residues 416–434 of the murine erythrocyte anion-transport protein (Fig. 1). When the sequence data from subsequent degradation cycles were re-examined in the light of this homology, qualitative assignments of the sequence from residues 20–34 (Ala-Leu-Ser-Pro-Ala-Ile-Thr-Phe-Gly-Phe-Leu-Leu-Gly-Glu-Lys) were obtained. This further sequence appeared to be highly homologous (13/15) with the corresponding residues from the murine protein. However, these assignments should be considered less reliable than those for residues 2–19.

Peptide PCB(795) was derived from a peptide contaminating a preparation of P5 that was cleaved with CNBr and fractionated as described by Brock *et al.* (1983). On gel filtration of the digest it was eluted from Bio-Gel P-10 in 80% (w/v) formic acid with P5CB2 and P5CB6 (Brock *et al.*, 1983). It was obtained in its most pure form from the trailing end of this peak. The peptide was homoserine lactone-coupled (Wachter *et al.*, 1975; Brock *et al.*, 1983) and sequenced unambiguously to completion. The contaminating peptides produced only minor distortions in the quantitative data (see the legend

to Fig. 1). Furthermore, its primary structure proved to be identical with residues 795–813 of the murine protein. Methionine is very likely to be the immediately preceding residue because of the specificity of the CNBr cleavage reaction.

Peptide Ch(672) was produced by exhaustive chymotryptic cleavage of erythrocyte ghosts. After stripping them with NaOH, the ghosts were redigested with chymotrypsin (1.5 mg/ml) in 5 mM-sodium phosphate, pH 8.0, at 37 °C for 10 min. The peptide was purified by taking the first peak eluted after gel filtration on Sephadex G-75 (Mawby & Findlay, 1982), which contained mainly the integral membrane glycopeptide, and rechromatographing it on a column (25 mm × 1000 mm) of Sephacryl S-300 (superfine grade) equilibrated with 1% SDS/1 mM-sodium EDTA, pH 8.0. Peptide Ch(672) was isolated from the trailing (low-*M_r*) end of the major peak eluted.

The peptides Ch(672) and GTh1 were DITC-coupled and sequenced. The amino acid PTH derivatives released from cycle 2 onwards in each case were identified qualitatively, but unambiguously, by t.l.c. (Cohen-Solal & Barnard, 1973) and back-hydrolysis followed by amino acid analysis (Mendez & Lai, 1975). The following sequences were identified: Ch(672), Xaa-Leu-Tyr-Pro-Leu-Arg-Pro; GTh1, Xaa-Asp-Leu.

DISCUSSION

In the discussion that follows, where a site from the human protein is identified by a number, that number refers to the residue in the murine protein with which that site is aligned by local homology (Fig. 1).

A number of proteolytic enzymes cleave the human erythrocyte anion-transport protein into its intracellular and integral membrane domains at closely related sites. Mawby & Findlay (1982) produced good quantitative evidence for the sequence at the chymotryptic cleavage site (372–373). The short *N*-terminal sequences presented here for GTh1 identify the thermolysin cleavage site as being two residues towards the *C*-terminus (374–375). In our hands this peptide became resistant to further cycles of Edman degradation after release of Leu-377, which is entirely consistent with the presence of Asn-378–Gly-379 in the sequence, as suggested by Mawby & Findlay (1982), and as is found in the murine protein.

This intracellular region sensitive to proteolysis is some distance from the first hydrophobic region in the murine sequence predicted to form a transmembrane α -helix. When erythrocyte ghosts were digested with pepsin in 1 M-acetic acid (which is likely to denature parts of the polypeptide chain not buried in the lipid bilayer), peptide P(416) was produced, which results from cleavage at a site very close to the membrane. It is noteworthy that, of the eight differences between the *N*-terminal sequence of P(416) and the murine protein, five are contiguous and their compositions in this region are almost identical. It has already been observed that the membrane-traversing segments of integral membrane proteins are highly conserved (Findlay *et al.*, 1981; Kopito & Lodish, 1985). Perhaps where changes do take place there is a tendency for complementary changes to occur nearby. Helical-wheel plots of these sequences show that homology extends to the secondary structural features with both forming amphipathic helices with no charged residues, but with polar side chains distributed radially over 120–140° and axially over the central one-third of the traverse.

Sequence information relating to the *C*-terminal half of the integral membrane domain has been much harder to obtain, because major peptides from this region carry the heterogeneous oligosaccharide moiety. The sequence of PCB(795) shows residues 794–813 to be completely conserved between the murine and human proteins. From residue 801 onwards this represents almost all of a very hydrophobic transmembrane α -helix predicted from hydropathy values (Kopito & Lodish, 1985).

The short *N*-terminal sequence from Ch(672) produced

by exhaustive chymotryptic digestion of erythrocyte ghosts locates another site accessible to a proteolytic enzyme and therefore at the membrane surface. Once again a sequence change in one position (Arg-675 → Pro) is located near a complementary change (Phe-677 → Arg).

With the data presented here the sequences of two more transmembrane segments of the human erythrocyte anion-transport protein have been determined. These, together with the information relating to the sites accessible to proteolysis, will contribute to the determination of the complete primary structure and its disposition in the membrane.

We are indebted to Miss J. Chapman-Scott, Mrs. K. Moreton and Mr. S. Peake for technical assistance. This work was supported by a grant from the Medical Research Council.

REFERENCES

- Brock, C. J., Tanner, M. J. A. & Kempf, C. (1983) *Biochem. J.* **213**, 577–586
- Cohen-Solal, M. A. & Barnard, J. L. (1973) *J. Chromatogr.* **80**, 140–143
- Dodge, J. T., Mitchell, C. & Hanahan, D. J. (1963) *Arch. Biochem. Biophys.* **100**, 119–130
- Findlay, J. B. C., Brett, M. & Pappin, D. J. C. (1981) *Nature (London)* **293**, 314–316
- Gates, F. T., Coligan, J. C. & Kindt, T. J. (1979) *Biochemistry* **18**, 2267–2272
- Gillies, R. J. (1982) *Trends. Biochem. Sci.* **7**, 41–42
- Grinstein, S., Ship, S. & Rothstein, A. (1978) *Biochim. Biophys. Acta.* **507**, 294–304
- Gunn, R. B. & Frohlich, O. (1978) *J. Supramol. Struct. Suppl.* **2**, 213
- Gunn, R. B. & Frohlich, O. (1979) *J. Gen. Physiol.* **74**, 351–374
- Kaul, R. K., Murthy, P. S. N., Reddy, A. G., Steck, T. L. & Kohler, H. (1983) *J. Biol. Chem.* **258**, 7981–7990
- Kempf, C., Brock, C., Sigrist, H., Tanner, M. J. A. & Zahler, P. (1981) *Biochim. Biophys. Acta.* **641**, 88–98
- Knauf, P. A. (1979) *Curr. Top. Membr. Transp.* **12**, 249–363
- Kopito, R. R. & Lodish, H. F. (1985) *Nature (London)* **316**, 234–238
- Laursen, R. A. (1971) *Eur. J. Biochem.* **20**, 89–95
- Mawby, W. J. & Findlay, J. B. C. (1982) *Biochem. J.* **205**, 465–475
- Mendez, E. & Lai, C. Y. (1975) *Anal. Biochem.* **68**, 47–53
- Wachter, E., Hofner, H. & Machleidt, W. (1975) in *Solid Phase Methods in Protein Sequence Analysis* (Laursen, R. A., ed.), pp. 31–46, Pierce Chemical Co., Rockford, IL
- Wieth, J. O., Anderson, O. S., Brahm, J., Bjerrum, P. J. & Borders, C. L. (1982) *Philos. Trans. R. Soc. London Ser. B* **299**, 383–399

Received 8 January 1986/21 February 1986; accepted 5 March 1986