Isolation and characterization of CD47 glycoprotein: a multispanning membrane protein which is the same as integrin-associated protein (IAP) and the ovarian tumour marker OA3

William J. MAWBY,*§ Christopher H. HOLMES,† David J. ANSTEE,‡ Frances A. SPRING‡ and Michael J. A. TANNER*
Departments of *Biochemistry, and †Obstetrics, University of Bristol, Bristol BS8 1TD, U.K. and ‡International Blood Group Laboratory, Southmead Road, Bristol, U.K.

The CD47 glycoprotein was isolated from human erythrocytes by immunoprecipitation using monoclonal antibody (mAb) BRIC-125. Enzymic deglycosylation of the protein showed it contained N-linked oligosaccharides, and trypsin proteolysis of the protein *in situ* in the erythrocyte membrane cleaved it into two portions, one of which was glycosylated. Both the intact protein and the glycosylated fragment had blocked N-termini. Amino acid sequence was obtained from several proteolytic fragments of CD47. Comparison with the sequence database showed the protein to be very similar to or identical with OA3, a multispanning membrane protein. The protein also appears to be the same as the integrin-associated protein, which has a role in cell adhesion in non-erythroid cells. CD47 has six potential N-glycosylation sites, five of which are in an Ig superfamily domain. We show that three of these sites carry N-glycans in erythrocytes. Immunocytochemical staining of human tissues showed that CD47 was broadly distributed on mesenchyme and epithelia at multiple sites. Reactivity was particularly prominent in surface and ductular epithelia, and in the brain. The possible roles of the CD47 glycoprotein are discussed.

INTRODUCTION

CD47 was originally defined as a component on erythrocytes [1] but is now known to be expressed on virtually all haematopoeitic cells, and in other tissues [2–4]. Monoclonal antibodies (mAbs) to CD47 immunoprecipitate an erythrocyte glycoprotein of molecular mass 47–52 kDa [4], but little is known about the structure or biological function of the CD47 protein. Interestingly, however, a panel of anti-CD47 mAbs show reduced reactivity with erythrocytes which totally lack the Rh antigen (Rh_{null} cells), suggesting that CD47 may be associated with the Rh complex [4].

Here we report the purification and characterization of the CD47 glycoprotein isolated from human erythrocytes and the determination of the amino acid sequence of derived peptide fragments. Comparison with known protein sequences reveals that CD47 displays a striking similarity to the ovarian tumour marker (OA3) identified by Campbell et al. [5] and to the integrin-associated protein (IAP) reported by Brown et al. [6]. Our studies further reveal that CD47 is very broadly distributed on normal adult tissues, as well as ovarian tumours, being especially abundant in some epithelia and the brain.

After this manuscript was prepared the complete cDNA sequence of IAP was reported [6a], confirming our conclusions about the identity between OA3 and IAP.

MATERIALS AND METHODS

Human erythrocytes were obtained from the South West Regional Blood Transfusion Centre, Bristol. NBTS/BRIC-125 (BRIC-125) and NBTS/BRIC-126 (BRIC-126) mAbs were produced as described in Avent et al. [4]; NBTS/BRIC-154 (BRIC-154) was produced as described in Wainwright et al. [7]; all were obtained as culture supernatants.

Purification of CD47 glycoprotein

Human erythrocytes were washed three times in 10 mM Na_2HPO_4, 140 mM NaCl, pH 8.0 (PBS8) and the buffy coat removed. The washed cells were then treated with BRIC-125 mAb for 16 h at 4 °C.

The agglutinated cells were washed gently with PBS8 and lysed in 5 mM Na_2HPO_4, pH 8.0 (5P8) at 4 °C. After several further washes at 4 °C the erythrocyte ghosts were solubilized with 10% (w/v) Triton X-100 in PBS8/2.5 mM EDTA/1 mM phenylmethanesulphonyl fluoride (PMSF). The insoluble material was removed by centrifugation at 48000 *g* for 1 h. The supernatant fraction was then passed through a 1.5 cm × 1.0 cm diam. Protein A–Sepharose (Bioprocessing Ltd, Consett, U.K.) or Protein A–agarose (Bio-Rad, Hemel Hempstead, U.K.) column. The eluate was recirculated for 16 h at 4 °C and the column was then washed with 5% (w/v) Triton X-100 in PBS8, followed by water. CD47 protein was eluted from the Protein A–Sepharose by incubating with 50 mM Tris, pH 6.8, 10% (w/v) SDS, 10 mM EDTA for 30 min at 37 °C.

For purification of CD47 material by column chromatography, material eluted from Protein A–Sepharose was centrifuged through a 100 kDa cut-off Ultrafree-MC membrane filter (Millipore, Watford, U.K.), then concentrated 4-fold in a 30 kDa cut-off Ultrafree-MC. This material was loaded on to a 45 cm × 0.6 cm diam. Sepharose G-150 column in 50 mM Na_2HPO_4, pH 8.0, 0.05% (w/v) SDS. The composition of the fractions was monitored by SDS/PAGE and fractions containing CD47 protein were pooled for further processing.

For purification of CD47 by SDS/PAGE, glycerol [(6%, w/v) final concentration] was added to the Protein A–agarose-eluted material and the samples were loaded on to 10% polyacrylamide gels using the Promega ChromaPhor buffer system (Promega, Southampton, U.K.). ChromaPhor dye was added to the top

Abbreviations used: mAb, monoclonal antibody; PBS8, 10 mM Na_2HPO_4, 140 mM NaCl, pH 8.0; 5P8, 5 mM Na_2HPO_4, pH 8.0; Lys-C, endoproteinase Lys-C; PMSF, phenylmethanesulphonyl fluoride; TLCK, tosyl-lysylchloromethane; TPCK, tosylphenylalanylchloromethane; PNGase-F, peptide-*N*-glycosidase F.

§ To whom correspondence should be addressed.

reservoir and the gels were run until the dye front was half way down the separating gel. The green bands visualized were excised and placed into 1.5 ml Eppendorf tubes. The gel pieces were crushed using a Teflon pestle in the presence of 50 mM NH_4HCO_3, 0.1% (w/v) SDS and processed according to the manufacturer's instructions (Promega, Southampton, U.K.). The material eluted from the gel was filtered through a 0.22 μm pore size Ultrafree-MC membrane filter and concentrated using 10 kDa or 30 kDa cut-off Ultrafree-MC membrane filters for further characterization.

Deglycosylation and lysyl endoproteinase (Lys-C) treatment of CD47 protein

Peptide-*N*-glycosidase F (PNGase F; Oxford Glycosystems, U.K.) (40 units) was added to column eluates containing CD47 or the gel-purified protein in 50 mM EDTA, 1% (v/v) 2-mercaptoethanol and 0.5% (w/v) β-octyl glucoside and the reaction allowed to proceed for 16 h at 37 °C. If the sample was to be further digested with proteases the deglycosylated protein was then concentrated with a 30 kDa cut-off Ultrafree-MC membrane filter to less than 50 μl. An aliquot (0.5 μl) of 100 μg/ml lysyl endoproteinase (Lys-C, Wako Chemicals GMBH, Germany) was added and the reaction allowed to proceed at 37 °C for 24–72 h.

Preparation of CD47 protein from protease-treated erythrocyte membranes

Human erythrocytes were washed as above and membranes were prepared without reacting with mAb. The washed membranes were incubated with tosylphenylalanylchloromethane (TPCK)-treated trypsin or tosyl-lysylchloromethane (TLCK)-treated chymotrypsin at a final concentration of 1 mg/ml for 1 h at 37 °C. The membranes were washed once with 5P8 and a cocktail of protease inhibitors was added. For trypsin, this cocktail was 1 mM PMSF, 0.5 mM TLCK and 0.3 mg/ml soybean trypsin inhibitor. For chymotrypsin, it was 1 mM PMSF, 0.5 mM TPCK and 20 mg/ml chymostatin. The membranes were washed twice more with 5P8. The centrifuged membranes were then incubated with the BRIC-125 mAb in the presence of 0.5 mM PMSF for 16 h at 4 °C with shaking. The membranes were centrifuged at 27500 *g* for 30 min, washed twice with 5P8, solubilized with 10% (w/v) Triton X-100 and purified as for the intact protein.

Tissue distribution of CD47 antigen

Adult human tissues were obtained after *post mortem* or from surgical specimens and snap frozen in liquid N_2-cooled isopentane (2-methylbutane). mAb were tested on 5 μm cryostat sections by an indirect immunoperoxidase staining technique as previously described [8]. Briefly, sections were incubated with mAb, washed, incubated with peroxidase-conjugated rabbit anti-(mouse IgG) antibody (Dako Ltd., Copenhagen, Denmark), and developed with 0.6 mg/ml diaminobenzidine tetrahydrochloride, 0.01% (w/v) H_2O_2. The slides were washed, counterstained with haematoxylin, dehydrated and mounted. Negative controls incubated in the absence of primary antibody or in the presence of an irrelevant mouse mAb were routinely included in these tests and were invariably negative.

Other methods

SDS/PAGE was performed using the Tricine buffer system [9]. Gels were stained with 0.1% (w/v) Coomassie Blue R in 40% (v/v) methanol, 5% (v/v) acetic acid, or were silver-stained using the method of Nielson and Brown [10] with minor modifications. Thioglycollic acid was added to the top gel buffer before electrophoresis to remove the background yellow coloration observed on silver staining. Samples for immunoblotting were transferred after electrophoresis to Immobilon-P (Millipore, Watford, U.K.) with an LKB semi-dry blotter at 0.8 mA/cm² for 1 h, using 39 mM glycine, 48 mM Tris, 0.04% (w/v) SDS, 20% (v/v) methanol as the transfer buffer. The membranes were blocked with 5% (w/v) skimmed milk powder, 0.2% (v/v) Tween-20 in PBS8, and immunoblotted as described [4].

Samples for sequencing were blotted on to ProBlott membrane (Applied Biosystems, Warrington, U.K.) at 1.6 mA/cm² for 1 h using 75 mM Tris/HCl, pH 8.5, 0.04% (w/v) SDS, 20% (v/v) methanol. The blotting membranes were stained with 0.1% (w/v) ServaBlue G in 40% (v/v) methanol/1% (v/v) acetic acid for 1 min, and destained in 50% (v/v) methanol. The stained bands were excised and sequenced on an Applied Biosystems 477A sequencer using the Blott cartridge and modified FastBlott cycles. Peptide sequences were compared with the Swissprot protein sequence and Genbank-EMBL DNA sequence databases available using the TFASTA program [11] on the SERC Seqnet facility at Daresbury, U.K.

RESULTS

Purification of CD47 glycoprotein

Immunoprecipitation of Triton X-100-solubilized membranes using mAb BRIC-125 and Protein A–Sepharose yields CD47 as a major diffuse component, migrating at 42–61 kDa on SDS/PAGE under non-reducing conditions (Figure 1b). This is similar in molecular mass to the protein previously immunoprecipitated from ^{125}I-labelled cells [4]. Protein-sequence studies confirmed that the minor 26 kDa component, which was also observed on SDS/PAGE, represented traces of IgG light chain, but no N-terminal sequence was obtained from the 42–61 kDa CD47 band, suggesting that CD47 has a blocked N-terminus.

The 42–61 kDa CD47 band was further purified by Sephadex G-150 column chromatography (Figure 1d). Deglycosylation of this material with PNGase-F generated components of

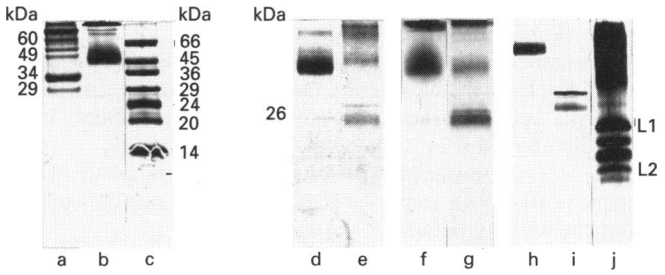

Figure 1 Purification of CD47 glycoprotein: deglycosylation and endoproteinase LysC digestion

Samples after SDS/PAGE. Lane a, Sigma SDS-7B prestained molecular-mass markers; lane b, CD47 protein material eluted from Protein A–Sepharose; lane c, Sigma SDS-7 molecular-mass markers; lane d, CD47 protein purified by Sepharose G-150 column chromatography; lane e, purified intact CD47 protein treated with PNGase-F prior to electrophoresis; lane f, immunoblot of deglycosylated purified CD47 protein with BRIC-125 antibody; lane g, immunoblot of CD47 protein material eluted from Protein A–Sepharose with BRIC-125 antibody; lane h, gel-purified CD47 protein; lane i, gel-purified 26 kDa deglycosylated CD47 band; lane j, Lys-C digestion of purified deglycosylated CD47 protein. Samples a–c; d and e; f and g; and h–j were run on four different gels. Lanes a–e were Coomassie Blue-stained, lanes h–j were silver-stained. Lanes a–c were run using non-reduced samples on SDS/PAGE. The bands which were sequenced after transfer to ProBlott membranes are indicated (L1, L2).

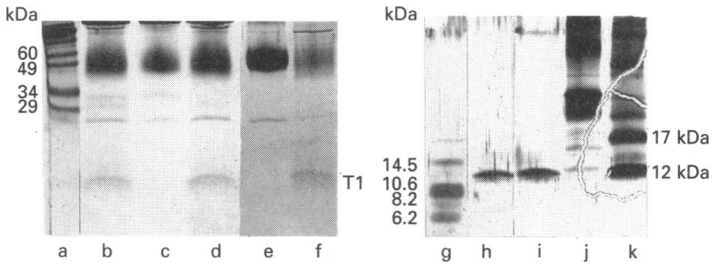

Figure 2 CD47 glycoprotein from protease-treated membranes: purification and deglycosylation of fragments

Samples after SDS/PAGE. Lane a, Sigma SDS-7B prestained molecular-mass markers; lanes b–d, CD47 protein material eluted from Protein A–Sepharose from (lane b) chymotrypsin-treated membranes, (lane c) untreated membranes and (lane d) trypsin-treated membranes; lanes e and f, CD47 from untreated and trypsin-treated membranes respectively, reacted with 10% 2-mercaptoethanol before electrophoresis; lane g, Sigma SDS-17S molecular-mass markers; lane h, gel-purified 12 kDa band from trypsin-treated membranes; lane i, PNGaseF-treated purified 12 kDa band; lane j, PNGaseF-treated purified CD47 from untreated membranes; lane k, PNGaseF-treated purified CD47 from trypsin-treated membranes. Lanes a–f were Coomassie-Blue-stained, and lanes g–k were silver-stained. Lanes a–d were run using non-reduced samples on SDS/PAGE on a separate gel to lanes e and f.

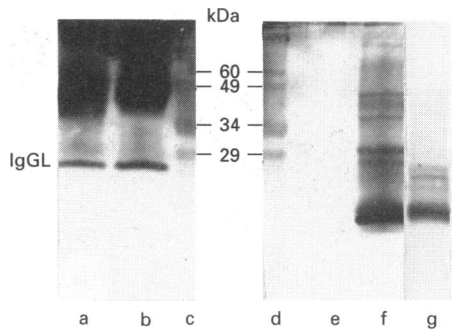

Figure 3 Immunoblotting of CD47 material from protease-treated membranes

Samples after SDS/PAGE. Lanes a, b and e–g, material immunoblotted with BRIC-125 antibody. Lane a, CD47 protein material from trypsin-treated membranes eluted from Protein A–Sepharose; lane b, CD47 protein material from untreated membranes eluted from Protein A–Sepharose; lane e, gel-purified 12 kDa band; lanes f and g, PNGaseF-treated purified CD47 band from trypsin-treated and chymotrypsin-treated membranes respectively; lanes c and d, Sigma SDS-7B prestained molecular-mass markers. Samples were run using non-reduced samples on SDS/PAGE.

26–26.5 kDa, together with a more sharply defined product of approx. 30 kDa (Figure 1e). These bands retain reactivity with mAb BRIC-125 on immunoblotting, suggesting that the BRIC-125-reactive epitope is not carbohydrate dependent (Figure 1g). Other higher-molecular-mass components (> 40 kDa) found after PNGase-F treatment appear to represent aggregates of the 26 kDa CD47 band since purification and further deglycosylation of these components had no effect on their mobility on SDS/PAGE (results not shown). Lys-C protease fingerprints of these purified components were the same as that obtained for the purified 26 kDa CD47 protein band (see below). Close examination of the 26 kDa CD47 products observed on both Coomassie Blue-stained gels and immunoblots suggested that this material was heterogeneous and present in the form of a doublet (Figures 1e and 1g respectively).

Lys-C protease digestion of the deglycosylated, purified CD47 protein generated multiple fragments which could be detected on silver staining of SDS/PAGE (Figure 1j). Only the 16.7 kDa (L1) and 8 kDa (L2) components were observed after transfer to ProBlott membrane and visualization with Coomassie Blue and these were sequenced. The 16.7 kDa band (L1) gave the N-terminal sequence SDAVSHTGDYTXEV and the 8 kDa band (L2) gave the N-terminal sequence SVEFTXDDXVVIP. Sequence was also obtained from a Lys-C protease fragment (L3) of the purified glycosylated protein. The major sequence obtained (L3A) was FKGRDIYTFDGALXK with a minor contaminating sequence NATGLGL (L3B).

Purification of CD47 from protease-treated membranes

Since intact CD47 protein appeared to have a blocked N-terminus and further sequence studies on the Lys-C-generated peptides were unsuccessful, an alternative strategy was used to obtain further protein sequence. Cleavage of the protein within the membrane environment was employed to generate new N-termini, and the derived CD47 protease fragments were subsequently purified by immunoprecipitation with mAb BRIC-125. Trypsin or chymotrypsin treatment of intact cells had no effect on CD47 protein (results not shown), but protease treatment of erythrocyte membranes yielded the fragments shown in Figures 2(b)–2(d). Treatment with trypsin (Figure 2d) or chymotrypsin

```
-18                 1
MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIP    22
                                SVEFTFXDDXVVIP
                                    (L2)

CFVTNMEAQNTTEVYVKWKFKGRDIYTFDGALNKSTVPTD    62
                        FKGRDIYTFDGALXK
                            (L3A)

FSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTELT    102
                            SDAVSHTGDYTXEV
                                (L1)

REGETIIELKYRVVSWFSPNENILIVIFPIFAILLFWGQF    142

GIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPG    182
        SGGMDEKTIALLVAGLVITVIVIVGAILF
                (T1)

EYSLKNATGLGLIVTSTGILILLHYYVFSTAIGLTSFVIA    222
     NATGLGL
      (L3B)

ILVIQVIAYILAVVGLSLCIAACIPMHGPLLISGLSILAL    262

AQLLGLVYMKFVASNQKTIQPPRKAVEEPLNAFKESKGMM    302

NDE    305
```

Figure 4 Amino acid sequence of CD47 glycoprotein

The amino acid sequence deduced from cDNA studies [5] is shown together with the peptide sequences obtained in this study. The sequence underlined with a broken rule is the presumed N-terminal signal sequence. Lined sequences are the putative transmembrane regions [19].

(Figure 2b) resulted in a slight shift in the leading edge of the 42–61 kDa CD47 band together with the appearance of an additional 12 kDa product, which was obtained in variable yield when the gels were run without 2-mercaptoethanol in the sample buffer. N-terminal amino acid sequencing of the 12 kDa product from trypsin digests (T1) gave the sequence SGGMDEKTIALL-VAGLVFTVIVIVGAILF. Surprisingly, the same N-terminal sequence, although against a much higher background of amino acids, was also obtained from the 42–61 kDa tryptic products in this preparation (results not shown). The tryptic 42–61 kDa band was purified under non-reducing conditions by SDS/PAGE. After treatment of this band with 2-mercaptoethanol, re-electrophoresis yielded more of the 12 kDa band, and also resulted in

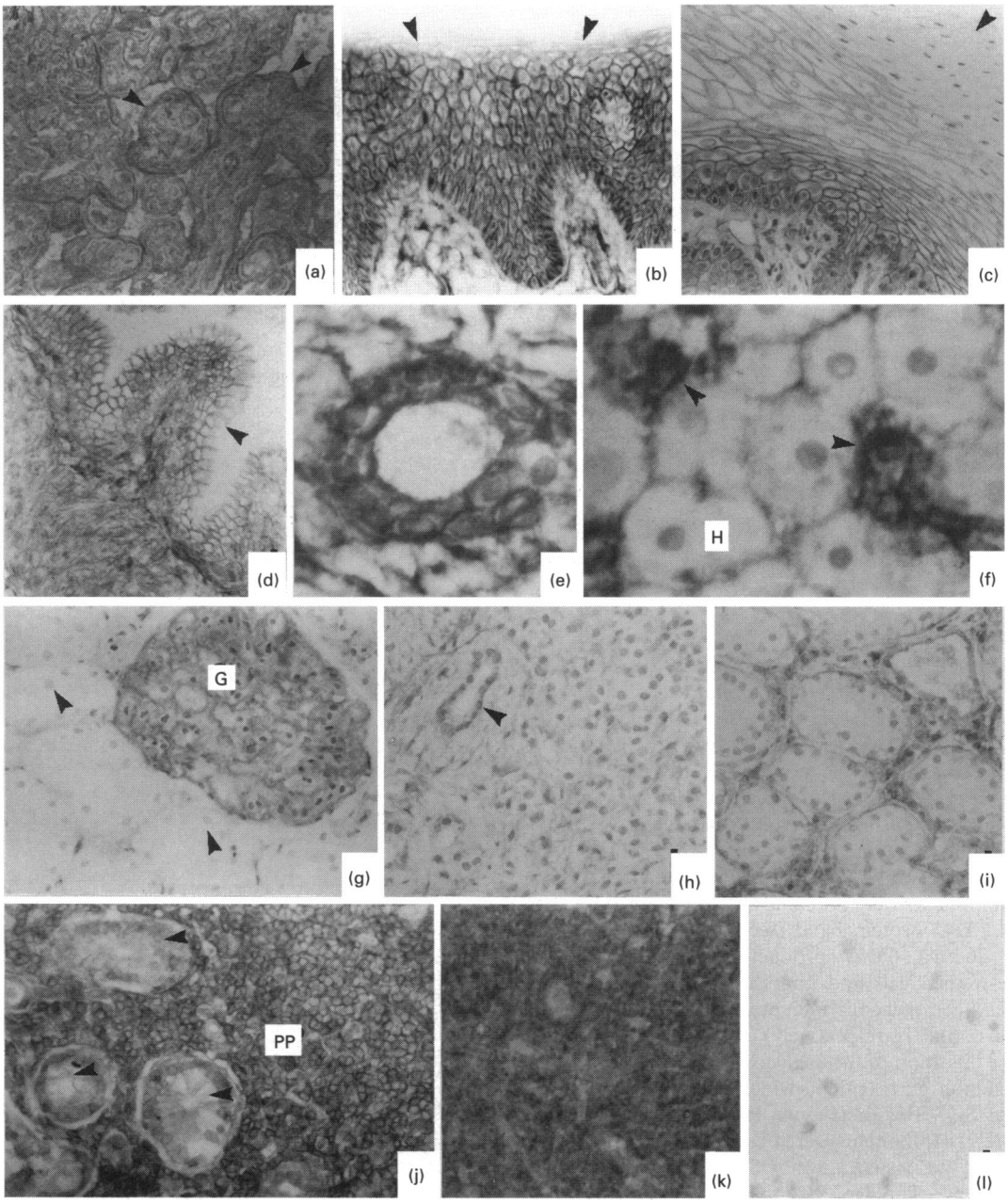

Figure 5 Immunoperoxidase staining of normal human tissues with mAb BRIC-125 to CD47

(**a**) Term placenta (× 40); (**b**) epidermis (× 100); (**c**) tongue (× 100); (**d**) gall bladder (× 100); (**e**) intrahepatic bile duct (× 400); (**f**) liver parenchyma (H, hepatocyte) (× 400); (**g**) renal cortex (G, glomerulus) (× 100); (**h**) pancreas (× 100); (**i**) stomach (× 125); (**j**) ileum (PP, Peyer's patch) (× 100); (**k**) brain cortex (× 160); (**l**) adjacent section to (**k**) (brain cortex) stained with the negative control mAb, BRIC-154 (× 160). Arrows identify: trophoblast epithelium covering chorionic villi in (**a**); BRIC-125-negative surface epithelial cells in (**b**) and (**c**); mucosa in (**d**); liver sinusoidal cells in (**f**); renal tubular epithelium in (**g**); pancreatic duct in (**h**); crypts in (**j**).

a shift in the mobility of the 42–61 kDa band (Figure 2f). The 12 kDa T1 band is therefore linked to the 42–61 kDa band by disulphide bonds since it is released by reduction.

The 42–61 kDa tryptic band is N-glycosylated as PNGase-F treatment caused a shift in its molecular mass to 17 kDa (Figure 2k). However, the molecular mass of the purified 12 kDa T1 band was not affected by PNGase-F treatment (Figure 2i). In common with the intact protein, the N-terminus of the 17 kDa deglycosylated tryptic fragment was blocked (results not shown), suggesting that this fragment represents the N-terminal portion of the protein. The higher-molecular-mass material seen in Figure 2(k) was most likely due to incomplete trypsin cleavage as well as to aggregation of the protein fragments.

Immunoblotting of the deglycosylated CD47 protein material from trypsin- or chymotrypsin-treated erythrocyte ghosts gave the results shown in Figure 3. BRIC-125 only detected the 17 kDa band (Figures 3f and 3g), which contains the deglycosylated N-terminal portion of the whole protein; the 12 kDa T1 fragment was unreactive (Figure 3e). The multiple higher-molecular-mass bands also observed may reflect partial deglyco-

sylation, as observed in the deglycosylation of the whole CD47 protein from untreated cells (Figure 1e).

The protein sequence of CD47 glycoprotein

The sequence obtained from the 12 kDa tryptic peptide fragment (T1) was scanned against the protein- and DNA-sequence databases and a perfect match was obtained with part of the predicted amino acid sequence of OA3, a human ovarian tumour marker cDNA isolated by expression cDNA cloning [5] (Figure 4). The other peptide sequences obtained in this study could also be aligned with the OA3 sequence (Figure 4). Brown et al. [6] have purified an integrin-associated protein (IAP) and reported the partial amino acid sequences of six peptides derived from this product. Three of these peptides match the cDNA-derived sequence for CD47 (Figure 4). IAP was demonstrated to be present in erythrocytes at an abundance similar to that of CD47 [12,13]. It seems clear therefore that IAP and CD47 are very closely related, if not identical, proteins (see the Discussion).

Tissue distribution studies

IAP is known to be widely distributed in human lymphoid and haemopoeitic cells and is expressed at the plasma membrane [6]. The determinant detected by mAb BRIC-125 was also shown to be broadly expressed on both haemopoeitic and non-haemopoeitic cells in the developing human liver [4]. By contrast, and despite its extensive distribution on ovarian carcinomas, the surface determinant OA3 has been reported to show a limited distribution on normal human tissues although, paradoxically, OA3 mRNA was found to be ubiquitous in normal tissues [5]. The distribution of CD47 was assessed by examining the reactivity of mAb BRIC-125 with a variety of normal human tissues by immunostaining (Figure 5). Reactivity was broadly distributed on both epithelial and mesenchymal elements at multiple sites. There was extensive reactivity in full-term placenta (Figure 5a); staining was present on trophoblast epithelium as well as on Hofbauer cells and endothelial cells within the chorionic villus cores. This was expected since IAP has previously been isolated from human placental membranes [6]. Reactivity was also observed on surface epithelia in skin (Figure 5b), tongue (Figure 5c), and trachea (results not shown). Although basal, parabasal and intermediate cells in these tissues were positive, superficial cells in both non-keratinizing (tongue) and keratinizing (skin) epithelia were consistently BRIC-125-negative (see arrow in Figure 5b). Staining of non-ductular epithelial cells in the liver and pancreas, and of tubular epithelial cells in kidney, was more variable; hepatocytes showed relatively weak peripheral membrane staining (Figure 5f) and renal tubular epithelial cells (Figure 5g) and pancreatic acinar cells (Figure 5h) were either weakly reactive or negative. In stomach (Figure 5i), reactivity appeared to be predominantly associated with the basal aspect of epithelial cells, and epithelial cells in the ileum, particularly those towards the base of the crypts, were also reactive (Figure 5j). Some of these tissues (particularly pancreas, kidney and stomach), appear to contain both BRIC-125-positive and BRIC-125-negative epithelial cells. Particularly intense staining was observed on sinusoidal cells in the liver (Figure 5f) and in aggregated lymphocytic nodules (Peyer's patches) in the ileum (Figure 5j). Finally, intense reactivity was also present in brain (Figure 5k); this is interesting because Campbell et al. [5] reported high levels of OA3 mRNA in human brain and these authors also reported that mAb OA3 bound to a neuroblastoma cell line. CD47 is therefore broadly distributed on mesenchyme and epithelia with prominent reactivity in surface and ductal epithelia, and in brain.

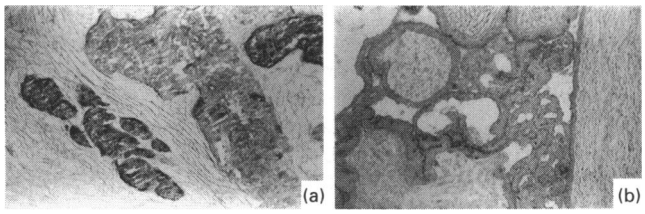

Figure 6 Immunoperoxidase staining of primary ovarian carcinomas with mAb BRIC-125 to CD47

Two different specimens are illustrated in (**a**) and (**b**) respectively, and both show extensive reactivity with mAb BRIC-125 (both × 40).

In view of the reported value of OA3 as an ovarian tumour marker [5] the reactivity of mAb BRIC-125 was also examined on a series of primary ovarian carcinomas. Two examples are shown in Figure 6. There was extensive reactivity in 12/12 tumours examined. In general, staining was very intense on tumour epithelial cells (Figure 6b) although in a minority of the specimens there was some heterogeneity in staining within a single tumour (for example, see Figure 6a).

DISCUSSION

Our interest in studying CD47 was because of its association with the Rh blood group system on human erythrocytes. Further studies have made it clear that it is a very broadly distributed protein. Our peptide-sequencing studies show that CD47 is the same as OA3, an ovarian tumour marker [5]. Campbell et al. [5] localized the OA3 gene to human chromosome 3, a location which is the same as that for CD47 [3,14]. CD47 is also closely similar to, or the same as, IAP described by Brown et al. [6], a molecule which is associated with a $\beta 3$ integrin in some tissues. IAP also migrates as a broad band with a similar molecular mass on SDS/PAGE to CD47 and has a comparable wide tissue distribution. The number of molecules of anti-IAP mAb bound to human erythrocytes at saturation is also similar to that reported for anti-CD47 mAbs [12,13]. The mRNA transcripts for OA3 are also found in most tissues [5]. The cDNA sequence of OA3 (CD47) predicts a multispanning membrane protein with a large N-terminal extracellular domain and six potential N-glycosylation sites [5]. Campbell et al. [5] have suggested that the N-terminus of the mature protein is glutamine-1 (see Figure 4). Our observation that the mature protein has a blocked N-terminus is consistent with the presence of pyroglutamic acid derived from this N-terminal glutamine.

The six potential N-glycosylation sites of CD47 are at Asn-5, Asn-16, Asn-32, Asn-55, Asn-93 and Asn-188 (Figure 4). Our sequence studies give some information on whether these sites are actually glycosylated in the erythrocyte form of CD47. PNGase-F treatment of the purified intact protein, or the 17 kDa trypsin fragment, produced more than one band on SDS/PAGE, suggesting the presence of multiple glycosylation sites. PNGase-F leaves an aspartic acid on the polypeptide chain in place of the glycosylated asparagine [15]. The sequences of two Lys-C peptides (L2 and L1) from the PNGase-F-treated protein gave aspartic acid instead of the predicted asparagines at residues 16 and 93 (Figure 4), showing that these two sites are glycosylated. The sequence of one of the peptides, obtained from the glycosylated protein (peptide L3A in Figure 4), spanned the N-glycosylation site at Asn-55 (Figure 4). No asparagine or other residue could be assigned at this sequence cycle, suggesting that

this site is also glycosylated. The glycosylation states of Asn-5 and Asn-32 are unknown. However, Gavel and Von Heijne [16] have suggested that N-glycosylation occurs preferentially at sites near the N-terminus of proteins and it might be expected that Asn-5 and Asn-32 are glycosylated. All the sites susceptible to PNGase-F are N-terminal to the trypsin cleavage site at Arg-150, which gives rise to the 17 kDa N-terminal trypsin fragment (Figure 2). Asn-188, which is C-terminal to this trypsin cleavage site, is therefore not likely to be glycosylated. The absence of glycosylation at Asn-188 is confirmed by the presence of asparagine in a peptide sequence obtained from Lys-C protease-digested glycosylated intact CD47 protein (peptide L3B in Figure 4).

Campbell et al. [5] pointed out that the extracellular domain of CD47 has significant homology with the V regions of the Ig superfamily present in a cartilage link protein and a poliovirus receptor. A search for common motifs showed the region around the cysteine at residue 96 had sequence similarities to other members of the Ig superfamily. Ig domains often contain a disulphide bond and, if this is also the case with CD47, Cys-96 might be expected to disulphide cross-link with either Cys-15 or Cys-23 [5], leaving one of the other cysteine residues free to disulphide cross-link elsewhere in the molecule. Our observation that the N- and C-terminal halves of the molecule are connected by a disulphide linkage suggests that this free cysteine residue (Cys-15 or Cys-23) may form a cross-link with Cys-241 or Cys-245 in the C-terminal portion of the molecule.

This paper shows that CD47 is present in many epithelial, mucosal and lymphoid tissues, including Peyer's patches. Springer [17] has pointed out that there are two types of circulating lymphocytes which differ in the types of antigen they encounter. One type circulate through the internal lymphatic system while the other moves through the peripheral lymph nodes draining the skin- and gut-associated lymph nodes (i-ncluding Peyer's patches) which drain the mucosal surfaces. Although we have not specifically studied the distribution of CD47 on the internal lymphatic system, it is clear that the abundance of CD47 parallels the second type of lymphocytic circulation system and CD47 may have a role in this circulatory process.

It is striking that the mAb OVTL-3 used by Campbell et al. [5] to define OA3 showed little or no reactivity with normal tissues but reacted with most ovarian carcinomas. OA3 (CD47) was reported to be an ovarian tumour marker on the basis of this OVTL-3 reactivity. In contrast the mAb used in this study show an extremely broad tissue distribution, and also reacted with all the ovarian carcinomas studied. The basis for this difference is unknown although it has been suggested that the reactivity of OVTL-3 is insensitive to carbohydrate differences in OA3 [5]. One possibility is that an unusual amino acid sequence is present in the CD47 isoforms expressed in ovarian cancer cells, and that OVTL-3 recognizes epitopes in this sequence. This clearly deserves further investigation.

Most receptors containing Ig domains have a single transmembrane segment [18]. It is possible that CD47 arises from the fusion of a gene for a protein with a single membrane-spanning segment as well as an Ig domain, and a multiple membrane-spanning transport protein, giving CD47 a dual function. Thus the protein could be involved in some form of cell signalling where binding of ligand to the Ig domain results in an alteration in the membrane transport function.

It is clear, both from our own immunohistochemical studies and those of Brown et al. [6], and from the Northern-blot studies of Campbell et al. [5], that CD47 has an unusually wide tissue distribution. This may reflect a widespread and general function for this protein. Alternatively the broad distribution could reflect a multifunctional role for the protein and may indicate that CD47 may have different functions in different tissues. Our initial interest in carrying out these studies was in the function of CD47 in the mature erythrocytes and its deficiency in Rh_{null} cells. The absence of integrins in mature erythrocytes makes it clear that CD47 does not have an integrin-associated role in erythrocytes. It is possible either that the transport function predominates in mature erythrocytes or that the protein has a role in immature erythrocytes during the erythropoetic events in the bone marrow which involve cell–cell interactions.

Part of this work was presented at the Fifth Conference on Leucocyte Differentiation Antigens, November 3–7, 1993, Boston, U.S.A. We wish to thank Dr. A. Gargaro for help with the computer analysis. This work was supported in part by grant no. 03219/1.5 from the Wellcome trust (C.H.H) and by the Central Blood Laboratories Authority, U.K.

REFERENCES

1 Knapp, W., Rieber, P., Dorken, B., Schmidt, R. E., Stein H. and Bourne, A. E. G. Kr. V. D. (1989) Immunol. Today **10**, 253–258
2 Hadam, M. R. (1989) in Leucocyte typing IV (Knapp, W., Dorken, B., Gilks, W. R., et al., eds.), p. 658–660, Oxford University Press, Oxford
3 Anstee, D. J., Holmes, C. H., Judson, P. A. and Tanner, M. J. A. (1992) in Protein Blood Group Antigens of the Human Red Cell. (Agre, P. and Cartron, J.-P., eds.), p. 170–181, Johns Hopkins Press, Baltimore, MD
4 Avent, N., Judson, P. A., Parsons, S. F., Mallinson, G., Anstee, D. J., Tanner, M. J. A., Evans, P. R., Hodges, E., Maciver, A. G. and Holmes, C. (1988) Biochem. J. **251**, 499–505
5 Campbell, I. G., Freemont, P. S., Foulkes, W. and Trowsdale, J. (1992) Cancer Res. **52**, 5416–5420
6 Brown, E., Hooper, L., Ho, T. and Gresham, H. (1990) J. Cell Biol. **111**, 2785–2794
6a Lindberg, F. P., Gresham, H. D., Schwarz, E. and Brown, E. J. (1993) J. Cell Biol. **123**, 485–496
7 Wainwright, S. D., Tanner, M. J. A., Martin, G. E. M., Yendle, J. E. and Holmes, C. H. (1989) Biochem. J. **258**, 211–220
8 Holmes, C. H., Simpson, K. L., Wainwright, S. D., Tate, C. G., Houlihan, J. H., Sawyer, I. H., Rogers, I. P., Spring, F. A., Anstee, D. J. and Tanner, M. J. A. (1990) J. Immunol. **144**, 3099–3105
9 Schagger, H. and Von Jagow, G. (1987) Anal. Biochem. **166**, 368–379
10 Nielson, B. L. and Brown, L. R. (1984) Anal. Biochem. **141**, 311–315
11 Lipman, D. J. and Pearson, W. R. (1985) Science **227**, 1435–1441
12 Rosales, C., Gresham, H. D. and Brown, E. J. (1992) J. Immunol. **149**, 2759–2764
13 Gardner, B., Anstee, D. J., Mawby, W. J., Tanner, M. J. A. and Bourne, A. E. G. Kr. V. D. (1991) Transfusion Med. **1**, 77–85
14 Anstee, D. J. and Spring, F. A. (1989) Transfusion Med. Rev. **3**, 13–23
15 Plummer, T. H., Jr., Elder, J. H., Alexander, S., Phelan, A. W. and Tarentino, A. L. (1984) J. Biol. Chem. **259**, 10700–10704
16 Gavel, Y. and von Heijne, G. (1990) Protein Eng. **3**, 433–442
17 Springer, T. A. (1990) Nature (London) **346**, 425–433
18 Williams, A. F. and Barclay, A. N. (1988) Annu. Rev. Immunol. **6**, 381–405
19 Kyte, J. and Doolittle, R. F. (1982) J. Mol. Biol. **157**, 105–132

Received 4 February 1994/29 June 1994; accepted 6 July 1994