Biochem. J. (1990) 271, 821–825 (Printed in Great Britain)

cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rhesus)-blood-group-antigen expression

Neil D. AVENT,*† Kay RIDGWELL,* Michael J. A. TANNER*‡ and David J. ANSTEE†

* Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, and
† International Blood Group Reference Laboratory, Southmead Road, Southmead, Bristol BS10 5ND, U.K.

The Rh-blood-group antigens (often described as Rhesus antigens) are associated with erythrocyte membrane proteins of approx. 30 kDa. We have determined the *N*-terminal 54 amino acid residues of the 30 kDa Rh D polypeptide (D_{30} polypeptide). We used primers based on these sequence data and the polymerase chain reaction (PCR) on human reticulocyte cDNA and genomic DNA to clone two types of PCR product of identical size. The two PCR products had related translated amino acid sequences between the 3′ ends of the primers, one of which was identical with that found for the D_{30} polypeptide. We designate the two related mRNA species which gave rise to the PCR products as Rh30A and Rh30B, the latter corresponding to the D_{30} polypeptide. We have isolated cDNA clones for the Rh30A protein which encode a hydrophobic membrane protein of 417 amino acids. The Rh30A protein has the same *N*-terminal 41 amino acids as the D_{30} polypeptide, but beyond this point the sequence differs, but is clearly related. The Rh30A protein probably corresponds to the $R6A_{32}$ polypeptide, another member of the Rh 30 kDa family of proteins, which may carry the C/c and/or E/e antigens. Hydropathy analysis suggests that the Rh30A protein has up to 12 transmembrane domains. Three of these domains are bordered by a novel cysteine-containing motif, which might signal substitutions at these cysteine residues. Information which supplements this paper (amino-acid-sequence-analysis histograms) is reported in Supplementary Publication SUP 50160 (4 pages), which has been deposited at the British Library Document Supply Centre, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in Biochem. J. (1990) **265**, 5.

INTRODUCTION

The Rh-blood-group system (often described as Rhesus) is of considerable significance in immunohaematology and probably second in importance only to the ABH system. The system is comprised of numerous antigens, the most polymorphic of which are the antigens C, c, E, e and D. The Rh antigens are of particular clinical importance because of their involvement in haemolytic disease of the newborn (Mollison *et al.*, 1987). Despite their importance, the structure and sequence of the proteins carrying the antigens remain poorly characterized. It has been suggested that the Rh polypeptides may be involved in maintaining the asymmetric distribution of phospholipids in the erythrocyte membrane (Kuypers *et al.*, 1984; de Vetten & Agre, 1988; Connor & Schroit, 1989). Immunoprecipitation studies using human polyclonal and monoclonal anti-D have identified a 30 kDa membrane polypeptide (D_{30} polypeptide) that appears not to be glycosylated (Moore *et al.*, 1982; Gahmberg, 1982, 1983). Anti-c and anti-E, as well as the murine monoclonal antibody R6A, immunoprecipitate related polypeptides (Moore *et al.*, 1982; Blanchard *et al.*, 1988; Moore & Green, 1987; Ridgwell *et al.*, 1983; Avent *et al.*, 1988*a*). Amino-acid-sequence analysis of the 32 kDa membrane component immunoprecipitated by R6A ($R6A_{32}$ polypeptide) showed that it had an *N*-terminal amino acid sequence that was identical with the *N*-terminal sequence of the D_{30} polypeptide (Avent *et al.*, 1988*b*; Bloy *et al.*, 1988; Saboori *et al.*, 1988). We describe here the sequence of a cDNA for a member of the $D_{30}/R6A_{32}$ group of proteins, which we denote the Rh30A protein. The cDNA probably corresponds to the $R6A_{32}$ polypeptide and may carry the C/c and/or E/e antigens. We also report the presence of an mRNA in erythroid cells for a second member of this group of proteins, the partial amino acid sequence of which corresponds with the D_{30} polypeptide. Information which supplements this paper (amino-acid-sequence-analysis histograms) is reported in Supplementary Publication SUP 50160 (4 pages), which has been deposited at the British Library Document Supply Centre, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in Biochem. J. (1990) **265**, 5.

MATERIALS AND METHODS

Purification and *N*-terminal sequencing of the D30 polypeptide

D_{30} polypeptide was purified and sequenced as described previously (Avent *et al.*, 1988*b*) from 500 ml of outdated human whole blood (presumptive Rh genotype cDE/cDE) using 1500 ml of anti-D culture supernatant. Approx. 7 nmol of D_{30} polypeptide was purified, of which approx. 1 nmol was loaded on to a pulsed liquid-phase protein sequencer (model 477A; Applied Biosystems International).

Polymerase-chain-reaction (PCR) amplification and cloning and sequencing of PCR products

PCR was done with human reticulocyte cDNA templates [prepared as described by Temple *et al.* (1977) and Maniatis *et al.* (1982)] using *Taq* polymerase (Perkin–Elmer/Cetus) according to

Abbreviations used: PCR, polymerase chain reaction; nt, nucleotide.

‡ To whom correspondence and reprint requests should be sent.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide sequence Databases under the accession number X54534.

(*a*)

```
          10        20        30        40        50  54
           :         :         :         :         :  :
SSKYPRSVRRXLPLWALTLEAALILLFYFFTHYDASLEDQKIRVLSYQDEQIND
```

(*b*)

```
5′-TTCCTTAGAGGATCAAAAGGGGCTCGTGGCATCCT-3′
    S  L  E  D  Q  K  G  L  V  A  S          Rh30PCR-A
    :  :  :  :  :  :     :     :
    S  L  E  D  Q  K  I  R  V  L  S          Rh30PCR-B
  CTCCCTGGAGGACCAGAAGATCCGGGTGCTGTCCT
```

Fig. 1. *N*-Terminal amino acid sequence of the D_{30} polypeptide and DNA sequence and translated protein sequence of cloned PCR products

(*a*) *N*-Terminal amino acid sequence of the D_{30} polypeptide. (*b*) Sequences of cloned Rh30 PCR products. The DNA and translated protein sequence of the Rh30PCR-A and Rh30PCR-B PCR products in the regions between the 3′ ends of the primers are shown.

the supplier's instructions. The degenerate sense PCR primer had the DNA sequence

TA(T/C)TTTTT(T/C)ACXCA(T/C)TA(T/C)GA(T/C)GC

and corresponded to amino acids 28–35 of the D_{30} polypeptide (Fig. 1*a*). The anti-sense PCR primer had the sequence

GAATTCCTC(G/A)TT(G/A/T)AT(T/C)-
TG(T/C)CT(A/G)TC(T/C)TG(A/G)TA

and corresponded to amino acids 47–54 of the D_{30} polypeptide (Fig. 1*a*) with the addition of an endonuclease-*Eco*RI site at the 5′ end. A total of 60 cycles of PCR were carried out. Although no product was detectable by PAGE analysis after the first 30 cycles of PCR, one-fifth of the reaction mixture at this stage was used as the template for a further 30 cycles of PCR, which yielded essentially only the 88 bp product. The 88 bp product was purified by PAGE, blunt-ended with T4 DNA polymerase and treated with polynucleotide kinase before ligation into phosphatase-treated *Eco*RV-cut Bluescript phagemid (Stratagene Cloning Systems, San Diego, CA, U.S.A.) using standard techniques (Maniatis *et al.*, 1982). Sequencing was done on single-stranded templates by the chain-termination method using Sequenase (U.S. Biochemical Corp., Cleveland, OH, U.S.A.) as described by the manufacturer. Three independent clones corresponding to Rh30PCR-A had the same DNA sequence between the 3′ ends of the primers, and the sequences of five independent clones corresponding to Rh30PCR-B were also the same in this region. Sequences were obtained from both strands of the PCR products.

Isolation and sequence of Rh30A cDNA clones

The total PCR product (containing both Rh30PCR-A and Rh30PCR-B) was prepared and purified as described above and labelled with a random-priming kit (BCL, Lewes, East Sussex, U.K.) using the oligoncleotides used for the PCR reaction as primers. This labelled probe was used to screen a human bone-marrow cDNA library in λgt10 (Clontech Laboratories, Palo Alto, CA, U.S.A.). Two positive clones were isolated, subcloned into Bluescript and sequenced after single-strand rescue as described above. Both strands were sequenced using synthetic oligonucleotides to prime progressively along the sequence.

```
                                                                -44 GATGCCTGGTGCTGGTGGAACCCCTGCACAGAGACGGACACAGG      -1
                    10                  20                  30                  40
 M  S  S  K  Y  P [R][S][V][R][R][C][L][P] L  W  A  L  T  L  E  A  A  L  I  L  L  F  Y  F  F  T  H  Y  D  A  S  L  E  D
ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCCTGCCCCTCTGGGCCCTAACACTGGAAGCAGCTCTCATTCTCCTCTTCTATTTTTTTACCCACTATGACGCTTCCTTAGAGGAT   120
                    50                  60                  70                  80
 Q  K  G  L  V  A  S  Y  Q  V  G  Q  D  L  T  V  M  A  A  L  G  L  G  F  L  T  S  N  F  R  R  H  S  W  S  S  V  A  F  N
CAAAAGGGGCTCGTGGCATCCTATCAAGTCGGCCAAGATCTGACCGTGATGGCGGCCCTTGGCTTGGGCTTCCTCACCTCAAATTTCCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC   240
                    90                  100                 110                 120
 L  F  M  L  A  L  G  V  Q  W  A  I  L  L  D  G  F  L  S  Q  F  P  P  G  K  V  V  I  T  L  F  S  I  R  L  A  T  M  S  A
CTCTTCATGCTGGCGCTTGGTGTGCAGTGGGCAATCCTGCTGGACGGCTTCCTGAGCCAGTTCCCTCCTGGGAAGGTGGTCATCACACTGTTCAGTATTCGGCTGGCCACCATGAGTGCT   360
                    130                 140                 150                 160
 M  S  V  L  I  S  A  G  A  V  L  G  K  V  N  L  A  Q  L  V  V  M  V  L  V  E  V  T  A  L  G  T  L  R  M  V  I  S  N  I
ATGTCGGTGCTGATCTCAGCGGGTGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTGGTGATGGTGCTGGTGGAGGTGACAGCTTTAGGCACCCTGAGGATGGTCATCAGTAATATC   480
                    170                 180                 190                 200
 F  N  T  D  Y  H  M  N  L  R  H  F  Y  V  F  A  A  Y  F  G  L  T  V  A  W [C][L][P][K][P][L][P][K] G  T  E  D  N  D  Q
TTCAACACAGACTACCACATGAACCTGAGGCACTTCTACGTGTTCGCAGCCTATTTTGGGCTGACTGTGGCCTGGTGCCTGCCAAAGCCTCTACCCAAGGGAACGGAGGATAATGATCAG   600
                    210                 220                 230                 240
 R  A  T  I  P  S  L  S  A  M  L  G  A  L  F  L  W  M  F  W  P  S  V  N  S  P  L  L  R  S  P  I  Q  R  K  N  A  M  F  N
AGAGCAACGATACCCAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGGCCAAGTGTCAACTCTCCTCTGCTGAGAAGTCCAATCCAAAGGAAGAATGCCATGTTCAAC   720
                    250                 260                 270                 280
 T  Y  Y  A  L  A  V  S  V  V  T  A  I  S  G  S  S  L  A  H  P  Q  R  K  I  S  M  T  Y  V  H  S  A  V  L  A  G  G  V  A
ACCTACTATGCTCTAGCAGTCAGTGTGGTGACAGCCATCTCAGGGTCATCCTTGGCTCACCCCCAAAGGAAGATCAGCATGACTTATGTGCACAGTGCGGTGTTGGCAGGAGGCGTGGCT   840
                    290                 300                 310                 320
 V  G  T  S  C  H  L  I  P  S  P  W  L  A  M  V  L  G  L  V  A  G  L  I  S  I  G  G  A [K][C][L][P] V  C  C  N [R] V  L
GTGGGTACCTCGTGTCACCTGATCCCTTCTCCGTGGCTTGCCATGGTGCTGGGTCTTGTGGCTGGGCTGATCTCCATCGGGGGAGCCAAGTGCCTGCCGGTGTGTTGTAACCGAGTGCTG   960
                    330                 340                 350                 360
 G  I  H  H  I  S  V  M  H  S  I  F  S  L  L  G  L  L  G  E  I  T  Y  I  V  L  L  V  L  H  T  V  W  N  G  N  G  M  I  G
GGGATTCACCACATCTCCGTCATGCACTCCATCTTCAGCTTGCTGGGTCTGCTTGGAGAGATCACCTACATTGTGCTGCTGGTGCTTCATACTGTCTGGAACGGCAATGGCATGATTGGC  1080
                    370                 380                 390                 400
 F  Q  V  L  L  S  I  G  E  L  S  L  A  I  V  I  A  L  T  S  G  L  L  T  G  L  L  L  N  L  K  I [W][K][A][P][H][V][A] K
TTCCAGGTCCTCCTCAGCATTGGGGAACTCAGCTTGGCCATCGTGATAGCTCTCACGTCTGGTCTCCTGACAGGTTTGCTCCTAAATCTCAAAATATGGAAAGCACCTCATGTGGCTAAA  1200
                    410                 417
 Y  F  D  D  Q  V  F [W][K][F][P][H][L][A] V  G  F  *
TATTTTGATGACCAAGTTTTCTGGAAGTTTCCTCATTTGGCTGTTGGATTTTAAGCAAAAGCATCCAAGAAAAACAAGGCCTGTTCAAAAACAAGACAACTTCCTCTCACTGTTGCCTGC  1320

ATTTGTACGTGAGAAACGCTCATGACAGCAAAGTCTCCTTATGTATAATGAAACAAGGTCAGAGACAGATTTGATATTAAAAAATTAAAAAAAAAAAAAAAA  1422
```

Fig. 2. DNA and inferred protein sequence of RH30A cDNA

The DNA sequence is numbered so that nt 1 corresponds with the initiating methionine codon for the protein. The sequence is a composite derived from two different cDNA clones. One clone contained the sequence from the 5′ end to nt 829, whereas the other clone contained the sequence from nt −41 to the 3′ end of the sequence shown. The CLP motifs and repeat sequence at the *C*-terminus of the protein are boxed.

RESULTS AND DISCUSSION

Cloning of PCR amplification products corresponding to 30 kDa Rh proteins

We were unsuccessful in obtaining cDNA clones for the 30 kDa Rh proteins using probes based on the previously published amino-acid-sequence data for the D_{30} and $R6A_{32}$ polypeptides (*N*-terminal sequence up to residue 32; Avent *et al.*, 1988*b*). We therefore determined further *N*-terminal amino acid sequence for the D_{30} polypeptide up to residue 54 (Fig. 1*a*). The $R6A_{32}$ polypeptide gave an identical amino acid sequence up to residue 41, but we were unable to obtain sequence data beyond this point. Mixed sequence oligonucleotides based on the D_{30} polypeptide amino acid sequence were prepared. By using PCR with these oligonucleotide primers and reticulocyte cDNA, PCR products of the predicted size (88 bp) were obtained (Fig. 1*b*). After cloning and sequencing, the PCR products were found to be of two types of identical size. The translated amino acid sequence between the 3′ ends of the primers of one type (Rh30PCR-B) was found to be identical with that of the D_{30} polypeptide amino acid sequence (residues 36–46; Fig. 1*b*). The sequence between the 3′ ends of the primers of the other type (Rh30PCR-A) was clearly related to that of Rh30PCR-B and showed conservation of eight of the 11 amino acid residues coded for in this portion of the sequence (Fig. 1*b*) and 60% DNA sequence similarity. This shows that at least two related mRNA species of the $D_{30}/R6A_{32}$ polypeptide family are present in erythroid cells. We designate the mRNAs which gave rise to the Rh30PCR-A and Rh30PCR-B products as Rh30A and Rh30B respectively.

Isolation and sequence of RH30A cDNA clones

The 88 bp products of PCR (containing both Rh30PCR-A and Rh30PCR-B) were used to screen a human bone-marrow cDNA library. Two positive cDNA clones of different size were isolated with the same DNA sequence where they overlapped. The translated amino acid sequence of the cDNA yielded the *N*-terminal amino acid sequence of the $D_{30}/R6A_{32}$ polypeptides up to residue 41 (amino acids 2–42 in Fig. 2). The DNA sequence of the clones corresponding to amino acid residues 37–47; Fig. 2) also perfectly matched the nucleotide sequence between the primers of the PCR product Rh30PCR-A (Fig. 1*b*), showing that the cDNA clones correspond to the Rh30A mRNA. The Rh30A cDNA predicted an open reading frame of 417 amino acids starting from the Met codon at nucleotide (nt) 1 (Fig. 2). This Met codon is preceeded by a purine residue at position −3, which is thought to be important in translation initiation (Kozak, 1986) and is likely to be the initiator Met codon. A Met codon is also present at nt −43 in a different reading frame upstream of this, but this is followed by an in-frame termination codon at nt 2. *N*-Terminal amino acid sequencing of the $D_{30}/R6A_{32}$ polypeptides (Avent *et al.*, 1988*b*; Bloy *et al.*, 1988; Saboori *et al.*, 1988) has shown that the sequences of the mature proteins start at Ser-2, so the initiating methionine residue is post-translationally cleaved. The composite 1466-bp cDNA contains 44 nts of 5′ non-coding region and 171 nts of 3′ non-coding region (Fig. 2). The 3′ end of the cDNA terminated with an A_{16} tract, but no sequence matching the consensus polyadenylation signal AATAAA was present close upstream of the A_{16} tract. It is possible that the 3′ end of the mRNA extends further than the 3′ end of our clones.

Structure of Rh30A protein

Hydropathy analysis of the predicted Rh30A sequence using a 20-amino-acid window showed that it contained two extended hydrophobic regions of sequence as well as five hydrophobic sections with the lengths expected for single membrane-spanning segments (Fig. 3*a*). The overall hydrophobic nature of the protein is apparent from the hydropathy plot and consistent with the properties of the $D_{30}/R6A_{32}$ polypeptide (de Vetten & Agre, 1988). The two long hydrophobic sections of sequence comprise residues 76–153 (77 amino acids) and 273–388 (115 amino acids). Examination of the sequence shows the presence of small clusters of polar amino acids within these long hydrophobic sections, which become evident as troughs in the hydropathy profile when the analysis is done using a smaller window of nine amino acids (Fig. 3*b*), but are obscured because of their short length when the analysis is done using the wider window. These polar clusters probably represent regions of polypeptide at the membrane surface and the first extended hydrophobic region could span the membrane three times, whereas the second could span the membrane four times. Overall the protein probably contains twelve membrane-spanning domains. The *N*- and *C*-terminal halves of the protein show similarities in their hydropathy profiles, suggesting that the protein may comprise two domains with similar structure, each containing six membrane-spanning regions. Analysis of the sequence using the DIAGON method (Staden, 1982) did not, however, yield any evidence for the two halves of the protein having similar amino acid sequences. However, the related sequences WKAPHVA and WKFPHLA are found adjacent to each other in the *C*-terminal tail of the protein and a 13-nt DNA sequence in the 5′ non-coding sequence (nts −38 to −26) is directly repeated in the coding region (nts 425–437). Four acidic residues (Glu-21, Asp-95, Glu-156 and Glu-340) and one basic residue (Arg-114) are present in putative membrane-spanning sequences and could be located within the membrane interior (Fig. 3*c*).

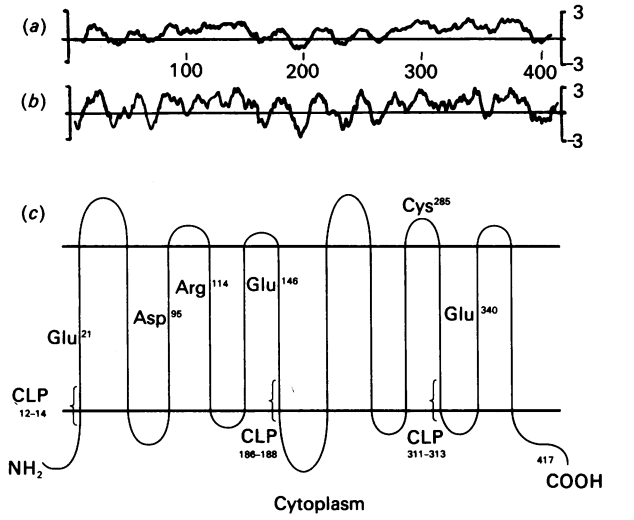

Fig. 3. Hydropathy analysis and schematic diagram of the possible topology of the Rh30A protein

(*a*) Kyte–Doolittle hydropathy analysis using a window of 20 amino acids. (*b*) Kyte–Doolittle hydropathy analysis using a window of nine amino acids. Hydrophobicity is shown as the vertical axis with the hydrophobic side of the plot having a positive value. The horizontal axis shows the amino acid residue number along the sequence. Analysis was done using the program PEPPLOT in the University of Wisonsin Genetics Computer Group package (Devereux *et al.*, 1984). (*c*) Possible topology of the Rh30A protein. This schematic diagram assumes the presence of 12 membrane-spanning segments in the protein and shows the possible locations of the CLP motifs, the extracellular cysteine residue and the acidic and basic residues which appear to be located within the membrane interior.

The lack of a cleaved *N*-terminal hydrophobic signal sequence suggests that the first membrane-spanning domain of the Rh30A protein acts as an uncleaved signal-anchor sequence. The sequences adjacent to this first membrane span fit the criteria described for signal-anchor sequences (Hartmann *et al.*, 1989; Dalbey, 1990; von Heijne, 1990). The sequence on the *N*-terminal side of the hydrophobic region is strongly basic [+5 using the criteria of Hartmann *et al.* (1989)], whereas that in the 15 amino acids flanking the *C*-terminal side of the hydrophobic region is predominantly acidic (−1.5). The charge difference between the flanking regions of +6.5 charges towards the *N*-terminus strongly predicts that the *N*-terminus is located on the cytoplasmic side of the membrane (Hartmann *et al.*, 1989). A schematic diagram of the possible topology of the protein based on this and the hydropathy analysis is shown in Fig. 3(*c*). Treatment of erythrocytes with thiol reagents abolishes Rh C and D antigen activity (Green, 1965, 1967). An extracellular thiol group is necessary for D antigen activity, since an impermeant maleimide (glutathione maleimide) inactivates the D antigen activity of intact erythrocytes (Abbott & Schachter, 1976). The model in Fig. 3(*c*) suggests that only one cysteine residue (Cys-285) is extracellular, and this may be the thiol required for Rh antigen activity.

Three of the six cysteine residues in the protein (Cys-12, Cys-186 and Cys-311 in Fig. 2) form part of the sequence motif CLP, and these are found at locations close to the points at where the polypeptide chain enters the membrane. A basic residue is located immediately adjacent to the *N*- or *C*-terminal side of the CLP motif, and another basic amino acid is present five residues away from the *N*- or *C*-terminal side of the motif in regions likely to be present at the membrane surface (Fig. 3*c*). The 30 kDa Rh polypeptides have been shown to be the major palmitoylated components of the erythrocyte membrane (Staufenbiel, 1988; de Vetten & Agre, 1988), and de Vetten & Agre (1988) suggest that each protein molecule is acylated several times with palmitoyl residues. Direct protein sequencing of the 30 kDa Rh polypeptides did not yield an identifiable residue at the position corresponding to Cys-12 (the cysteine in the first CLP motif), even after pyridylethylation of the polypeptide (Avent *et al.*, 1988*b*), suggesting that this cysteine residue is substituted in some way, and it is possible that it is palmitoylated. The other cysteine residues located in CLP motifs may be similarly substituted. However, it is noteworthy that, in another polytopic membrane protein, namely bovine rhodopsin, the two cysteine residues in the sequence CCGK (residues 323–324) are both palmitoylated (Ovchinnikov *et al.*, 1988). This sequence occurs at the cytoplasmic surface of the membrane on the *C*-terminal side of a membrane-spanning domain. A similar sequence CCNR occurs in the Rh30A protein (residues 315–318), and is also likely to be at the membrane surface. These cysteine residues could alternatively be the sites of palmitoylation. The asymmetric distribution of phospholipids in the erythrocyte membrane results from an ATP-dependent transport system (Seigneuret & Devaux, 1984), and it has been suggested that the 30 kDa Rh polypeptides are involved in this process (Kuypers *et al.*, 1984; de Vetten & Agre, 1988; Connor & Schroit, 1989). The CLP motifs in the Rh30A protein may have a functional role in this process.

Relationship of Rh30A protein to the Rh antigens

The 41 residues of the *N*-terminal amino acid sequence of the $R6A_{32}$ polypeptide which we have been able to establish by protein sequencing are identical with that predicted from the Rh30A cDNA (residues 2–42). This portion of the sequence is also the same as the D_{30} polypeptide sequence (Fig. 1*a*). Immediately adjacent to this sequence are the three amino-acid-sequence differences which are so far known to distinguish the D_{30} from the Rh30A protein. The Rh30A protein therefore corresponds to the $R6A_{32}$ polypeptide rather than the D_{30} polypeptide. There are no potential *N*-glycosylation sites (NXS/T) in the predicted Rh30A protein sequence. It has been shown that the D_{30} polypeptide is not glycosylated (Gahmberg, 1983; Avent *et al.*, 1988*a*), and it is likely that the Rh30A protein is similar in this respect. We previously suggested that the $R6A_{32}$ polypeptide was *N*-glycosylated, on the basis of a shift in M_r of the purified protein on treatment with endoglycosidase preparations (Avent *et al.*, 1988*a*). A likely explanation for this observation is that it was due to cleavage resulting from proteinase contamination of the partially purified endoglycosidase mixture used in these experiments.

We have shown that mRNAs for two members of the $D_{30}/R6A_{32}$ polypeptide family are present in erythroid cells. The portion of the Rh30B mRNA cloned in the Rh30PCR-B clone (Fig. 1*b*) corresponds in translated amino acid sequence to that found for the D_{30} polypeptide (Fig. 1*a*). It is possible that other members of this family are also present in erythrocytes, and it has been suggested that the C and E antigens are present on different polypeptides (Blanchard *et al.*, 1988). Although the Rh30A polypeptide does not give rise to the Rh D antigen, the C/c and/or E/e antigens may arise from polymorphic structures in this protein. The availability of cDNA clones for members of this family should provide a basis for elucidating the structures responsible for determining the different antigens of the Rh system.

Note added in proof (received 10 September 1990)

After the submission of this paper a report appeared describing the isolation of a cDNA clone which is the same as the Rh30A cDNA (Cherif-Zahar *et al.*, 1990).

This work was supported in part by grants from the Medical Research Council, the Central Blood Laboratory Authority and the Wellcome Trust. We are grateful to Professor L. Luzzatto for a sample of reticulocyte-rich blood.

REFERENCES

Abbott, R. E. & Schachter, D. (1976) J. Biol. Chem. **251**, 7176–7183
Avent, N. D., Judson, P. A., Parsons, S. F., Mallinson, G., Anstee, D. J., Tanner, M. J. A., Evans, P. R., Hodges, E., Maciver, A. G. & Holmes, C. (1988*a*) Biochem. J. **251**, 499–505
Avent, N. D., Ridgwell, K., Mawby, W. J., Tanner, M. J. A., Anstee, D. J. & Kumpel, B. M. (1988*b*) Biochem. J. **256**, 1043–1046
Blanchard, D., Bloy, C., Hermand, P., Cartron, J.-P., Saboori, A. M., Smith, B. L. & Agre, P. (1988) Blood **72**, 1424–1427
Bloy, C., Blanchard, D., Dahr, W., Beyreuther, K., Salmon, C. & Cartron, J.-P. (1988) Blood **72**, 661–666
Cherif-Zahar, B., Bloy, C., Le van Kim, C., Blanchard, D., Bailly, P., Hermand, P., Salmon, C., Cartron, J.-P. & Colin, Y. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6243–6247
Connor, J. & Schroit, A. J. (1989) Biochemistry **28**, 9680–9685
Dalbey, R. E. (1990) Trends Biochem. Sci. **15**, 253–257
de Vetten, M. P. & Agre, P. (1988) J. Biol. Chem. **263**, 18193–18196
Devereux, J., Haeberli, P. & Smithies, P. (1984) Nucleic Acids Res. **12**, 387–395
Gahmberg, C. G. (1982) FEBS Lett. **140**, 93–97
Gahmberg, C. G. (1983) EMBO J. **2**, 223–227
Green, F. A. (1965) Vox Sang. **10**, 32–53
Green, F. A. (1967) Immunochemistry **4**, 247–257
Hartmann, E., Rapoport, T. A. & Lodish, H. F. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 5786–5790
Kozak, M. (1986) Cell (Camridge, Mass.) **44**, 283–292
Kuypers, F., van Linde-Sibenius-Trip, M., Roelofsen, B., Tanner, M. J. A., Anstee, D. J. & Op den Kamp, J. A. F. (1984) Biochem. J. **221**, 931–934

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Mollison, P. L., Engelfriet, C. P. & Contreras, M. (1987) Blood Transfusion in Clinical Medicine, 8th edn., Blackwell, Oxford

Moore, S. & Green, C. (1987) Biochem. J. **244**, 735–741

Moore, S., Woodrow, C. F. & McClelland, D. B. L. (1982) Nature (London) **295**, 529–531

Ovchinnikov, Y. A., Abdulaev, N. G. & Bogachuk, A. S. (1988) FEBS Lett. **230**, 1–5

Ridgwell, K., Roberts, S. J., Tanner, M. J. A. & Anstee, D. J. (1983) Biochem. J. **213**, 267–269

Saboori, A. M., Smith, B. L. & Agre, P. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 4042–4045

Seigneuret, M. & Devaux, P. F. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 3751–3755

Staden, R. (1982) Nucleic Acids Res. **10**, 2951–2961

Staufenbiel, M. J. (1988) J. Biol. Chem. **263**, 13615–13622

Temple, G. F., Chang, J. & Kan, Y. W. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 3047–3051

von Heijne, G. (1990) J. Membr. Biol. **115**, 195–201

Received 9 August 1990; accepted 5 September 1990