Biochem. J. (1992) 287, 223–228 (Printed in Great Britain)

Isolation of cDNA clones for a 50 kDa glycoprotein of the human erythrocyte membrane associated with Rh (Rhesus) blood-group antigen expression

Kay RIDGWELL,* Nigel K. SPURR,† Bimbola LAGUDA,† Catriona MACGEOCH,† Neil D. AVENT‡ and Michael J. A. TANNER*§

*Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, U.K.,
†Clare Hall Laboratories, Blanche Lane, South Mimms, Potters Bar, Herts. EN6 3LD, U.K.,
and ‡International Blood Group Reference Laboratory, Southmead Road, Bristol BS10 5ND, U.K.

The Rh blood-group antigens are associated with human erythrocyte membrane proteins of approx. 30 kDa (the Rh_{30} polypeptides). Heterogeneously glycosylated membrane proteins of 50 and 45 kDa (the Rh_{50} glycoproteins) are coprecipitated with the Rh_{30} polypeptides on immunoprecipitation with anti-Rh-specific mono- and poly-clonal antibodies. We have isolated cDNA clones representing a member of the Rh_{50} glycoprotein family (the Rh50A glycoprotein). We used PCR with degenerate primers based on the *N*-terminal amino acid sequence of the Rh_{50} glycoproteins and human genomic DNA as a template and cloned and sequenced three types of PCR product of the expected size. Two of these products, Rh50A and Rh50B, gave the same translated amino acid sequence which corresponded to the expected Rh_{50} glycoprotein sequence but had only 75% DNA sequence similarity. The third product (Rh50C) contained a single base deletion, and the translated amino acid sequence contained an in-frame stop codon. We have isolated cDNA clones containing the full coding sequence of the Rh50A glycoprotein. This sequence predicts that it is a 409-amino acid *N*-glycosylated membrane protein with up to 12 transmembrane domains. The Rh50A glycoprotein shows clear similarity to the Rh30A protein in both amino acid sequence and predicted topology. Our results are consistent with the Rh30 and Rh50 groups of proteins being different subunits of an oligomeric complex which is likely to have a transport or channel function in the erythrocyte membrane. We mapped the Rh50A gene to human chromosome 6p21-qter, showing that genetic differences in the Rh30 rather than the Rh50 genes specify the major polymorphic forms of the Rh antigens.

INTRODUCTION

The human Rh blood-group antigens are of importance in blood transfusion and immunohaematology because of their high antigenicity in man and their involvement in haemolytic disease of the newborn (Mollison *et al.*, 1987). The detailed structure of the many antigens of the Rh blood group system is still unknown, but recent identification of some of the polypeptides involved in the expression of these antigens has allowed progress in their sequence and structural characterization.

Immunoprecipitation studies using anti-c, D or E sera and Rh-related mouse monoclonal antibodies of the R6A type (which react with erythrocytes of all phenotypes except Rh_{null}) have identified two Rh-related erythrocyte membrane polypeptides of 30 and 32 kDa (Moore *et al.*, 1982; Gahmberg, 1982; Ridgwell *et al.*, 1983; Moore & Green, 1987; Avent *et al.*, 1988*a*; Blanchard *et al.*, 1988), denoted the D_{30} and $R6A_{32}$ polypeptides respectively. These 30 kDa integral membrane proteins are not glycosylated (Moore *et al.*, 1982; Gahmberg, 1982). The *N*-terminal amino acid sequence of both polypeptides has been determined (Avent *et al.*, 1988*b*; Bloy *et al.*, 1988; Saboori *et al.*, 1988) and cDNA clones corresponding to one of them (designated Rh30A, which most likely represents the $R6A_{32}$ polypeptide) have been isolated (Avent *et al.*, 1990; Cherif-Zahar *et al.*, 1990). The Rh30A polypeptide contains 417 amino acids and is a very hydrophobic protein with up to 12 membrane-spanning segments.

In addition to the 30 kDa polypeptides described above, immunoprecipitates obtained using anti-Rh-specific antibodies or the R6A type of mouse monoclonal antibodies also contain higher-molecular-mass glycoproteins which migrate as broad bands on SDS/PAGE because of heterogeneity in their glycan chains (Moore & Green, 1987; Avent *et al.*, 1988*a*,*b*). A glycoprotein of molecular mass 45–100 kDa (the D_{50} glycoprotein) is obtained together with the D_{30} polypeptide on immunoprecipitation with anti-D antibodies. Similarly, a glycoprotein of 35–52 kDa (the $R6A_{45}$ glycoprotein) is immunoprecipitated together with the $R6A_{32}$ polypeptide when the R6A type of antibody is used. The amino acid sequence of the D_{50} polypeptide has been determined up to residue 39 although residue 2 could not be assigned (Avent *et al.*, 1988*b*). The $R6A_{45}$ glycoprotein gave the same *N*-terminal amino acid sequence up to residue 30 except that amino residue 2 was assigned as cysteine (Avent *et al.*, 1988*b*). The *N*-terminal amino acid sequence of the $D_{50}/R6A_{45}$ glycoproteins (collectively described here as the Rh_{50} glycoproteins) differed from that of the $D_{30}/R6A_{32}$ polypeptides (collectively described here as the Rh_{30} polypeptides), although the sequences of the two groups of proteins showed some similarity (Avent *et al.*, 1988*b*). The coprecipitation of these two groups of proteins suggested that they may be associated in the erythrocyte membrane.

We describe here the sequence of a cDNA clone for a member of the Rh_{50} group of glycoproteins which we denote the Rh50A glycoprotein. The cDNA predicts that the Rh50A glycoprotein is a very hydrophobic integral membrane protein which is likely to be glycosylated at two of three potential *N*-glycan addition sites. The amino acid sequence of the protein is related to that of the Rh30A polypeptide and it is likely to have a structure and topology in the membrane which is very similar to the Rh30A polypeptide (Avent *et al.*, 1990; Cherif-Zahar *et al.*, 1990). The association between the two proteins and their similarity in

§ 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 X64594.

structure are consistent with earlier suggestions that they are subunits of a functional complex in the erythrocyte membrane (Avent *et al.*, 1988*a*; Anstee & Tanner, 1988; Agre & Cartron, 1991).

MATERIALS AND METHODS

PCR amplification, and cloning and sequencing of PCR products

D_{50} and $R6A_{45}$ polypeptides were purified and sequenced as described by Avent *et al.* (1988*b*). PCR was carried out using human genomic DNA as a template and Amplitaq Taq polymerase (Perkin–Elmer/Cetus Corp., Norwalk, CT, U.S.A.) according to the manufacturer's instructions. The forward degenerate primer corresponded to amino acid residues 1–9 of the D_{50} glycoprotein sequence (Avent *et al.*, 1988*b*) and had the DNA sequence 5′-ATGTGCTTCACXTTC/TCCXC/TTXATGGC-3′. The reverse primer corresponded to amino acid residues 23–21 of the D_{50} glycoprotein and had the DNA sequence 5′-GTCTGGTCXGTC/TTCA/GTAC/TTCXACG/AAA-3′. The expected 92 bp product was obtained using 30 cycles of amplification, and was purified by PAGE, ligated into the Bluescript phagemid (Stratagene Cloning Systems, San Diego, CA, U.S.A.) and sequenced using the Sequenase Version 2.0 DNA sequencing kit (U.S.B. Corp., Cleveland, Ohio, U.S.A.) according to the manufacturer's instructions.

Isolation and sequencing of Rh50 cDNA clones

Two oligonucleotides were prepared corresponding to nucleotides 4–31 of the unique sequence found for the Rh50A PCR product and nucleotides 7–31 of the unique sequence of Rh50B PCR product shown in Fig. 1. Each was separately radiolabelled at the 5′ end using polynucleotide kinase (Amersham International, Amersham, Bucks., U.K.) and purified by elution from a 20% denaturing polyacrylamide gel (Sambrook *et al.*, 1989). The labelled oligonucleotides were used separately to screen a number of human cDNA libraries. Only the Rh50A-specific oligonucleotide yielded positive clones and these were obtained from a bone marrow λgt10 cDNA library and a fetal liver λgt11 cDNA library (Clontech Laboratories, Palo Alto, CA, U.S.A.). The inserts in the λgt10 clones were excised with restriction endonucleases as described in the text, subcloned into Bluescript and sequenced as described above. The insert in the fetal liver library λgt11 clone was subcloned after PCR amplification using λgt11 forward and reverse primers (New England Biolabs Inc., Beverly, MA, U.S.A.) and blunt-end ligation into Bluescript.

Double-stranded sequencing of the insert in the λB2 cDNA clone

Phage DNA was isolated and purified by caesium chloride centrifugation using standard techniques (Sambrook *et al.*, 1989). The double-stranded template was sequenced using a variation of the method of Zimmerman *et al.* (1990). Template (2 μl; approx. 5 ng of DNA) was mixed with 1 μl of λgt10 reverse primer (5 ng/μl, New England Biolabs) and 1 μl of 0.5 M-NaOH, heated at 80 °C for 5 min and then cooled to 37 °C for 5 min before addition of 1 μl of 0.5 M-HCl and 5 μl of 80 mM-Tris/HCl (pH 7.5)/40 mM-$MgCl_2$ containing 10% dimethyl sulphoxide. Annealing was continued for a further 10 min at 37 °C before the DNA sequence reactions were carried out using Sequenase as described above.

PCR amplification of the 5′ upstream region of Rh50A cDNA, and single-stranded sequencing of asymmetric PCR products

Thirty cycles of PCR amplification were carried out with human reticulocyte cDNA as template and AmpliTaq Taq DNA polymerase (Perkin–Elmer/Cetus) according to the manufacturer's instructions. The forward primer corresponded to nucleotides −2 to −19 of the Rh50A cDNA (see Fig. 2). Three different reverse primers were used and these had sequences corresponding to the reverse complement of nucleotides 98–117, 196–215 or 360–380 of the Rh50A cDNA (see Fig. 2). PCR products of the predicted sizes were obtained, purified by PAGE and used as templates for 40 cycles of asymmetric PCR (Gyllensten, 1989) with a 50-fold excess of the forward primer. The single-stranded asymmetric PCR products were purified by ultrafiltration (Centricon 30, Amicon, Beverly, MA, U.S.A.) before DNA sequencing.

Chromosomal localization of the Rh50A gene

Genomic DNA was isolated on an Applied Biosystems 340A DNA extractor according to the supplied protocols. Standard procedures were used for restriction enzyme digestion and Southern blotting of the genomic DNA and for hybridization with radiolabelled probes (Sambrook *et al.*, 1989). The 1.87 kb B1A *Eco*RI fragment was used as a probe and was radiolabelled by random oligonucleotide priming (Feinberg & Vogelstein, 1984). The culture and characterization of human–rodent somatic cell hybrids have been previously described (Spurr *et al.*, 1986).

RESULTS AND DISCUSSION

Cloning and sequencing of Rh_{50} glycoprotein partial cDNAs using PCR

We previously reported the *N*-terminal amino acid sequences of the D_{50} and $R6A_{45}$ polypeptides up to residues 39 and 30 respectively (Avent *et al.*, 1988*b*). Mixed-sequence degenerate oligonucleotides based on this amino acid sequence were used for PCR amplification, and products of the predicted 92 bp size were isolated. When these products were cloned and sequenced, three different DNA sequences were found in the region between the 3′ ends of the primers. Two of these (Rh50A and Rh50B, Fig. 1) had only 75% DNA sequence similarity but when translated gave the same amino acid sequence and this corresponded to that found for the Rh50 glycoproteins. The third PCR product (Rh50C, Fig. 1) had the same sequence as Rh50B but a single nucleotide change and a single nucleotide deletion introduced an in-frame stop codon at amino acid residue 13 which would result in a truncated amino acid sequence. This DNA sequence may result from PCR misincorporation errors or it may be the product of a pseudogene in the Rh50 gene family. These data suggest that there may be at least two genes for the Rh50 family of glycoproteins.

Isolation and sequencing of Rh50A cDNA clones

Oligonucleotides corresponding to the unique sequence be-

```
10            15            20
 I  V  L  E  I  A  M  I  V  L  F  G  L
TATAGTCCTGGAAATTGCCATGATTGTTTTATTTGGATTA    Rh50A
 :: :: ::::: :::::::::::::::  : :::::  :
CATTGTGCTGGAGATTGCCATGATTGTGCTGTTTGGCCTG    Rh50B
 I  V  L  E  I  A  M  I  V  L  F  G  L
::::::::::  ::::::::::::::::::::::::::::
CATTGTGCTGT-GATTGCCATGATTGTGCTGTTTGGCCTG    Rh50C

 I  V  L   *
10    12
```

Fig. 1. DNA sequence and translated amino acid sequence of cloned PCR products

The DNA sequence and translated amino acid sequence of the cloned Rh50A, Rh50B and Rh50C PCR products in the region between the 3′ ends of the primers. At least three independent clones of each type were sequenced.

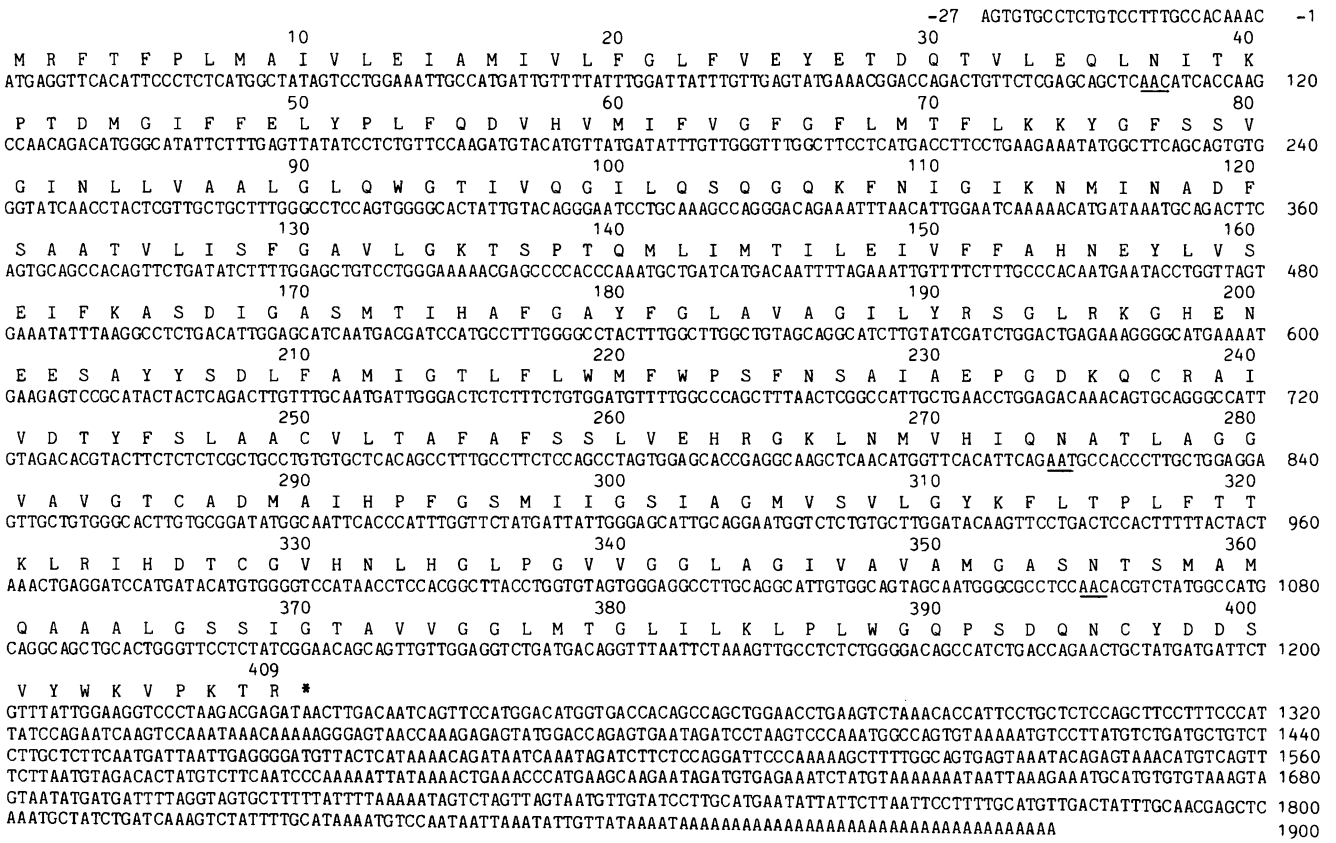

Fig. 2. DNA and predicted protein sequence of Rh50A cDNA

The DNA sequence shown is derived from three different cDNA clones. The F1 clone contains the sequence from nucleotide 22 to the 3′ poly(A) tail, the B1A clone from nucleotide 34 to the 3′ poly(A) tail, and the B2 clone contains the whole coding sequence and 3′ non-coding region, and 27 nucleotides of 5′ upstream sequence. The DNA sequence is numbered so that nucleotide 1 corresponds to the initiator methionine codon of the protein. Potential *N*-glycan addition sites are underlined.

tween the primers of the Rh50A and Rh50B PCR products were used to screen a number of human cDNA libraries in λgt10 and λgt11. No positive clones were obtained with the Rh50B-specific probe, but three positive clones were isolated with the Rh50A-specific probe, two (λB1 and λB2) from a human bone marrow cDNA library and one (λF1) from a human fetal liver cDNA library. Treatment of the λB1 clone with *Eco*RI yielded two inserts of 1.87 kb (B1A) and 0.48 kb (B1B). Subcloning and sequencing showed that the 1.87 kb B1A cDNA had a DNA sequence corresponding to nucleotides 34–1900 in Fig. 2. The translated sequence of B1B was not related to the Rh_{50} glycoprotein and its presence is probably the result of concatenation during construction of the library. The λF1 cDNA insert could not be excised from the vector with *Eco*RI. This insert (F1) was subcloned and sequenced after PCR amplification using λgt11 forward and reverse primers. F1 contained the sequence corresponding to nucleotides 22–1854 (Fig. 2) with an A_{13} tract at the 3′ end. The DNA sequence of the F1 insert extended 160 nucleotides further on the 5′ side of nucleotide 22 in Fig. 2. However, the translated protein sequence of the F1 cDNA on the 5′ side of nucleotide 22 did not correspond to the Rh_{50} glycoprotein sequence, unlike the sequence on the 3′ side. We conclude that this 5′ upstream sequence is an artefact, perhaps the result of concatenation with an unrelated cDNA. F1 had the same DNA sequence as B1A in the region where they overlapped, except at nucleotide 724 which was A in the F1 clone and G in the B1A cDNA, yielding Asn-242 or Asp-242 respectively in the translated amino acid sequence (Fig. 2). The translated amino acid sequence of the combined B1A and F1 cDNAs (which extends from nucleotide 22 in Fig. 2) corresponded almost exactly to the known *N*-terminal amino acid sequence of the Rh_{50} polypeptides from Met-8 to Thr-39 except that residues 37 and 38 are predicted to be Asn-Ile from the cDNA sequence but were assigned as Pro-X on amino acid sequencing of the isolated D_{50} glycoproteins (Avent *et al.*, 1988*b*). Asn-37 is likely to be *N*-glycosylated (see below) and the presence of the glycosylated residue may have given rise to the amino acid misassignment.

Although a 2.2 kb insert (B2) was obtained on *Eco*RI digestion of the λB2 clone, we were unsuccessful in several attempts to subclone this *Eco*RI fragment in two different plasmid vectors and using two different bacterial hosts. We confirmed that the λB2 insert was related to the B1A and F1 cDNAs by sequencing of the 3′ end of the insert in the intact λB2 clone, and this showed that the insert contained nucleotides 1624–1854 of the Rh50A cDNA with an A_{10} tract at the 3′ end. The 1680 nucleotides of the 5′ end of the B2 insert were subcloned and sequenced as part of a *Hin*dIII fragment of the intact λB2 phage DNA. The nucleotide sequence of the B2 cDNA was the same as that of the B1A cDNA where they overlapped, with G at nucleotide 724 predicting Asp-242, as was found in the B1A cDNA (Fig. 2). It is not clear whether the asparagine predicted at this point in the F1 cDNA is due to a polymorphism in the protein or results from a PCR misincorporation error during cloning.

The Rh50A *N*-terminal amino acid sequence derived from the B2 cDNA corresponded to that of the Rh_{50} glycoproteins except that amino acid 2 was predicted to be arginine instead of cysteine. Residues 37 and 38 were predicted to be Asn-Ile instead of the Pro-Xaa assigned from protein sequencing (as was also

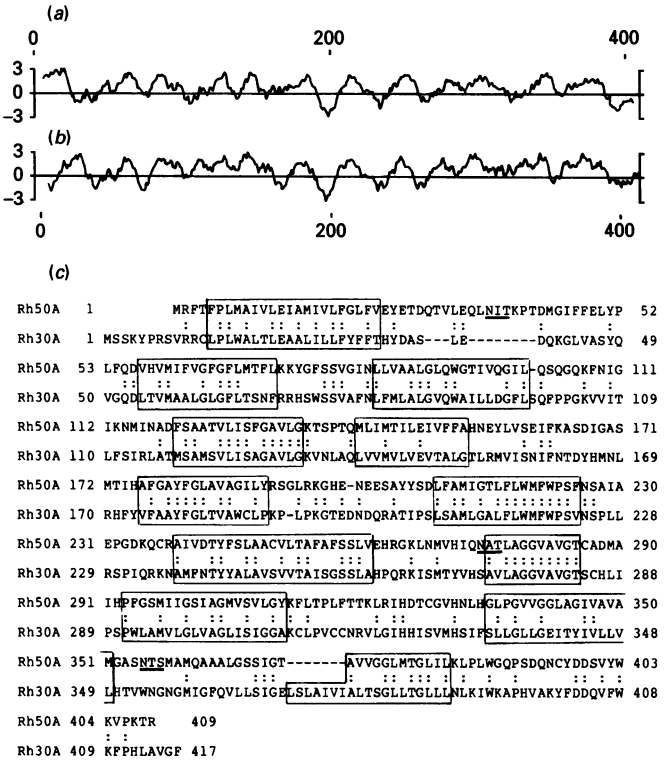

Fig. 3. Hydropathy analysis and amino acid sequence comparison of the Rh50A and Rh30A proteins

(*a*, *b*) Kyte–Doolittle hydropathy analysis of the Rh50A protein and comparison with that of the Rh30A protein. Hydropathy analysis was carried out using a window of nine amino acids. Hydrophobicity is shown on the vertical axis with the hydrophobic side of the plot having a positive value. The horizontal axis shows amino acid sequence residue. Analysis was done using the PEPPLOT program (Devereux *et al.*, 1984). (*a*) Rh50A protein. (*b*) Rh30A protein. (*c*) Protein sequence predicted by the Rh50A cDNA clones, and alignment with the Rh30A protein sequence. The 12 predicted hydrophobic membrane-spanning regions are boxed. Potential *N*-glycosylation sites are underlined. Alignment was done using the PRTALN program (Wilbur & Lipman, 1983).

found in the other cDNAs). The 5′ region of the B2 cDNA was found to contain additional sequence which extended from the 5′ side of that shown in Fig. 2. This sequence comprised an inverted repeat of nucleotides 1–122 in Fig. 2 and a poly(T) region at the extreme 5′ end, both of which were probably the result of cloning artefacts. The Rh50A cDNA sequence compiled from the B1A, B2 and F1 cDNAs is shown in Fig. 2.

In order to confirm that amino acid residue 2 was arginine in the Rh50A glycoprotein, a sense oligonucleotide primer was synthesized corresponding to nucleotides −19 to −2 (Fig. 2). This was employed in the PCR with one of several different reverse primers which matched the coding sequence in Fig. 2 at nucleotides 98–117, 196–215 and 360–380 with human reticulocyte cDNA as the template. Products of the predicted size were obtained with each of these primers and these were used as templates for asymmetric PCR to generate single-stranded DNA for DNA sequencing. All the DNA sequences obtained predicted the presence of arginine at amino acid residue 2 in the protein. This experiment also confirmed that at least nucleotides −19 to −2 of the 5′ upstream region of the BM4 clone form part of the Rh50A cDNA.

The combined coding sequence of the three Rh50A cDNA clones predicts an open reading frame of 409 amino acids from the methionine codon which corresponds to the *N*-terminus of the protein. The 3′ non-coding region terminates with a poly(A) tract. However, this sequence is also part of a polyadenylation signal AATAAA (nucleotides 1862–1867) so that the exact site of polyadenylation is not clear. Another copy of the polyadenylation signal AATAAA is located in the 3′ non-coding region at nucleotides 1339–1344.

The cDNA clones predict a protein (Fig. 2) which corresponds to the Rh_{50} glycoprotein sequence from residue 1 to 39 except that there is an arginine instead of a cysteine residue at position 2, and Asn-Ile instead of Pro-Xaa at residues 37 and 38. We attempted to establish the amino acid at residue 2 by further *N*-terminal amino acid sequencing using the Edman degradation on an additional three samples of the $R6A_{45}$ glycoprotein. One sample of $R6A_{45}$ immunoprecipitate obtained using monoclonal antibody BRIC 69 (Avent *et al.*, 1988*a*) gave arginine at amino acid residue 2. We were unable to assign the amino acid at residue 2 in the other two samples but no evidence for the presence of cysteine was found.

Structure of the Rh50A glycoprotein and relationship to Rh30A polypeptide

The 409-amino acid Rh50A glycoprotein sequence predicted from the cDNA contains three potential *N*-glycan addition sites at Asn-37, Asn-274 and Asn-355. The D_{50} and $R6A_{45}$ polypeptides are known to be highly glycosylated, giving diffuse bands on SDS/PAGE and showing a marked decrease in size to a sharp band of approx. 30 kDa after treatment with endoglycosidase-F preparations (Moore & Green, 1987; Avent *et al.*, 1988*a*). The deglycosylated $R6A_{45}$ protein was similar in size to the Rh_{30} polypeptide.

Hydropathy analysis suggests that the Rh50A glycoprotein may have 12 membrane-spanning domains (Fig. 3*a*) and shows the overall hydrophobic nature of the polypeptide. Alignment of the amino acid sequence of the Rh50A glycoprotein with that of the Rh30A polypeptide shows a high degree of sequence similarity between the proteins (Fig. 3*c*). The hydropathy plots for the two proteins are also very similar (Figs. 3*a* and 3*b*), with each being predicted to have the same number of membrane-spanning domains in similar locations. The regions of highest sequence similarity between the proteins tend to be located in these putative membrane-spanning domains (boxed in Fig. 3*c*), and it is clear that, as well as being sequence-related, the two proteins are very similar in overall structure and topology. A model for the Rh30A polypeptide has been suggested which predicts 12 membrane-traversing segments with the *N*-terminus located in the cytoplasm (Avent *et al.*, 1990; Cherif-Zahar *et al.*, 1990). On the reasonable assumption that the Rh50A glycoprotein has the same topology, two of the potential *N*-glycosylation sites (Asn-37 and Asn-355) are predicted to be on the extracellular side of the membrane and are thus likely to carry *N*-glycan chains, whereas the site at Asn-274 is located within the cytoplasm.

The Rh50A polypeptide does not contain any of the Cys-Leu-Pro (CLP) motifs found in the Rh30A polypeptide (Avent *et al.*, 1990). These motifs are all likely to be on the cytoplasmic surface of the membrane and are suggested sites for palmitoylation in the Rh30A polypeptide. Another possible palmitoylation site in the Rh30A polypeptide is at the CCNR motif (Avent *et al.*, 1990), but no similar motif occurs in the Rh50A polypeptide.

The Rh50A glycoprotein and Rh30A polypeptide clearly belong to the same protein family. No other related proteins were found in sequence similarity searches of the Genbank and EMBL DNA sequence databases and Swissprot protein sequence database. Insertions and deletions in the two proteins make them differ most in the regions that are predicted to form the first and last extracellular loops and which contain the glycosylation sites

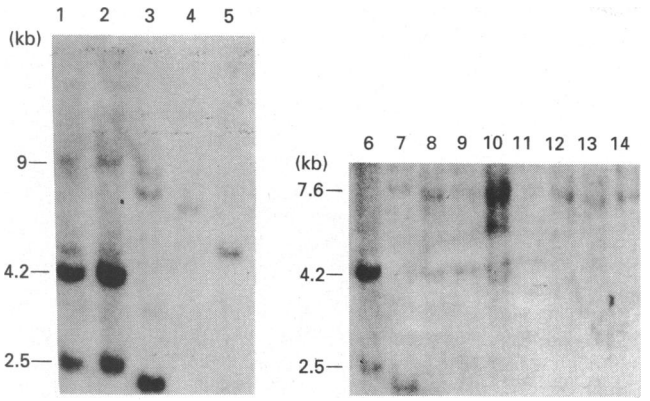

Fig. 4. Chromosomal mapping of the Rh50A glycoprotein gene

Genomic DNA samples were digested with *Eco*RI, transferred to nylon membranes and hybridized with radiolabelled cDNA probe. The autoradiographs show human DNA (tracks 1, 2 and 6), mouse DNA (tracks 3 and 7), rat DNA (track 4), hamster DNA (track 5) and DNA from seven somatic cell hybrids (tracks 8–14). The somatic cell hybrids are SIF4A31 (track 8), PLTI.S (track 9), MCP6 (track 10), 5647C122 (track 11), clone21E (track 12), FIR5R3 (track 13) and PCTBA1.8 (track 14). References to the hybrids are given in Table 1.

in the Rh50A glycoprotein. The differences in the last extracellular loop appear to result in the Rh30A protein having a longer final hydrophobic membrane-spanning segment. It is interesting that known sequence differences between the Rh30A and the Rh30B proteins also occur in the first extracellular loop (Avent *et al.*, 1990).

The Rh_{50} glycoproteins are coprecipitated with the Rh_{30} polypeptides on immunoprecipitation with anti-D and -R6A antibodies suggesting that the two proteins are associated with each other in the erythrocyte membrane (Moore & Green, 1987; Avent *et al.*, 1988*a*,*b*). This association taken together with the sequence similarity of the two proteins suggests that the two proteins are different subunits of an oligomeric complex in the erythrocyte membrane. Anti-Rh D, c and E recognize different sequence epitopes and there is no evidence for sequence differences in the Rh_{50} glycoproteins in different Rh genotype cells (see below). Since anti-Rh D, c and E each coprecipitate the Rh_{50} glycoprotein with different members of the Rh_{30} family of proteins (Moore *et al.*, 1982), it is extremely unlikely that the coprecipitation of the Rh_{50} and Rh_{30} polypeptides is due to reaction of the antibodies with common epitopes on both proteins. The deficiency of several different polypeptides in Rh_{null} erythrocytes has also led to the view that the Rh antigens are part of a multimeric complex in the erythrocyte membrane (Anstee & Tanner, 1988; Avent *et al.*, 1988*b*; Agre & Cartron, 1991). The functional role of this complex is not clear but it has been suggested that it may be associated with a membrane transport process (Anstee & Tanner, 1988), since Rh_{null} erythrocytes have an abnormal cation permeability (Ballas *et al.*, 1984). Consistent with this view, the Rh30A and Rh50A proteins have highly polytopic structures typically found in membrane transporters and channels. It is also of interest that two acidic residues present in the centre of the putative first and fifth membrane spans of the Rh50A glycoprotein (Glu-13 and Glu-146, Fig. 3*c*) are also conserved in the same locations in the Rh30A polypeptide (Glu-21 and Glu-146). This may be a further indication that the proteins are involved in the transfer of cationic material between the two surfaces of the membrane.

The $R6A_{45}$ and D_{50} glycoproteins

The $R6A_{45}$ and D_{50} glycoproteins are characterized by their different molecular-mass distributions in immunoprecipitates obtained using antibodies of the R6A type or anti-D respectively. This appears to result from differences in carbohydrate heterogeneity and no amino acid sequence differences have so far been detected between the two proteins. The two proteins are known to be associated with distinct polypeptides of the Rh30 group (the Rh30A and Rh30B polypeptides, the latter of which is involved in the expression of the D antigen, Avent *et al.*, 1990), and this raises the question of whether they are distinct gene products, for example the Rh50A and Rh50B glycoproteins, or result from differences in the processing of the *N*-glycan chains of a single polypeptide (Rh50A) when it is associated with one or other of the Rh30 polypeptides. Further studies on the cloning and expression of the Rh30B polypeptide and the putative glycoprotein corresponding to the Rh50B DNA sequence will be required to resolve this further.

Table 1. Chromosomal assignment of the Rh50A glycoprotein gene

Spaces indicate not tested; * = translocation, t = trace; R = reactivity with probe B1A; Y = reactive with probe B1A; N = not reactive with probe B1A.

Hybrid name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	R	Reference
MOG2	+	−	+	+	+	+	+		+		+	+		+	+	+	+	+	+	−	+	−	+	Y	Povey *et al.* (1980)
PLTI.S	−	+	−	+	−	+	+	+	−	−	+	+	−	+	−	−	+	+	−	+	+	+	−	Y	Martin *et al.* (1987)
CTP34B4	+	+	+	−	+	+	+	+	−	−	−	+	−	+	−	+	+	+	−	−	−	−	+	Y	Jones *et al.* (1976)
CTP41A2	−	+	−	−	−	+	+	−	−	−	−	−	−	+	−	−	+	−	−	−	−	−	+	Y	Jones *et al.* (1976)
SIF4A31	−	−	+	+	+	+	−	−	−	−	−	−	−	+	−	+	−	−	−	−	−	−	−	Y	Edwards *et al.* (1985)
TWIN19-D12	+	−	+	+	−	+	−	+	−	−	+	+	−	+	−	+	+	+	−	+	−	−	−	Y	Philips *et al.* (1985)
LSR34S49	+	+	+	−	+	+	−	−	−	+	−	−	+	+	+	+	−	+	+	+	+	−	+	Y	Spurr *et al.* (1986)
MCP6	−	−	−	−	−	*	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	Y	Goodfellow *et al.* (1982)
5647C122		−	−	+		*	+		−	−	+	−	+	−	+	+	+	−	−	+	+	−	+	N	Nagarajan *et al.* (1986)
DT1.2.4.	−	−	+	−	−	−	−	−	−	+	+	−	−	−	+	−	+	+	−	+	+	−	+	N	Swallow *et al.* (1970)
DUR4.3	−	−	+	−	+	−	−	−	−	+	+	+	+	+	+	−	+	+	−	+	+	+	+	N	Solomon *et al.* (1976)
FIR5R3	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	−	−	−	+	N	Hobart *et al.* (1981)
CLONE21E	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	N	Croce & Koprowski (1974)
PCTBA1.8	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	N	Bai *et al.* (1982)
FG10	−	+	−	−	+	−	−	+	−	+	−	−	−	−	+	−	−	+	−	−	+	−	+	N	Kielty *et al.* (1982)
3W4CL5	−	−	−	−	−	−	+	−	−	+	+	+	−	+	+	−	+	−	−	−	+	−	+	N	Nabholz *et al.* (1969)
SIR74ii	+	+	+	+	−	−	−	−	−	−	−	+	+	+	−	−	+	t	−	−	+	t	+	N	Whitehead *et al.* (1982)
F4Scl3C112	*	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	N	Heisterkamp *et al.* (1982)
HORL9X	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	N	Heisterkamp *et al.* (1982)

Chromosomal localization of the Rh50A gene

The 1.87 kb B1A cDNA insert was used to determine the chromosomal localization of the Rh50A protein gene using DNA from a panel of human–mouse somatic cell hybrids with appropriate human and rodent controls. The probe hybridized to human *Eco*RI-digested genomic DNA to give two strongly hybridizing fragments of 2.5 kb and 4.2 kb and a weakly hybridizing fragment of 9 kb (Fig. 4). In *Eco*RI-digested rodent control DNA, the probe detected 2.3 kb, 7.6 kb and 8.0 kb fragments in mouse (Fig. 4, tracks 3 and 7), a single 7.0 kb fragment in rat (Fig. 4, track 4) and a 5.4 kb fragment in hamster (Fig. 4, track 5). The human fragments were clearly distinguishable from the rodent fragments, allowing a number of somatic cell hybrids to be analysed. Using *Eco*RI-digested DNA from a panel of 19 human–mouse somatic cell hybrids, concordance was seen with hybrids containing human chromosome 6 (Table 1). The assignment was confirmed using two hybrids containing fragments of chromosome 6. MCP6 contains an X:6 translocation (Xqter-Xq13:6p21-qter) as the only human contribution on a mouse background (Goodfellow *et al.*, 1982). The hybrid 5647C122 contains a translocation involving chromosomes 6 and 17 [6pter-6p21:17p13-17qter (Nagarajan *et al.*, 1986]. MCP6 was positive (Fig. 4, track 10) and 5647C122 was negative (Fig. 4, track 11) confirming the assignment of the Rh50A locus to chromosome 6, between 6p21-qter.

Expression of Rh blood-group antigens and the Rh_{30} and Rh_{50} proteins

The Rh30A gene has been assigned to chromosome 1p (Cherif-Zahar *et al.*, 1991; MacGeoch *et al.*, 1992) and this correlates with the localization of the Rh blood-group locus to the same region of chromosome 1 by cytogenetic analysis (Bruns & Sherman, 1989). Therefore the major different polymorphic forms of the Rh antigens (Cc, D and Ee) are specified by genetic differences in the Rh30 rather than the Rh50 proteins. It is not clear at this stage whether the structures that give rise to the Rh antigens depend solely on the Rh30 proteins or also require the association of the Rh50 glycoproteins with the Rh30 polypeptides. The Rh antigen structures could involve portions of both proteins or the Rh50 glycoproteins could be required for the correct presentation of the antigens on the Rh30 proteins.

This work was supported in part by grants from the Medical Research Council, the Wellcome Trust and the Central Blood Laboratory Authority.

REFERENCES

Agre, P. & Cartron, J.-P. (1991) Blood **78**, 551–563
Anstee, D. J. & Tanner, M. J. A. (1988) Blood Rev. **2**, 115–120
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
Avent, N. D., Ridgwell, K., Tanner, M. J. A. & Anstee, D. J. (1990) Biochem. J. **271**, 821–825
Ballas, S. K., Clark, M. R., Mohandas, N., Colfer, H. F., Caswell, M. S., Bergren, M. O., Perkins, H. A. & Shohet, S. B. (1984) Blood **63**, 1046–1055
Bai, Y., Sheer, D., Hiorns, L., Knowles, R. W. & Tunnacliffe, A. (1982) Ann. Hum. Genet. **46**, 337–347
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., Beyruther, K., Salmon, Ch. & Cartron, J.-P. (1988) Blood **72**, 661–666
Bruns, G. A. P. & Sherman, S. L. (1989) Cytogenet. Cell Genet. **51**, 67–90
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
Cherif-Zahar, B., Mattei, M. G., le van Kim, C., Bailly, P., Cartron, J.-P. & Colin, Y. (1991) Hum. Genet. **86**, 398–400
Croce, C. M. & Koprowski, H. (1974) J. Exp. Med. **140**, 1221–1229
Devereux, J., Haeberli, P. & Smithies, P. (1984) Nucleic Acids Res. **12**, 387–395
Edwards, Y. H., Parker, M., Povey, S., West, L. F., Parrington, J. M. & Solomon, E. (1985) Ann. Hum. Genet. **49**, 101–109
Feinberg, A. P. & Vogelstein, B. (1984) Anal. Biochem. **137**, 266–267
Gahmberg, C. G. (1982) FEBS. Lett. **140**, 93–97
Goodfellow, P. N., Banting, G., Trowsdale, J., Chambers, S. & Solomon, E. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 1190–1194
Gyllensten, U. (1989) in PCR Technology, Principles and Applications for DNA Amplification (Ehrlich, H. A., ed.), pp. 45–60, Stockton Press, New York
Heisterkamp, N., Groffen, J., Stephenson, J. R., Spurr, N. K., Goodfellow, P., Solomon, E., Carritt, B. & Bodmer, W. F. (1982) Nature (London) **299**, 747–749
Hobart, M. J., Rabbitts, T. H., Goodfellow, P. N., Solomon, E., Chambers, S., Spurr, N. K. & Povey, S. (1981) Ann. Hum. Genet. **45**, 331–335
Jones, E. A., Goodfellow, P. N., Kennett, R. H. & Bodmer, W. F. (1976) Somat. Cell Genet. **2**, 483–496
Kielty, C. M., Povey, S. & Hopkinson, D. A. (1982) Ann. Hum. Genet. **46**, 307–327
MacGeoch, C., Avent, N. D., Ridgwell, K., Tanner, M. J. A. & Spurr, N. K. (1992) Cytogenet. Cell Genet. **59**, 261–263
Martin, D., Tucker, D. F., Gorman, P., Sheer, D., Spurr, N. K. & Trowsdale, J. (1987) Ann. Hum. Genet. **51**, 145–152
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
Nabholz, M., Miggiano, V. & Bodmer, W. (1969) Nature (London) **223**, 358–363
Nagarajan, L., Louie, E., Tsujimoto, Y., Ar-Rushdi, A., Huebner, K. & Croce, C. M. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2556–2560
Phillips, J. R., Shepard, E. A., Povey, S., Davis, M. B., Kelsey, G., Monteiro, M., West, L. F. & Cowell, F. (1985) Ann. Hum. Genet. **49**, 267–274
Povey, S., Jeremiah, S. J., Barker, R. F., Hopkinson, D. A., Robson, E. B., Cook, P. J. L., Solomon, E., Bobrow, M., Carritt, B. & Buckton, K. (1980) Ann. Hum. Genet. **43**, 241–248
Ridgwell, K., Roberts, S. J., Tanner, M. J. A. & Anstee, D. J. (1983) Biochem. J. **213**, 267–269
Saboori, A. M., Smith, B. & Agre, P. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 4042–4045
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
Solomon, E., Bobrow, M., Goodfellow, P. N., Bodmer, W. F., Swallow, D. M., Povey, S. & Noel, B. (1976) Somat. Cell Genet. **2**, 125–140
Spurr, N. K., Durbin, H., Sheer, D., Parkar, M., Bobrow, L. & Bodmer, W. F. (1986) Int. J. Cancer **38**, 631–636
Swallow, D. M., Solomon, E. & Pajunen, L. (1970) Cytogenet. Cell Genet. **18**, 136–148
Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W. F., Povey, S. & Fey, G. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 5012–5025
Wilbur, W. J. & Lipman, D. J. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 726–730
Zimmermann, J., Voss, H., Schwager, C., Stegemann, J., Erfle, H., Stucky, K., Kristensen, T. & Ansorge, W. (1990) Nucleic Acids Res. **18**, 1067–1068

Received 9 January 1992/19 March 1992; accepted 31 March 1992