*Proc. Natl. Acad. Sci. USA*
Vol. 87, pp. 6243–6247, August 1990
Medical Sciences

# Molecular cloning and protein structure of a human blood group Rh polypeptide

(erythrocyte membrane/Rh antigen/cDNA/RNA blot-hybridization analysis)

BAYA CHÉRIF-ZAHAR, CHRISTIAN BLOY, CAROLINE LE VAN KIM, DOMINIQUE BLANCHARD, PASCAL BAILLY, PATRICIA HERMAND, CHARLES SALMON, JEAN-PIERRE CARTRON*, AND YVES COLIN

Institut National de la Santé et de la Recherche Médicale Unité U76, Institut National de Transfusion Sanguine, 6 rue Alexandre Cabanel, 75015 Paris, France

*Communicated by Victor A. McKusick, May 29, 1990*

ABSTRACT **cDNA clones encoding a human blood group Rh polypeptide were isolated from a human bone marrow cDNA library by using a polymerase chain reaction-amplified DNA fragment encoding the known common N-terminal region of the Rh proteins. The entire primary structure of the Rh polypeptide has been deduced from the nucleotide sequence of a 1384-base-pair-long cDNA clone. Translation of the open reading frame indicates that the Rh protein is composed of 417 amino acids, including the initiator methionine, which is removed in the mature protein, lacks a cleavable N-terminal sequence, and has no consensus site for potential N-glycosylation. The predicted molecular mass of the protein is 45,500, while that estimated for the Rh protein analyzed in $NaDodSO_4$/polyacrylamide gels is in the range of 30,000–32,000. These findings suggest either that the hydrophobic Rh protein behaves abnormally on $NaDodSO_4$ gels or that the Rh mRNA may encode a precursor protein, which is further matured by a proteolytic cleavage of the C-terminal region of the polypeptide. Hydropathy analysis and secondary structure predictions suggest the presence of 13 membrane-spanning domains, indicating that the Rh polypeptide is highly hydrophobic and deeply buried within the phospholipid bilayer. In RNA blot-hybridization (Northern) analysis, the Rh cDNA probe detects a major 1.7-kilobase and a minor 3.5-kilobase mRNA species in adult erythroblasts, fetal liver, and erythroid (K562, HEL) and megakaryocytic (MEG01) leukemic cell lines, but not in adult liver and kidney tissues or lymphoid (Jurkat) and promyelocytic (HL60) cell lines. These results suggest that the expression of the Rh gene(s) might be restricted to tissues or cell lines expressing erythroid characters.**

The Rh antigens are of wide interest in clinical medicine, since they are involved in hemolytic reactions of immune origin due to transfusion incompatibilities, hemolytic disease of the newborn, and autoimmune diseases (for reviews, see refs. 1–3). Moreover, although the physiological role of these antigens is still unclear, they are probably important structures for erythrocyte membrane integrity, as evidenced by the mild hemolytic anemia (4, 5), abnormalities in cation permeability (6, 7), and membrane lipid asymmetry (8) that accompany the lack or severe deficiency of Rh antigens. Genetic analyses have demonstrated that a polymorphic Rh haplotype producing the Rh D antigen, as well as antigens of the Cc and Ee series, is inherited *en bloc* from one generation to another (1, 2, 9), but whether these antigens are the products of one gene or several closely linked genes has not been clarified (review in ref. 1). Our knowledge of the Rh biochemistry is largely incomplete. It is known, however, that Rh antigens are sulfhydryl-containing polypeptides (10, 11) carried by 30- to 32-kDa integral membrane proteins (12, 13) that are not significantly glycosylated or phosphorylated (14); the Rh proteins, also, do represent major fatty acid-acylated proteins (15) associated with the membrane skeleton (16–18). Several lines of evidence (19–21), including two-dimensional iodopeptide map analysis, indicated that the Rh D, c, and E antigens are most likely carried by three distinct but homologous membrane proteins (56) and that all share a common N-terminal protein sequence (19, 22, 23).

As a further step to characterize better the structure and function of the Rh antigens at a molecular level, this report describes the isolation, sequence analysis, and tissue-specific expression of a mRNA encoding one of the human Rh polypeptides.†

## MATERIALS AND METHODS

**Materials.** Restriction enzymes, bacterial alkaline phosphatase, and pUC vectors were from Appligene (Strasbourg, France). T4 DNA ligase and T4 polynucleotide kinase were from Biolabs (Northbrook, IL), and radiolabeled nucleotides were from Amersham. pUC sequencing kits were from Pharmacia. Avian myeloblastosis virus (AMV) reverse transcriptase and *Thermus aquaticus* polymerase (*Taq* polymerase) were from Perkin–Elmer.

**Polymerase Chain Reaction (PCR) Amplification of mRNA and Genomic Sequences.** Primers used for PCR amplification were synthesized on a Milligen Biosearch 8700 DNA synthesizer and purified on a 20% acrylamide/urea gel. The nucleotide sequence of these primers was deduced from the previously determined N-terminal protein sequence (Ser[1]-Ser-Lys-Tyr-Pro-Arg-Ser-Val-Arg-Arg[10]-Xaa-Leu-Pro-Leu-Trp-Ala-Leu-Thr-Leu-Glu[20]-Ala-Ala-Leu-Ile-Leu-Leu-Phe-Tyr-Phe-Phe[30]-Thr-His), which is common to all Rh polypeptides (19, 22, 23).

Primer A (14-mer) was specific for amino acids 3–7 of the mature protein (Lys-Tyr-Pro-Arg-Ser) and was a mixture of 16 oligonucleotides, 5′-AARTACCCIMGIWS (coding strand, where R = G or A, M = A or C, W = A or T, and S = G or C), in which deoxyinosine (I) was incorporated in positions where codon degeneracy exceeded 2. Primer B (18-mer) was specific for amino-acids 32–27 of the mature protein (His-Thr-Phe-Phe-Tyr-Phe) and was a mixture of the following 16 oligonucleotides: 5′-GTGGGTRAARAARTARAA (noncoding strand). First cDNA strands were synthesized as described (24) after hybridization of 70 ng of primer B with 2.5 μg of poly(A)$^+$ RNA from splenic erythroblasts of an adult β-thalassemic patient, followed by elongation with 1 mM of each dNTP in the presence of 7 units of avian myeloblastosis virus reverse transcriptase. One-third

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PCR, polymerase chain reaction.

*To whom reprint requests should be addressed.

†The nucleotide sequence reported in this paper has been deposited in the GenBank database (accession no. M34015).

of the cDNA products or 1 μg of human blood leukocyte genomic DNA was enzymatically amplified between primers A and B (3 μg each) by using the *Taq* polymerase (25). The reaction was carried out in a Perkin–Elmer thermal DNA cycler for 35 cycles, each cycle comprising denaturation at 94°C for 1 min, primer annealing at 51°C for 2 min, and chain extension at 72°C for 2 min. Amplification products of the expected sizes were purified on agarose gels and subcloned in pUC18 vectors.

**cDNA Library Screening.** Purified PCR products specific for the N-terminal region of the Rh polypeptide were labeled with [$\alpha$-$^{32}$P]dCTP by the random priming method (Boehringer Mannheim) to a specific activity of $10^9$ cpm/μg and were used as probes to screen a human bone marrow λgt11 cDNA library (Clontech). About $1.2 \times 10^6$ plaques were transferred to Hybond N filter (Amersham) and screened as recommended by the supplier. The *Eco*RI inserts from positive clones were subcloned into pUC18 vectors.

**DNA Sequencing.** Inserts from recombinant pUC18 vectors were sequenced on both strands by the dideoxy chain-termination method (26) with a Pharmacia T7 sequencing kit; universal primers or specific oligonucleotides were used as internal primers.

**RNA (Northern) Blot-Hybridization Analysis.** Total RNAs or poly(A)$^+$ RNAs from cultured cell lines or from milligram quantities of fresh tissues were prepared by published methods (27, 28), resolved by electrophoresis on 6% (wt/vol) formaldehyde/1% (wt/vol) agarose gels, and then transferred onto Hybond N nylon filters. Hybridization with nucleic acid probes and stringent washes were performed as described by Thomas (29).

## RESULTS AND DISCUSSION

**PCR Amplification of the Coding Sequence for the Human N-Terminal Rh Polypeptide.** Several highly degenerate oligonucleotides deduced from the known 32 amino acids common to the N-terminal region of the human Rh D, c, and E polypeptides (19, 22, 23) were synthesized and used as probes for screening three human cDNA libraries (from fetal liver, reticulocytes, and HEL cells) under a variety of hybridization conditions. Since these attempts to isolate an Rh cDNA clone were unsuccessful, a PCR approach using deoxyinosine-containing primers (30) was developed.

Two primers specific for amino acids 3–7 (primer A) and 32–27 (primer B), respectively, were used to enzymatically amplify genomic or mRNA sequences coding for the N-terminal region of the mature Rh polypeptides. The use of deoxyinosine at the highly degenerate positions in primer A (see *Materials and Methods*) and of a stringent annealing temperature (51°C) during the PCR reaction were critical for successful amplification of specific products. Indeed, only under these experimental conditions was a clear and single amplification product of the expected size (90 bp) obtained when genomic DNA or cDNA derived from erythroblast mRNA was used as templates. In contrast, no sequence could be amplified between primers A and B with cDNA derived from mRNA of adult liver. Sequence analysis of the 90-base-pair (bp) DNA fragment amplified from either genomic or mRNA templates revealed that it encodes a protein sequence identical to the published N-terminal primary structure of the Rh polypeptide (19, 22, 23) and that a cysteine residue was present at the previously undetermined position 11.

**Isolation and Nucleotide Sequence of the Rh cDNA Clone.** The 90-bp PCR fragment was used as probe to screen a human bone marrow cDNA library. Of $1.2 \times 10^6$ recombinant λgt11 phages, 5 positive clones were obtained, of which the clone RhIXb carries the largest insert (1384 bp). The nucleotide sequence of this insert is presented in Fig. 1. It contains 1281 bases of open reading frame and 103 bases of 3′ untranslated sequence, which combined do not extend down to the end of the mRNA, since there is no poly(A) tail or polyadenylylation signal. We assumed that the ATG codon at positions 31–33 is the translation initiation site because (*i*) it precedes the N-terminal amino acid sequence determined by Edman degradation of the Rh polypeptides (19, 22, 23), (*ii*) it closely conforms to the Kozak consensus translation initiation sequence (31), and (*iii*) the amino acids potentially encoded by the first nucleotides of this clone do not match with a typical N-terminal cleavable peptide signal (32), although a peptide segment between positions 13 and 33 exhibits a strong hydrophobic character suitable for a transmembrane domain. Therefore, starting from this first methionine residue, the polypeptide encoded by the RhIXb cDNA clone would be composed of a single polypeptide chain of 417 amino acids, which is probably processed cotranslationally by limited proteolysis to remove the initiator methionine, as it occurs in many proteins (33), yielding a 416-amino-acid-long mature protein. Further evidence that the clone RhIXb was encoding the Rh protein was provided by a comparison of the deduced protein sequence (Fig. 1) with the directly sequenced peptide fragments 38–44, 194–199, 203–208, and 286–292, which were purified by reverse-phase HPLC after chemical or enzyme degradation of the 30- to 32-kDa Rh protein (C.B., unpublished data). As the major transcript detected on Northern blot is 1.7 kb (see below), the RhIXb clone represents ≈80% of the mRNA and includes only 30 nucleotides of the 5′ untranslated region.

**Amino Acid Sequence and Membrane Organization of the Rh Protein.** The predicted translation product of the RhIXb clone is a basic protein of isoelectric point 9.4 and of relative molecular mass of ≈45,500. However, the estimated molecular mass of the immunoprecipitated Rh protein analyzed in $NaDodSO_4$/polyacrylamide gels is in the range of 30,000–32,000 (12, 13). Since nucleotide sequencing of the RhIXb clone was performed twice on both DNA strands, this discrepancy may arise from (*i*) the presence of an unspliced intron in the cDNA sequence leading to a wrong C-terminal sequence prediction, (*ii*) the abnormal behavior of the Rh protein in $NaDodSO_4$ gels, or (*iii*) the proteolytic cleavage of the mature Rh protein. To rule out possibility *i*, a PCR analysis using two oligonucleotide primers (nucleotides 872–888 and 1162–1178) flanking the nucleotide region encoding the authentic amino acids 286–292, deduced from protein sequencing, was performed. Since the same expected PCR product of 294 bp was amplified, either from the cDNA clone RhIXb or from the poly(A)$^+$ RNA preparation isolated from human erythroblasts (not shown), it is concluded that there is no retained intron in the RhIXb clone. Possibility *ii* would depend upon the hydrodynamic properties of the Rh protein–$NaDodSO_4$ complexes. These complexes take usually a rod-like configuration (34), but for reasons presently unclear (amino acid composition or aggregation state of the Rh protein in solution), this structure might be less extended than expected, therefore modifying the proportional relationship between the polypeptide chain length and its molecular mass (34). Alternatively, according to possibility *iii*, the RhIXb cDNA clone may encode a protein that is further matured by a proteolytic cleavage of its C-terminal region. This C-terminal region of the protein carries several stretches of apolar amino acids; therefore, it is possible that the postulated proteolytic product may remain bound to the membrane which raises the question of its potential role and antigenic properties.

That the Rh protein may run anomalously on gels cannot be ruled out at the moment, but because of the large size difference of the predicted protein, a posttranslational cleavage of the Rh protein may appear more likely. Further studies based on carboxyl-terminal sequencing of the Rh protein and

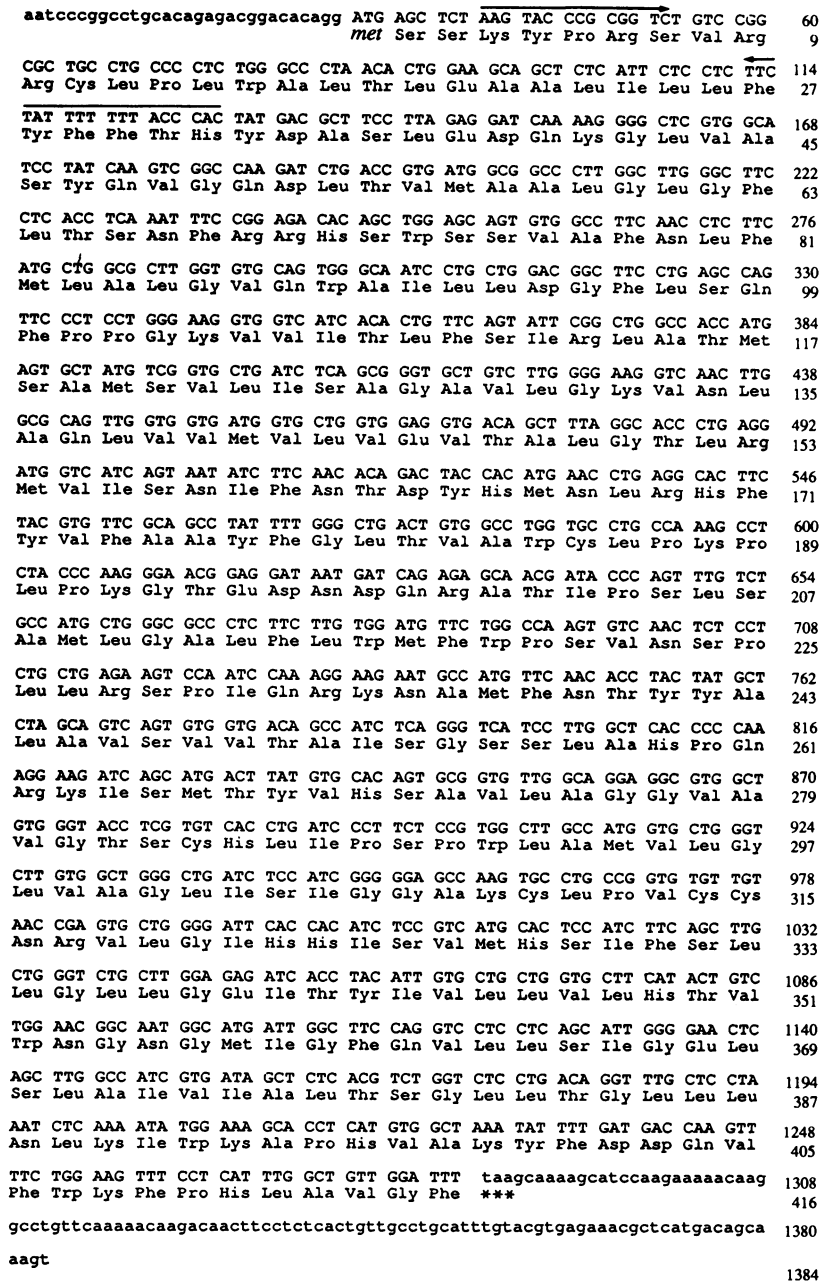


FIG. 1. Nucleotide sequence of the human Rh cDNA clone and predicted amino acid sequence of the Rh protein. Nucleotides encoding the open reading frame are given in capital letters, and those flanking the coding region are in lower case letters. The stop codon is indicated by asterisks. Arrows show the regions chosen for enzymatic amplification of mRNA and genomic sequences specific for the known N-terminal sequence of the Rh protein(s) (see text). Amino acids (three-letter code) are numbered with respect to the N-terminal amino acid residue determined by protein sequencing (19, 22, 23).

on transfection of eukaryotic cells by the Rh cDNA cloned in expression vectors, followed by an analysis of the mature recombinant Rh polypeptide present at the cell surface, should help to address these issues. Examination of the protein structure shown in Fig. 1 indicates that the Rh protein carries six cysteine residues, a finding that correlates well with the observation that functional sulfhydryl groups are required for Rh antigenic activity (10). However, which Rh antigen (D, C, c, E, e, etc. . . .) is encoded by this protein cannot be determined yet, since the Rh antibodies recognize only the native membrane proteins. The protein also lacks the canonical sequence for N-glycosylation to asparagine residues (35), again in agreement with previous investigations indicating that the Rh protein is not significantly glycosylated (14).

The identification in the Rh protein of membrane-associated $\alpha$-helices was carried out by analysis of their hydrophobicity according to the method of Engelman *et al.* (36) using a scanning window of 20 amino acids (Fig. 2 *Upper*). This analysis predicts that the Rh protein contains as many as 13 transmembrane domains that are regularly spaced on the polypeptide chain. A hypothetical model of the protein topology based on the hydrophobicity plot is presented in Fig. 2 *Lower*. The overall hydrophobicity of the Rh protein is striking, and the hydrophilic loops that link the membrane-spanning segments are short, not exceeding 20 amino acids. This suggests that this protein is deeply buried in the phospholipid bilayer, a fact that may explain its relative insensitivity to proteolytic degradation at the cell surface, as previously noted in several reports (1, 14, 19, 37). Since the Rh

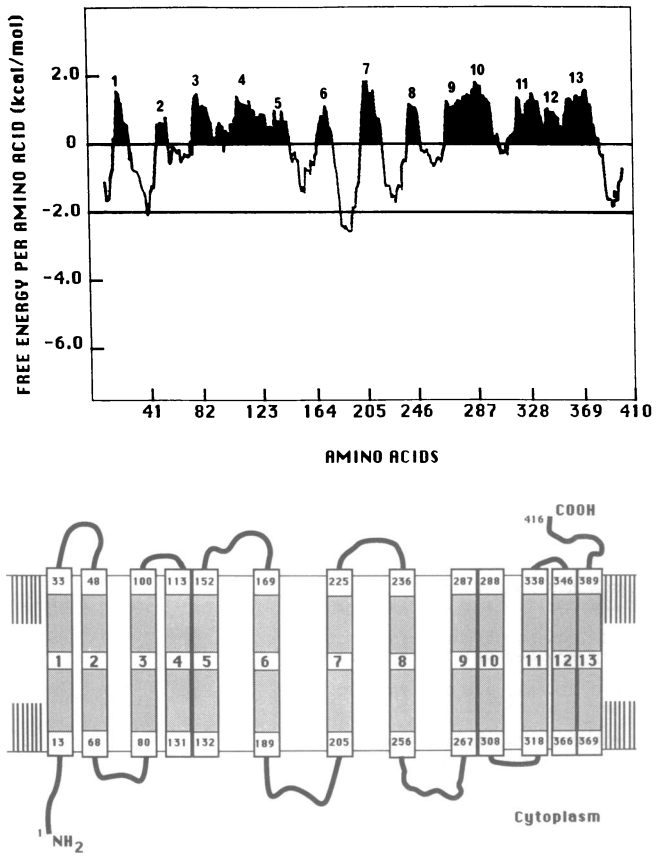


FIG. 2. Hydropathy plot of the Rh protein and hypothetical secondary structure model. (*Upper*) Average hydrophobicity values for spans of 20 residues were determined as described by Engelman *et al.* (36). Positive values are filled in black and corresponding postulated membrane-penetrating segments are numbered from 1 to 13. (*Lower*) Proposed transmembrane topology of Rh protein based on predictions of the hydropathy profile and on the intracellular orientation of the N terminus predicted from the absence of a cleavable signal peptide (see Fig. 1). Putative transmembrane $\alpha$-helices are shaded.

protein is devoid of a cleavable signal peptide and as the first hydrophobic segment conforms to the charge-difference rule proposed by Hartmann *et al.* (38), the N terminus of the protein is predicted on the cytoplasmic side of the membrane. The second membrane helix would then be inserted in orientation opposite to the first, and so on (38, 39), so that the C terminus of the Rh protein should be exposed at the cell surface. However, the deduced location of the C terminus to the extracellular side of the membrane and the proposed membrane orientation of the polar loops connecting the putative helices should be further established by direct biochemical investigations such as carboxypeptidase digestion experiments and immunochemical analysis with anti-peptide antibodies, respectively. It should provide also information to define which region of the protein interacts directly or indirectly with the membrane skeleton (16–18).

The pairs of peaks 4 and 5, 9 and 10, and 12 and 13 on the hydropathy plot (Fig. 2 *Upper*) define broad regions of largely apolar residues that may not cross the membrane or must do it twice. They might be arranged as coupled anti-parallel helices with $\beta$-turn-forming amino acids (proline, glycine, and serine) and charged residues located at the interface between the phospholipid bilayer and the cytoplasm or the outside of the cell, as proposed in Fig. 2 *Lower*. Construction of helical "wheel diagrams" (40) indicates also that, with the possible exception of helices 3 and 5, none of the helices exhibit an amphipathic character due to the presence of charged residues. In addition, the packing of the $\alpha$-helices in the membrane is largely unknown.

The transmembrane topology of the Rh protein resembles strongly that of several multispanning membrane proteins that are associated with transport (41, 42) or channel (43–46) functions. In contrast to these proteins, however, the Rh polypeptide has no significant cytoplasmic domains, such as those occasionally important for biological function or signal transduction (47). In addition, a search in the GenBank data base and National Biomedical Research Foundation Protein Identification Resource files did not reveal any homology with an already known gene or protein structure, including the $Na^+/H^+$-antiporter or the glucose transporter, which are located close to the Rh locus on chromosome 1 (48–50). Whether the Rh protein may have an as-yet-unrecognized transport function or may derive from such a transport protein remains an open question. However, recent studies based on biochemical and physical similarities between the Rh polypeptides and the aminophospholipid transporter of human erythrocytes suggest that Rh proteins might be responsible for the transbilayer movement of phosphatidylserine in the membrane (A. J. Schroit, C.B., J. Connor, and J.-P.C., unpublished data).

**Northern Blot Analysis of Rh Expression.** On Northern blot analysis (Fig. 3), the RhIXb cDNA clone detected a major 1.7-kilobase (kb) mRNA species in adult spleen erythroblasts; this mRNA is present in much lower amounts in K562 erythroleukemic and MEG01 megakaryocytic leukemic cell lines. In addition to the 1.7-kb signal, a faint band at 3.5 kb was observed, but whether it corresponds to another Rh transcript or to the expression of a Rh-related gene cannot be determined yet. The 1.7- and 3.5-kb mRNA species also were detected in human fetal liver, which contains a large number of erythroid precursors, and in the HEL (erythroleukemic) cell line (not shown) but were absent from adult kidney and liver tissues and from Jurkat (T-lymphoblastic) and HL60 (promyelocytic) cell lines. These results strongly suggest that the expression of Rh is restricted to tissues or cell lines exhibiting erythroid features. The presence of the Rh mRNA in a megakaryocytic cell line (MEG01) does not contradict

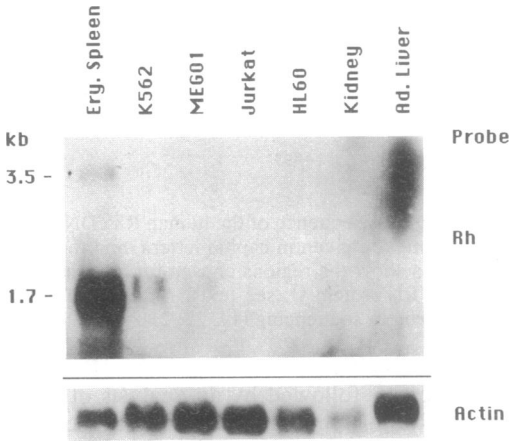


FIG. 3. Northern blot analysis and tissue specificity of the Rh mRNA. One microgram of poly(A)$^+$ RNA from adult spleen erythroblasts or 15 $\mu$g of total RNA from the erythroleukemic (K562), megakaryocytic (MEG01), T-lymphoblastic (Jurkat), and promyelocytic (HL60) cell lines and from adult kidney and liver tissues were resolved on a denaturating agarose gel, transferred to Hybond N filters, and hybridized with the $^{32}$P-labeled Rh cDNA probe and autoradiographed for 72 hr at −80°C. On longer exposure, the faint signal at 3.5 kb seen with erythroblast messengers was also detectable in K562 and MEG01 mRNA preparations. The actin probe at the bottom was used as control of the integrity of RNA samples and constancy of loading.

this proposal, since it has been shown recently that some trans-acting factors (NF-E1, NF-E2, and NF-E6) involved in the regulation of tissue-specific gene expression are found in nuclear extracts from both erythroid and megakaryocytic cells (refs. 51 and 52; Y.C., V. Joulin, C.L.V.K., P.-H. Roméo, and J.-P.C., unpublished data). Such findings suggest that these cells may derive from a common bipotent stem cell (52). Therefore, it is plausible that Rh expression appears very early during hematopoietic differentiation, even before commitment to the erythroid and megakaryocytic lineages, and that the level of expression and the antigen density on the pluripotent cells might be very low. After commitment, Rh antigens might behave as differentiation markers with progressively increased expression during erythroid differentiation (53). Indeed, recent reports indicate that these antigens are either absent or present in very low quantity on human platelets (54) but can be detected by flow cytometry analysis on mature (CFU-E, colony-forming unit-erythroid) but not on more immature (BFU-E, burst-forming unit-erythroid) erythroid progenitors (55). The availability of the Rh probe should help to reinvestigate the biosynthesis and expression of the Rh antigens during hematopoietic differentiation to sort out the gene(s) organization of the complex Rh locus, to analyze the potential function of the Rh protein(s), and to clarify the molecular basis of the defects occurring in Rh-deficient individuals.

We thank J. L. Guillaume and A. D. Strosberg (Institut Pasteur, Paris) for protein sequence analysis, C. Hattab and S. Chrétien (Institut National de Transfusion Sanguine, Paris) for the supply of synthetic oligonucleotides, and F. Kruger for technical assistance. We thank also J. Conboy (Cancer Research Institute, San Francisco) and J. P. Rosa (INSERM U150, Paris) for the generous gifts of the human reticulocytes and HEL cDNA libraries, respectively. Computer calculations were performed at the CITI2 (Paris). The authors acknowledge the support of North Atlantic Treaty Organization (Grant 0556/88), of the Institut National de la Santé et de la Recherche Médicale, and the Caisse Nationale d'Assurances Maladie des Travailleurs Salariés.

1. Race, R. R. & Sanger, R. (1975) *Blood Groups in Man* (Blackwell, Oxford), 6th Ed.
2. Issitt, P. (1985) *Applied Blood Group Serology* (Montgomery Scientific, Miami), 3rd Ed.
3. Mollison, P. L., Engelfriet, C. P. & Contreras, M. (1987) *Blood Transfusion in Clinical Medecine* (Blackwell, Oxford), 8th Ed.
4. Sturgeon, P. (1970) *Blood* **36,** 310–320.
5. Nash, R. & Shojania, A. M. (1987) *Am. J. Hematol.* **24,** 267–275.
6. Lauf, P. K. & Joiner, C. H. (1976) *Blood* **48,** 457–468.
7. 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.
8. Kuypers, P., 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.
9. Issitt, P. (1989) *Transfusion Med. Rev.* **3,** 1–12.
10. Green, F. A. (1967) *Immunochemistry* **4,** 247–257.
11. Green, F. A. (1983) *Mol. Immunol.* **20,** 769–775.
12. Moore, S., Woodrow, C. F. & McClelland, D. B. F. (1982) *Nature (London)* **295,** 529–531.
13. Gahmberg, C. G. (1982) *FEBS Lett.* **140,** 93–97.
14. Gahmberg, C. G. (1983) *EMBO J.* **2,** 223–227.
15. de Vetten, M. P. & Agre, P. (1988) *J. Biol. Chem.* **263,** 18193–18196.
16. Gahmberg, C. G. & Karhi, K. K. (1984) *J. Immunol.* **133,** 334–337.
17. Ridgwell, K., Tanner, M. J. A. & Anstee, D. J. (1984) *FEBS Lett.* **174,** 7–10.
18. Paradis, G., Bazin, R. & Lemieux, R. (1986) *J. Immunol.* **137,** 240–244.
19. Bloy, C., Blanchard, D., Dahr, W., Beyreuther, K., Salmon, C. & Cartron, J. P. (1988) *Blood* **72,** 661–666.
20. Hughes-Jones, N. C., Bloy, C., Gorick, B., Blanchard, D., Doinel, C., Rouger, P. & Cartron, J. P. (1988) *Mol. Immunol.* **25,** 931–936.
21. Suyama, K. & Goldstein, J. (1990) *Blood* **75,** 255–260.
22. Saboori, A. M., Smith, B. & Agre, P. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4042–4045.
23. Avent, N. D., Ridgwell, K., Mawby, W. J., Tanner, M. J. A., Anstee, D. J. & Kumpel, B. (1988) *Biochem. J.* **256,** 1043–1046.
24. Colin, Y., Le Van Kim, C., Tsapis, A., Clerget, M., d'Auriol, L., London, J., Galibert, F. & Cartron, J. P. (1989) *J. Biol. Chem.* **264,** 3773–3780.
25. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Nature (London)* **324,** 163–166.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
28. Favaloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65,** 718–749.
29. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
30. Knoth, K., Roberds, S., Poteet, C. & Tankun, M. (1988) *Nucleic Acids Res.* **16,** 10932.
31. Kozak, M. (1987) *Nucleic Acids Res.* **87,** 8125–8131.
32. von Heijne, G. (1984) *EMBO J.* **3,** 2315–2318.
33. Arfin, S. M. & Bradshaw, R. A. (1988) *Biochemistry* **27,** 7979–7984.
34. Reynolds, J. A. & Tanford, C. (1970) *J. Biol. Chem.* **245,** 5161–5165.
35. Marshall, R. D. (1982) *Biochem. Soc. Symp.* **40,** 17–56.
36. Engelman, D., Steitz, T. & Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem.* **15,** 321–353.
37. Agre, P., Saboori, A. M., Asimos, A. & Smith, B. L. (1987) *J. Biol. Chem.* **262,** 17497–17503.
38. Hartmann, E., Rapoport, T. M. & Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5786–5790.
39. Wessels, P. H. & Spiess, M. (1988) *Cell* **55,** 61–70.
40. Schiffer, M. & Edmundson, A. B. (1967) *Biophys. J.* **7,** 121–131.
41. Mueckler, M., Caruso, C., Baldwin, S. A., Panico, M., Blench, P., Morris, H. R., Allard, N. J., Lienhard, G. E. & Lodish, H. F. (1985) *Science* **229,** 941–945.
42. Hediger, M. A., Coady, M. J., Ikeda, T. S. & Wright, E. M. (1987) *Nature (London)* **330,** 379–381.
43. Kopito, R. & Lodish, H. F. (1985) *Nature (London)* **316,** 234–238.
44. Shull, G. E., Schwartz, A. & Lingrel, J. B. (1985) *Nature (London)* **316,** 691–695.
45. Caterall, W. A. (1988) *Science* **242,** 50–61.
46. Sardet, C., Franchi, A. & Pouyssegur, J. (1989) *Cell* **56,** 271–280.
47. Strader, C. D., Sigal, I. S. & Dixon, R. A. (1989) *FASEB J.* **3,** 1825–1832.
48. Marsh, W. L., Changati, R. S. K., Gardner, F. H., Mayer, K., Nowell, P. C. & German, J. (1974) *Science* **183,** 966–968.
49. Mattei, M. G., Sardet, C., Franchi, A. & Pouyssegur, J. (1988) *Cytogenet. Cell Genet.* **48,** 6–8.
50. Shows, T. B., Eddy, R. L., Byers, M. G., Fukushima, Y., DeHaven, C. R., Murray, J. C. & Bell, G. I. (1987) *Diabetes* **36,** 546–549.
51. Roméo, P.-H., Prandini, M. H., Joulin, V., Mignotte, V., Prenant, M., Vainchenker, W., Marguerie, G. & Uzan, G. (1990) *Nature (London)* **344,** 447–449.
52. Martin, D., Zon, L., Mutter, G. & Orkin, S. H. (1990) *Nature (London)* **344,** 444–447.
53. Rearden, A. & Masouredis, S. P. (1977) *Blood* **50,** 981–986.
54. Dunstan, R. A., Simpson, M. B. & Rosse, W. F. (1984) *Transfusion* **24,** 243–246.
55. Falkenburg, J. H. F., Fibbe, W. E., van der Vaart-Duikerken, N., Nichols, M. E., Rubinstein, P. & Jansen, J. (1985) *Blood* **66,** 660–663.
56. Blanchard, D., Bloy, C., Hermand, P., Cartron, J. P., Saboori, A. M., Smith, B. & Agre, P. (1988) *Blood* **72,** 1424–1427.