

Molecular Cloning of RhD cDNA Derived From a Gene Present in RhD-Positive, But Not RhD-Negative Individuals

By Miguel A. Arce, E. Scott Thompson, Steven Wagner, Katharine E. Coyne, Barbara A. Ferdman, and Douglas M. Lublin

The Rh blood group system plays a major role in immune and nonimmune hemolytic states. Although an Rh cDNA has been previously cloned, there is no information on which Rh antigenic protein it encodes. Using polymerase chain reaction (PCR) amplification, we have identified this original Rh clone, here designated Rh21, and an additional Rh cDNA clone, Rh13, that is 96% nucleotide- and 92% amino acid-identical to Rh21, with the substitutions scattered throughout the sequence. A molecular genetic approach was used to match this Rh clone with an Rh specificity. The mRNA transcript for Rh13 was present in reticulocytes from RhD-positive individuals, but was ab-

sent from the reticulocytes of RhD-negative individuals. Using conventional screening of genomic libraries, as well as PCR cloning, partial genomic clones for these two Rh cDNAs were obtained. Based on PCR analysis and Southern blots, the Rh21 gene was present in all individuals, but an intact Rh13 gene was only present in RhD-positive and not RhD-negative individuals. Thus, by correlating the presence of Rh mRNA and gene sequences with individual Rh phenotypes, we were able to establish that the new Rh13 cDNA clone represents the RhD protein.

© 1993 by The American Society of Hematology.

THE Rh BLOOD GROUP system plays a role of central importance in immunohematology and transfusion medicine.¹ The Rh antigens are the most immunogenic red blood cell antigens in humans (outside the preformed antibodies of the ABO system). Immunization through prior blood transfusion or pregnancy can lead to hemolytic transfusion reactions or hemolytic disease of the newborn. In addition, the target of many of the autoantibodies in warm-type autoimmune hemolytic anemia is an Rh polypeptide. A nonimmune hemolytic anemia occurs for the red blood cells of the rare Rh_{null} phenotype.² Thus, the Rh blood group system is associated with hemolysis in several disease states.

The major Rh antigens D, C, c, E, and e are carried on membrane proteins of approximately 32,000 *M_r*,^{3,4} with distinct but homologous proteins carrying the D, C/c, and E/e specificities.^{5,6} A major breakthrough in the molecular analysis of the Rh system occurred with the cloning of an Rh cDNA by two independent groups.^{7,8} This was done by using oligonucleotides based on the amino-terminal sequence of purified Rh protein, but the common amino-terminal sequence did not permit any assessment of which Rh antigen might be associated with this clone. The most direct approach to this problem would be to express the cDNA clone and assess antigenic reactivity, but our attempts to date have been unsuccessful.⁹ Therefore, we have undertaken a molecular genetic approach to the assignment of clones. In this report, we describe a new Rh cDNA and gene and the use of genetic approaches to demonstrate that this Rh clone represents the D antigen. At the time of the final preparation of this report, a report appeared describing the isolation of the same RhD cDNA.¹⁰ The present work is an independent verification of the identity of RhD.

MATERIALS AND METHODS

Rh phenotypes. Red blood cells were phenotyped for Rh antigens by standard agglutination techniques¹ using human monoclonal antibodies to Rh kindly provided by John Moulds (Gamma Biologicals, Houston, TX).

Polymerase chain reaction and DNA sequencing. Oligonucleotides based on the published Rh cDNA clone^{7,8} were used to amplify DNA from a human bone marrow cDNA library in λ gt11 (Clontech Laboratories, Palo Alto, CA), from genomic DNA, or from first-strand cDNA derived by reverse transcription of RNA prepared from reticulocytes as described.¹¹ Oligonucleotides in-

cluded the following (numbering based on position +1 representing A of the initiation ATG codon): R11C, nt -10 to 10; R24C, nt 521 to 540; R10N, nt 1330 to 1311; R13N, nt 670 to 651; R24N, nt 623 to 604. A standard 50- μ L polymerase chain reaction (PCR) contained 1 μ L viral stock at 5×10^9 plaque-forming units (pfu)/mL, 1 μ g genomic DNA, or 5 μ L cDNA, and used AmpliTaq DNA polymerase (Perkin-Elmer, Norwalk, CT) for 35 cycles at 94°C for 1 minute, 50°C for 1 minute, and 72°C for 3 minutes. DNA was subcloned, mapped, and sequenced as previously described.¹²

Screening genomic DNA library. A human genomic DNA library (Clontech) was screened with Rh cDNA at high stringency. All techniques for screening, plaque purification, subcloning, mapping, and DNA sequencing have been described previously.¹³

Southern blot analysis. Ten micrograms of genomic DNA was digested overnight at 37°C with *Eco*RI restriction enzyme (Promega, Madison, WI) in a 100- μ L vol. The DNA was then ethanol precipitated, brought back up in 10 μ L water, separated on a 0.8% agarose gel in $1 \times$ Tris-acetate/EDTA (TAE) buffer, and blotted onto a nylon membrane (Oncor, Gaithersburg, MD). The membrane was hybridized overnight, washed at high stringency,¹³ and exposed to x-ray film for 48 hours.

RESULTS

Two distinct but homologous Rh cDNA clones. Because there are several Rh proteins, we began by cloning additional Rh cDNA so that we could then correlate Rh phenotype with individual cDNA species. Multiple cDNA clones were produced by PCR amplification of a human bone marrow cDNA library using oligonucleotides based on the one known Rh cDNA clone, designated RhIXb⁷ or

From the Division of Laboratory Medicine, Departments of Pathology and Internal Medicine, Washington University School of Medicine, St Louis, MO.

Submitted December 2, 1992; accepted March 12, 1993.

Address reprint requests to Douglas M. Lublin, MD, PhD, Department of Pathology, Washington University School of Medicine, 660 S Euclid Ave, Box 8118, St Louis, MO 63110.

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. section 1734 solely to indicate this fact.

© 1993 by The American Society of Hematology.

0006-4971/93/8203-0033\$3.00/0

Rh30A.⁸ The oligonucleotide pair R11C/R10N was chosen to produce a full-length coding region. Subcloning and sequencing of many PCR products demonstrated two types of cDNA: one, exemplified by clone Rh21, was identical with the RhIXb/Rh30A clone; the second, exemplified by clone Rh13, had a distinct but homologous cDNA sequence. Of 14 randomly selected clones arising from several independent PCR amplifications, nine were of the Rh21 type and five of the Rh13 type. The DNA and derived protein sequence of the new Rh13 cDNA (Fig 1) demonstrate that it is the same size as the Rh21 cDNA. Comparison of the two cDNAs shows that there are no insertions or deletions, but there are 44 nucleotide substitutions, 35 of which result in amino acid substitutions (Fig 2). The two cDNAs are 96% identical at the nucleotide level and 92% identical at the protein level.

Expression of Rh13 RNA transcripts correlates with RhD-positive phenotype. We next looked for any correlation between either of the two Rh cDNAs (Rh21 and Rh13) and the known Rh polymorphisms. One of the nucleotide substitutions from Rh21 to Rh13 eliminates a *Bcl* I restriction site, so this could be used to identify the two cDNAs. Reticulocyte RNA was used to prime first-strand cDNA synthesis and this material was then amplified by PCR using the oligo-

nucleotide pair R24C/R24N that amplifies both Rh21 and Rh13, producing a DNA fragment that includes the *Bcl* I site. Amplified DNA was digested with *Bcl* I and analyzed by nondenaturing acrylamide gel electrophoresis and ethidium bromide staining. All Rh phenotypes possessed the Rh21 RNA (as evidenced by a band that cut with *Bcl* I), whereas Rh13 RNA was found in all RhD-positive, but no RhD-negative individuals (as evidenced by a band that fails to cut with *Bcl* I) (Fig 3). The C/c or E/e polymorphisms did not correlate with the presence of Rh13. These results indicate that the Rh13 clone represents the RhD protein.

Isolation and sequencing of parts of two Rh genes corresponding to two Rh cDNAs. To confirm and extend this conclusion, we next obtained partial genomic clones corresponding to these two cDNAs. The two Rh cDNA clones Rh21 and Rh13 contain nucleotide substitutions scattered throughout the entire coding region, so clearly they do not arise by alternative splicing, but must correspond to separate genes. The Rh21 cDNA was used as a probe to screen a λ bacteriophage human genomic DNA library, and a genomic clone of 15 kb was isolated. Restriction mapping and DNA sequencing produced the exon/intron structure of this gene (Fig 4A), and the sequence of the exons showed that it was the Rh21 gene. Next, genomic DNA from an RhD-nega-

```

ATGAGCTCTAAGTACCGCGGTCTGTCCGGCGCTGCCCTCTGGGCCCTAACACTGGAAGCAGCTCTCATTCTCCTTTCTATTTT 90
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 30
TTTACCACATGACGCTTCTTAGAGGATCAAAGGGGCTCGTGGCATCCTATCAAGTTGGCCAAGATCTGACCGTGATGGCGGCCATT 180
F T H Y D A S L E D Q K G L V A S Y Q V G Q D L T V M A A I 60
GGCTTGGGCTTCCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCCTTCAACCTCTTCATGCTGGCGCTTGGTGTGCAGTGG 270
G L G F L T S S F R R H S W S S V A F N L F M L A L G V Q W 90
GCAATCCTGCTGGACGGCTTCTGAGCCAGTCCCTTCTGGGAAGGTGGTCATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360
A I L L L D G F L S Q F P S G K V V I T L F S I R L A T M S A 120
TTGTCCGGTCTGATCTCAGTGGATGCTGTCTTGGGAAGGTCAACTTGGCGCAGTTGGTGGTGTGTTGGTGGTGGAGGTGACAGCTTTA 450
L S V L I S V D A V L G K V N L A G L V V M V L V E V T A L 150
GGCAACCTGAGGATGGTCATCAGTAATATCTTCAACACAGACTACCACATGAACATGATGCACATCTACGTGTTCCGAGCCTATTTTGGG 540
G N L R M V I S N I F N T D Y H M N M M H I Y V F A A Y F G 180
CTGTCTGTGGCCTGGTGCCTGCCAAAGCCTCTACCCGAGGGAACGGAGGATAAAGATCAGACAGCAACGATACCCAGTTTGTCTGCCATG 630
L S V A W C L P K P L P E G T E D K D Q T A T I P S L S A M 210
CTGGGCGCCCTCTTCTTGTGGATGTTCTGGCCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAATCGAAAGGAAGAATGCCGTGTTCAAC 720
L G A L F L W M F W P S F N S A L L R S P I E R K N A V F N 240
ACCTACTATGCTGTAGCAGTCAGCGTGGTGACAGCCATCTCAGGGTCACTCTGGCTCACCCCCAAGGGAAGATCAGCAAGACTTATGTG 810
T Y Y A V A V S V V T A I S G S S L A H P Q G K I S K T Y V 270
CACAGTGGCGTGTGGCAGGAGGCGTGGCTGTGGGTACCTCGTGTACCTGATCCCTTCTCGGTGGCTTGGCATGGTGTGGTCTTGTG 900
H S A V L A G G V A V G T S C H L I P S P W L A M V L G L V 300
GCTGGGCTGATCTCCGTCGGGGGAGCCAAGTACCTGCCGGGGTGTGTAACCGAGTGTGGGGATTCCCACAGCTCCATCATGGGCTAC 990
A G L I S V G G A K Y L P G C C N R V L G I P H S S I M G Y 330
AACTTCAGCTTGTGGTCTGCTTGGAGAGATCATCTACATTTGTTGCTGGTGTGTTGATACCGTGGAGCCGGCAATGGCATGATTGGC 1080
N F S L L G L G L E I I Y I V L L V L D T V G A G N G M I G 360
TTCAGGCTCCTCCTCAGCATTGGGGAAGTCAAGTCTGGCCATCGTATAGCTCTCAGTCTGGTCTCCTGACAGGTTTGGCTCCTAAATCTT 1170
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 390
AAAAATGGAAGCACCTCATGAGGCTAAATATTTGATGACCAAGTTTCTGGAAGTTTCTCATTGGCTGTTGGATTTTAA 1254
K I W K A P H E A K Y F D D Q V F W K F P H L A V G F 417

```

Fig 1. DNA and amino acid sequence of coding region of Rh13 cDNA clone. The Rh13 cDNA was derived by PCR amplification from a human bone marrow cDNA library using primers based on the sequence of the RhIXb/Rh30A cDNA. The DNA sequence of the full-length coding region together with the derived amino acid sequence (using standard single-letter abbreviations) is shown. The nucleotide sequence has been deposited in the GenBank database under accession no. L08429.

Rh13	MSSKYPRSVRRCLPLWALTLEAALILLFYFFTHYDASLEDQKGLV	45
Rh21	
Rh13	ASYQVGQDLTVMAAIGLGFLTSSFRHSWSSVAFNLFMLALGVQW	90
Rh21L.....N.....	
Rh13	AILLDGFLSQFSPGKVVITLFSIRLATMSALSVLISVDAVLGKVN	135
Rh21P.....M.....AG.....	
Rh13	LAQLVVMVLVEVTALGNLRMVISNIFNTDYHMNMHIYVFAAYFG	180
Rh21T.....LR-F.....	
Rh13	LSVAWCLPKLPEGTEDKDQTATIPSLSAMLGALFLWMFWPSFNS	225
Rh21	-T.....K.....N.....R.....V..	
Rh13	ALLRSPIERKNAVFNTYYAVAVSVVTAISGSSLAHPQGKISKTYV	270
Rh21	P.....Q.....M.....L.....R.....M...	
Rh13	HSAVLAGGVAVGTSCHLIPSPWLAMVLGLVAGLISVGGAKYLPGC	315
Rh21C.....V.....I.....C.....V..	
Rh13	CNRVLGIPHSSIMGYNFSLGLLGEIIVLVLLDVTGAGNGMIG	360
Rh21H-I-V-HSI.....T.....H.....WN.....	
Rh13	FQVLLSIGELSLAIVIALTSGLLTGLLLNLKIWKAPHEAKYFDDQ	405
Rh21V.....	
Rh13	VFWKFPHLAVGF	417
Rh21	

Fig 2. Comparison of amino acid sequences of two Rh cDNA clones. The complete derived amino acid sequences for the new Rh13 cDNA clone and the original Rh21 cDNA clone (also designated Rh1Xb or Rh30A) are aligned. A dash is shown wherever the two sequences are identical.

tive and RhD-positive individual was amplified by PCR using an oligonucleotide pair R24C/R13N that spans the first two exons of the partial Rh21 genomic clone and hence includes the intron between cDNA nt 634 and 635 (Fig 4A). Agarose gel analysis of this reaction showed the expected Rh21 genomic fragment of 1,200 bp in both samples, but the RhD-positive sample also produced an additional fragment of 600 bp (Fig 4B). Each of these fragments was subcloned and sequenced: the 1,200-bp fragment from both samples corresponded to the Rh21 gene that had already been cloned (Fig 4A), and the 600-bp fragment from the RhD-positive individual contained exonic sequences that identified it as the Rh13 gene. DNA sequence analysis

showed that the Rh13 gene had a smaller intron due to a large deletion in the middle of the intron (Fig 4C).

Finally, these findings were confirmed by direct Southern blot analysis of genomic DNA from multiple Rh phenotype individuals probed with the 600-bp Rh13 gene fragment. This blot demonstrated a common Rh gene in all individuals, plus a second gene present only in RhD-positive individuals (Fig 4D). The common upper band of 6,000 bp represents the Rh21 gene based on the size of the *EcoRI* fragment from the Rh21 genomic clone (Fig 4A), and thus the lower band represents the Rh13 gene. The Rh13 gene is present only in RhD-positive individuals. Hence, based on both RNA transcript expression (Fig 3) and the actual exist-

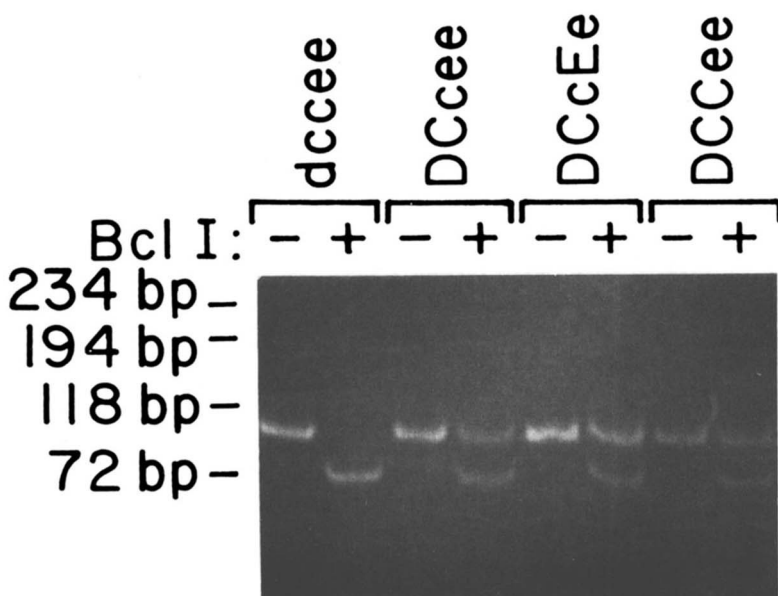
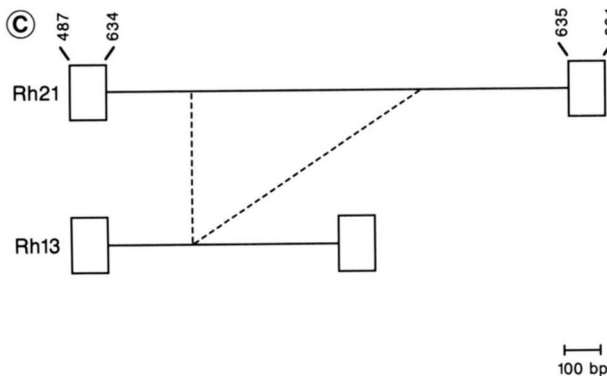
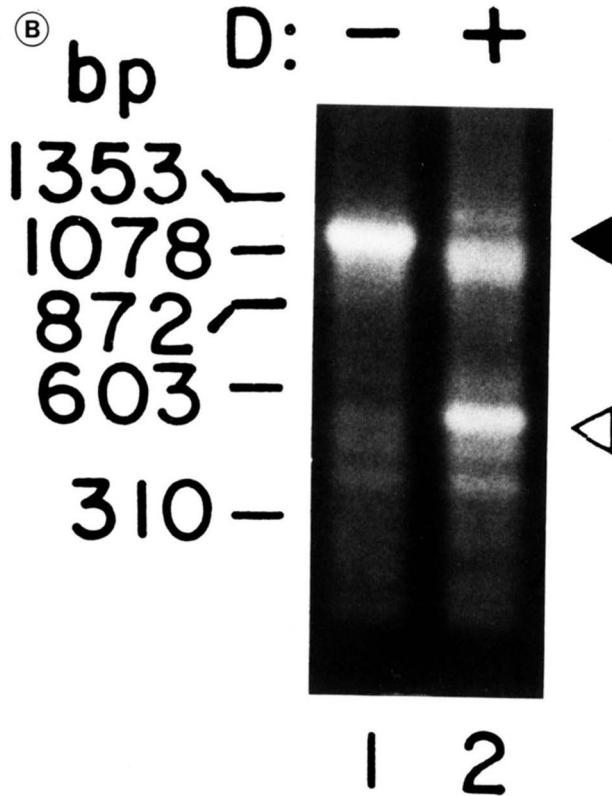
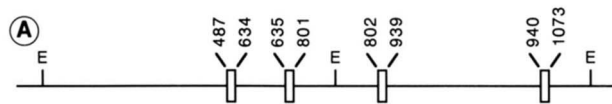


Fig 3. Correlation between expression of Rh RNA transcripts and Rh antigenic phenotypes. RNA from reticulocytes of individuals of differing Rh phenotypes was reverse transcribed into first-strand cDNA and amplified by PCR using the R24C/R24N primers. Products were digested with *Bcl I* or control buffer, and analyzed on a non-denaturing 12% acrylamide gel followed by ethidium bromide staining. Two other unrelated RhD-negative individuals (not shown) demonstrated the same pattern of complete digestion with *Bcl I* as shown with the RhD-negative sample in the two left lanes.



tence of an intact Rh13 gene only in RhD-positive individuals (Fig 4B and D), we conclude that the Rh13 cDNA and gene encode the RhD protein. It is formally possible that Rh13 represents another Rh protein that is present only in RhD-positive but not RhD-negative individuals. However, this is extremely unlikely, since in that case the RhD-negative individuals should produce an antibody against this putative protein, and there are no unexplained antibodies produced only by RhD-negative individuals. A long amino-

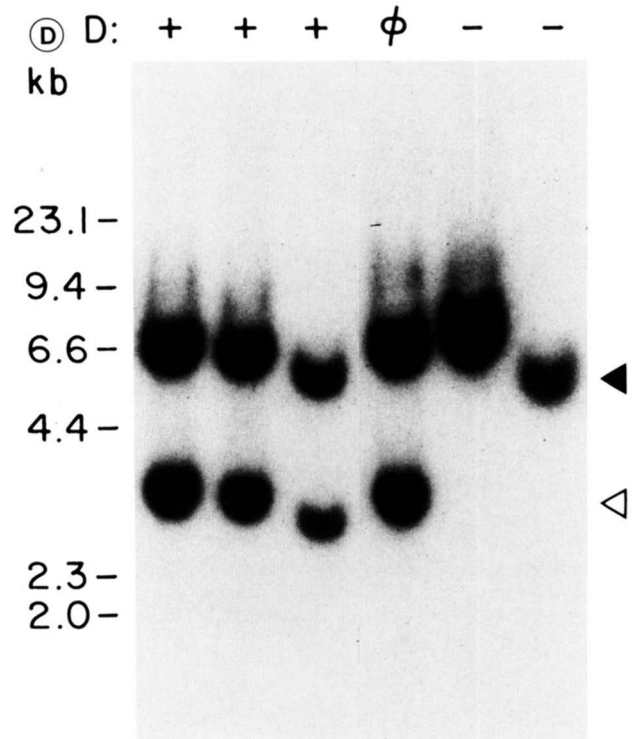


Fig 4. Analysis of two Rh genes corresponding to the Rh21 and Rh13 cDNA clones. (A) By screening a λ bacteriophage human genomic library with the Rh21 cDNA, a partial Rh21 genomic clone was derived. The exon/intron structure of the gene plus an *EcoRI* (E) map are shown. (B) PCR amplification of genomic DNA from an RhD-positive or RhD-negative individual using the R24C/R13N primers that span the intron between nt 634 and 635, analyzed by agarose gel electrophoresis and ethidium bromide staining. Solid arrowhead marks the 1,200-bp band seen in both samples, and Open arrowhead marks the 600-bp band seen only in the RhD-positive sample. (C) Subcloning and sequencing of the 1,200- and 600-bp products from (B) gave the partial gene structure for the Rh21 and Rh13 genes shown. Dotted lines mark the portion of the Rh21 gene intron that are deleted in the corresponding Rh13 gene intron. (D) Southern blot of *EcoRI* digest of 10 μ g genomic DNA from individuals of varying Rh phenotypes, probed at high stringency with the 600-bp fragment of Rh13 gene. Solid arrowhead marks the upper band seen in all individuals, and Open arrowhead marks the lower band seen only in RhD-positive individuals. Rh phenotypes for the samples in the lanes are (1) DccEE, (2) DCcEe, (3) DCcEe, (4) Rh_{null}, (5) dccee, (6) dccee.

terminal sequence for purified RhD protein⁸ agrees with the Rh13 cDNA clone for the first 41 amino acids, but has a few discrepancies past there. However, this is quite far into the protein for highly accurate amino acid sequence determination and likely represents amino acid sequencing errors.

This specific assignment of the Rh13 clone as RhD was performed in the same manner as the work assigning the individual alleles of the ABO blood group system that established a correspondence between the presence of each

cDNA or genomic sequence and the ABO phenotype.¹⁴ Attempts to express any of the Rh cDNA clones have been unsuccessful,⁹ so this RhD clone cannot be confirmed at present by expression of the RhD antigen. This makes the strength of the molecular genetic evidence especially crucial, and the fact that we and another group¹⁰ independently discovered and verified this RhD clone is therefore important. Both reports demonstrated that the RhD sequences could only be amplified from genomic DNA of RhD-positive individuals, and in the present study this was supported by analysis of mRNA transcripts from reticulocytes (Fig 3). The assignment of the RhD clone was further confirmed by analysis of partial genomic fragments from the RhD and Rh21 (presumed RhC/c and/or RhE/e) genes. There is only one amino acid difference between the two reported sequences: amino acid 218 is Met in this report (unchanged from clone Rh21) but Ile in the other report.¹⁰ Both groups used the same bone marrow cDNA library, and one would not expect a polymorphism in the RhD protein sequence. We sequenced five clones and all had Met at this position, so the explanation of this discrepancy is not clear.

DISCUSSION

Although the present work solves the identity of the RhD cDNA, there are many questions remaining in the molecular analysis of the Rh gene family. The failure to express any RhD-related antigen in RhD-negative individuals is due to a large (or complete) deletion of the RhD (Rh13) gene in the genome of these individuals. Southern blot analysis with the RhIXb (Rh21) cDNA had led to the suggestion that the RhD-negative phenotype represents a deleted gene,¹⁵ but that conclusion could not be definitive without cloning the RhD cDNA or gene as has been done here. The actual PCR analysis used for this work (Fig 4B) can form the basis for a simple RhD phenotyping of individuals. The Rh21 gene present in all common Rh phenotypes most likely encodes the C/c and/or E/e antigens, but further analysis will be needed, in particular molecular genetic analysis correlating Rh clones and Rh phenotypes. Whether these two Rh genes, RhD (Rh13) and Rh21, encode all the major Rh antigenic proteins versus the possible existence of a third Rh gene is also to be determined.

This work has also led to some information on the structure of the Rh genes in Rh_{null} individuals. The Southern blot of the genomic DNA from one of the first identified Rh_{null} individuals² probed with the 600-bp RhD gene fragment (Fig 4D) or with the full-length RhD cDNA (data not shown) demonstrates that the Rh genes are grossly intact. Finally, although we have uncovered the sequence of the RhD polypeptide, future work will require expression of this RhD cDNA (perhaps as part of an Rh complex) to decipher the function of this protein.

ACKNOWLEDGMENT

We thank John Moulds for providing human monoclonal antibodies to Rh, and Claire Kruppe and H.H.H. for providing an Rh_{null} sample.

REFERENCES

1. Mollison PL, Engelfreit CP, Contreras M: Blood Transfusion in Clinical Medicine (ed 8). Oxford, UK, Blackwell, 1987
2. Sturgeon P: Hematological observations on the anemia associated with blood type Rh_{null}. *Blood* 36:310, 1970
3. Moore S, Woodrow CF, McClelland DBL: Isolation of membrane components associated with human red cell antigens Rh₀(D), (c), (E) and Fy^a. *Nature* 295:529, 1982
4. Gahmberg CG: Molecular identification of the human Rh₀(D) antigens. *FEBS Lett* 40:93, 1982
5. Bloy C, Blanchard D, Dahr W, Beyreuther K, Salmon C, Cartron JP: Determination of the N-terminal sequence of human red cell Rh(D) polypeptide and demonstration that the Rh(D), (c), and (E) antigens are carried by distinct polypeptide chains. *Blood* 72:661, 1988
6. Blanchard D, Bloy C, Hermand P, Cartron JP, Saboori AM, Smith BL, Agre P: Two-dimensional iodopeptide mapping demonstrates that erythrocyte Rh D, c, and E polypeptides are structurally homologous but nonidentical. *Blood* 72:1424, 1988
7. Cherif-Zahar B, Bloy C, Le Van Kim C, Blanchard D, Bailly P, Hermand P, Slamon C, Cartron JP, Colin Y: Molecular cloning and protein structure of a human blood group Rh polypeptide. *Proc Natl Acad Sci USA* 87:6243, 1990
8. Avent ND, Ridgwell K, Tanner MJA, Anstee DJ: cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rhesus)-blood-group-antigen expression. *Biochem J* 271:821, 1990
9. Arce MA, Ferdman BA, Thompson ES, Lublin DM: Identification of additional Rh cDNA clones. *Transfusion* 32:47S, 1992 (suppl, abstr)
10. Le Van Kim C, Mouro I, Cherif-Zahar B, Raynal V, Cherrier C, Cartron JP, Colin Y: Molecular cloning and primary structure of the human blood group RhD polypeptide. *Proc Natl Acad Sci USA* 89:10925, 1992
11. Temple GF, Chang JC, Kan YW: Authentic β -globin mRNA sequences in homozygous β^0 -thalassemia. *Proc Natl Acad Sci USA* 74:3047, 1977
12. Lublin DM, Thompson ES, Green AM, Levene C, Telen MJ: Dr(a-) polymorphism of decay accelerating factor: biochemical, functional, and molecular characterization and production of allele-specific transfectants. *J Clin Invest* 87:1945, 1991
13. Post TW, Arce MA, Liszewski MK, Thompson ES, Atkinson JP, Lublin DM: Structure of the gene for human complement protein decay accelerating factor. *J Immunol* 144:740, 1990
14. Yamamoto F, Clausen H, White T, Marken J, Hakomori S: Molecular genetic basis of the histo-blood group ABO system. *Nature* 345:229, 1990
15. Colin Y, Cherif-Zahar B, Le Van Kim C, Raynal V, Van Huffel V, Cartron JP: Genetic basis of the RhD-positive and RhD-negative blood group polymorphism as determined by Southern analysis. *Blood* 78:2747, 1991