The Rh (Rhesus) protein family comprises Rh50 glycoprotein and Rh30 polypeptides, which form a complex essential for Rh antigen expression and erythrocyte membrane integrity. This article describes the structural organization of Rh50 gene and identification of its associated splicing defect causing Rhnull disease. The Rh50 gene, which maps at chromosome 6p11–21.1, has an exon/intron structure nearly identical to Rh30 genes, which map at 1p34–36. Of the 10 exons assigned, conservation of size and sequence is confined mainly to the region from exons 2 to 9, suggesting that RH50 and RH30 were formed as two separate genetic loci from a common ancestor via a transchromosomal insertion event. The available information on the structure of RH50 facilitated search for candidate mutations underlying the Rh deficiency syndrome, an autosomal recessive disorder characterized by mild to moderate chronic hemolytic anemia and spherostomatocytosis. In one patient with the Rhnull disease of regulator type, a shortened Rh50 transcript lacking the sequence of exon 7 was detected, while no abnormality was found in transcripts encoding Rh30 polypeptides and Rh-related CD47 glycoprotein. Amplification and sequencing of the genomic region spanning exon 7 revealed a G → A transition in the invariant GT motif of the donor splice site in both Rh50 alleles. This splicing mutation caused not only a total skipping of exon 7 but also a frameshift and premature chain termination. Thus, the deduced translation product contained 351 instead of 409 amino acids, with an entire different C-terminal sequence following Thr315. These results identify the donor splicing defect, for the first time, as a loss-of-function mutation at the RH50 locus and pinpoint the importance of the C-terminal region of Rh50 in Rh complex formation via protein-protein interactions.

The Rh (Rhesus) protein family is currently known to consist of three erythroid-specific integral membrane proteins, the Rh50 glycoprotein and two Rh30 (RhD and RhCE) polypeptides (1–4). Although their genetic loci are mapped on chromosomes 6p11–21.1 and 1p34–36, respectively, Rh50 and Rh30 share a clear sequence homology (36% overall identity) and a similar 12-transmembrane (TM) topology (50% identity in the putative α-helices) (5–8). As nonglycosylated and palmitoylated proteins, RhD and RhCE each contain 417 amino acids, serving as the carriers of D and CcEe blood group antigens (5–7). By contrast, the 409-amino acid Rh50 glycoprotein in itself does not carry Rh antigens but rather interacts with Rh30 polypeptides to form a protein complex, thereby functioning as a coexpressor to facilitate Rh antigen disposition in the erythrocyte membrane (8–10).

Apart from being a structural unit of Rh antigen expression, the Rh50 and Rh30 proteins appear to possess some hitherto undefined roles essential for the function and integrity of plasma membranes. This proposal is highlighted primarily by the occurrence of Rh deficiency syndrome, a rare autosomal recessive disorder characterized by a chronic hemolytic anemia of varying severity, a hereditary spherostomatocytosis, and multiple membrane abnormalities (1–3). The Rh deficiency syndrome exists in two conditions in which a complete absence of all Rh antigens defines the Rhnull status and a barely detectable presence defines the Rhmod phenotype (11, 12). Both conditions exhibit an absence or weakened expression of several other membrane glycoproteins or associated antigens, including Rh50, CD47, LW, Duffy (Fy5), and glycoporphlin B (GPB for SsU) (1–3). Therefore, the Rh deficiency syndrome can be regarded as a disorder of impaired protein-protein interactions.

As shown by family studies, Rh deficiency is almost invariably associated with consanguinity and can occur on different genetic backgrounds (11, 12). The amorph type of Rhnull is thought to arise by silencing mutations at the RH50 locus encoding RhD and RhCE polypeptides, but its underlying molecular defect has remained to be determined (13–15). In contrast, the regulator Rhnull and Rhmod phenotypes are considered to result from suppressor or "modifier" mutations independent of the RH30 locus (16). The genuine interaction of Rh50 with Rh30 proteins in Rh complex formation points to RH50 locus as a primary candidate responsible for the suppressor forms of Rh deficiency. To facilitate the identification of such suppressor mutations, the organization of Rh50 gene has now been delineated. Here, I describe the exon/intron structure of the Rh50 gene and identification of its associated splicing defect as a loss-of-function mutation in one Rhnull patient. The findings reported herein correlate the disease phenotype with an impaired Rh complex formation and provide evidence for the importance of the C-terminal region of Rh50 participating in protein-protein interactions.

EXPERIMENTAL PROCEDURES

Blood Samples—Blood samples from normal human blood donors with RhD-positive (RhD+) and RhD-negative (RhD−) phenotypes (de-
located in the 3' untranslated region (3'UTR), leading to a potential open reading frame in the cDNA and upstream ATG codons were obtained by two rounds of PCR using AP1 + GSP1 and AP2 + GSP2 (Table I).

M, λ (HindIII) and dX174 (HaeIII) DNA markers. B, organization of the Rh50 gene and comparison with the Rh30 gene. Exons are denoted by open bars and introns by broken lines (not to scale). The size of coding sequence for each exon (in base pairs) is shown; exon 1 is counted from ATG and exon 10 ends before the TAA codon (marked by asterisks).

Direct Nucleotide Sequencing and Sequence Analysis—All amplified cDNAs and genomic DNA products were purified from total genomic DNA, which encompassed the whole intron 4 and intron 6, respectively. Bands seen in other lanes each were obtained by two rounds of PCR using amplimers. All nucleotide (nt) positions of sense (s) and antisense (a) direction are indicated. Note that lanes 8 and 13 are products amplified directly from total genomic DNA, which encompass the whole intron 4 and intron 6, respectively. Bands seen in other lanes each were obtained by two rounds of PCR using amplimers or a gene-specific primer (GSP1), whereas in the second step, nested AP2 and GSP2 were used. The resultant products were analyzed by agarose gel and sequenced after purification by 5% polyacrylamide gel electrophoresis.

Characterization of Exon/Intron Structure of the Rh50 Gene—To determine the structural organization of the Rh50 gene, genomic DNA from a normal person was digested separately with restriction endonucleases EcoRV, HindII, PvuII, SmaI, SspI, and StuI. The total digests of each restriction enzyme were ligated to the same adaptor to generate the genomic library containing the Rh50 gene and comparison with the Rh30 gene.

Exons are denoted by broken lines or open bars and introns by broken lines (not to scale). The size of coding sequence for each exon (in base pairs) is shown; exon 1 is counted from ATG and exon 10 ends before the TAA codon (marked by asterisks).
RESULTS

Organization of Rh50 Gene and Comparison with Rh30 Gene—To delineate the structural organization of the Rh50 gene, a bidirectional walking approach was taken to retrieve unknown sequences (Fig. 1A). 40 synthetic primers that cover various coding sequences (Table I) were used in combination to amplify the adaptor-ligated, restriction enzyme-specific unknown sequences (Fig. 1). This global organization is strikingly similar to that of the Rh30 genes (23, 24) and is essentially conserved in the Rh50 homologues from the mouse and Caenorhabditis elegans.\(^2\) Comparison of Rh50 with Rh30 showed that their sequence homology is confined mainly to exons 2–9, whereas their 5’ or 3’ regions share little or no sequence similarity. The size of all internal exons except exons 7 and 8 was conserved, and exon 2 of Rh50 was missing codon AGT for Ser99, which is present in Rh30 genes (5–7). Thus, Rh50 and Rh30 show the molecular genetics of Rhnull disease.

### Table I

| Designation | Sequence (5’-3’) | Nucleotide location | Location |
|-------------|------------------|---------------------|----------|
| E1–1a       | CTCAAGAAATGGCCACATGCTTG | 147-125               | Exon 1   |
| E1–2a       | CCATGTCGCTGACGCTCGTACG | 120-97                | Exon 1   |
| E1–1a       | AGTGGCTGATGTCAGCGCTGACGAC | –25 ~ –4               | 5’-UTR   |
| E1–2a       | ATGAGGTGTCATCTTCTCCCTGAT | 1-24                  | Exon 1   |
| E2–1a       | CCAAACGCGCCACGCTGACGCTG   | 268-246               | Exon 2   |
| E2–1a       | TGATACCCACGTCTCGACGCTGACGAC | 247-224               | Exon 2   |
| E2–1a       | TGTCAGAACATGTCAGCCAGCTG   | 158-180               | Exon 2   |
| E2–2a       | TGGTCGGGTTGCTCCCTCCCTGAT | 268-246               | Exon 2   |
| E3–1a       | CACATATGTCGATGTCAGCGCTGACGAC | 481-458               | Exon 3   |
| E3–2a       | ATGGTGTCATCTTCTCCCTGAT | 457-414               | Exon 3   |
| E3–1a       | GAAATCCAAAAACGATGATCGACGACGAC | 331-357               | Exon 2/3 |
| E3–2a       | ATATGTCGACATGTCAGCCAGCTG   | 346-360               | Exon 3   |
| E4–1a       | CGATGTTCGATGTCAGCGCTGACGAC | 627-644               | Exon 4   |
| E4–1a       | CCTCTTCTTATCATCCCTCCCTTTC | 606-584               | Exon 4   |
| E4–2a       | GCTCTCTGACATGTCAGCGCTGACGAC | 493-516               | Exon 4   |
| E4–2a       | AAGTCCAACTCATACTTCAACAGACGAC | 602-624               | Exon 4   |
| E5–1a       | GACTCTTCTCATTAGATGTCAGCGCTGACGAC | 732-710               | Exon 5   |
| E5–2a       | CTGTTCTTCTTCCCGCTTTGACGACGAC | 708-685               | Exon 5   |
| E5–1a       | AGTCTTCTTCTGAGGATGTCAGCGCTGACGAC | 670-693               | Exon 5   |
| E5–1a       | GATACATGTGGGGTCCATAACCTCAGACGACGAC | 703-729               | Exon 5   |
| E6–1a       | AGTCAAGGACATCTTCTGACGCTGACGACGAC | 945-922               | Exon 6   |
| E6–2a       | AGACTTCTTCTGACGCTGACGCTGACGACGAC | 919-897               | Exon 6   |
| E6–2a       | GTTCAGACATAGTGTGACGCTGACGCTGACGAC | 808-831               | Exon 6   |
| E6–1a       | CTGAGGAGCGACATGTCAGCGCTGACGCTGACGAC | 856-879               | Exon 6   |
| E7–1a       | GTAGTCTTCTTCTGACGCTGACGCTGACGACGAC | 1057-1040             | Exon 7   |
| E7–2a       | GGGGCTCTTCTTCTGACGCTGACGCTGACGACGAC | 1026-1005             | Exon 7   |
| E7–2a       | AAGTCCCACTCATACTTCAACAGACGACGACGAC | 946-968               | Exon 7   |
| E7–3a       | CACTTCTTCTTCTTCTGACGCTGACGCTGACGAC | 976-1002              | Exon 7   |
| E7–4a       | GACTATGTCGCTGACGCTGACGCTGACGACGAC | 1135-1112             | Exon 7   |
| E7–4a       | AGATGTCGCTGACGCTGACGCTGACGCTGACGAC | 1088-1066             | Exon 7   |
| E8–1a       | CTATGTCGCTGACGCTGACGCTGACGCTGACGAC | 1069-1092             | Exon 8   |
| E8–2a       | GCTCTTCTTCTGACGCTGACGCTGACGCTGACGAC | 1097-1120             | Exon 8   |
| E8–1a       | GCTCTTCTTCTGACGCTGACGCTGACGCTGACGAC | 1140-1165             | Exon 8   |
| E9–1a       | ATGAGGTGTCATCTTCTCCCTGAT | 1214-1191             | Exon 9   |
| E9–2a       | TGGTCGGGTTGCTCCCTCCCTGAT | 1266-1243             | Exon 9   |
| E9–1a       | TGTCAGAACATGTCAGCCAGCTG   | 1281-1298             | 3’-UTR   |
| E9–2a       | CAGGTCATCTTCTCCTCCCTGAT | 1256-1243             | 3’-UTR   |
| E10–1a      | AAGTGGACAAAGGACGCTGAGGACGACGACGAC | 1321-1298             | 3’-UTR   |
| E10–2a      | GCTGTGTCATCTTCTCCCTGAT | 1237-1294             | 3’-UTR   |
| E10–2a      | TGGTCGGGTTGCTCCCTCCCTGAT | 1297-1303             | 3’-UTR   |

\(^a\) Suffixes "a" and "s" denote antisense (reverse) and sense (forward) primers, whereas numbers 1 and 2 indicate GSP1 and GSP2, respectively. E10 primers are all located in the 3´-untranslated region downstream of TAA stop codon.

\(^b\) All nucleotide positions are accounted from the first base of ATG initiation codon (Fig. 2).

Sequence of Splice Sites and Exon/Intron Junctions in the Rh50 Gene

Comparison of Rh50 with Rh30 showed that their sequence homology is confined mainly to exons 2–9, whereas their 5’ or 3’ regions share little or no sequence similarity. The size of all internal exons except exons 7 and 8 was conserved, and exon 2 of Rh50 was missing codon AGT for Ser99, which is present in Rh30 genes (5–7). Thus, Rh50 and Rh30 show the same assignment of exon/intron junctions except for a difference in their exon 7/exon 8 boundaries (Fig. 1B). The 5’ region of Rh50 has several putative cis-acting elements (Fig. 2), including the TATA boxes that are absent from the proximal promoter of both RhD and RhCE genes (23, 24). Multiple transcription initiation sites occur between the two putative Ets binding sites (see footnote 3). The first position of ATG codon assigned for translation initiation of the erythropoietic-specific Rh50 protein (8) is denoted. The encoded amino acids of exon 1 and exon 2 (partial) are shown below the nucleotide sequence.
Of the 10 exons identified, only splice sites conform to the "GT-AG" rule and possess the junctions in the Rh50 gene. All the 5' splice sites in the Rh50 gene are distributed in 10 exons (boxed). The nucleotide positions marking the beginning and end of each exon are numbered: nt 1 denotes the first nucleotide of erythroid ATG initiation codon and nt 1230 the third base of TAA stop codon. "aataaa" indicates one of the polyadenylation signals present in the 3' -UTR. Exon sequences are denoted by uppercase letters, whereas intron sequences, including the 3' -acceptor and 5' -donor splice sites, are indicated by lowercase letters. Interval exon sequences are omitted (shown by dots). Amino acids encoded by the respective exon/exon boundaries are indicated below the middle position of the triplet code.

**Rh50 Gene**—Fig. 3 schematically shows the nucleotide sequence of splice sites as well as the structure of exon/intron junctions in the Rh50 gene. All the 5' donor and 3' acceptor splice sites conform to the "GT-AG" rule and possess the consensus splicing signals (25). Of the 10 exons identified, only exon 6 is symmetrical, having intraxen codons GTT (Val270) and ACT (Thr315) at its 5' and 3' ends, respectively, whereas the other exons have either one or two split interexon codons (Fig. 3). One potential consequence of this type of exon/intron arrangement is that skipping of any single internal exon, except exon 6, during the splicing of Rh50 pre-mRNA would result in a shift in open reading frame and, therefore, alter the encoded amino acid sequence downstream of the skipped exon.

**Expression of Rh50, Rh30, and CD47 mRNAs in Normal and Rhnull Cells**—To identify the molecular defect underlying the Rhnull disease, the expression of candidate genes encoding the Rh50, Rh30, and CD47 proteins was characterized by RT-PCR and nucleotide sequencing. The full-length cDNA of Rh30 or CD47 was readily detectable in normal and Rhnull erythroid cells (gels not shown), indicating a comparable expression of the corresponding mRNA. Sequencing showed that the Rh30 or CD47 cDNA from Rhnull was normal and that the Rh30 cDNA contained both RhD and RhCe, indicating that the patient was a Dce/Dce homozygote. Definition of this Rh genotype by transcript analysis was in full agreement with the result of DNA typing by SpI polymorphisms (15). These data showed that the Rh30 or CD47 locus itself is not responsible for the disease phenotype.

However, RT-PCR analysis of Rh50 gene expression in erythroid cells revealed an important difference between the normal and Rhnull patient. Although there was no apparent change in size of the 5' portion of Rh50 cDNA encompassing exons 1–5, the 3' portion of Rh50 cDNA encompassing exons 4–10 always showed a truncation in the Rhnull patient (Fig. 4A). This finding indicated that the Rh50 mRNA from Rhnull could be an aberrantly spliced form lacking a portion of the 3' sequence. Indeed, sequencing showed that the 122-bp sequence of exon 7 was excluded from the truncated cDNA, resulting in the connection of exon 6 to exon 8 (Fig. 4B). To determine whether the skipping was complete or partial, a 3' RACE reaction was carried out using 7 s and 3'-UTR primers. A cDNA product of expected size (376 bp) was found in normal controls but not in the Rhnull patient (Fig. 4C), indicating that no splicing of exon 7 occurred for the Rh50 primary transcript. Further studies showed that this exon skipping was not seen in 15 normal subjects nor in other Rhnull Patients examined; thus, it could not be a constitutive splicing or regulated alternative splicing event.

**Identification of Rhnull-associated Donor Splice Site Mutation in Rh50 Gene**—The complete absence of exon 7 associated with Rh50 cDNA suggested strongly that either a splicing defect or a genomic deletion was present in the cognate gene. To define the nature of the underlying mutation, amplification from Rhnull genomic DNA of a segment encompassing exon 7 of the Rh50 gene was attempted. A fragment of 354 bp in size was detected, excluding the possibility of gene deletion. Sequencing of this fragment on both strands led to the identification of a single G → A mutation in the invariant GT element (+1 position) of the 5' donor splice site attached to exon 7 (Fig. 5A). Sequencing of other exon/intron junctions amplified with intron-specific primers (data not shown) confirmed this mutation to be the only structural alteration in the Rh50 gene.

Because the mutation abolished a PmlI restriction site (CAC ↓ GTG) (Fig. 3), and produced the novel NlaIII site (↓ CATG), a direct diagnostic assay was performed on amplified exon 7-containing fragments. The two enzymes showed an opposite cleavage pattern in normal and Rhnull fragments (Fig. 5B), confirming the mutation at the splicing junction. To demonstrate that the PmlI site was not caused by PCR spurious mutations, Southern blot of native genomic DNAs was hybridized with a probe spanning the exon7/intron 7 junction. As shown, the PmlI specific band was seen in normal but not in Rhnull (Fig. 5C). Given the observation of no dosage reduction in RH50, these results confirmed that the patient is homozygous for the G → A splicing mutation. Such a genotype assignment is consistent with the inheritance of Rhnull syndrome in an autosomal recessive fashion.

**Deduced Primary Sequence and Predicted Membrane Topology of Rh50 Mutant Protein**—To gain information on the primary structure of Rh50 glycoprotein, the Rhnull-associated Rh50 cDNAs were sequenced to completion. Compared with normal Rh50, no point mutation other than an absence of the sequence encoded by exon 7 was observed in the Rhnull patient (Fig. 6A). Because exon 7 is asymmetric in codon distribution at the 5' side (Fig. 3), its complete skipping and the subsequent joining of exon 6 with exon 8 inevitably resulted in an open reading frame shifting (Fig. 6A). In turn, the deduced translation product would be truncated and prematurely terminated, containing only 351 amino acid residues. This includes the loss of 41 amino acid residues encoded in exon 7 and gain of an entirely new sequence of 36 residues following Thr315 (Fig. 6A).

Compared with the wild-type Rh50 protein (8), hydrophathy plot analysis of the mutant form suggested two possible alterations in membrane organization of the C-terminal region.
Structure of RH50 and Molecular Genetics of Rh_null Disease

Fig. 4. Analysis of Rh50 transcript expression in normal and Rh_null erythroid cells. RT-PCR analysis of Rh50 transcript was carried out using 3'-UTR primer for cDNA synthesis and two pairs of amplimers for cDNA amplification. The location, direction, and designation of primers with respect to the structure of Rh50 are specified. A, agarose gel electrophoresis of amplified Rh50 cDNA products from RhD+, RhD−, and Rh_null. The size of segment 4s-10a from Rh_null is smaller than that of controls, indicating a deletion in the region spanning exons 5–10. Note that the Rh_null lanes were overloaded. B, nucleotide profiles of the exon/exon boundary associated with exon skipping. Exon boundary is indicated by a vertical arrow. In normal, exon 6 is joined to exon 7, whereas in Rh_null exon 7 is absent, resulting in exon 6 to exon 8 connection. C, 3' RACE assay for the functional splicing of exon 7 in Rh50 pre-mRNAs. A primer anchored in exon 7, 7s, was coupled with 3'-UTR primer for 3' RACE reaction. The expected cDNA product of 376 bp is clearly seen in control lanes but not the Rh_null lane, confirming a complete exclusion of exon 7 from the latter.

Discussion

Rh50 glycoprotein is a critical coexpressor of Rh30 polypeptides, the carriers of erythrocyte Rh antigens (1–4). Here, the exon/intron structure of Rh50 gene has been delineated, which should facilitate identification of mutations underlying the suppressor forms of Rh deficiency syndrome. A homology-based approach coupling with bidirectional walking revealed that Rh50 is a single copy gene with 10 exons and has a global organization strikingly similar to its related Rh30 members (23, 24). Both the structural conservation and sequence homology of the two genes are confined mainly to exons 2–9, while their 5' and 3' regions, including the promoter and untranscribed sequences, share little or no similarity. Since Rh50 and Rh30 genes are located on different chromosomes (5–8), these findings suggest that the two genetic loci might be formed by a rare transchromosomal insertion event. Our recent studies suggest that Rh50 and Rh30 genes originated from a common ancestor and were linked to each other following their initial duplication; later, one was translocated and diverged as the independent locus on a separate chromosome.4 Comparative analysis of the Rh50 and Rh30 gene orthologues in lower organisms should help decipher the evolutionary pathway ultimately leading to the establishment of two genetic loci encoding the Rh family proteins in Homo sapiens.

The extreme rareness, recessive nature, and consanguineous background of Rh deficiency syndrome (11, 12) point to a heterogeneous spectrum of the underlying mechanisms. At present, the molecular defect at RH30 locus responsible for the amorph type of Rhnull remains unknown (13–15). Nevertheless, several lines of evidence suggest that the RH50 locus is the

4 C.-H. Huang, J. Cheng, Y. Chen, and Z. Liu, unpublished observations.
prime target of suppressor mutations resulting in the regulator Rhnull disease. (i) Rh50 is thought to directly interact with Rh30, and the deficiency of the two proteins in the plasma membrane occurs in parallel (9, 26). (ii) Despite a close link of Rhnull with absence or deficiency in GPB, Duffy, or LW, the erythrocytes lacking these glycoproteins per se exhibit no change in the Rh antigen expression and no apparent perturbations in membrane physiology and cell morphology (27–30). Presumably these proteins are casually associated components not essential for the interaction and membrane assembly of Rh family proteins. (iii) Although CD47 is also reduced in Rhnull state, its low level of expression is restricted to erythroid cells but not to other hematopoietic cells (31, 32), suggesting that CD47 deficiency occurs as the consequence of, rather than the cause for, the defect in Rh complex formation. (iv) More recently, two small DNA deletions causing frameshift in the Rh50 gene have been found to be associated with the regulator Rhnull phenotype in unrelated patients (16).

Our previous studies showed that this Rhnull patient had a grossly intact RH30 locus occurring in the form of DCe/DCe haplotype combination (15). The present study confirmed this assignment and showed further that the RH30 locus gave rise to expression of both RhD and RhCe transcripts with sequences identical to that from normal subjects. These results, together with the identification of a normal CD47 gene, exclude the involvement of mutations of RH30 or CD47 locus in this Rhnull patient. However, transcript analysis showed that there was no expression in the Rh null cells of any full-length form of Rh50 mRNAs except the shortened one specifically lacking the sequence of exon 7. Genomic sequencing revealed the occurrence of a homozygous G → A mutation in the invariant GT element of 5’ donor splice site as the only alteration in the Rh50 gene. These findings establish the pre-mRNA splicing defect, for the first time, as the suppressor mutation of RH50 leading to a loss-of-function phenotype characteristic of the regulator form of Rhnull disease.

Mutations in the GT and AG motifs of the donor and acceptor splice sites, the cis-acting elements essential for pre-mRNA splicing (33), portray an important mechanism for the origin of human genetic diseases (34). The donor splice site mutation described here has caused a complete skipping of exon 7 from the mature form of Rh50 mRNA in the Rhnull patient. Signifi-
interest to note that such mutations all target the TM domains in the C-terminal half that are conserved in the Rh50 homologues from the mouse to C. elegans. Currently, little is known about how the disruption of the Rh protein complex causes the multiple facets of structural and functional abnormalities in the Rh-deficient erythrocytes. There is also a lack of general information regarding the involvement and coordination of possible intracellular factor(s) in the functioning of the Rh membrane complex. A full description of Rh-null disease mutations and assessment of their phenotypic effects in model systems, such as C. elegans, should lead to a better understanding of the membrane assembly and structure/function relations of the Rh family of proteins.

Acknowledgment—I am particularly grateful to Y. Okubo and M. Reid for providing and typing the Rh-null blood sample used in this investigation. I thank Y. Chen for technical assistance, and T. Ye for help in the construction of human Marathon genomic libraries. I also thank O. O. Blumenfeld and C. Redman for comments on the manuscript.

REFERENCES
1. Huang, C.-H. (1997) Curr. Opin. Hematol. 4, 94–103
2. Cartron, J.-P., and Agre, P. (1995) Blood Cell Biochem. 6, 189–225
3. Cartron, J.-P. (1994) Blood Rev. 8, 199–211
4. Avent, N. D., Liu, W., Warner, K. M., Mawby, W. J., Jones, J. W., Ridgwell, K., Hermand, P., Salmon, C., Cartron, J.-P., and Colin, Y. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6243–6247
5. Avent, N. D., Ridgwell, K., Tanner, M. J. A., and Anstee D. J. (1996) Biochem. J. 321, 821–825
6. Le Van Kim, C., Moiro, J., Cherif-Zahar, B., Raynal, V., Cherrier, C., Cartron, J.-P., and Colin, Y. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 10925–10929
7. Ridgwell, K., Spurr, N. K. Laguda, E. MacGee, C., Avent, N. D., and Tanner, J. M. A. (1992) Biochem. J. 287, 223–236
8. Evers, S. A. S., Ridgwell, K., Mawby, W. J., and Tanner, M. J. A. (1994) J. Biol. Chem. 269, 6417–6420
9. Avent, N. D., Liu, W., Warner, K. M., Mawby, W. J., Jones, J. W., Ridgwell, K., and Tanner, M. J. A. (1996) J. Biol. Chem. 271, 14233–14239
10. Rac, R., and Sanger, R. (1975) Blood Groups in Man, 6th Ed., pp. 78–260, Blackwell Scientific Publications, Oxford, London
11. Nash, R., and Shojania, A. M. (1987) Am. J. Hematol. 26, 401–422
12. Cherif-Zahar, B., Raynal, V., Gane, P., Mattei, M.-G., Bailly, P., Gibbs, B., Colin, Y., and Cartron, J.-P. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10925–10929
13. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. (1986) Science 230, 487–491
14. Huang, C.-H., Reid, M. E., Chen, Y., Coghlan, G., and Okubo, Y. (1996) Am. J. Hum. Genet. 58, 133–142
15. Cherif-Zahar, B., Raynal, V., Gane, P., Mattei, M.-G., Bailly, P., Gibbs, B., Colin, Y., and Cartron, J.-P. (1996) Nat. Genet. 12, 168–173
16. Guussens, M., and Kan, Y. W. (1981) Methods Enzymol. 76, 805–817
17. Huang, C.-H., Guizzzi, M. L., McCreary, J., Leigh, E. M., and Blumenfeld, O. O. (1991) Blood 77, 381–386
18. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. (1988) Science 239, 487–491
19. Huang, C.-H., Chen, Y., Reid, M. E., and Ghosh, S. (1996) Am. J. Hum. Genet. 59, 825–833
20. Campbell, I. G., Freemont, P. S., Foulkes, W., and Trowsdale, J. (1992) Cancer Res. 52, 5416–5420
21. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
22. Huang, C.-H. (1996) Blood 88, 2326–2333
23. Cherif-Zahar, B., Le Van Kim, C., Bouillac, C., Raynal, V., Cartron, J.-P., and Colin, Y. (1994) Genomics 19, 68–74
24. Shapiro, M. B., and Senanayake, P. (1987) Nucleic Acids Res. 15, 7156–7174
25. Mallinson, G., Aanette, D. J., Avent, N. D., Ridgwell, K., Tanner, M. J. A., Daniels, G. L., and Tippett, P., and von dem Borne, A. E. G. (1990) Blood 77, 841–848
26. Lindberg, F. P., Lublin, D., Melen, J. M., Veile, R. A., Miller, Y. E., Denis-Keller, H., and Brown, E. J. (1994) J. Biol. Chem. 269, 1567–1570
27. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986) Annu. Rev. Biochem. 55, 1119–1150
28. Krawczak, M., Reiss, J., and Cooper, D. N. (1992) Hum. Genet. 90, 41–54

FIG. 6. Amino acid sequence and predicted membrane topology of mutant Rh50 in Rh_null disease. A, comparison of the primary structure between the wild-type (wt) and mutant (mt) Rh50 glycoproteins. The mutant lacks 41 amino acids (dashes) encoded by exon 7, but gains a new 36-amino acid sequence (bold) due to a frameshift and premature termination (preterm). Note that in both the mutant and wild types, the amino acid at position 242 is occupied by Asn (N) with an asterisk (*) but not by Asp (D) as reported (8). This Asn is seen in all unrelated normal and Rh_null individuals examined (n > 15). B, model for membrane topology of the mutant Rh50 protein. Also shown is the hydrophobic profile of wild-type Rh50 with 12 TM domains connected by short loops on either side of the lipid bilayer. "Y" indicates the N-glycan on Asn7 (9). The mutant Rh50 protein lacks the last two TM domains and carries an extended C-terminal sequence most likely facing the cytoplasmic space. Its features, including the underlying defect and associated structural alterations, are summarized at right margin.