Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 5306–5310, June 1994
Medical Sciences

The LW blood group glycoprotein is homologous to intercellular adhesion molecules

(erythrocyte membrane/LW antigen/cDNA/CD4)

PASCAL BAILLY*, PATRICIA HERMAND*, ISABELLE CALLEBAUT†, HANS H. SONNEBORN‡, SAMIR KHAMLICHI*, JEAN-PAUL MORNON†, AND JEAN-PIERRE CARTRON*§

*Institut National de la Santé et de la Recherche Médicale Unité U76, Institut National de Transfusion Sanguine, 6 rue Alexandre Cabanel, F-75739 Paris Cedex 15, France; ‡Biotest AG Research Department, Geleitsstrasse 103, 6050 Offenbach, Germany; and †Département des Macromolécules Biologiques, Laboratoire de Minéralogie et de Cristallographie, Centre National de la Recherche Scientifique Unité de Recherche Associée, UP6/UP7, 4, Place Jussieu, F-75252 Paris Cedex 05, France

Communicated by Victor A. McKusick, February 3, 1994

ABSTRACT **The LW blood group antigens reside on a 42-kDa erythrocyte membrane glycoprotein that was purified by immunoaffinity and partially sequenced. From this information, a specific PCR-amplified DNA fragment was used to screen a λgt11 human bone marrow cDNA library. Two forms of cDNA were isolated; the first encoded a single spanning transmembrane protein of 270 amino acids, including a 29-amino acid peptide signal and four potential N-glycosylation sites, and the second encoded a shortened protein form of 236 residues devoid of transmembrane and cytoplasm domains. A rabbit antibody raised against the 15 N-terminal amino acids of the predicted protein reacted on immunoblots with authentic LW glycoprotein and in indirect agglutination test with all human erythrocytes except those from LW(a−b−). This showed that the protein encoded by these clones was LW gene product and suggested that the N terminus of the LW protein is oriented extracellularly. Most interestingly, the LW protein was found to exhibit sequence similarities (with ≈30% identity) with intercellular adhesion molecules ICAM-1, -2, and -3, which are the counterreceptors for the lymphocyte function-associated antigens LFA-1. The extracellular domain of LW consists, like that of ICAM-2, of two immunoglobulin-like domains, and the critical residues involved in the binding of LFA-1 to ICAMs were partially conserved in LW.**

The LW and Rh (rhesus) blood group systems were discovered simultaneously and were confused for a long time (reviews, refs. 1–3). It is now clear that these systems are genetically independent but are closely associated at the phenotypic level, since erythrocytes that are deficient for Rh antigens are also deficient for LW antigens (see ref. 4). However, rare individuals who lack LW antigens have been found among Rh-positive individuals. These observations served as a basis for a genetic theory suggesting that Rh and LW might have evolved from the same substrate (5), although no biochemical evidence was provided.

Biochemical investigations indicate that the LW antigens are carried by a 40- to 42-kDa glycoprotein that is linked to the membrane skeleton (6–8) and requires intramolecular disulfide bonds for antigenic reactivity (9). When deglycosylated, the LW protein is reduced to a 25-kDa apoprotein that is still reactive with anti-LW antibodies (8, 10). Moreover, the LW antigens were inactivated by EDTA (not EGTA) but could be restored by addition of Mg^{2+} cations (10). In contrast with the genetic theory discussed above, comparative analysis by two-dimensional iodopeptide mapping of the Rh and LW proteins suggested that LW was not a glycosylated form of Rh nor is Rh a precursor of LW (11).

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.

To obtain further information on the structure and function of the LW glycoprotein, this molecule was immunopurified, partially sequenced, and cloned.¶ We found that LW exhibited a striking similarity with intercellular adhesion molecules (ICAMs), which are the counterreceptors for the lymphocyte function-associated antigens LFA-1.

MATERIALS AND METHODS

Reagents. Common blood samples and Rh_{null} sample (donor Fri.) were from the Institut National de Transfusion Sanguine (Paris). LW(a−b−) erythrocytes (Mil.) were a gift from V. Taliano (Canadian Red Cross, Montreal) and LW(a−b+) erythrocytes (Bis.) were from L. Mannessier (Centre de Transfusion Sanguine, Lille, France). Murine monoclonal anti-LW^{ab} antibody (BS46) has been described (12).

Affinity Purification of the LW Protein. Membranes from two units of LW(a+b−) red cells were solubilized with 1% (wt/vol) Triton X-100 in phosphate-buffered saline (PBS) and applied to a specific affinity matrix column, prepared by binding 9 mg of purified murine monoclonal IgG antibody anti-LW (BS46) to 2 ml of protein A-agarose followed by cross-linking of the complex with dimethylpimelimidate (ImmunoPure IgG orientation kit, Pierce). After washing, the LW antigenic material bound was eluted with a glycine buffer (pH 2.8) and immediately brought to near neutrality.

Oligonucleotide Primers and Probes. Deoxyinosine (I) was incorporated where codon degeneracy exceded three. Sense primers LW.6c and LW.7c (5′-ATG TCI CCI GAR TTY GT-3′ and 5′-ATG AGI CCI GAR TTY GT-3′, respectively) encoded amino acids MSPEFV (peptide 5). Antisense primer LW.13 (3′-TAP TGI CGI ATR TTY GG-5′) encoded ITAYKP (peptide 13) and antisense LW.14 primer (3′-ATR TTY GGI GGI GTR-5′) encoded YKPPH of the same peptide (see Table 1). In the primers, P = T, G, or A, R = G or A, and Y = T or C. Poly(A)+ RNAs from spleen erythroblasts of a β-thalassemic patient were prepared as described (13) and purified on oligo(dT)-cellulose column. First cDNA strands were synthetized with primer LW.14. From this template, fragments amplified by PCR (annealing temperature, 45°C; 35 cycles) between primers LW.6c, LW.7c, and LW.13 (1 μg each) were analyzed by Southern blot with the LW.10c probe (5′-GAA/G TTT/C GTG/C GCI GTG/C CAA/G CC-3′) deduced from the internal sequence (EFVAVQPGK) of peptide 5.

Abbreviations: ICAM, intercellular adhesion molecule; LFA-1, lymphocyte function-associated antigen 1.
§To whom reprint requests should be addressed.
¶The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L27670 and L27671).

5′ End Determination by PCR. First strand cDNA was synthesized with primer LW.K (nt 540–517) and used as template in the 5′-Ampli-Finder Race kit from Clontech. After ligation of a single-stranded oligonucleotide anchor directly to the 3′ end of the first strand cDNA, PCR amplification was carried out between a primer complementary to the anchor and the antisense primer LW.D (nt 135–112). Positive clones were identified by Southern blot analysis and hybridization with the internal probe, LW.P (nt 99–70).

Antiserum Production. An N-terminal 15-amino acid peptide of the mature LW protein was synthetized and coupled to keyhole limpet hemocyanin (Neosystem, Strasbourg, France). Rabbits were immunized as described (14).

Sequence and Structure Analysis. The FASTA program (15) was used to perform initial searches in the sequence data banks. Three-dimensional manipulations were realized with the program MANOSK (16). Crystallographic data were taken from the Protein Data Bank (17).

RESULTS

Purification and Microsequencing of the LW Polypeptide. The LW polypeptide was purified from red cell membrane lysates by immunoaffinity with a murine monoclonal anti-LW antibody (BS46) covalently bound to protein A-agarose. The LW protein was specifically absorbed and no LW-positive material was detected in the flow through (Fig. 1, lane b). The material eluted from the affinity matrix was analyzed on a SDS/polyacrylamide gel and stained by Coomassie blue (not shown), silver, and immunoblot with the BS46 antibody (Fig. 1). A strong band of 42 kDa and a faint band at 85 kDa were detected. Both bands were immunostained by BS46 and most likely represented LW monomer and dimer, respectively (Fig. 1, lane c). No additional bands were detected after silver staining (Fig. 1, lane d). The purified fractions containing the LW protein were pooled, concentrated, and used for N-terminal and internal peptide sequence determination following trypsin cleavage in the presence of detergent (18). The N-terminal sequence of 24 residues was derived with three provisional determinations and one undetermination, whereas the tryptic peptides 3, 5, and 13 each were composed of 4, 12, and 15 identified residues, respectively (Table 1).

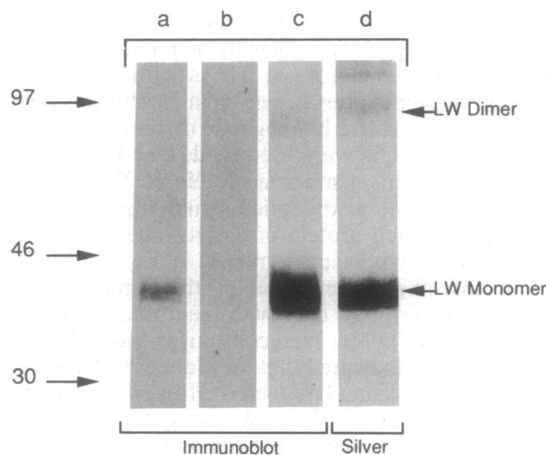

FIG. 1. Isolation of LW antigen by immunoaffinity. Fractions were separated by SDS/PAGE (12% acrylamide) under nonreducing conditions, blotted to nitrocellulose, and incubated with the BS46 antibody (5 μg/ml). Specifically bound antibody was detected by the alkaline phosphatase-conjugated substrate kit. Lane a, red cell lysate before immune absorption; lane b, lysate after immune absorption; lane c, LW polypeptide eluted from the BS46 matrix; lane d, silver staining of purified LW polypeptide. Arrows indicate the migration position of protein markers (left; in kDa) and LW polypeptide (right).

Table 1. N-terminal and tryptic peptide sequences from the LW glycoprotein

Peptide	Amino acid sequence
N-terminal	AQSPKGSPLAPSG(G)SVPFXVRM(S)(P)
Peptide 3	WATS(R)
Peptide 5	MSPEFVAVQPGK
Peptide 13	ITAYKPPHSVILEPP

Tryptic peptides were separated by HPLC on a DEAE C_{18} column and sequenced with a gas/liquid solid-phase sequenator (Applied Biosystems, model 470 A). N-terminal sequence analysis was carried out by protein transfer onto ProBlott [poly(vinylidene difluoride) sheet, Applied Biosystems] using the Applied Biosystems protein sequencer (model 473A). ( ), Provisional determination; X, undetermined. The underlined amino acid sequence represents the peptide sequence used to produce a rabbit polyclonal antibody.

Peptide 5 most likely represented an extension of the N-terminal peptide.

Isolation and Characterization of the LW cDNA. Four primers designed from peptides 5 and 13 were used in reverse transcription-PCR with mRNAs prepared from human adult spleen erythroblasts. Sequence analysis of a 245-bp PCR product detected by Southern hybridization included structural information derived from peptide 3, thus indicating that this fragment was specific for the purified LW protein. Four clones (I, II, III, and IV) were isolated from a human bone marrow cDNA library (1.6 × 10^6 recombinant λgt11 phages) screened with the 245-bp probe. Clones III and IV carried the largest inserts (1.0 kb and 1.3 kb, respectively).

As the cDNA insert from clone III could not be excised, it was subcloned and sequenced after PCR amplification using λgt11 forward and reverse primers. This cDNA contains nt 238–1256 (Fig. 2*A*). The 3′ end was terminated with a poly(A) tract and several potential polyadenylylation signals localized between nt 1170–1192 and nt 1219–1224.

Digestion of clone IV with *Eco*RI yielded two inserts of 0.2 kb and 1.1 kb that were subcloned and sequenced. Clone IV sequence corresponded to nt 40–1192 (Fig. 2*A*) and exhibited the same sequence as clone III in their overlapping region (nt 238–1192), except at position 704 where an insertion of 147 nt was found (Fig. 2*B*). This additional sequence altered the reading frame and generated a premature stop codon at position 718. The predicted translated product of the clone IV cDNA corresponds exactly to the N-terminal sequence of the LW polypeptide determined by Edman degradation, except for residues 14 and 19, which were predicted as Gly and Xaa and found as Thr and Trp, respectively, from the nucleotide sequence.

5′ End Determination of the LW Message. The 5′ end sequence encoding the full length mRNA was cloned by a modified rapid amplification of cDNA ends technique (19). Accordingly, a cDNA segment of 184 bp was generated that represented the first 135 nt at the 5′ end from primer LW.D, in addition to the 49 bp derived from the Ampli-Finder anchor and anchor primers (see *Materials and Methods*). This fragment hybridized with the internal probe LW.P and exhibited a complete sequence identity at its 3′ end with the expected 96-base overlap of clone IV. The 5′ end region was found to contain 9 bp of 5′ untranslated region, the initiating ATG codon at position 10, and the beginning of the signal peptide, which was missing in clone IV.

Amino Acid Sequence of the LW Protein. The combined nucleotide sequence of clones III and IV, including the 5′ end region, predicted a first open reading frame of 810 nt for clone III (Fig. 2*A*) and a second of 708 nt for clone IV (Fig. 2*B*).

In clone III, the longest open reading frame encoded a 270-amino acid polypeptide chain initiated by the ATG codon at nt 10 and terminated by the stop codon TAA at nt 820 (Fig. 2*A*). Protein sequencing indicated that the N terminus of the

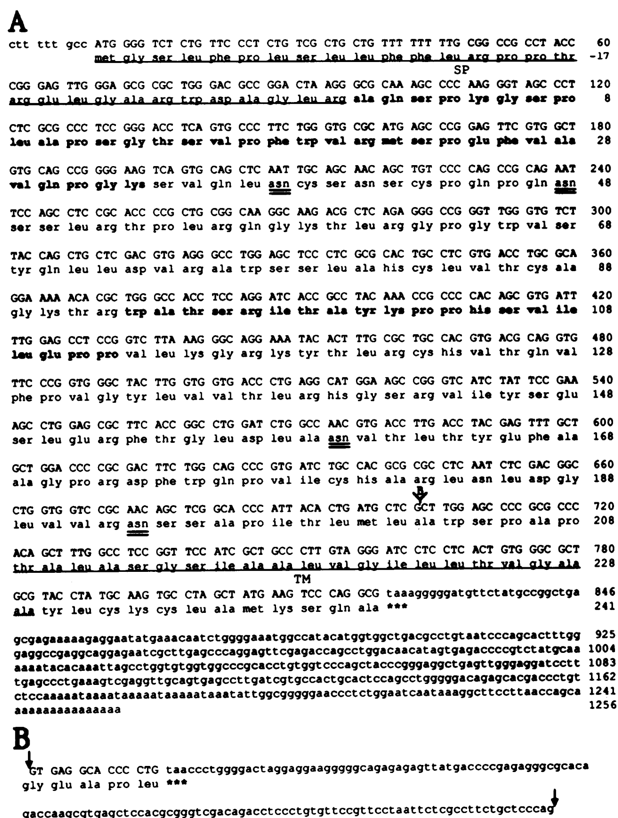

FIG. 2. Nucleotide sequence and predicted amino acid sequence of LW. (*A*) The DNA sequence is a composite derived from clones III and IV and from 39 nt of the 5′ upstream sequence (see text). Clone III contains the sequence from nt 238 down to the 3′ poly(A) tail, and clone IV contains the sequence from nt 40 to 1192, with an insertion at position 704 (arrow) of 147 nt shown in *B*. Amino acid sequence given in three-letter code and numbered beginning with 1 for the first amino acid of the mature protein. Hydrophobic putative signal peptide (SP) and transmembrane sequence (TM) are underlined. The N terminus and tryptic peptides of the mature protein are shown in boldface type. Potential N-glycosylation sites at Asn residues are underlined by double lines.

mature LW protein corresponded to Ala encoded by nt 97–99, thus suggesting that the first 29 amino acids of the predicted LW protein belong to a signal peptide with a typical hydrophobic character, as confirmed by hydrophobicity analysis (20). Secondary structure predictions indicated that the mature polypeptide of 241 residues (calculated molecular mass, 26.5 kDa) encoded by clone III was a transmembrane protein that consists of an extracellular domain of 208 amino acids, followed by a single hydrophobic domain of 21 residues (209–229) and by a cytoplasmic tail of 12 residues (230–241). The protein carried four predicted N-linked glycosylation sites (Asn-38, -48, -160, -193), which would result in a glycoprotein of 38 to 46 kDa if substituted by biantennary glycan chains. This is in agreement with size determinations of the LW glycoprotein on SDS gels in the native form (42 kDa) and after O-glycanase and N-glycanase digestion (25 kDa), as reported (10). Moreover, the presence of disulfide bonds between pairs of cysteine (Cys-39/Cys-83, Cys-123/Cys-180, and Cys-43/Cys-87), as predicted below, correlated well with the loss of BS46 antigenic activity when LW was reduced with dithiothreitol (7).

Clone IV encoded a shortened polypeptide of 236 amino acids sharing the same N-terminal sequence with the LW membrane protein but ending with a premature stop codon at position 718 (Fig. 2*B*). In the shortened form, the transmembrane and cytoplasmic domains were missing and replaced by a different and shorter C-terminal sequence resulting from a frameshift and premature termination generated by the 147-bp insertion.

Similarities of the LW Protein with ICAMs. Search in sequence data banks indicated that LW was structurally related to the ICAM-2 molecule (21) and to the two first immunoglobulin-like domains of ICAM-1 (22) and ICAM-3 (23–25) (Fig. 3*A*, Table 2). By analogy with the predicted ICAM-2 structure (26), the extracellular domain of LW would consist of two immunoglobulin-like domains (Fig. 3*B*). The first two domains of ICAMs have been assigned to the C2 set, having the C-type fold but showing sequence patterns in the second half of the domain that are more similar to V set than to C1 set sequences (26). For both immunoglobulin-like domains of LW, the current disulfide linkage between strands B and F (Cys-39 with Cys-83, Cys-123 with Cys-180) was present (29). An extra disulfide link (between Cys-43 and Cys-87) predicted between strands B and F at the top of domains 1 of ICAM-1, -2, and -3 (26, 30) was also conserved (Fig. 3*B*).

The interdomain arrangements of the LW polypeptide and CD4 three-dimensional structure would be similar because of stretches of similar sequences taking part in the packing (26) (Fig. 3*B*). Residues D26QPK, G46NN, P70DGQ, Q73, and E34 shown to be involved in the binding of ICAM-1 to LFA-1 (26–28) were only partially conserved between LW and other ICAMs (Fig. 3*A*).

Immunochemical Analysis. A rabbit antiserum raised against the N terminus of the mature LW protein reacted in the indirect agglutination test (titer of 1:100) with all human erythrocytes except those from LW(a−b−) individuals, thus suggesting a specificity related to LW. On immunoblot (Fig. 4), this antibody strongly reacted with monomeric and dimeric forms of the 42-kDa protein isolated from LW(a+b−) erythrocytes by immunoaffinity. In addition, it reacted with a 42-kDa protein present in all membrane proteins prepared from LW-positive red cells but was unreactive with those from an LW(a−b−) individual who lacked the LW protein. The antibody also reacted better with Rh-positive than Rh-negative membrane proteins. LW(a−b+), Rh-positive cells that lack only the LW^a antigen even reacted weakly.

DISCUSSION

The LW polypeptide from human erythrocytes was immunopurified with a murine monoclonal antibody directed against LW^{ab} antigens and partially sequenced. PCR amplification of human erythroblast RNAs with the primers deduced from these peptides generated a 245-bp specific probe that was used to screen a human bone marrow cDNA library. Two forms of cDNAs were identified. One form encoded a single spanning transmembrane protein of 270 amino acids, including a 29 amino acid peptide signal, and a second form encoded a shortened protein of 236 residues without transmembrane and cytoplasm domains. As a rabbit antibody raised against the N terminus of these mature proteins reacted on Western blot only with membrane proteins from LW(a−b+) or LW(a+b−) red cells but not from LW(a−b−) erythrocytes, this indicated that the cloned proteins were the direct products of the LW gene. However, whether the cloned protein encodes the LW^a or LW^b antigen is still unknown. Since the rabbit antibody agglutinated native red cells, it is most likely that the N terminus of the LW glycoprotein is exposed extracellularly. This is in contrast with previous studies based on carboxypeptidase digestion (10), but it is believed that the discrepancy relies on the use of large amounts of a carboxypeptidase possibly contaminated by trace amount of proteases (14).

The molecular characterization of the LW glycoprotein is of interest since this glycoprotein is absent from erythrocyte membranes of Rh-deficient individuals who suffer a generally

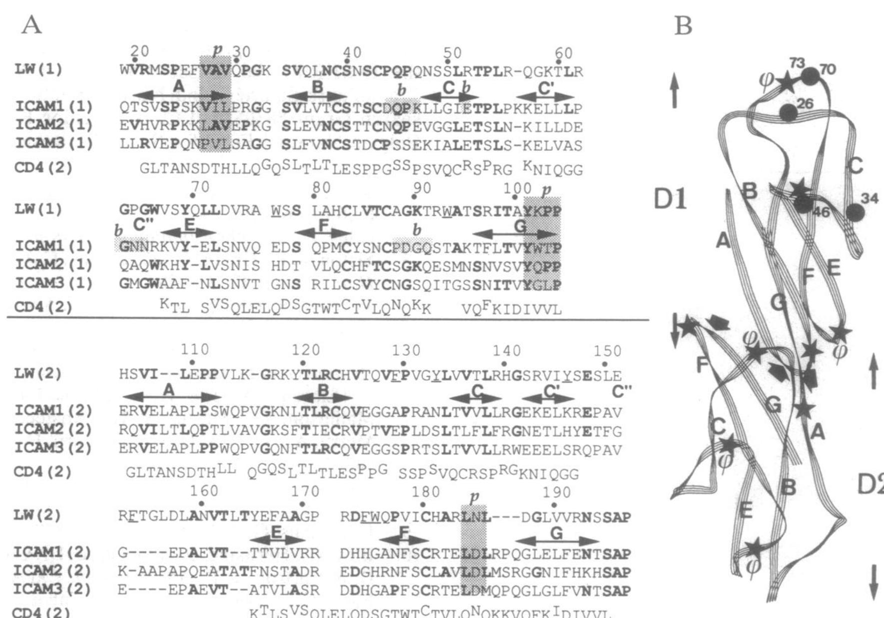

FIG. 3. Sequence alignment of the two immunoglobulin-like domains of LW and ICAMs and ribbon representation of the LW three-dimensional organization. (*A*) Amino acid numbering is only shown for LW, considered without its signal peptide. Alignment with ICAMs and CD4 sequences (second domain) is according to Berendt *et al.* (26). Sequence identities between ICAM and LW are shown in boldface type. Identities between ICAM or LW sequences and CD4 are shown with uppercase letters within the CD4 sequence. Positions of β strands (A, B, C, C′, E, F, G) are indicated above the ICAM-1 sequence. Sequences similar to those of CD4 that take part in the interdomain packing arrangement are boxed and highlighted by the letter *p* for "packing." Residues shown to be involved in the binding of ICAM-1 to LFA-1 (26–28) are boxed and highlighted by the letter *b* for "binding." Aromatic amino acids (F, Y, W) of LW that substitute hydrophilic amino acids of ICAM-1 and ICAM-2 are underlined. (*B*) The two first CD4 immunoglobulin-like domains were used to illustrate the major structural features of those of LW. The first CD4 domain (V set) has, however, been substituted by the second one (C2 set), as the two immunoglobulin-like domains of LW would both belong to the C2 set. The ribbon diagram depicts the strands of the model, labeled following the normal convention for immunoglobulin folds with A, B, and E forming one β sheet and C, F, and G forming the other. The sequence included between strands C and E is poorly conserved between the members of the ICAM family and was thought not to be included in either β sheet. Stars indicate regions where insertions or deletions have to be made on the CD4 structure in relation with the LW sequence. Predicted positions of LW corresponding to ICAM-1 or ICAM-2 residues involved in LFA-1 binding are shown by filled circles (●). Regions probably involved in interdomain packing (see above) are shown by thick arrows. Aromatic amino acids that substitute hydrophilic amino acids of the ICAM sequences are shown with the symbol ϕ.

well-compensated hemolytic anemia of varying severity with morphological and functional abnormalities (for reviews see refs. 5 and 31). Shortly after their discovery, Rh-deficient cells that lack all Rh structures were found to lack LW antigens as well (1), and it was speculated that the cumulative defects might be explained if Rh structures were the biochemical precursors for LW (4). That Rh proteins are probably not precursor of LW was suspected earlier following two-dimensional fingerprint analysis of purified Rh and LW proteins (11). However, the present results provide the definitive proof that Rh and LW are not biochemically related, since no similarity between the sequences of these molecules could be detected. The favored explanation for the multiple protein deficiencies in the Rh-deficient syndrome assumes that the Rh proteins are assembled within the membrane as a complex of unrelated proteins, including LW, maintained together by noncovalent linkages (31). The function of this complex, if any, is still unknown. Expression studies carried out by cotransfection of eukaryotic cells with the cDNA encoding each protein of the Rh complex will be crucial to address this issue.

Table 2. Identities between LW and ICAM immunoglobulin-like domains

		ICAM-1	ICAM-2	ICAM-3	CD4
LW	D1	30.2	30.4	30.4	9.7
	D2	25.8	27.2	26.9	8.1
ICAM-1	D1	—	33.9	36.3	13.3
	D2	—	34.5	77.4	10.5
ICAM-2	D1		—	36.5	12.2
	D2		—	34.6	11.3
ICAM-3	D1			—	13.4
	D2			—	11.6

Percent amino acid identity between immunoglobulin-like domains (D1 and D2) was determined with the FASTA program; only the first two immunoglobulin-like domains of ICAM-1 and ICAM-3 have been considered. *Z* scores between D1/D2 domains are highly significant: LW/ICAM-1, 17.4 SD above the mean; LW/ICAM-2, 28.1 SD above the mean; LW/ICAM-3, 21.5 SD above the mean.

The finding of two LW protein isoforms, a membrane-bound and a secreted form, was unexpected and raises new questions regarding the nature of the LW proteins and antigens within the cell membrane. Preliminary studies suggest that the secreted isoform might result from an aberrant splicing event (unspliced intron) rather than from a typical alternative splicing. Whether the secreted form is released or remains attached to the cell surface by protein interaction with the membrane-bound LW or with another membrane protein is presently unknown. However, immune precipitation and Western blot analysis of human red cell membrane proteins with anti-LW antibodies (human or murine) consistantly detected a single band of 42–46 kDa, without any trace of a related protein of lower mass (6–8). This may suggest, but not prove, that the secreted isoform of the LW protein is released from the cells.

Another interesting discovery from these studies is the striking structural similarity of LW with ICAM-1, -2, and -3, a series of well-characterized proteins involved in adhesion processes, which act as receptors for LFA-1 (32). Since the structural genes encoding LW and ICAM-1 colocalized to

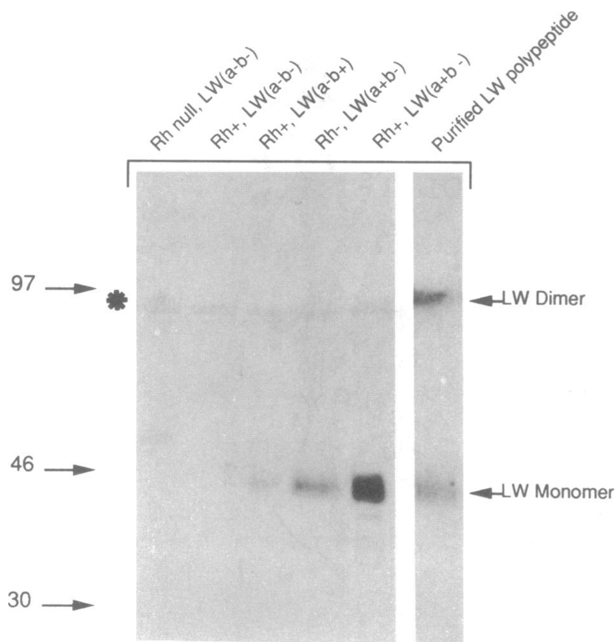

FIG. 4. Immunochemical analysis with a rabbit antibody directed against the N-terminal peptide of LW. Total membrane proteins from common and rare LW phenotypes were analyzed by SDS/PAGE and immunoblotted with the rabbit antibody to the N terminus of the LW protein (1:1000). Antibody molecules specifically bound to the LW polypeptide (42 kDa) were visualized as described in the legend to Fig. 1. The asterisk indicates an artifactual band in all membrane preparations, not present in the purified LW material.

chromosome 19p13 (33, 34), they may have evolved from a common ancestral gene. Whether the LW protein exhibits adhesion properties that play a role in the cohesion of the Rh membrane complex deserves further investigation.

Recent studies (26) suggested that LFA-1 interacts with the end and with one face of the first domain of ICAM-1 formed predominantly by the CFG strands of one of the β sheets (box highlighted by the letter *b* for "binding" in Fig. 3*A* and filled circles in Fig. 3*B*), whereas erythrocytes infected by *Plasmodium falciparum* interacts with the opposite face. Whether LW could be a ligand for LFA-1 is uncertain since the residues involved in LFA-1/ICAM interaction are partially conserved. The same interrogation holds for a potential interaction of LW with the as yet unidentified ligand of *P. falciparum*-infected erythrocytes. Two crucial positions for LFA-1 binding display significant changes. Position E34, which is common to all ICAMs, was replaced by R52 in LW (Fig. 3*A*). It is known that the E34A mutation abolishes LFA-1 binding to ICAM-1 (30), but the effect of the E to R substitution is unknown. Position Q73 is also common to ICAMs but was substituted by T91 in LW. Interestingly, a Q73T mutation in ICAM-1 reduced LFA-1 binding by only 60% (30). It is possible, therefore, that LW may still bind LFA-1. The inability to detect LFA-1 binding to human red cells in previous studies (35, 36) might result both from a weak binding affinity and from the low copy number of LW molecules. Alternatively, aromatic amino acids (F, Y, W) within loops of LW connecting the β strands (Fig. 3*B*), which substituted hydrophilic amino acids of ICAMs, might play a role in interaction with other receptors that remain to be identified.

We thank M. Huet (Institute National de Transfusion Sanguine, Paris) for technical assistance, P.-H. Roméo (Institut National de la Santé et de la Recherche Médicale Unité U91, Créteil, France) for the gift of β-thalassemic erythroblasts, and J. D'Alayer (Institut Pasteur, Paris) for protein sequencing. This work was supported by the INSERM, the Caisse Nationale d'Assurances Maladies des Travailleurs Salariés, and North Atlantic Treaty Organization Research Grant 88/0556. I.C. is a Senior Research Assistant of the Belgian National Fund for Scientific Research.

1. Race, R. R. & Sanger, R. (1975) *Blood Groups in Man* (Blackwell, Oxford), 6th Ed.
2. Giles, C. M. (1980) *Immunol. Commun.* **9,** 225–242.
3. Sistonen, P., Green, C. A., Lomas, C. G. & Tippett, P. (1983) *Ann. Hum. Genet.* **47,** 277–284.
4. Race, R. R. (1965) *Ann. NY Acad. Sci.* **127,** 884–891.
5. Tippett, P. (1990) *Transf. Med. Rev.* **4,** 56–68.
6. Moore, S. (1983) in *Red Cell Membrane Glycoconjugates and Related Genetic Markers*, eds. Cartron, J. P., Rouger, P. & Salmon, C. (Librairie Arnette, Paris), pp. 96–106.
7. Mallinson, G., Martin, P. G., Anstee, D. J., Tanner, M. J. A., Merry, A. H., Tills, D. & Sonneborn, H. H. (1986) *Biochem. J.* **234,** 649–652.
8. Bloy, C., Blanchard, D., Hermand, P., Kordowicz, M., Sonneborn, H. H. & Cartron, J. P. (1989) *Mol. Immunol.* **26,** 1013–1019.
9. Konighaus, G. J. & Holland, T. L. (1984) *Transfusion* **24,** 536–537.
10. Bloy, C., Hermand, P., Blanchard, D., Chérif-Zahar, B., Goossens, D. & Cartron, J. P. (1990) *J. Biol. Chem.* **265,** 21482–21487.
11. Bloy, C., Hermand, P., Chérif-Zahar, B., Sonneborn, H. H. & Cartron, J. P. (1990) *Blood* **75,** 2245–2249.
12. Sonneborn, H. H., Uthemann, H., Tills, D., Lomas, C. G., Shaw, M. A. & Tippett, P. (1984) *Biotest. Bull.* **2,** 145–148.
13. Chirgwin, J. M., Praybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
14. Hermand, P., Mouro, I., Bloy, C., Suyama, K., Goldstein, J., Cartron, J. P. & Bailly, P. (1993) *Blood* **82,** 669–679.
15. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
16. Cherfils, J., Vaney, M. C., Morize, I., Surcouf, E., Colloc'h, N. & Mornon, J. P. (1988) *J. Mol. Graphics* **6,** 155–160.
17. Berstein, F. C., Koetzle, T. F., William, G. J. M., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, I., Shimanouchi, T. & Tasumi, M. (1987) *J. Mol. Biol.* **112,** 532–542.
18. Rosenfeld, J., Capdeville, J., Guillemot, J.-C. & Ferrara, P. (1992) *Anal. Biochem.* **203,** 173–179.
19. Edwards, J. B. D. M., Delort, R. & Mallet, J. (1991) *Nucleic Acids Res.* **19,** 5227–5232.
20. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
21. Staunton, D. E., Dustin, M. L. & Springer, T. A. (1989) *Nature (London)* **339,** 61–64.
22. Simmons, D., Makgoba, M. & Seed, B. (1988) *Nature (London)* **331,** 624–627.
23. Fawcett, J., Holness, C. L. L., Needham, L. A., Turley, H., Gatter, K. C., Mason, D. Y. & Simmons, D. L. (1992) *Nature (London)* **360,** 481–484.
24. Vazeux, R., Hoffman, P. A., Tomita, J. K., Dickinson, E. S., Jasman, R. L., St John, T. & Gallatin, W. M. (1992) *Nature (London)* **360,** 485–488.
25. de Fougerolles, A. R., Klickstein, L. B. & Springer, T. A. (1993) *J. Exp. Med.* **177,** 1187–1192.
26. Berendt, A. R., McDowall, A., Craig, A. G., Bates, P. A., Sternberg, M. J. E., Marsh, K., Newbold, C. I. & Hogg, N. (1992) *Cell* **68,** 71–81.
27. Diamond, M. S., Staunton, D. E., Marlin, S. D. & Springer, T. A. (1991) *Cell* **65,** 839–847.
28. Staunton, D. E., Dustin, M. L., Erickson, H. P. & Springer, T. A. (1990) *Cell* **61,** 243–254.
29. Williams, A. F. & Barclay, A. N. (1988) *Annu. Rev. Immunol.* **6,** 381–405.
30. Giranda, V. L., Chapman, M. S. & Rossmann, M. G. (1990) *Proteins* **7,** 227–233.
31. Cartron, J. P. & Agre, P. (1993) *Semin. Haematol.* **30,** 193–208.
32. Springer, T. A. (1990) *Nature (London)* **346,** 425–434.
33. Sistonen, P. (1984) *Ann. Hum. Genet.* **48,** 239–242.
34. Trask, B., Fertitta, A., Christensen, M., Youngblom, J., Bergmann, A., Copeland, A., de Jong, P., Mohrenweiser, H., Olsen, A., Carrano, A. & Tynan, K. (1993) *Genomics* **15,** 133–145.
35. Beatty, P. G., Ledbetter, J. A., Martin, P. J., Price, T. H. & Hansen, J. A. (1983) *J. Immunol.* **131,** 2913–2918.
36. Campana, D., Sheridan, B., Tidman, N., Hoffbrand, A. V. & Janossy, G. (1986) *Eur. J. Immunol.* **16,** 537–542.