

IDENTIFICATION OF DISTINCT C3b AND C4b  
RECOGNITION SITES IN THE HUMAN C3b/C4b  
RECEPTOR (CR1, CD35) BY DELETION MUTAGENESIS

BY LLOYD B. KLICKSTEIN,\*<sup>||</sup> THOMAS J. BARTOW,<sup>||</sup> VOJISLAV MILETIC,<sup>||</sup>  
LAUREL D. RABSON,<sup>||</sup> JOHN A. SMITH,<sup>‡,§</sup> AND DOUGLAS T. FEARON<sup>||</sup>

*From the \*Program in Cell and Developmental Biology and the †Department of Pathology,  
Harvard Medical School, Boston, Massachusetts 02115; the §Departments of Molecular Biology and  
Pathology, Massachusetts General Hospital, Boston, Massachusetts 02114; and the  
<sup>||</sup>Division of Molecular and Clinical Rheumatology, Department of Medicine,  
The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205*

Complement receptor type 1 (CR1, CD35) is a membrane glycoprotein that is present on erythrocytes, leukocytes, glomerular podocytes, and splenic follicular dendritic cells, and mediates the binding by these cells of particles and immune complexes that have activated complement (1, 2). This function of CR1 is dependent on its capacity to bind reversibly the C3b and C4b fragments of C3 and C4 that are covalently attached to activators of complement. CR1 also can inhibit complement activation by impairing the formation and function of the alternative and classical pathway C3/C5 convertases, and by serving as a cofactor for the cleavage by factor I of C3b to iC3b, C3c and C3d,g, and of C4b to C4c and C4d.

Four molecular weight allotypes of CR1 have been described that vary by increments of 40,000–50,000, and each is able to mediate binding of C3b (1, 3). The most frequently occurring F or A allotype has an  $M_r$  after reduction of 250,000 on SDS-PAGE. The receptor is comprised of a single polypeptide chain and has an estimated six to eight *N*-linked complex oligosaccharides and no *O*-linked carbohydrate. The amino acid sequence of ~75% of the extracellular region, the single 25-amino acid membrane spanning domain, and the 43-amino acid cytoplasmic sequence has been determined by sequence analysis of overlapping cDNA clones (4). The extracellular domain consists of a series of tandemly arranged short consensus repeats (SCRs)<sup>1</sup> of 60–70 amino acids, each SCR having four conserved cysteines and a consensus sequence involving ~40% of the residues. Every eighth SCR is a highly homologous repeat, such that SCR-1, -8, and -15, SCR-2, -9, and -16, etc. are 65–100% identical. Thus, seven SCRs constitute a long homologous repeat (LHR). This earlier study presented the sequence of three LHRs, and a fourth NH<sub>2</sub>-terminal LHR was predicted for the F allotype (4).

Although the LHR appears to be unique to CR1, the basic SCR structural element has been found in other C3/C4-binding proteins such as factor H, C4b-binding

This work was supported by grants AI-22833 and GM-07753 from the National Institutes of Health, the Arthritis Foundation Biomedical Research Center, and Hoechst AG, Frankfurt am Main, Federal Republic of Germany.

<sup>1</sup> Abbreviations used in this paper: LHR, long homologous repeat; SCR, short consensus repeat.

protein (C4bp), decay accelerating factor, complement receptor type 2, factor B, C1r, C1s, C2, and membrane cofactor protein (1, 5). The presence of the SCR in the noncomplement proteins, the IL-2-R, factor XIIIb, and  $\beta$ -2-glycoprotein I, indicates that the SCR, although capable of forming a C3/C4-binding site, does not necessarily have this activity.

The present study is an analysis of the SCRs of CR1 necessary to form the C3b/C4b-binding sites of CR1. The sequence of the NH<sub>2</sub>-terminal LHR, LHR-A, has been deduced from cDNA clones encoding this region of the receptor. A full-length cDNA clone containing all of the coding sequence of the F allotype of CR1 has been constructed and expressed in L cells and in COS cells. The function of this recombinant CR1 and of CR1-derived proteins encoded by a series of deletion mutants has been assessed and has permitted the mapping of functional domains.

## Materials and Methods

**Construction of a cDNA Library.** A selectively primed cDNA library,  $\lambda$ HH, was constructed from 3  $\mu$ g of poly(A)<sup>+</sup> RNA purified from DMSO-induced HL-60 cells (6-8) as described with the modification that LK35.1, a 35-mer oligonucleotide (5'-TGAAGTCATC ACAG-GATTTT ACTTCACATG TGGGG-3'), was used in place of oligo (dT)<sub>12-18</sub>, and 40  $\mu$ Ci of  $\alpha$ -[<sup>32</sup>P]dCTP was added during second strand synthesis. One third of the cDNA was cloned in  $\lambda$ gt11 and 750,000 independent recombinants were obtained. The size-selected human tonsil library,  $\lambda$ S2T, has been previously described (4, 9) (American Type Culture Collection accession No. 37546).

**Isolation of Clones, Probes, and DNA Sequence Analysis.** The probes used for screening cDNA libraries study were CR1-1 (9) (American Type Culture Collection, accession No. 57331), CR1-2 (9), CR1-4 (10), and CR1-18, a 252-bp Sau 3AI fragment from the 0.5-kb Eco RI fragment of cDNA clone  $\lambda$ H3.1 corresponding to nucleotides 101-352 in Fig. 1. Under conditions of high stringency, CR1-18 hybridizes only to cDNA clones encoding either the NH<sub>2</sub>-terminal SCR of LHR-A or the signal peptide. The inserts of the cDNA clones were sequenced by the dideoxynucleotide technique (11) after subcloning fragments into M13mp18 and M13mp19 (12). Deletion mutants for sequencing were made by the exonuclease III method (13). Deletion constructions for expression were sequenced directly from the double-stranded plasmid as described (14). DNA sequences were aligned and analyzed with the UWGCG package (15) or the MicroGenie software (Beckman Instruments, Inc., Fullerton, CA).

**Strains and Plasmids.** *Escherichia coli* MC1061/P3 and the CDM8 expression vector (16), were gifts from Dr. Brian Seed (Massachusetts General Hospital, Boston, MA). The expression vector, pMT.neo.1, was a gift from Dr. Keith Peden (The Johns Hopkins University School of Medicine, Baltimore, MD). *E. coli* GM48 and GM271, *dam*<sup>-</sup> and *dcm*<sup>-</sup>, respectively, were gifts from Dr. Elizabeth Rayleigh (New England Biolabs, Beverly, MA) and *E. coli* DK1, a *recA* derivative of MC1061, was a gift from Dr. David Kurnit (University of Michigan Medical School, Ann Arbor, MI). Strains DK1/P3 and GM271/P3 were prepared by transforming the indicated parental strain with plasmid DNA isolated from MC1061/P3 and selecting for kanamycin resistance. Strain DH5 $\alpha$  (Bethesda Research Laboratories, Bethesda, MD), pBluescript KS<sup>+</sup> (Stratagene, Madison, WI), and pGEM3b (Promega Biotec, La Jolla, CA) were obtained as indicated.

**Construction of pBSABCD and piABCD.** Restriction fragments derived from the cDNA clones,  $\lambda$ T109.1,  $\lambda$ H10.3,  $\lambda$ H7.1,  $\lambda$ T8.3 (9),  $\lambda$ T6.1 (4, 10),  $\lambda$ T50.1 (4), and  $\lambda$ T8.2 (4) were ligated and inserted into pBluescriptKS<sup>+</sup> to form pBSABCD (Fig. 1). The 6.9-kb Xho I/Not I fragment containing the entire CR1 cDNA coding sequence was ligated to the 4.4-kb Xho I/Not I fragment of CDM8 (Fig. 4), the ligation mixture used to transform DK1/P3, and the clone, piABCD, containing the CR1 cDNA insert was selected.

**Construction of CR1 Deletion Mutants.** The construction of the deletion mutants utilizes the four Bsm I sites in homologous positions near the sequence encoding the NH<sub>2</sub> terminus of each of the four LHRs and the absence of Bsm I sites elsewhere. 10  $\mu$ g of pBSABCD DNA

was partially digested with 50 U of Bsm I for 45 min, and the restriction fragments of 8.55, 7.20, and 5.85 kb that corresponded to linear segments of the parent plasmid lacking sequence encoding one, two, or three LHRs, respectively, were purified, self ligated, and used to transform competent DH5 $\alpha$ .

The 8.55-kb fragment was generated from pBSABCD by releasing any one of three 1.35-kb Bsm I fragments corresponding to the coding sequence for 92% of LHR-A, -B, or -C. Thus, three distinct plasmids distinguishable by restriction mapping with Sma I were generated after ligation: pBSBCD, pBSACD, and pBSABD, in which the capital letters following pBS represent the LHRs present. The 5.6-kb insert of each of these constructs was released by digestion with Xho I/Not I and ligated to the expression vector, CDM8, to form piBCD, piABD, and piACD. The 7.20-kb fragment from the partial digestion of pBSABCD resulted from Bsm I digestion at three adjacent sites or at two sites separated by a single uncut site. The two possible products obtained after ligation and transformation, pBSAD, and pBSCD, were distinguished by digestion with Xho I/Pst I. The 4.2-kb insert was released from each by digestion with Xho I/Not I and subcloned into CDM8 to yield piAD and piCD. The 5.85-kb fragment from the Bsm I digestion of pBSABCD represented a product of complete digestion and was self ligated to form pBSD. The 2.9-kb insert of pBSD was released with Xho I/Not I and ligated into CDM8 to yield piD.

The plasmid, pBSBD, was prepared by Bsm I partial digestion of pBSBCD. The linear 7.2-kb fragment corresponding to cleavage of two adjacent Bsm I sites was self-ligated and used to transform DH5 $\alpha$ . The 4.2-kb insert containing 1.2 and 6.0 kb Sma I fragments was released with Xho I/Not I and transferred to CDM8 to yield piBD.

The plasmid piABCD was digested to completion with Bst EII, and a doublet of 1.35 kb and a single fragment of 8.6 kb were ligated and the mixture was used to transform DK1/P3. Colonies were identified by hybridization with the CR1 cDNA probe, CR1-4, and the DNA of positive clones was digested with Sma I. The plasmid, piE-2, was identified as containing a weakly CR1-4<sup>+</sup> clone having a single 8.6-kb Sma I fragment.

The plasmid, piA/D, was prepared by digesting piABCD to completion with Pst I and partially with Apa I. The 3' overhangs were removed with the Klenow fragment of *E. coli* DNA polymerase I. A 7.5-kb fragment was ligated and used to transform DK1/P3.

*Transfection of Recombinant CR1 Plasmids.* Each 30–50% confluent 10-cm dish of COS cells or L cells was transfected with 8  $\mu$ g of DNA and 1.6 mg of DEAE-dextran in the presence of 100  $\mu$ M chloroquine diphosphate (8). The transfected cells were shocked with 10% DMSO (17) and cultured for 2–3 d in DMEM supplemented with 10% FCS, 2 mM glutamine, 50 U/ml penicillin, 50  $\mu$ g/ml streptomycin, and 1 mM sodium pyruvate.

*RIA, Immunofluorescence, and Immunoprecipitation of Recombinant CR1.* Replicate samples of  $3 \times 10^5$  transfected cells in 0.1 ml PBS containing 1% BSA and 0.02% sodium azide were incubated at 0°C for 60 min with 3  $\mu$ g/ml YZ1 IgG1 anti-CR1 (18) or with 90  $\mu$ g/ml rabbit IgG anti-CR1. The cells were washed and resuspended in 0.1 ml of buffer containing 1–2  $\mu$ Ci/ml of <sup>125</sup>I-F(ab)<sub>2</sub> goat anti-mouse IgG or <sup>125</sup>I-protein A. After 1–2 h at 0°C, the cells were washed and assayed for <sup>125</sup>I.

For immunofluorescent detection of CR1, transfected cells were sequentially incubated with YZ1 anti-CR1 and affinity-purified FITC-labeled sheep F(ab)<sub>2</sub> anti-mouse IgG (Cappel Laboratories, Cochranville, PA).

CR1 was immunoprecipitated from detergent lysates of <sup>125</sup>I-surface-labeled (19) cells with Sepharose-YZ1 or with rabbit IgG anti-CR1/protein A-Sepharose (18). Immunoabsorbed proteins were assessed by SDS-PAGE (20) and autoradiography.

*Assay of Recombinant CR1 Function.* Sheep erythrocytes sensitized with rabbit antibody (EA) and limited amounts of C4b (EAC4b [lim]) and 12,000 <sup>125</sup>I-C3b/cell (EAC4b [lim], 3b) were prepared by sequential treatment of EAC4b (lim) (Diamedix, Miami, FL) with C1, C2, and <sup>125</sup>I-C3 followed by incubation for 60 min at 37°C in gelatin veronal-buffered saline containing 40 mM EDTA. Alternatively, methylamine-treated C3 (C3 [ma]) and (C4 [ma]) were covalently attached to sheep erythrocytes treated with 3-(2-pyridyldithio) propionic acid *N*-hydroxysuccinimide ester (Sigma Chemical Co., St. Louis, MO) (21). EAC4b were prepared with purified C4 (22).

The C3- and C4-binding functions of recombinant CR1 were assayed by rosette formation. Transfected cells,  $1-4 \times 10^6$ /ml, were incubated with C3- or C4-bearing erythrocytes,

$2-6 \times 10^8$ /ml, in 0.02 ml for 60 min at 20°C. The percentage of transfected cells forming rosettes was evaluated microscopically with a transfected cell scored as a rosette if there were at least five adherent erythrocytes.

The factor I-cofactor function of recombinant CR1 was assessed by incubating transfected COS cells with 0.5  $\mu$ g  $^{125}$ I-C3(ma) and 0.2  $\mu$ g factor I (23) in 0.04 ml for 60 min at 37°C followed by SDS-PAGE and autoradiography. Alternatively, detergent lysates of  $10^6$  COS cells were immunoprecipitated sequentially with Sepharose-UPC10 anti-levan and Sepharose-YZ1. The washed beads were incubated for 60 min at 37°C with  $^{125}$ I-C3(ma) and factor I in 0.05 ml PBS-0.5% NP-40.

## Results

*Isolation and Sequence of cDNA Clones Encoding the Signal Peptide and LHR-A of CR1.* A specifically primed  $\lambda$ gt11 cDNA library,  $\lambda$ HH, that contained  $7.5 \times 10^5$  recombinants was prepared with cDNA synthesized from poly(A)<sup>+</sup> RNA from DMSO-induced HL-60 cells. These cells express only the F allotype of CR1 (3) which is predicted to have four LHRs (4). The primer, LK35.1, was an antisense 35-mer corresponding to nucleotides 896-930 of the previously published partial cDNA sequence of CR1 (4). This oligonucleotide was shown to hybridize to LHR-B, LHR-C, and LHR-D under the conditions of reverse transcription. 250 positive clones were identified in a plating of  $3.8 \times 10^5$  unamplified recombinant phage screened with a mixture of the CR1 cDNA probes, CR1-1 and CR1-4. 38 positive clones were picked and plaque purified. Southern blots of Eco RI-digested DNA from these clones were screened with the 23-mer oligonucleotide, KS23.1 (5'-CTGAGCGTAC CCAAA-GGGAC AAG-3') corresponding to nucleotides 763-785 of the partial CR1 cDNA sequence (4). This probe hybridizes under conditions of high stringency at a single site in the sequence encoding LHR-B but not to sequences encoding LHR-C or LHR-D. The insert of clone  $\lambda$ H7.1 (Fig. 1) contained three Eco RI fragments of 1.0, 0.9, and 0.4 kb, and the two larger fragments hybridized to KS23.1, indicating that this clone contained sequences coding for the 3' five SCRs of LHR-A and all of LHR-B. This finding confirmed the prediction that LHR-A would be highly homologous to LHR-B (4). Clone  $\lambda$ H3.1 (Fig. 1) contained a single KS23.1<sup>+</sup> Eco RI fragment of 1.0 kb and a 5', 0.5-kb fragment that hybridized weakly with CR1-4 at high stringency. This clone was considered to contain the additional 5' sequence completing LHR-A, including SCR-1 and -2 and 0.1 kb of upstream sequence. None

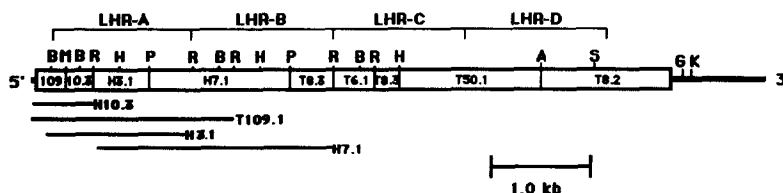


FIGURE 1. Restriction map of the insert of the plasmid, pBSABCD, encoding human CR1. Indicated within the box delineating the region containing the coding sequence are the nine fragments of eight cDNA clones that were ligated to form the CR1 construct. The brackets designate the positions of LHR-A, -B, -C, and -D, respectively. The lines below the box represent the positions of the newly isolated 5' cDNA clones. The restriction sites are abbreviated: A, Apa I; B, Bam HI; G, Bgl II; H, Hind III; K, Kpn I; M, Bsp MII; P, Pst I; R, Eco RI; and S, Sac I.

of the remaining 36 clones, all of which hybridized with CR1-1, were detected with the probe, CR1-18, a 252-bp *Sau* 3AI fragment from the 0.5-kb *Eco* RI fragment of clone  $\lambda$ H3.1 that does not hybridize to sequences encoding LHR-B, -C, or -D.

DNA sequence analysis of  $\lambda$ H3.1 revealed that the open reading frame continued to the 5' end of the cDNA, indicating that the clone did not extend to the translational start site. Therefore, the cDNA libraries,  $\lambda$ HH and  $\lambda$ S2T (4, 9), were rescreened with the probe CR1-18 to identify one clone from each  $\lambda$ H10.3 and  $\lambda$ T109.1, respectively. The *Eco* RI fragments of these clones that hybridized with CR1-18 were sequenced as were the inserts from the clones  $\lambda$ H3.1 and  $\lambda$ H7.1. The composite sequence is presented such that the nucleotide following 1531 in Fig. 2 is nucleotide 1 in Fig. 2 of the previously published sequence (4). The overlapping sequences of the cDNA clones from the HL-60 and tonsillar libraries are identical.

Immediately upstream of LHR-A, clones  $\lambda$ H10.3 and  $\lambda$ T109.1 contain identical putative hydrophobic leader sequences (24) encoding 41 amino acids, including an ATG matching the consensus NNA/GNNATGG proposed for eukaryotic translation initiation sites (Fig. 3) (25). A second ATG, located six codons upstream of the chosen ATG and just downstream of an in-frame stop codon, is a poor match

```

1  CGTAGGTTGTAGATGTGCTTGGGGAGATGGGGGCCCTCTTCTCCAGAGGCCCGAGCCCT 60
61  GTCCGGCCCGCCGGCCCGCCGCTCCGCCCTTCTGCTGGGGAGGCCCTGCGCGGTATTG 120
121  GTCCCTGCTTGGCCCTGCGCCGCTGGCCCTGGGCTCAGTCCAGATGCCCGAGATGCCCTTCCATT 180
181  GCCAGGCCCTCCAGCCCTACTGATGATTTGAGTTCCGCCCTTGGGCTATATCTGACTAT 240
241  GATTCGGCCCGCTGGTATTCGGGAGGCCCGCTTTCTATCTATCTGCGCTATATATCTGATC 300
301  TGGATCTGGCTTAGAGGCTAGGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 360
361  AATGGCTGGCTGGCTGGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 420
421  ACTAAGGGGTACCGACTCATTTGGTTCCTGCTGGCCATGCTCATCTCAGGTGATACT 480
481  GTCTATTGGGATATGAAACACCTATTCTTGGTCAAGATTCCTTGTGGCTACCCCCACCC 540
541  ATCAACATGGGATTTCTATTGACCAACACAGAGAGATTTTCACTATGGATCAGTGGTG 600
601  ACCCCCGCTCCATCCCTGGATGGGCGGGGAGAAAGATTTATGACTTGTGGCTGGGCTGG 660
661  TCCATGACTGCAACCGATGAGGATCAGATGGCCATCTGGAGGGCCCGCCCGCCCTCAG 720
721  TGCATTTACCTAACCAATGCAAGCCCTCCAAATGTGGAAATGGAAATTTGGGTATCTGAC 780
781  AATCGAAGCTATTTTCTCTTAAATGAGATTTGGAGTTTGGGTATCAGCCCTGGCTTTGTC 840
841  ATGAAAGGCCCGCCCGCTGTGAGTCCAGGCCCTGGACAAATGGAGCCCGGAGCTACCA 900
901  AGCTGCTCCAGGGTATGTCAGCCACCTCCAGATGCTGCTGCTGAGCGTACCCAAAGG 960
961  GTCAGAGCAACTTTTCACTGGCCAGGATGTTCTTACGCTGTGAGCCCGCTACGAC 1020
1021  CTCAGAGGGGCTGGGCTGATGGCTGACCAACCCCGGGGACTGGAGCCCTCCAGCCCGCC 1080
1081  AATGATGATGAAATCCATGATGACTTCAATGGCCACTTCTTAAATGGCCGCTGTGCTA 1140
1141  TTTCCAGTAAATCTCCGCTTGGAGCAAAATGGATTTTGTGTTGGATGAAAGATTTCAA 1200
1201  TTTAAAGGCACTCTGCTGATTAAGTGTCTTGGCTGGATGGAAAGCCCTTGGAAATGAC 1260
1261  AGTTTCCAGTGTGTAACAAATCTTTTCCAGTCCCTCCAGTATTTCTAATGGGGA 1320
1321  CACAGAGAAATCCCTCGATATCTTTCCCTTGGAAAGCCATAAATACACATGGAC 1380
1381  CCCCACCCAGCAGGGGAGCCCTTCCAGCTCATTGGAGGAGCAACATCCGCTCCCA 1440
1441  AATGACCTCAAGGGATGGGGTATTGGAGCCAGCCCTGCGCCCTGGCTGTGGATTTCTGGGT 1500
1501  CACTGTCAGCCCGATGATTTTCTGTTG 1531
    
```

FIGURE 2. Nucleotide sequence of the 5' CR1 cDNA clones. The composite sequence begins with the first nucleotide after the octamer *Eco* RI linker in clone  $\lambda$ T109.1. Nucleotide 1,531 of this sequence is the first nucleotide 5' of nucleotide 1 of the sequence depicted in Fig. 2 of reference 4. The proposed initiation codon is underlined and an upstream stop codon is indicated by the overbar. These sequence data have been submitted to the EMBL/GenBank under the accession number Y00816.

AA #	LHR	Signal
1	MCLGRGASSPSPSPGPPAGLPPGCGSLLAVVLLALFVNG	
47	<u>Q</u> QAP <u>P</u> HL <u>P</u> FP <u>P</u> IN <u>L</u> DE <u>P</u> FF <u>P</u> IG <u>T</u> IL <u>M</u> Y <u>H</u> Q <u>P</u> G <u>S</u> GP <u>P</u> SL <u>I</u> CL <u>N</u> SW <u>I</u> G <u>A</u> ND <u>R</u>	A
497	H Q <u>D</u> R <u>F</u> L <u>K</u> L <u>R</u> I <u>Q</u> <u>N</u> A <u>S</u> D S K E Y T D L S S P V K	B
947	H Q <u>D</u> R <u>F</u> L <u>K</u> L <u>R</u> I <u>Q</u> <u>N</u> A <u>S</u> D S K E Y T D L S S P V K	C
1400	H <u>R</u> T <u>Q</u> F S <u>I</u> F <u>I</u> N <u>D</u> V S F <u>N</u> M S E L S S V E N	D
107	<u>N</u> S <u>G</u> N <u>P</u> D <u>F</u> P <u>A</u> W <u>H</u> V <u>I</u> N <u>G</u> I <u>O</u> F <u>S</u> Q <u>I</u> Q <u>S</u> C <u>I</u> N <u>G</u> I <u>L</u> I <u>G</u> S <u>S</u> A <u>C</u> I <u>S</u> C <u>V</u> I <u>M</u> E <u>P</u> I <u>C</u> R	A
557	<u>K</u> T T D V R N T H H E L N A H S T K P Q	B
1007	<u>K</u> T T D V R N T H H E L N A H S T K P Q	C
1460	<u>G</u> P E F I <u>N</u> I <u>D</u> T T V N N E F P T I V N N T K G A E I	D
169	<u>I</u> P <u>G</u> I <u>P</u> P <u>I</u> L <u>I</u> N <u>G</u> I <u>F</u> I <u>S</u> I <u>N</u> S <u>R</u> L <u>H</u> S <u>F</u> I <u>S</u> I <u>N</u> E <u>V</u> E <u>P</u> F <u>P</u> Q <u>G</u> F <u>V</u> M <u>G</u> E <u>P</u> R <u>V</u> K <u>Q</u> A <u>I</u> N <u>E</u> P <u>E</u> L <u>P</u> S <u>G</u> S <u>R</u>	A
619	A	B
1069	A L R	C
1522	S E P S Y N T S N T Q H T P D E Q L R K V S P R	D
239	<u>I</u> N <u>S</u> I <u>P</u> P <u>V</u> I <u>N</u> G <u>I</u> L <u>S</u> D <u>N</u> S <u>R</u> L <u>H</u> S <u>F</u> I <u>S</u> I <u>N</u> E <u>V</u> E <u>P</u> F <u>P</u> Q <u>G</u> F <u>V</u> M <u>G</u> E <u>P</u> R <u>V</u> K <u>Q</u> A <u>I</u> N <u>E</u> P <u>E</u> L <u>P</u> S <u>G</u> S <u>R</u>	A
689		B
1139		C
1592	S T A E A R P G F T I I R V S H T Q T N G R G K H	D
301	<u>V</u> Q <u>P</u> P <u>P</u> D <u>V</u> L <u>H</u> A <u>R</u> K <u>I</u> G <u>R</u> K <u>I</u> N <u>P</u> R <u>Q</u> E <u>V</u> F <u>P</u> S <u>G</u> E <u>P</u> D <u>I</u> R <u>G</u> A <u>S</u> A <u>E</u> P <u>C</u> T <u>P</u> Q <u>D</u> M <u>S</u> P <u>A</u> P <u>T</u> C <u>E</u> V	A
751		B
1201	E I G H P E R Q I H E R A	C
1654	E I G H I S R Q S I H E R T	D
361	<u>N</u> S <u>G</u> N <u>P</u> D <u>F</u> P <u>A</u> W <u>H</u> V <u>I</u> N <u>G</u> I <u>O</u> F <u>S</u> Q <u>I</u> Q <u>S</u> C <u>I</u> N <u>G</u> I <u>L</u> I <u>G</u> S <u>S</u> A <u>S</u> V <u>C</u> V <u>L</u> A <u>S</u> H <u>E</u> S <u>I</u> N <u>E</u> S <u>V</u> E <u>P</u> F <u>P</u> Q <u>G</u>	A
811		B
1261	L P H L S R V H V R N H	C
1714	L P H L L S R R H N A	D
424	<u>I</u> P <u>G</u> I <u>P</u> P <u>V</u> I <u>N</u> G <u>I</u> H <u>I</u> G <u>R</u> K <u>I</u> E <u>V</u> F <u>P</u> P <u>G</u> A <u>V</u> V <u>I</u> T <u>Q</u> D <u>H</u> P <u>R</u> O <u>D</u> S <u>F</u> L <u>I</u> G <u>E</u> S <u>P</u> I <u>D</u> A <u>C</u> T <u>S</u> O <u>P</u> Q <u>G</u> N <u>G</u> M <u>S</u> P <u>A</u> P <u>T</u> C <u>E</u> V	A
874		B
1324	N A L T S G D I Y E I S M T N H	EL C
1777	N A L T P G D I Y E I S A T M T N S	EL D

FIGURE 3. The deduced amino acid sequence of the 5' cDNA clones encoding the seven SCRs of LHR-A, and alignment of this sequence with the corresponding SCRs of LHR-B, -C, and -D. The four cysteines that are conserved in each SCR are underlined. A residue is shown for LHR-B, -C, and -D only where it is different from that in LHR-A. These sequence data have been submitted to the EMBL/GenBank under the accession number Y00816.

for this consensus sequence. The first three amino acids of this leader sequence for CR1, MGA, are the same as those reported for CR2 (1). The sequences of these two clones diverge upstream of the ATG, and that from clone  $\lambda$ H10.3 is believed to represent a portion of an intervening sequence (data not shown) as has been described earlier for other CR1 cDNA clones (4).

The signal peptidase cleavage is predicted (24) to occur between glycine-46 and glutamine-47, suggesting that the blocked  $\text{NH}_2$  terminus of CR1 (26, 27) may be due to the presence of a pyrrolidone amide. The first two SCRs of the  $\text{NH}_2$ -terminal LHR-A contained in these clones are only 61% identical to the corresponding region of LHR-B, whereas SCRs 3-7 of LHR-A are 99% identical to the corresponding SCRs of LHR-B (Fig. 3). Comparison of LHR-A with LHR-C reveals that only the third and fourth SCRs of each are highly homologous (99% identical). LHR-A and -D have only 68% overall identity, with maximal identity of 81% between the sixth SCR of each LHR. Thus, completion of the 5' cDNA sequence of CR1 indicates that the F allotype is composed of 2,039 amino acids, including a 41-amino acid signal peptide, four LHRs of seven SCRs each, two additional COOH-terminal SCRs, a 25-residue transmembrane region and a 43-amino acid cytoplasmic domain. There are 25 potential N-linked glycosylation sites.

**Expression of Recombinant CR1 Protein.** Restriction fragments of eight cDNA clones were ligated to form the plasmid, pBSABCD, having the entire coding sequence of the F allotype of human CR1 (Fig. 1). The 6.9-kb Xho I-Not I fragment of pBSABCD containing this sequence was subcloned into the eukaryotic expression vectors, CDM8, to form piABCD and pMT.neo.1 to form pMTABCD (Fig. 4). Expression of the insert is driven from a cytomegalovirus promoter in piABCD and

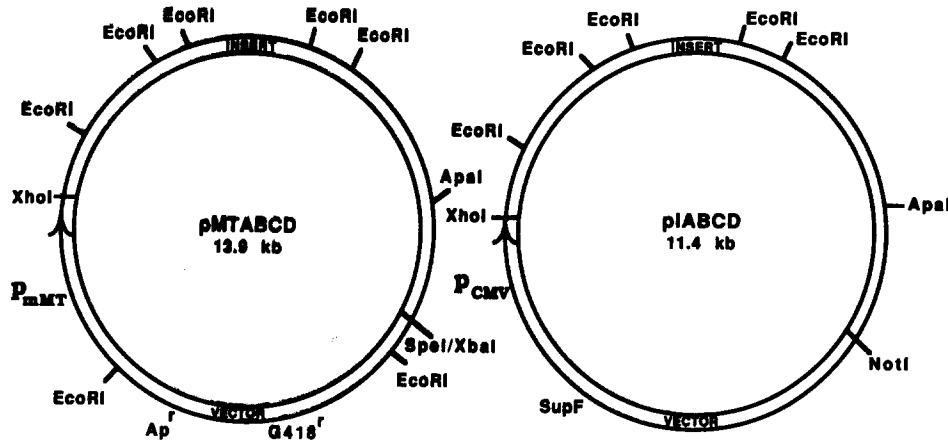


FIGURE 4. Restriction maps of the expression plasmids piABCD and pMTABCD. P<sub>mMT</sub> and P<sub>CMV</sub> represent the murine metallothionein and cytomegalovirus immediate early promoters, respectively.

from the mouse metallothionein promoter in pMTABCD; both plasmids contain an SV40 polyadenylation signal. The “ABCD” notation refers to the LHRs that are present in the expression constructs.

Murine L cells were cotransfected by the DEAE-dextran method (8, 17) in duplicate with 0, 2, or 4 μg of either piABCD or pMTABCD and 2 μg of pXGH5, a reporter plasmid that directs the expression of growth hormone (28). The cells were harvested after 2 d and assayed for expression of CR1 by binding of YZ1 monoclonal anti-CR1. There was a dose-response relationship between recombinant plasmid DNA and the expression of CR1 antigen (Table I). The plasmid, piABCD, directed the expression of nearly threefold more CR1 antigen than did pMTABCD. The growth hormone concentration in the culture medium varied by less than twofold with the exception of plate 5. Additional experiments revealed that piABCD directed the tran-

TABLE I  
Dose Response of Recombinant CR1 and Human Growth Hormone  
in Cotransfected L Cells

Plate	pXGH5	pMTABCD	piABCD	YZ1 RIA	Growth hormone
	μg	μg	μg	cpm	ng/ml
1	2	0	0	1,444	120
2	2	0	2	6,058	130
3	2	0	2	6,531	140
4	2	0	4	10,620	180
5	2	0	4	9,898	80
6	2	2	0	3,111	180
7	2	2	0	2,747	160
8	2	4	0	3,547	160
9	2	4	0	3,337	140

sient expression of threefold more CR1 antigen in COS cells than in L cells (data not shown).

CR1 antigen present on the surface of the transfected COS cells was distributed in clusters when assessed by indirect immunofluorescence of cells stained with YZ1 anti-CR1 (Fig. 5). This distribution of recombinant CR1 on COS cells resembles that of wild-type CR1 on human leukocytes (29).

The  $M_r$  of the recombinant CR1 was determined by surface iodination of COS cells transfected with piABCD, immunoprecipitation of cell lysates with Sepharose-YZ1, SDS-PAGE, and autoradiography. The recombinant CR1 had an  $M_r$  of 190,000 unreduced, which was equivalent to that of the F allotype and less than that of the S allotype of erythrocyte CR1 (Fig. 6).

The C3b-binding and C4b-binding function of recombinant CR1 was assayed by the formation of rosettes between the transfected COS cells and EAC4b or EAC4b(lim),3b. In 31 separate transfections, 5–50% of COS cells transfected with the plasmid, piABCD, bound five or more EAC4b or EAC4b(lim),3b (Fig. 7). The COS cells expressing CR1 did not form rosettes with EAC4b(lim),3bi, although this intermediate did form rosettes with Raji B lymphoblastoid cells expressing CR2 (data not shown).

The factor I-cofactor activity of recombinant CR1 immunoadsorbed from detergent lysates of transfected COS cells with Sepharose-YZ1 was evaluated by incubation with 0.5  $\mu$ g of  $^{125}$ I-C3(ma) and 200 ng of factor I. Factor I cleaved the  $\alpha$  chain

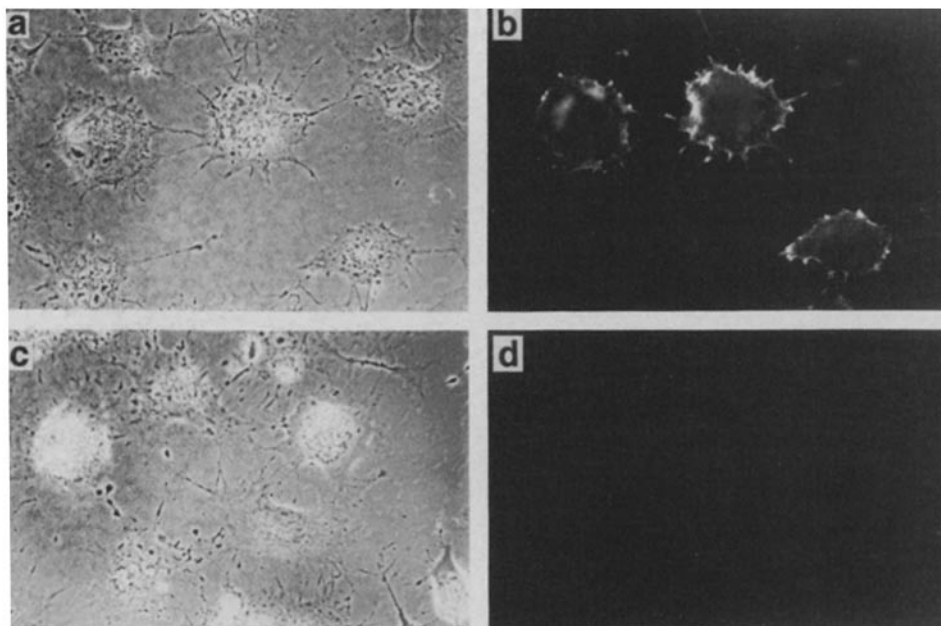


FIGURE 5. Analysis by phase contrast (*a* and *c*) and immunofluorescent (*b* and *d*) microscopy of COS cells transfected with piABCD (*a* and *b*) and CDM8 vector alone (*c* and *d*), respectively, and indirectly stained with YZ1 monoclonal anti-CR1 and fluorescein-labeled goat anti-mouse F(ab)<sub>2</sub>.



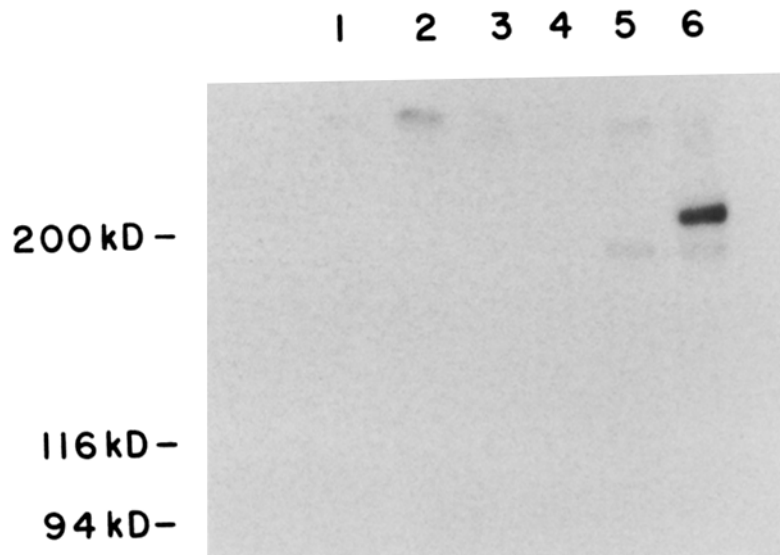


FIGURE 6. Analysis of recombinant CR1 expressed by transfected COS cells by SDS-PAGE. COS cells transfected with the CDM8 vector alone (lanes 1 and 4) and with piABCD (lanes 2 and 5), respectively, and erythrocytes from an individual having the F and S allotypes of CR1 (lanes 3 and 6) were surface labeled with  $^{125}\text{I}$ . Detergent lysates of the cells were sequentially immunoadsorbed with Sepharose-UPC10 (lanes 1-3) and Sepharose-YZ1 (lanes 4-6), and the eluates were analyzed by SDS-PAGE under nonreducing conditions and autoradiography.

of C3(ma) into fragments of 76,000 and 46,000  $M_r$  only in the presence of immunoimmobilized recombinant CR1 or factor H (Fig. 8). The regions corresponding to bands from the autoradiogram were excised from the gel and assayed for  $^{125}\text{I}$  to determine the amount of the  $\alpha$  chain cleaved. In the presence of factor H, 91% of the  $\alpha$  chain was cleaved, while in the presence of increasing amounts of recombinant CR1, 26%, 41%, and 55%, respectively, was cleaved.

*Identification and Localization of Multiple C3b/C4b-binding Sites in CR1.* Deletion mutagenesis of recombinant CR1 was performed to determine whether multiple distinct C4b/C3b-binding sites are present in the receptor. The clones piBCD, piAD, piBD, piCD, and piD were prepared from partial digests of the full coding sequence of CR1 cDNA with Bsm I, which restricted the DNA at single sites near the sequence encoding the second cysteine of the first SCR of each LHR (Fig. 9). After ligation of the various restriction fragments, deletion mutants lacking one, two, or three LHRs were generated.

The clone piA/D was prepared by digesting the CR1 cDNA with Pst I and Apa I, which restricted the DNA at a site between the codons for cysteine-3 and -4 of the fifth SCR of LHR-A and cysteine-3 and -4 of the fourth SCR of LHR-D, respectively (Fig. 9). Ligation of the appropriate fragments formed a hybrid construct containing the  $\text{NH}_2$ -terminal four and three quarters SCRs of LHR-A and the COOH-terminal three and one quarter SCRs of LHR-D.

The clone piE-2 was prepared by digesting the CR1 cDNA with Bst EII, which

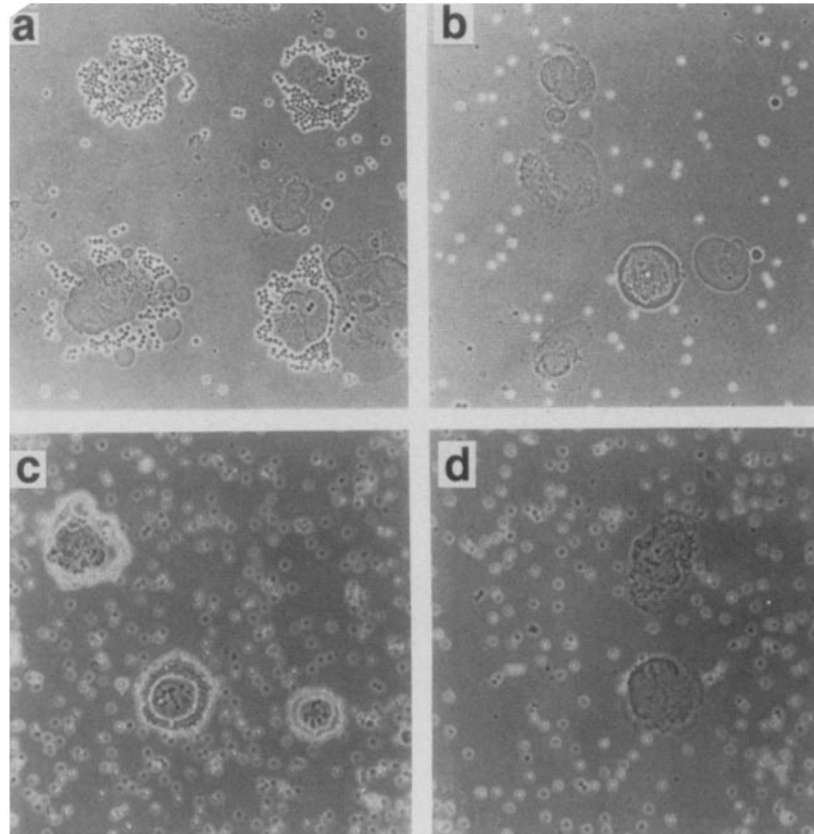


FIGURE 7. Analysis of C3b and C4b binding by COS cells expressing recombinant CR1. COS cells transfected with piABCD (*a* and *c*) or with the CDM8 vector alone (*b* and *d*) were incubated with EAC4b(1im),3b (*a* and *b*) or with EAC4b (*c* and *d*) and examined for formation of rosettes by phase contrast microscopy.

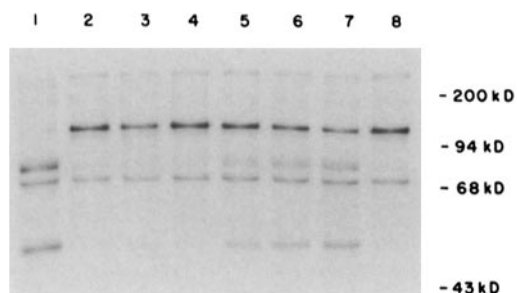


FIGURE 8. Cleavage of  $^{125}\text{I}$ -C3(ma) by factor I in the presence of immobilized recombinant CR1. Replicate samples of  $^{125}\text{I}$ -C3(ma) were treated with factor I in the presence of factor H (lane 1), Sepharose-UPC10 preincubated with the lysate of COS cells transfected with the CDM8 vector alone (lane 2), Sepharose-UPC10 preincubated with the lysate of piABCD-transfected COS cells (lane 3), Sepharose-YZ1 preincubated with the lysate of CDM8-transfected COS cells (lane 4), and 6  $\mu\text{l}$  (lane 5), 12  $\mu\text{l}$  (lane 6), and 25  $\mu\text{l}$  (lane 7) of Sepharose-YZ1 that had been preincubated with the lysate of piABCD-transfected COS cells. Samples of  $^{125}\text{I}$ -labeled C3(ma) were also treated in the absence of factor I with 25  $\mu\text{l}$  of Sepharose-YZ1 that had been preincubated with the lysate of piABCD-transfected COS cells (lane 8). After reduction, the  $^{125}\text{I}$ -C3(ma) was analyzed by SDS-PAGE and autoradiography.

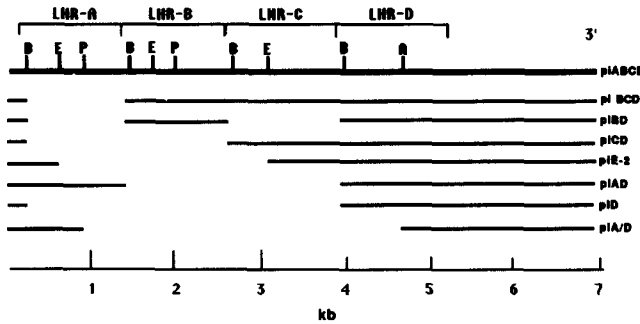


FIGURE 9. The cDNA constructs encoding the CR1 deletion mutants. The positions of the cDNA segments encoding the four LHRs are indicated by the brackets above the full-length piABCD construct on which are shown the restriction sites used for preparation of the deletion mutants. The cDNA restriction fragments remaining in each of the mutants are indicated by the solid lines. The restriction sites are abbreviated: A, Apa I; B, Bsm I; E, Bst EII; and P, Pst I.

restricted the DNA at a site between the codons for cysteine-1 and -2 of the third SCR of LHR-A, -B, and -C, respectively (Fig. 9). Ligating the 5' LHR-A-encoding fragment to the fragment encoding the 3' end of LHR-C and the rest of the 3' sequence created a construct in which SCR-1 and -2 of LHR-A were substituted for the corresponding SCRs of LHR-C, with deletion of the rest of LHR-A and all of LHR-B.

COS cells transiently expressing the piABCD, piBCD, piCD, and piD constructs, respectively, were surface labeled with <sup>125</sup>I and immunoprecipitated with anti-CR1. On SDS-PAGE after reduction, the product of the piABCD construct comigrated with the F allotype of CR1, while the deletion mutants demonstrated stepwise decrements of ~45,000 M<sub>r</sub>, indicative of the deletion of one, two, and three LHRs, respectively (Fig. 10).

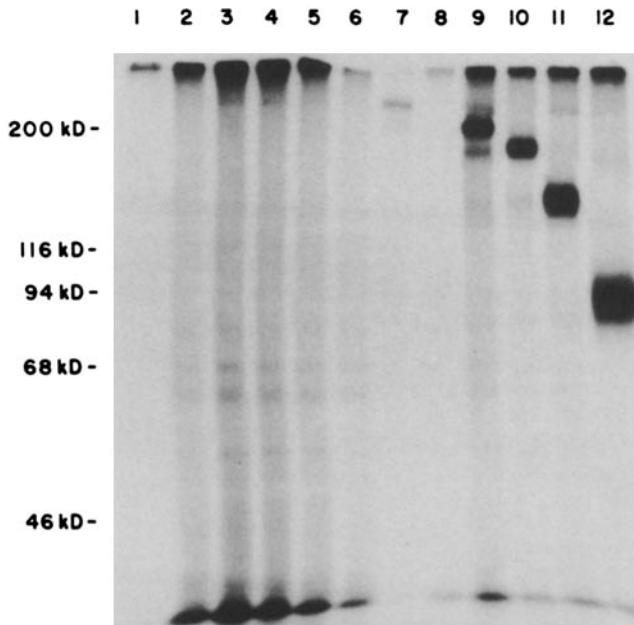


FIGURE 10. Comparison of recombinant deletion mutants of CR1 with the wild-type F and S allotypes of CR1. Detergent lysates of <sup>125</sup>I-surface-labeled erythrocytes (lanes 1 and 7) and COS cells transfected with CDM8 vector alone (lanes 2 and 8), piABCD (lanes 3 and 9), piBCD (lanes 4 and 10), piCD (lanes 5 and 11), and piD (lanes 6 and 12), respectively, were immunoprecipitated with Sepharose-UPC10 anti-levan (lanes 1-6), Sepharose-YZ-1 anti-CR1 (lanes 7-11), and rabbit anti-CR1 and protein A-Sepharose (lane 12), respectively. The eluates were subjected to SDS-PAGE under reducing conditions and autoradiography.

In each of three separate experiments, the proportion of COS cells expressing the full-length piABCD construct that formed rosettes with the EC3(ma) was similar to the fraction having detectable recombinant receptor, as assessed by immunofluorescence using either YZ1 monoclonal anti-CR1 or rabbit anti-CR1 (Table II). In contrast, cells expressing piD did not form rosettes, indicating that a C3-binding site(s) must reside in or require the presence of LHR-A, -B, or -C. A site was shown to be present in both LHR-B and -C by demonstrating that cells expressing either the piBD or piCD constructs formed rosettes with EC3(ma). Cells expressing piAD, piA/D, or piE-2 did not have equivalent C3-binding function. As the piE-2 construct differs from piCD only in having SCR-1 and -2 of LHR-A instead of the first two SCRs of LHR-C, the function of the C3-binding site in LHR-C must require these NH<sub>2</sub>-terminal SCRs.

The proportion of COS cells expressing the full-length piABCD recombinant that formed rosettes with EC4(ma) was less than the fraction rosetting with EC3(ma), perhaps reflecting fewer C4(ma) per erythrocyte (Table II) or fewer C4-binding sites per receptor. Deletion mutants having all or part of LHR-A, the piAD, piA/D, and piE-2 constructs, bound EC4(ma) better than did the deletion mutants, piBD and piCD; piD lacked this function. Thus, the C4-binding site of CR1 resides primarily in LHR-A, although secondary sites may be present in LHR-B and -C. The improved rosetting capability of the piE-2 construct relative to that of piCD suggests that SCR-1 and -2 of LHR-A are involved in the C4-binding site.

RIA of the binding of YZ1 monoclonal anti-CR1 indicated significant uptake by COS cells expressing the piABCD, piAD, piBD, and piCD constructs (Table III). Cells transfected with piD or piA/D, which is composed of the five NH<sub>2</sub>-terminal SCRs of LHR-A and the three COOH-terminal SCRs of LHR-D, did not bind YZ1 anti-CR1, although the products of these constructs bound polyclonal anti-CR1 (Table

TABLE II  
*Formation of Rosettes between COS Cell Transfectants  
Expressing Recombinant Forms of CR1 and Sheep  
Erythrocytes Bearing C3(ma) or C4(ma)*

COS cell transfectant	Percent of transfectants forming rosettes/ percent of transfectants fluorescent with anti-CR1	
	EC3 (ma)*	EC4 (ma)†
piABCD	109§ (3)¶	62 (2)
piAD	8 (3)	107 (2)
piBD	107 (3)	12 (2)
piCD	127 (3)	32 (2)
piD	0 (3)	0 (2)
piA/D	11 (2)	83 (2)
piE-2	0 (1)	102 (1)

\* The numbers of C3 (ma)/E were 60,000, 350,000, and 900,000, respectively, in the three experiments using this intermediate.

† The number of C4 (ma)/E were 160,000 and 140,000, respectively, in the two experiments using this intermediate.

§ Mean of separate experiments.

¶ Number of experiments.

TABLE III  
*Binding of Monoclonal and Polyclonal Anti-CR1 to COS Cell  
 Transfectants Expressing Recombinant Forms of CR1*

COS cell transfectant	YZ1 mAb	Rabbit polyclonal antibody
piABCD	2,362*	12,277*
piAD	2,879	19,891
piBD	3,646	21,922
piCD	2,189	19,926
piA/D	410	23,052
piD	404	16,386
CDM8	428	4,886

\* Mean of duplicate determinations, cpm/3 × 10<sup>5</sup> COS cells.

III). Thus, the YZ1 epitope is repeated in LHR-A, -B, and -C, is not present in the NH<sub>2</sub>-terminal SCRs of LHR-A, and is not present or is inaccessible in LHR-D.

*Factor I Cofactor Function of the CR1 Deletion Mutants for Cleavage of C3.* COS cells transfected with piABCD, piAD, piBD, piCD, and piD, respectively, were incubated with <sup>125</sup>I-C3(ma) and factor I to evaluate factor I-cofactor activity. In contrast to the experiment depicted in Fig. 8, the recombinant CR1 was assayed on intact COS cells because the product of the piD construct cannot be immunoadsorbed by Sepharose-YZ1 (Table III). Although the COS cells transfected with the CDM8 vector alone contained some endogenous factor I-cofactor activity, an increase in this function was evident with COS cells transfected with piABCD, piBD, and piCD (Fig. 11). No enhanced cleavage of <sup>125</sup>I-C3(ma) was seen with COS cells transfected with piAD or piD. Thus, among these constructs, only the deletion mutants, piBD and piCD, which conferred on COS cells a capacity for binding C3, also had factor I-cofactor activity for cleavage of C3.

### Discussion

The primary structure of the NH<sub>2</sub> terminus and the signal peptide of the F allotype of CR1 has been deduced by the isolation and sequencing of 5' cDNA clones. The highly repetitive nature of the CR1 sequence made critical the development of an appropriate strategy for the preparation and identification of cDNA clones encoding this region of the receptor. A cDNA library was prepared using as a primer a 35-mer oligonucleotide known to hybridize under the conditions of reverse transcription to LHR-B, -C, and -D; the possibility was considered that this primer might hybridize also to LHR-A that had been predicted to be highly homologous to LHR-B (4). Appropriate cDNA clones would be identified by the use of another oligonucleotide, KS23.1, which would hybridize only to LHR-B under stringent conditions, thereby increasing the probability of finding 5' cDNA clones. Two clones were found that encompassed almost all of the residual sequence of CR1, and a Sau 3AI fragment of one of these, CR1-18, had sequence sufficiently unique to permit its use in the identification of the remaining 5' clones (Figs. 1-3).

Amino acid sequencing studies of CR1 purified from erythrocyte membranes had revealed that the NH<sub>2</sub> terminus was blocked (26, 27). Therefore, unambiguous demonstration that the full-length sequence of CR1 had been completed required

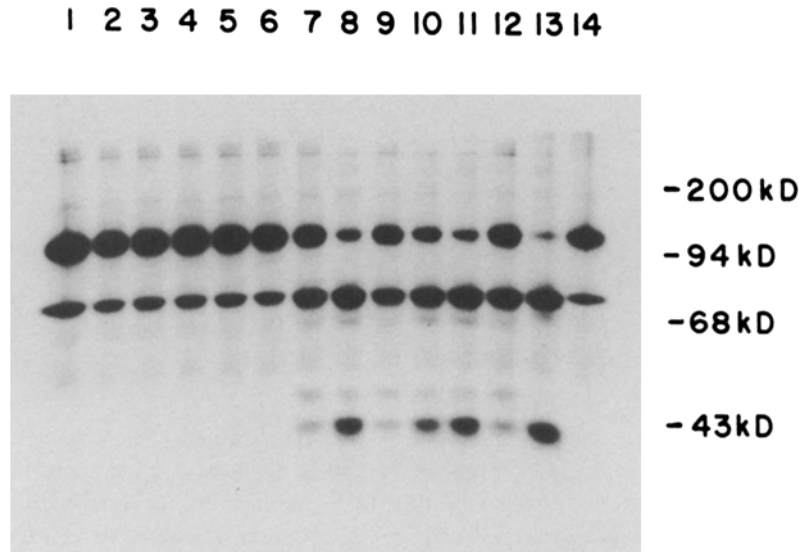


FIGURE 11. Cleavage of  $^{125}\text{I}$ -C3(ma) by factor I in the presence of COS cells expressing full-length and deletion mutants of CR1. Replicate samples of  $^{125}\text{I}$ -C3(ma) were incubated with COS cells transfected with the CDM8 vector alone (lanes 1 and 7), piABCD (lanes 2 and 8), piAD (lanes 3 and 9), piBD (lanes 4 and 10), piCD (lanes 5 and 11), and piD (lanes 6 and 12), respectively, in the absence (lanes 1-6) or presence of factor I (lanes 7-12). Samples of  $^{125}\text{I}$ -C3(ma) also were incubated with factor H and factor I (lane 13) and with factor I alone (lane 14), respectively. After reduction, the  $^{125}\text{I}$ -C3(ma) was analyzed by SDS-PAGE and autoradiography.

expression of recombinant receptor and comparison with wild-type CR1. Restriction fragments of eight cDNA clones were ligated and the putative full-length construct was inserted into two expression vectors (Fig. 4) to direct the synthesis of the recombinant receptor detectable by immunofluorescence on COS cells (Fig. 5). The recombinant CR1 was indistinguishable from the F allotype of erythrocytes by SDS-PAGE (Fig. 6), it mediated the binding of sheep erythrocytes bearing either C4b or C3b, reproducing the ligand specificity of CR1 (Fig. 7), and it exhibited factor I-cofactor activity for cleavage of the  $\alpha$  polypeptide of C3(ma) (Fig. 8). Thus, all of the coding sequence of the F allotype CR1 had been cloned and, after cleavage of the signal peptide of 41 amino acids, the mature receptor contained 1,998 amino acids, including an extracellular domain of 1,930 residues that forms 30 SCRs, 28 of which are organized into LHR-A, -B, -C, and -D (Fig. 3), a single membrane-spanning domain of 25 amino acids and a relatively short cytoplasmic domain of 43 amino acids (4).

Among the C3/C4-binding proteins that contain multiple SCRs, CR1 is unique in having groups of SCRs organized into LHRs. Comparison of the four LHRs of CR1 reveals that each is a composite of four types of SCRs: types a, b, c, and d (Fig. 12). For example, the sequences of SCR-1 and -2 of LHR-A are only 62%, 62%, and 57% identical to the first two SCRs of LHR-B, -C, and -D, respectively. However, SCR-3 through SCR-7 differ from the corresponding SCRs of LHR-B at only a single position, and SCR-3 and -4 differ from those of LHR-C at only three positions (Fig. 3). Thus, some of the type "a" SCRs of LHR-A are also present in

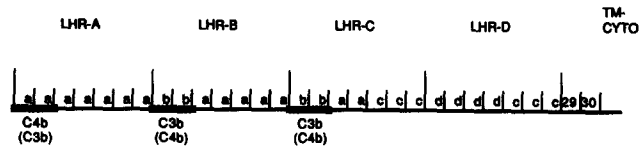


FIGURE 12. Schematic model depicting the types of SCRs composing each LHR of CR1, and the predicted sites determining the specificities of the receptor for C3b and C4b. The secondary binding specificities of these are indicated by the parentheses.

LHR-B and -C. The first two SCRs of LHR-B, which differ from those of LHR-A, are 99% identical to the corresponding SCRs of LHR-C, so that LHR-B and -C share the type "b" SCR at these positions. The fifth, sixth, and seventh SCRs of LHR-C are only 77% identical to the type "a" SCRs in LHR-A and -B at these positions, and are considered as type "c" SCRs. The first through fourth SCRs of LHR-D are relatively unique and are type "d", while the fifth through seventh SCRs are ~93% identical to the "c" type found in LHR-C. This mosaic composition of the LHRs, which may have arisen through gene conversion or homologous recombination with unequal crossover, provided an experimental approach for determining those SCRs involved in the C4 and C3 specificity of CR1.

The conserved Bsm I site found midway through the coding sequence of the first SCR of each LHR permitted the construction of a series of deletion mutants that corresponded closely to the boundaries of the LHRs, and maintained the open reading frame and the appropriate positions of the four cysteines necessary for the presumed disulfide bond formation (Fig. 9). Comparison of the C3(ma)- and C4(ma)-binding functions of these deletion mutants would distinguish not only the LHRs having these specificities, but also those SCRs critical for determining the ligand specificity. Thus, the capacity of piAD, piA/D, and piE-2 forms of the receptor, but not the piD form, to mediate rosette formation between the transfected COS cells and EC4(ma) indicated that the NH<sub>2</sub>-terminal two SCRs of LHR-A contained a site for interaction with this complement protein (Table II). This site was only relatively specific for C4(ma) because transfectants expressing piAD and piA/D also were capable of binding EC3(ma) (Table II). The C3(ma)-binding function of the receptors encoded by the piBD and piCD constructs, demonstrated by rosette assay and factor I-cofactor function for cleavage of C3(ma) (Table II; Fig. 11), indicated the presence of sites specific for C3(ma) in the first two SCRs of these LHRs. These sites also were capable of interacting with C4(ma) (Table II). The finding of preferential, but overlapping, C4- and C3-binding activities in LHR-A, -B, and -C is not unprecedented, as both factor H and C4b-binding protein have been found to have secondary specificities for C4b and C3b, respectively (30, 31), a finding that perhaps reflects shared structural features in C3 and C4. Alternatively, the capacity of the COS cells expressing the piBD and piCD constructs to bind EC4(ma) may have been caused by the transfer of nucleotides encoding the NH<sub>2</sub>-terminal 36 amino acids from SCR-1 of LHR-A to LHR-B, and -C through the ligation of the Bsm I fragments. However, these 36 amino acids alone did not confer on the piD product C4-rosetting function. We cannot exclude a secondary function of LHR-D in these reactions because this LHR was present in all the constructs assayed for function. The finding of three distinct ligand recognition sites in CR1, two for C3b and one for C4b (Fig. 12), indicates that each

receptor molecule may be capable of effectively binding complexes bearing multiple C4b and C3b molecules despite having a relatively low affinity for monovalent ligands (32). This finding also provides an explanation for the inability of soluble C4b to inhibit formation of rosettes between erythrocytes bearing C3b and a human B lymphoblastoid cell line (33). The recent description of covalent complexes between C4b and C3b (34), and perhaps between C3b molecules, suggests that possible ligands for which CR1 would be especially adapted may be the molecular complexes C4b/C3b and C3b/C3b, which are generated during activation of the classical and alternative pathways, respectively. The presence of distinct binding sites in three of the four LHRs also suggests that the CR1 structural allotypes differing by their number of LHRs may have significant functional differences caused by variations in the number of ligand-binding sites. Although *in vitro* studies have not reported differing binding activities of the F, S, and F' (A-C, respectively) allotypes, the smaller F' allotype, presumably having only three LHRs, has been reported possibly to be associated with systemic lupus erythematosus (3), perhaps reflecting an impaired capability of this allotype to participate in the clearance of immune complexes.

The demonstration that the epitope recognized by YZ1 monoclonal anti-CR1 was present in LHR-A, -B, and -C indicates that prior estimates based on the binding of this antibody of the number of CR1 molecules expressed by various cell types may be excessive (35, 36). Supporting this conclusion is an earlier study in which the concentration of CR1 in a purified preparation of receptor estimated by an assay using YZ1 was 2.7-fold greater than that estimated by protein determination (26). In addition, studies have reported variation in the number of erythrocyte epitopes recognized by different mAbs (37, 38). However, the low number of CR1 sites on erythrocytes of patients with lupus cannot be accounted for by abnormal expression of epitopes as patients and normals have the same frequency of the F and S allotypes, and diminished numbers of receptors were also found when assayed by the binding of soluble C3b (39).

### Summary

Complementary DNA clones encoding the NH<sub>2</sub>-terminal region of human CR1 have been isolated and sequenced. The deduced complete amino acid sequence of the F allotype of human CR1 contains 2,039 residues, including a 41-residue signal peptide, an extracellular domain of 1,930 residues, a 25-amino acid transmembrane domain, and a 43-amino acid cytoplasmic region. The extracellular domain is composed exclusively of 30 short consensus repeats (SCRs), characteristic of the family of C3/C4-binding proteins. The 28 NH<sub>2</sub>-terminal SCRs are organized as four long homologous repeats (LHRs) of seven SCRs each. The newly sequenced LHR, LHR-A, is 61% identical to LHR-B in the NH<sub>2</sub>-terminal two SCRs and >99% identical in the COOH-terminal five SCRs. Eight cDNA clones were spliced to form a single construct, piABCD, that contained the entire CR1 coding sequence downstream of a cytomegalovirus promoter. COS cells transfected with piABCD transiently expressed recombinant CR1 that comigrated with the F allotype of erythrocyte CR1 on SDS-PAGE and that mediated rosette formation with sheep erythrocytes bearing C4b and C3b. Recombinant CR1 also had factor I-cofactor activity for cleavage of C3(ma). Analyses of six deletion mutants expressed in COS cells indicated that



the NH<sub>2</sub>-terminal two SCRs of LHR-A contained a site determining C4 specificity and the NH<sub>2</sub>-terminal two SCRs of LHR-B and -C each had a site determining C3 specificity. The presence of these three distinct sites in CR1 may enable the receptor to interact multivalently with C4b/C3b and C3b/C3b complexes generated during activation of the classical and alternative pathways.

The authors thank Dr. Brian Seed for providing the COS cells, the CDM8 vector, and the strain, MC1061/P3; Dr. Richard Selden for providing the plasmid, pXGH5, and L cells; Drs. Luigi Bergamaschini and Michael Carroll for providing EAC4b; Dr. Keith Peden for providing the vector, pMT.neo.1; and Drs. Joseph Ahearn and Winnie Wong for advice and assistance in the preparation of cellular intermediates.

*Received for publication 30 June 1988 and in revised form 11 August 1988.*

### References

1. Ahearn, J. M., and D. T. Fearon. The molecular biology of complement receptors type 1 (CR1, CD35) and type 2 (CR2; CD21). *Adv. Immunol.* In press.
2. Ross, G. D., and M. E. Medof. 1985. Membrane receptors specific for bound fragments of C3. *Adv. Immunol.* 37:217.
3. van Dyne, S., V. M. Holers, D. M. Lublin, and J. P. Atkinson. 1987. The polymorphism of the C3b/C4b receptor in the normal population and in patients with systemic lupus erythematosus. *Clin. Exp. Immunol.* 68:570.
4. Klickstein, L. B., W. W. Wong, J. A. Smith, J. H. Weis, J. G. Wilson, and D. T. Fearon. 1987. Human C3b/C4b receptor (CR1): demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristic of C3/C4 binding proteins. *J. Exp. Med.* 165:1095.
5. Kristensen, T., P. D'Eustachio, R. T. Ogata, L. P. Chung, K. B. Reid, and B. F. Tack. 1986. The superfamily of C3b/C4b-binding proteins. *Fed. Proc.* 46:2463.
6. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
7. Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligo-thymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA.* 69:1408.
8. Current Protocols in Molecular Biology. F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl, editors. John Wiley & Sons, New York. 1987.
9. Wong, W. W., L. B. Klickstein, J. A. Smith, J. H. Weis, and D. T. Fearon. 1985. Identification of a partial cDNA clone for the human receptor for complement fragments C3b/C4b. *Proc. Natl. Acad. Sci. USA.* 82:7711.
10. Wong, W. W., C. A. Kennedy, E. T. Bonacchio, J. G. Wilson, L. B. Klickstein, J. H. Weis, and D. T. Fearon. 1986. Analysis of multiple restriction length polymorphisms of the gene for the human complement receptor type 1: duplication of genomic sequences occurs in association with a high molecular weight receptor allotype. *J. Exp. Med.* 164:1531.
11. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
12. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene.* 33:103.
13. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted break-points for DNA sequencing. *Gene.* 28:351.

14. Chen, E. Y., and P. H. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA (NY)*. 4:165.
15. Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387.
16. Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature (Lond.)*. 329:840.
17. Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb. 1984. High-level expression of a chloramphenicol acetyltransferase gene by DEAE-dextran-mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucleic Acids Res.* 12:5707.
18. Changelian, P. S., R. M. Jack, L. A. Collins, and D. T. Fearon. 1985. PMA induces the ligand-independent internalization of CR1 on human neutrophils. *J. Immunol.* 134:1851.
19. Fraker, P. J., and J. C. Speck, Jr. 1978. Protein and cell membrane iodinations with a sparingly soluble chloroamide, 1, 3, 4, 6-tetrachlor-3a, 6a-diphenylglycoluril. *Biochem. Biophys. Res. Commun.* 80:849.
20. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)*. 227:680.
21. Lambris, J. D., O. Scheiner, T. F. Schulz, and M. P. Dierich. 1983. Coupling of C3b to erythrocytes by disulfide bond formation: preparation of EC3b for hemolytic and complement receptor assays. *J. Immunol. Methods.* 65:277.
22. Hammer, C. H., G. H. Wirtz, L. Renfer, H. D. Gresham, and B. F. Tack. 1981. Large scale isolation of functionally active components of the human complement system. *J. Biol. Chem.* 256:3995.
23. Fearon, D. T. 1977. Purification of C3b inactivator and demonstration of its two polypeptide chain structure. *J. Immunol.* 119:1248.
24. Von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683.
25. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 44:283.
26. Wong, W. W., R. M. Jack, J. A. Smith, C. A. Kennedy, and D. T. Fearon. 1985. Rapid purification of the human C3b/C4b receptor (CR1) by monoclonal antibody affinity chromatography. *J. Immunol. Methods.* 82:303.
27. Holers, V. M., T. Seya, E. Brown, J. J. O'Shea, and J. P. Atkinson. 1986. Structural and functional studies on the human C3b/C4b receptor (CR1) purified by affinity chromatography using a monoclonal antibody. *Complement.* 3:63.
28. Selden, R. F., K. Burke-Howie, M. E. Rowe, H. M. Goodman, and D. D. Moore. 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol. Cell. Biol.* 6:3173.
29. Fearon, D. T., I. Kaneko, and G. G. Thomson. 1981. Membrane distribution and adsorptive endocytosis by C3b receptors on human polymorphonuclear leukocytes. *J. Exp. Med.* 153:1615.
30. Fujita, T., and V. Nussenzweig. 1979. The role of C4-binding protein and  $\beta$ 1H in proteolysis of C4b and C3b. *J. Exp. Med.* 150:267.
31. Pangburn, M. K., R. D. Schreiber, and H. J. Muller-Eberhard. 1977. Human complement C3b inactivator: isolation, characterization, and demonstration of an absolute requirement for the serum protein  $\beta$ 1H for cleavage of C3b and C4b in solution. *J. Exp. Med.* 146:257.
32. Arnaout, M. A., N. Dana, J. Melamed, R. Medicus, and H. R. Colten. 1983. Low ionic strength or chemical cross-linking of monomeric C3b increases its binding affinity to the human complement C3b receptor. *Immunology.* 48:229.
33. Gaither, T. A., I. T. Magrath, M. Berger, C. H. Hammer, L. Novikovs, M. Santaella,

- and M. M. Frank. 1983. Complement receptor expression by neoplastic and normal human cells. *J. Immunol.* 131:899.
34. Takata, Y., T. Kinoshita, H. Kozono, J. Takeda, E. Tanaka, K. Hong, and K. Inoue. 1987. Covalent association of C3b with C4b within C5 convertase of the classical complement pathway. *J. Exp. Med.* 165:1494.
  35. Wilson, J. G., W. D. Ratnoff, P. H. Schur, and D. T. Fearon. 1986. Decreased expression of the C3b/C4b receptor (CR1) and the C3d receptor (CR2) on B lymphocytes and of CR1 on neutrophils of patients with systemic lupus erythematosus. *Arthritis Rheum.* 29:739.
  36. Wilson, J. G., W. W. Wong, E. E. Murphy, III, P. H. Schur, and D. T. Fearon. 1987. Deficiency of the C3b/C4b receptor (CR1) of erythrocytes on systemic lupus erythematosus: analysis of the stability of the defect and of a restriction fragment length polymorphism of the CR1 gene. *J. Immunol.* 138:2706.
  37. Edberg, J. C., E. Wright, and R. P. Taylor. 1987. Quantitative analyses of the binding of soluble complement-fixing antibody/dsDNA immune complexes to CR1 on human red blood cells. *J. Immunol.* 139:3739.
  38. Iida, K., R. Mornaghi, and V. Nussenzweig. 1982. Complement receptor (CR1) deficiency in erythrocytes from patients with systemic lupus erythematosus. *J. Exp. Med.* 155:1427.
  39. Wilson, J. G., W. W. Wong, P. H. Schur, and D. T. Fearon. 1982. Mode of inheritance of decreased C3b receptors on erythrocytes of patients with systemic lupus erythematosus. *N. Engl. J. Med.* 307:981.