Molecular identification of Knops blood group polymorphisms found in long homologous region D of complement receptor 1

Joann M. Moulds, Peter A. Zimmerman, Ogobara K. Doumbo, Lalla Kassambara, Issaka Sagara, Dapa A. Diallo, John P. Atkinson, Malgorzata Krych-Goldberg, Richard E. Hauhart, Dennis E. Hourcade, David T. McNamara, Daniel J. Birmingham, J. Alexandra Rowe, John J. Moulds, and Louis H. Miller

Complement receptor 1 (CR1) has been implicated in rosetting of uninfected red blood cells to *Plasmodium falciparum*–infected cells, and rosette formation is associated with severe malaria. The Knops blood group (KN) is located on CR1 and some of these antigens, ie, Mc-Coy (McC) and Swain-Langley (Slª), show marked frequency differences between Caucasians and Africans. Thus, defining the molecular basis of these antigens may provide new insight into the mechanisms of *P falciparum* malaria. Monoclonal antibody epitope mapping and serologic inhibition studies using CR1 deletion constructs localized McC and Slª to long homologous repeat D of CR1. Direct DNA sequencing of selected donors identified several single nucleotide polymorphisms in exon 29 coding for complement control protein modules 24 and 25. Two of these appeared to be blood group specific: McC associated with K1590E and Slª with R1601G. These associations were confirmed by inhibition studies using allele-specific mutants. A sequence-specific oligonucleotide probe hybridization assay was developed to genotype several African populations and perform family inheritance studies. Concordance between the 1590 mutation and McC was 94%; that between Slª and 1601 was 88%. All but 2 samples exhibiting discrepancies between the genotype and phenotype were found to be due to low red cell CR1 copy numbers, low or absent expression of some alleles, or heterozygosity combined with low normal levels of CR1. These data further explain the variability observed in previous serologic studies of CR1 and show that DNA and protein-based genetic studies will be needed to clarify the role of the KN antigens in malaria. (Blood. 2001;97:2879-2885)

© 2001 by The American Society of Hematology

Introduction

Complement receptor 1 (CR1) is an integral membrane protein found on red blood cells (RBCs), macrophages, neutrophils, lymphocytes, follicular dendritic cells, and kidney podocytes. The major function of CR1 is the binding to and removal of C4b- and C3b-bearing immune complexes, although it also serves as a complement regulatory protein.[1] Recently, CR1 has been implicated in the rosetting of *Plasmodium falciparum*–infected RBCs to uninfected red cells through its interactions with *P falciparum* erythrocyte membrane protein 1 (PfEMP-1).[2] Because several studies[3-6] have shown that in vitro rosette formation is associated with severe malaria, the potential exists for CR1 polymorphisms to play a role in reducing susceptibility to malaria pathogenesis.

The extramembrane portion of CR1 is composed of 30 complement control protein modules (CCPs) or short consensus repeats. These CCPs are further arranged into 4 long homologous regions (LHRs) having arisen by duplication of 7 CCPs (Figure 1). The CR1 protein exhibits 3 known polymorphisms, including molecular weight, RBC expression levels, and the Knops (KN) blood group. CR1 molecular weights differ by approximately 30 kd, ranging from the smallest form of 190 kd (CR1*3) to the largest

form of 280 kd (CR1*4). The basis of this molecular-weight polymorphism is the deletion and duplication by unequal crossover events of the CCPs, mostly contained in LHR-B.[7] The CR1 structural variants differ in their number of CCPs/LHRs and subsequently the number of C3b-binding sites, with the smallest form having only 1 binding site and the largest having 4 sites. In all populations studied to date, the 2 most frequent alleles are CR1*1 (220 kd) and CR1*2 (250 kd).[8]

Unlike most other cells, RBC-CR1 levels can vary as much as 10-fold among healthy individuals.[1] This accounts for the variability observed in serologic tests for the Knops blood group described below.[9] Furthermore, RBCs having low levels of CR1 (approximately 10% of normal) have the serologic null phenotype known as the "Helgeson phenotype."[9,10] Although RBC expression appears to be under genetic control, the exact mechanism has not been determined. In Caucasians, the RBC-CR1 copy number is linked to a *Hin*dIII restriction fragment length polymorphism (RFLP).[11] This association, however, was not apparent in a study of African Americans.[12] Rowe et al[2] reported that RBCs with CR1 levels less than approximately 100 copies/cell, ie, Helgeson phenotype,
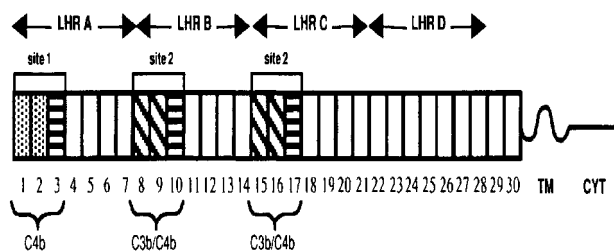
Figure 1. Diagram of the most common allotypic variant of CR1. The 30 complement control protein repeats (CCPs) constitute the extramembranous portion of CR1. These modules are followed by a hydrophobic transmembrane domain (TM) and a 43–amino acid cytoplasmic tail (CYT). There are 2 distinct functional sites (site 1 in LHR-A and 2 nearly identical copies of site 2 in LHR-B and -C), each composed of 3 CCPs. Functional differences between the sites are determined by sequence differences. The first 2 CCPs in site 1 (CCP 1 and 2) are 39% different from the first 2 CCPs in site 2 (CCPs 8 and 9), as well as from CCPs 15 and 16, and they are marked by different shading. The third CCP in site 1 is nearly identical to the third CCP (10 or 17) in site 2. CCPs 3, 10, and 17 are represented by boxes with horizontal lines.

showed markedly reduced rosetting in in vitro tests using the R29R$^+$ *P falciparum* parasite strain.

The Knops blood group localized to CR1[13,14] includes several presumed allelic pairs: Knops a and b (Kn^a, Kn^b), McCoy a and b (McC^a, McC^b), Swain-Langley (Sl^a), and Villien (Vil).[15] Similar to the in vitro rosetting observations associated with the Helgeson phenotype, preliminary studies using RBCs from individuals serologically phenotyped to be Sl(a−) consistently showed the formation of fewer rosettes with R29R$^+$/PfEMP1-transfected COS7 cells. RBCs with high CR1 copy numbers can type as either Sl(a+) or Sl(a−), indicating that this phenotype is not solely dependent on CR1 copy number. The frequency of the Sl(a−) phenotype is increased among African Americans and blacks residing in Africa, where malaria is endemic.[16] Results from these earlier studies suggested that defining the molecular basis of the Knops blood group antigens may provide new insights into the mechanisms of severe malaria caused by *P falciparum.* This study was therefore conducted to identify DNA sequence polymorphisms associated with the Sl(a−) and McC^b phenotypes, both of which occur at elevated frequencies in African-derived populations.

Materials and methods

Study area and population

All samples were obtained under protocols approved by the appropriate institutional review boards, and donors gave informed consent before participating. Samples representing North American ethnic groups were obtained from the University of Texas–Houston Medical School and the American Red Cross. The sampling included 250 African Americans and 101 samples from Mali, including 16 related individuals from 9 families. Additional blood samples from West African individuals (n = 182) from 17 villages distributed between Senegal and Ghana were collected as part of studies conducted in conjunction with the Onchocerciasis Control Programme, Ouagadougou, Burkina Faso.[17] Donor samples containing the various Knops system antibodies were obtained through the Serum, Cell, and Rare Fluid Exchange as well as from individual immunohematology reference laboratories.

Blood samples

Five to 20 mL of blood was collected by venipuncture into tubes containing either acid-citrate-dextrose or EDTA. The samples were centrifuged and the plasma was separated into aliquots for frozen storage at −20°C. For some donors, the buffy coat was removed and used to prepare genomic DNA. A modified Alsever's solution (Gamma Biologicals, Houston, TX) was added

to the RBCs for preservation until testing was complete. When possible, the remaining RBCs were frozen in 40% sucrose and stored in liquid nitrogen.

Serologic phenotyping

The RBCs were tested within 36 hours of collection in a blinded manner. The samples were typed for ABO using commercial antisera, and a direct Coombs test was performed using rabbit polyspecific antihuman globulin (Gamma Biologicals). Because commercial antisera were unavailable, phenotyping of the Knops blood group antigens (including Kn^a, McC^a, McC^b, Sl^a, and Vil) was done using antisera obtained from sensitized donors. The initial testing was performed using single-source antisera, and samples discrepant with the genotype were studied further with additional examples of the antibody specificity. For the KN typings, 50 μL of a 4% solution of washed RBCs was added to 75 μL of antisera and allowed to ncubate for 60 minutes at 37°C. The cells were washed 3 times in phosphate-buffered saline (PBS, pH 7.2), and 2 drops of antiglobulin reagent (Gamma Biologicals) were added, followed by centrifugation. The test was read with an agglutination viewer and scored by standard methods.[18]

RBC-CR1 quantification

RBC ghosts were prepared by hypotonic lysis in the presence of protease inhibitors, and the proteins were solubilized in 1% NP-40/PBS. The samples were kept frozen at −80°C until testing. The number of CR1 copies per RBC was determined as described previously using J3D3 for capture and E11 as the detection antibody.[9,16] The binding site for E11 has been shown previously to reside in a nonduplicated region of CR1 and is not affected by the CR1 size polymorphism.[19,20] A standard curve was prepared for each assay from which test values were interpolated. A single donor previously studied as part of the VIIth International Complement Genetics Workshop was used for the standard.[21]

Western blotting of CR1 proteins

A second portion of the RBC ghosts was solubilized in sodium dodecyl sulfate (SDS) loading buffer and boiled for 10 minutes. The proteins were separated by nonreducing SDS–polyacrylamide gel electrophoresis with a 3% stacking gel and a 5% resolving gel. The proteins were transferred electrophoretically to nylon membranes that were blocked overnight in 5% milk powder prepared in PBS. The blots were washed and incubated with one of several anti-CR1 monoclonal antibodies (MAbs) for 1 hour at 24°C. The blot was washed again and similarly incubated with sheep antimouse conjugated with horseradish peroxidase. Following a final wash, the proteins were visualized by enhanced chemiluminescence (Amersham Pharmacia Biotech, Piscataway, NJ) and a 1-minute exposure to film.[16] Molecular-weight determinations were made in comparison with known standards (Bio-Rad, Hercules, CA). A total of 20 MAbs to CR1 were screened for blood group specificity, and 2 (6B1.H12 and 3C6.D11) were chosen for further studies of the allelic expression of the McC^a and Sl^a antigens.

CR1 constructs

CR1 deletion constructs for LHR-A, LHR-B, and LHR-D$^+$ (CCPs 22-30) have been described previously by Krych-Goldberg et al.[22] Constructs introducing amino acid substitutions associated with McC^b (K1590E; single-letter amino acid code) and Sl^a (R1601G) were made as reported previously[22] by incorporating into the LHR-D$^+$ cDNA the single nucleotide polymorphisms c4795A>G and c4828A>G, respectively. An additional construct introduced amino acid substitution K1590R. For the inhibition studies,[23] 1 μg of soluble CR1 LHR-D$^+$ was used to inhibit 75 μL of the selected KN antisera for 15 minutes at 24°C. Following the inhibition incubation period, 50 μL of a 4% solution of antigen-positive RBCs from a donor having moderate to high RBC-CR1 was added as an indicator cell, and testing proceeded as described above. Inhibition was defined as the lack of hemagglutination after addition of the indicator RBCs.

Table 1. Polymerase chain reaction primers and complement receptor 1 amplification strategies

Exon	CCP	Primer	LHR	Sequence	Annealing temperature	Amplicon length (bp)
22/23	19/20a	19/20aUP	C	5′-ACC TCT GAC TAG CTA TGA GGT CTT C-3′	65°C	518
		19/20aDN		5′-TTA AGC TCA AGA GTC ACT CTG CAG-3′		
24	20b	20bUP	C	5′-AAT GGT GCA TTC ATC CAG CCA CTA-3′	65°C	164
		20bDN		5′-AAA CAG AAA GAT CCA GAC TGA ACA AGT-3′		
25	21	21UP	C	5′-AGT ATG ACA ACT GTA GAA TCA CCT TG-3′	65°C	336
		21DN		5′-CCA TTT AGG ATA TGT GGA AGT GGA TA-3′		
26	22	22UP	D	5′-GAG GTA GGG TGG AAG TCT CTC TG-3′	65°C	279
		22DN		5′-GCT TTA ACA GAC ATG AAG TAG TTC C-3′		
27	23a	23aUP	D	5′-ACT CTA TAC TTA ATC CCA AAT TCT GC-3′	65°C	184
		23aDN		5′-TCT GCT GTT CTG GAA CTC ATA TTC-3′		
28	23b	23bUP	D	5′-ATG GTG AGT TAA ATG GGA AAT ATG-3′	60°C	200
		23bDN		5′-AGG GGG CCA GGC AGG ATC ACA C-3′		
29	24/25	24/25UP	D	5′-TAA AAA ATA AGC TGT TTT ACC ATA CTC-3′	60°C	476
		24/25DN		5′-CCC TCA CAC CCA GCA AAG TC-3′		
30/31	26/27a	26/27aUP	D	5′-TTT TAA GCC ATC TGG TAA GCA TAA G-3′	65°C	505
		26/27aDN		5′-GAT CTC AAG AGA GTG ATT CTC CAC AC-3′		
32	27b	27bUP	D	5′-AGA TGG GAA TTG CTC ACA CAT TTG G-3′	58°C	185
		27bDN		5′-CTG GAA AAC GAA AAT AAG GAG AGC-3′		
33	28	28UP	D	5′-GTC ACA GGT CAC TAT TGT TTC AG-3′	65°C	332
		28DN		5′-CCA TTT AGG ATA TGT GGA AGC AGG-3′		
34	29	29UP		5′-TGT GTT TGT GTG GGA ACT TGT TCT TAG-3′	65°C	247
		29DN		5′-CCA TAC AAC CCA GCA AGA ATA AGG-3′		
35	30	30UP		5′-TAA TTT CAG AAG TAA ATT GTA GCT TCC-3′	63°C	233
		30DN		5′-CAT GAC CGG AAG ATC CAC ATT C-3′		
36	Tma	TmaUP		5′-ATG CCA GAG TGA TGT TTT GTG AC-3′	58°C	86
		TmaDN		5′-ATC AAA ACT TTC ATA AAA CTT ACC-3′		
37	Tmb	TmbUP		5′-CTT GTA AGT TAT ATT TTC CAT GC-3′	58°C	200
		TmbDN		5′-TGT CTT TTA CTT TCT TTT GAC TTC C-3′		
38	Cyto	CytoUP		5′-GAC AGT TTT TCT ATT TTT TTC TCT GC-3′	65°C	180
		CytoDN		5′-GAG TTG TTG AAA ACA TTT TGA ATT CAG-3′		

CCP indicates complement control protein repeat; LHR, long homologous repeat (Figure 1); UP, sense; DN, antisense; Tm, transmembrane; Cyto, cytoplasmic.

DNA extraction and polymerase chain reaction amplification

DNA extraction was performed using the QIAmp blood extraction protocol (Qiagen, La Jolla, CA). Strategies to characterize DNA sequence polymorphisms associated with KN blood group polymorphisms McC^a, McC^b, and Sl^a were based upon LHR-D, exon-specific amplification (exons 26-35, CCP 22-30; CR1*1 nomenclature) from genomic DNAs. Overall, DNA sequence analysis performed in this study covered exons 22 to 38, CCPs 19 through the cytoplasmic tail. Polymerase chain reaction (PCR) amplification was performed in a final volume of 25 μL containing 2.5 pmol of the appropriate positive-strand and negative-strand primers (Table 1); 67 mM Tris-HCl (pH 8.8); 6.7 mM MgSO$_4$; 16.6 mM (NH$_4$)$_2$SO$_4$; 10 mM 2-mercaptoethanol; 100 μM dATP, dGTP, dCTP, and dTTP; 2.5 U of thermostable DNA polymerase; and 10 to 50 ng of purified human genomic DNA. The thermocycling program used for each of the exon-specific amplifications was 30 seconds at 94°C, 30 seconds at the appropriate primer annealing temperature (Table 1), and 30 seconds at 72°C for 40 cycles.

Post-PCR cloning and DNA sequence analysis

Following amplification of LHR-D–specific exons from human genomic DNA templates, PCR products were either sequenced directly or cloned into pCR2.1TOPO (Invitrogen, Carlsbad, CA). Multiple clones from subjects of interest were sequenced using fluorescence-based sequencing protocols on an ABI377 automated sequencer (PerkinElmer, Boston, MA).

Sequence-specific oligonucleotide probe hybridization analysis

To prepare PCR products for dot blotting, we heated 10 μL of amplicon solution to 95°C for 2 minutes, followed by addition of an equal volume (10 μL) of 20 × SSC (3.0 M NaCl, 0.3 M Na$_3$-citrate, pH 7.0). Two microliters of this solution was then spotted in duplicate onto Hybond N$^+$ filters

(Amersham Pharmacia Biotech). After drying, the nylon filters were bathed first in denaturing buffer (0.4 N NaOH) for 10 minutes and second in neutralizing buffer (1.8 M NaCl, 0.1 M NaH$_2$PO$_4$, 0.01 M EDTA) for 1 minute. The filters were air-dried and cross-linked by UV light exposure in the Stratagene UV cross-linker (La Jolla, CA). Filters prepared in this manner were incubated in 10 mL of hybridization buffer (0.75 M NaCl, 0.075 M Na$_3$-citrate, 0.1% SDS, 5% liquid block [supplied by Amersham Pharmacia Biotech], 30% dextran sulfate) for 1 hour before adding labeled sequence-specific oligonucleotide probes (SSOPs). Appropriate fluorescein isothiocyanate–labeled SSOPs (Table 2) were added to the hybridization buffer for overnight incubation at 40°C. After hybridization, the filters were twice washed in 5 × SSC at room temperature for 5 minutes each, followed by 2 15-minute high-stringency washes optimized for specific hybridization of each individual SSOP (Table 2). Detection was performed using the enhanced chemifluorescence signal amplification kit (Amersham Pharmacia Biotech) following the supplier's recommended protocol. Fluorescence was detected using the Storm 860 (Molecular Dynamics, Sunnyvale, CA).

Table 2. CCP 25 SSOPs and hybridization conditions

Probe	Sequence	Stringency temperature	Wash condition salt
4795A	5′-CTA CTA ATA AAT GCA C-3′	37°C	0.8x
4795G	5′-CTA CTA ATG AAT GCA C-3′	40°C	0.8x
4828A	5′-TGC AAT TAG AGT ACC-3′	39°C	1.0x
4828C	5′-GGT ACT CCA ATT GCA-3′	44°C	0.4x
4855T	5′-TTT CTT TTC CCT CAC-3′	37°C	1.0x
4855A	5′-TTT CTT TAC CCT CAC-3′	37°C	0.8x
4803	5′-AGC TCC AGA AGT TGA-3′	44°C	1.0x

SSOP indicates sequence-specific oligonucleotide probe. Underlined base pair indicates the polymorphic site.

Statistical analysis

Fisher's exact test was calculated using standard methods and computed using the programs available through Simple Interactive Statistical Analysis (SISA; http://members.aol.com/johnp71/javastat.html). Associations between each of the CR1-CCP 25 single nucleotide polymorphisms (snps) and McC(a+), McC(b+), and Sl[a] phenotypes were analyzed following the procedures described by Weir.[24]

Results

Localization of McC[a], McC[b], and Sl[a] epitopes to LHR-D of CR1

Twenty different CR1 MAbs were screened for evidence of blood group specificity. Two CR1 MAbs, previously identified as being specific for LHR-D,[20] were shown to have KN blood group specificity. One antibody (3C6.D11) did not react with most cells that were either McC(a−) or Sl(a−) (Figure 2). The MAb 6B1.H12 was nonreactive with McC(a−) RBCs. Inhibition tests using soluble CR1 containing only LHR-A, LHR-B, LHR-D+, or the full-length LHR-ABCD were performed to confirm LHR-D as the candidate region (Table 3). The combined data indicated that McC and Sl[a] antigens were located in LHR-D and that further analysis of this region of the CR1 gene might reveal DNA sequence polymorphisms associated with these KN antigens.

DNA sequence survey of CR1–LHR-D encoding exons

Survey of the LHR-D of CR1 for associations between McC and/or Sl[a] phenotypes and DNA sequence polymorphisms was performed by exon-specific PCR amplification followed by DNA sequence analysis. Nucleotide sequence comparisons were made between phenotypically different Caucasian Americans, African Americans, and study subjects from Mali. These included McC(a+b−)/Sl(a+), McC(a+b+)/Sl(a+), McC(a+b−)/Sl(a−), McC(a+b+)/Sl(a−), and McC(a−b+)/Sl(a−) individuals. In the region covering exons 22 through 38, initial comparisons between Caucasian and African Americans identified several snps. These snps were evaluated to determine whether they led to an amino acid substitution, and, if so, whether genotypic and phenotypic differences
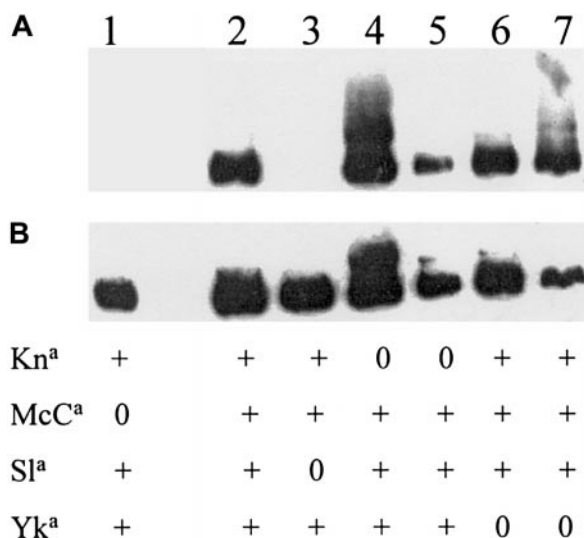

Figure 2. Localizing KN antigens to LHR-D. Western blots of red cell CR1 from donors selected for their KN phenotype (shown below each lane). (A) The MAb 3C6.D11 raised against LHR-D. To demonstrate that protein was present in lanes 1 and 3, the same blot was washed and incubated with another CR1 MAb (J3D3) that is known to react with all size and blood group variants of CR1 (B).

Table 3. Inhibition of Knops antisera using soluble CR1

CR1 peptide	McC[a]-1	McC[a]-2	McC[b]	Sl[a]-1	Sl[a]-2	Sl[a]-3	Sl[a]-4	Vil
PBS control	NI	NI	NI	NI	NI	NI	NI	NI
LHR-ABCD*	I	I	NI	I	I	I	I	NI
LHR-A	NI	NI	NI	NI	NI	NI	NI	NI
LHR-B	NI	NI	NI	NI	NI	NI	NI	NI
LHR-D+	I	I	NI	I	I	I	I	NI

NI indicates no inhibition; I, inhibition.

*We have previously shown that LHR-D+ has the phenotype McC(a+b−), Sl(a+). Therefore, only anti-McC[a] and Sl[a] are inhibited and not the antithetical sera McC[b] or Vil.

Two different anti-McC[a] and 4 anti-Sl[a] sera were used in the inhibition tests.

between the Caucasian and African American study subjects could be associated. The 9 snps identified through these comparisons are summarized in Table 4 according to cDNA (nucleotide and amino acid) positions corresponding to GenBank Y00816. Because of the consistency with which snps at positions 4795 were observed for McC(a+)/McC(b+), and 4828 were observed for Sl(a+)/Sl(a−) individuals, further analysis of DNA sequence polymorphism in the 99 Mali study subjects focused upon exon 29 encoding CCPs 24 and 25. Direct sequencing of exon 29 from Mali study subjects revealed that the snps at c4795A>G (K1590E) and c4828A>G (R1601G) were frequently observed. Individual clones from Caucasian Americans, African Americans, and Mali study subjects revealed 5 different CCP 25 allelic DNA sequences and their predicted allelic amino acid sequences. No polymorphisms have been observed in the DNA sequence encoding CCP 24.

Development of PCR-SSOP genotyping system

Following a comparison of single-strand conformational polymorphism, heteroduplex, restriction endonuclease, sequence-specific priming, and SSOP hybridization analyses, the SSOP hybridization assay was chosen to perform post-PCR, high-throughput genotyping of snps identified in CCP 25. Using the probe hybridization and washing conditions described earlier (Table 2), we evaluated SSOP hybridization specificity on previously cloned and sequenced CCP 25 alleles. Results from control experiments illustrated that hybridization specificities were successfully developed for each of the SSOPs used in this genotyping system (data not shown).

In a subset of 44 individuals analyzed by DNA sequence analysis and SSOP hybridization assays, genotyping results were 100% concordant between methodologies at each snp position. Because some related Malians were sampled to enable linkage analysis between snps and serologic phenotypes, the Mali study group is not a random sample. To provide a population-based perspective of the frequency and distribution of the CCP 25 snps, a wider survey of other West Africans was performed. These results (Table 5) show a similar representation of genotype frequencies between Malian and other West African study groups.

Table 4. Single nucleotide polymorphisms (snps) identified in exons 22 through 38 (LHR-D+) of the CR1 gene

Exon	CCP	BP change	AA change
22	19	c3650A > G	H1208R
25	21	c4250C > T	T1408I
26	22	c4041A > C	Silent
29	25	c4795A > G	K1590E
29	25	c4828A > G	R1601G
29	25	c4855T > A	S1610T
29	25	c4870A > G	I1615V
33	28	c5507C > G	P1827R
34	29	c5575G > C	D1850H

Table 5. CR1-CCP 25 genotype frequencies at snp positions in West African study subjects

Study group	n	4795[a]			4828[b]		
		A/A	A/G	G/G	A/A	A/G	G/G
Mali	99	49	40	10	9	30	60
West Africa*	182	81	89	12	7	61	114

a indicates McC locus; b, Sl[a] locus.

*West Africa includes Senegal, Guinea, Sierra Leone, Ivory Coast, and Ghana.

Pedigree-based linkage of DNA sequence polymorphism with McC and Sl[a] epitopes

Of the 9 Malian families chosen for further studies, 5 were informative for the inheritance of McC and 6 were informative at the Sl[a] locus. In these families, the inheritance of the McC[b] allele correlated with an E at position 1590, whereas Sl[a] correlated with R at 1601. The family shown in Figure 3 demonstrates inheritance at both the McC and Sl[a] loci. When Sl(a−) family members were typed for the putative allele Vil, they were Vil[+] with the exception of those samples having low RBC-CR1.

Allele-specific inhibition of anti-McC[a], McC[b], Sl[a], and Vil

Biochemical confirmation of the CR1 amino acid and blood group antigen association was pursued using soluble CR1 LHR-D modified by site-directed mutagenesis (Table 6). In these studies, as expected, substitution of K to E at amino acid 1590 changed the serologic reactivity from McC[a] to McC[b], but inhibition studies using anti-Sl[a] were more complicated. Without changing the amino acid at position 1601(Sl[a]), the 1590E mutant unexpectedly could no longer inhibit the anti-Sl[a] sera. However, the reverse situation was not true; the change of R to G at 1601 had no effect on the McC (1590) antigens. Because the K (positive charge) to E (negative charge) substitution at 1590 represents a significant physical change and may disrupt the domain structure, we tried an alternative strategy using a 1590R-1601R construct. This neutral charge substitution at 1590 affected only the McC epitope and did not influence the reactivity with the Sl[a] antisera. Because the 1590E-1601R haplotype is not observed in the Mali study subjects, it was not possible to assess further the observation from these in vitro
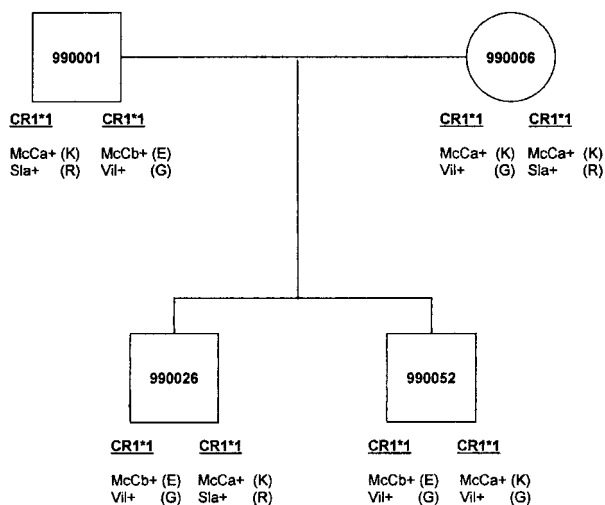

Figure 3. An example of genetic inheritance of KN. Malian family shows inheritance of McC[a] with amino acid 1590K, McC[b] with 1590E, Sl[a] with 1601R, and Vil with 1601G. The father passes the CR1*1 allele bearing McC[b] (E)/Vil (G) to both sons. The mother passes the McC[a] (K)/Vil (G) haplotype to son 990052 and the McC[a] (K)/Sl[a] (R) haplotype to son 990026.

Table 6. Inhibition of Knops antisera using various mutants of soluble CR1 LHR-D[+]

CR1 mutant	McC[a]-1	McC[a]-2	McC[b]	Sl[a]-1	Sl[a]-2	Sl[a]-3	Sl[a]-4	Vil
1590K-1601R*	I	I	NI	I	I	I	I	NI
1590**E**-1601R	NI	NI	I	NI	NI	NI	NI	NI
1590**R**-1601R	NI	NI	NI	I	I	I	I	NI
1590K-1601**G**	I	I	NI	NI	NI	NI	NI	I
1590**E**-1601**G**	NI	NI	I	NI	NI	NI	NI	I

I indicates inhibition; NI, no inhibition.

*Wild-type LHR-D[+]; same as in Table 3. Mutations at 1590 and 1601 are noted in bold type. 1590K = McC[a]; 1590E = McC[b] (allele to McC[a]); 1590R = no change in net charge; 1601R = Sl[a]; 1601G = Vil (allele to Sl[a]).

inhibition studies suggesting that the K to E substitution at amino acid 1590 may influence the Sl[a] phenotype. Finally, the double mutant (1590E-1601G) changed the inhibition of serologic reactivity from McC[a] to McC[b] and Sl[a] to Vil.

Correlating DNA sequence polymorphisms with serologic reactivity

DNA sequence and serologic polymorphisms were correlated for the Mali study subjects only. Analyses of allelic and phenotypic associations were performed individually for 1590K with McC(a+), 1590E with McC(b+), and 1601R/G with Sl[a]/Vil. Excluding the 2 donors having the Helgeson phenotype, overall genotype to phenotype concordance was 94% (93 of 99 cases) for the 1590/McC serology and 88% (87 of 99 cases) for the1601/Sl[a] serology. These results (Table 7) show statistically significant associations between 1590K and McC(a+), 1590E and McC(b+), and 1601R/G and the Sl[a] phenotype (Fisher's exact test, $P < .0001$). There were 18 instances in which the genotype did not correlate with the KN phenotype in the studies of the Mali samples (Table 7). Only 6 of 99 samples were discordant for the McC typings, and all were heterozygous KE by DNA typing. Three of the 6 discordant McC samples were heterozygous for the molecular-weight polymorphism, with one allele being expressed at lower levels than the other (Table 8; nos. 011, 017, 008). In these cases, Western blots could be used to determine which antigens were on each size variant. For example, using sample 011, the MAb 6B1.H12 reacted with the 220-kd protein (Figure 4; lane 1), which was expressed in higher quantity (lane 2). Thus, the lower expressed 190-kd form (lane 2) must have McC[b] but tested negative by serology. The other 3 discordant McC samples appeared homozygous for molecular weight and were not studied for individual allele expression at this time. Thus, the presence of either a hypomorph or an amorph could not be ruled out as a cause of these discrepancies.

Discordance between the 1601 genotype and Sl[a] phenotype included 11 heterozygous (RG) individuals who typed Sl(a−). Of these, 6 exhibited heterozygosity for the CR1 structural alleles and were studied as described earlier. In 5 of the 6 cases (Table 8; nos. 008, 003, 035, 041, 080), a low-expression allele was observed,

Table 7. Associations between CR1 genotype and Knops phenotype in Mali subjects

Nucleotide*	Position 4828		Position 4795		
	Sl(a+)	Sl(a−)	McC(a+b−)	McC(a+b+)	McC(a−b+)
A/A	9	0	49	0	0
A/G	19	**11**	**4**	33	**2**
G/G	**1**	59	0	0	11

*At both positions 4828 (Sl[a]) and 4795 (McC), the A nucleotide codes for the wild type and G for the mutation. Bold numbers indicate the number of samples exhibiting a typing discrepancy between genotype and phenotype.

Table 8. Samples exhibiting discrepancies between genotype and phenotype of the KN antigens

Number	McCᵃ	McCᵇ	1590	Slᵃ	1601	Haplotype*	CR1#	Mᵣ allele†	6B1	3C6
011	+	0	KE	+	RG	↓ EG/KR	453	↓ 190/220	220	—
017	+	0	KE	0	GG	↓ EG/KG	424	↓ 190/220	220	—
008	0	+	KE	0	RG	EG/KR ↓	186	220/250 ↓	250	250
003	+w	+	KE	0	RG	EG/KR ↓	240	220/250 ↓	—	250
035	+	0	KK	0	RG	↓ KR/KG	254	↓ 220/250	—	220
041	+	0	KK	0	RG	KG/KR ↓	384	220/250 ↓	—	220/250
080	+	0	KK	0	RG	↓ KR/KG	392	↓ 220/250	—	220
063	+	0	KK	0	RG	KR/KG	236	220/280	—	220

6B1 indicates MAb 6B1.H12; 3C6, MAb 3C6.D11; ↓, allele exhibiting decreased expression.

*Haplotype as determined by cloning of individual amplicons and sequencing of multiple clones.

†Molecular-weight alleles were initially determined using MAb J3D3 and nonreduced gels.

similar to the findings for McC. Figure 4 shows a sample (no. 008) that was discordant for both McCᵃ and Slᵃ. Western blots showed that the less-expressed 250-kd allele (lane 3) carried the McCᵃ (K) and Slᵃ (R) genotype, resulting in false-negative typings for both antigens. In addition, the total RBC-CR1 level was in the low normal range (186, versus normal mean 301), which may also contribute to the negative serologic typings. One sample (Table 8; no. 063) showed equal expression of the structural alleles and normal total RBC-CR1 (total 236). However, because the sample was heterozygous for Slᵃ (RG), only 50% of the CR1 protein (118 copies) would carry Slᵃ, which is not sufficient for hemagglutination and results in a false-negative phenotype.

Discussion

CR1 and malaria

The ability of *P falciparum*–infected red cells to form rosettes correlates with an increase in disease severity (ie, profound anemia or coma) and may relate to the high mortality of this type of malaria.[3-6] Hence, Rowe et al[2] studied the proteins involved in rosetting and reported that CR1 interacted with PfEMP-1. Their specific findings showed that low RBC-CR1 expression (Helgeson phenotype) and the Sl(a−) phenotype were associated with lower in vitro rosette formation compared with normal RBC-CR1 expres-

Figure 4. McC/Slᵃ haplotypes were determined by Western blot to resolve typing discrepancies. Donor no. 011 (Table 8) exhibits weaker expression of the 190-kd form and normal expression of the 220-kd protein (lane 2) when probed with the standard anti-CR1 (J3D3). Probing with the monoclonal anti-McCᵃ (6B1.H12) shows that only the 220-kd protein carries McCᵃ (lane 1). Therefore, the 190-kd form carries McCᵇ, and because of its weak expression, a false-negative result is obtained for serologic typing. Donor no. 008 (Table 8) has a low total RBC-CR1 level (186 copies/RBC) and exhibits weaker expression of the 250-kd form compared with the 220-kd form (lane 4). With the use of MAb 3C6.D11, the 250-kd protein is shown to carry both McCᵃ and Slᵃ (lane 3). The low expression of this allele contributes to the false-negative serologic typing for both McCᵃ and Slᵃ. Molecular weights are noted to the side of each allele, and the genotypes for positions 1590 (McC) and 1601 (Sla) are in parentheses.

sion and the Sl(a+) phenotype. Additional work has demonstrated that the frequency of Sl(a−) and McC(b+) is greatly increased among West Africans.[16] These results suggest that defining the molecular polymorphisms associated with the KN blood group antigens might identify specific molecular interactions that contribute to adhesion of *P falciparum*–infected and –uninfected RBCs.

Localization and identification of McC and Slᵃ

Earlier studies of several ethnic groups provided no evidence that the KN blood group antigens were based on the CR1 molecular-weight polymorphism. Comparisons of the McC and Slᵃ phenotypes with the molecular-weight polymorphisms among the Malians also showed no consistent trend associating KN antigens and size polymorphisms.[16] Additionally, a donor homozygous for the CR1*3 allele, presumed to be due to a deletion of LHR-B, expressed all of the high-frequency KN antigens, including Knᵃ, McCᵃ, Slᵃ, and Ykᵃ (J.M.M., unpublished data, November 1996).

In the present study, we localized the McC and Slᵃ blood group antigens on CR1 to LHR-D using epitope mapping by MAbs, inhibition tests with CR1 deletion constructs, and direct DNA sequencing. Furthermore, we found a cluster of 3 snps within exon 29 encoding CCPs 24 and 25, all of which led to amino acid substitutions within the first half of CCP 25. Two of these snps have now been identified as antigens in the KN blood group system: the K1590E polymorphism is associated with the McCᵃ and McCᵇ antigens, whereas the R1601G is associated with the Slᵃ and Vil phenotypes. The alleles McCᵇ and Vil were not assigned initially to the KN blood group system because they lacked the biochemical and/or genetic information required for inclusion. Based on data provided in this study, McCᵇ has recently been assigned the KN number 022006 and Vil the number 022007 by the International Society of Blood Transfusion. Furthermore, because these have been shown to be separate genetic mutations, the McCᶜ (Slᵃ) and McCᵈ (Vil) terminology previously proposed[25] should be discarded.

A third snp, c4870A>G (I1615V), has also been reported by Xiang et al.[26] This snp and other previously reported mutations in LHR-D (including the P1827R and the *Hin*dIII RFLP) have not been associated with any of the known KN antigens.[12] The P1827R mutation creates a *Mnl*I restriction enzyme site that has been used to further study RBC-CR1 expression. Both the *Hin*dIII and *Mnl*I RFLPs correlate with CR1 copy number among Caucasians, but a similar association has not been found in African Americans[12] or Malians (J. Middleton and L. Hernandez, unpublished data, August 1998).

Comparison of typing data

As reported previously,[9] we found that the RBC-CR1 level affects the ability to serologically phenotype the cells for the KN blood group antigens. Two of the Malian samples had the Helgeson
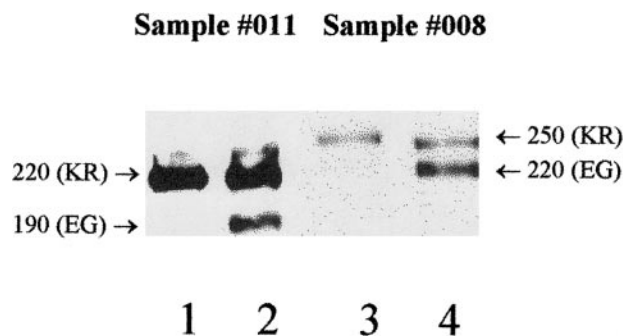
phenotype, which is known to result in false-negative serologic typing for all of the KN blood group antigens. Furthermore, most of the discordant typings for either McC or Sl[a] were found among individuals who were heterozygous by DNA genotyping and had either low-normal levels of RBC-CR1 or low expression of one allele. In the first situation, because only half of the CR1 proteins carried the respective blood group antigens, false-negative serologic types were obtained even though the total RBC-CR1 was in normal range. In these cases, the amount of antigen-positive CR1 protein on the RBC membrane is insufficient for hemagglutination to occur. When 2 CR1 proteins differing in molecular weight were present, we demonstrated low expression of one allele and assigned the KN antigens by Western blot (Figure 4). In these cases, the low-expressing alleles carried the blood group epitopes, but the reduced levels of that protein resulted in the false-negative serologic phenotype. Interestingly, reduced RBC-CR1 has been reported to be an acquired phenomenon due to immune-complex–mediated diseases such as systemic lupus erythematosus (SLE).[27,28] However, in 2 of the Malian families analyzed for inheritance, the low-expressing CR1 structural allele was passed from parent to child, showing that this is an inherited characteristic.[15] Similar results were reported by Van Dyne et al,[29] who investigated CR1 expression in Caucasian families with SLE. The exact genetic mechanism for low CR1 expression is presently unknown.

Finally, the remaining discrepant samples could be explained by the presence of an amorph, which has been reported for all of the major blood group systems.[15] Alternatively, there may be other mutations in CCP 25 or LHR-D that contribute to the Sl(a−) phenotype. For example, anti-Chido was originally believed to recognize a single epitope on the fourth component of complement, known as Ch[a]. Following molecular investigation by Yu et al,[30] the Ch[a] antigen was subdivided into 6 antigens, named Ch1 to Ch6, which result from 4 amino acid (bp) changes. A similar situation may exist for Sl[a].

In summary, we have identified the molecular basis for the McC[a], McC[b], Sl[a], and Vil blood group antigens and have developed DNA-based typing methods for their ascertainment. This methodology, in combination with the serologic typing, quantification of RBC-CR1 levels, and Western blots of the CR1 proteins, will facilitate the investigation of the KN antigens in the rosetting phenomenon as well as malaria pathophysiology.

Acknowledgments

References

1. Ahearn JA, Fearon DT. Structure and function of the complement receptors CR1 (CD35) and CR2 (CD21). Adv Immunol. 1989;46:183-219.

2. Rowe A, Moulds JM, Newbold CI, Miller LH. P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. Nature. 1997;388:292-295.

3. Carlson J, Helmby H, Hill AVS, Brewster D, Greenwood BM, Wahlgren M. Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. Lancet. 1990;336:1457-1460.

4. Kun JFJ, Schmidt-Ott RJ, Lehman LG, et al. Merozoite surface antigen 1 and 2 genotypes and rosetting Plasmodium falciparum in severe and mild malaria in Lambarene, Gabon. Trans R Soc Trop Med Hyg. 1998;92:110-114.

5. Rowe JA, Obeiro J, Newbold CI, Marsh K. Plasmodium falciparum rosetting is associated with malaria severity in Kenya. Infect Immun. 1995;63: 2323-2326.

6. Treutiger CJ, Hedlund I, Helmby H, et al. Rosette formation in Plasmodium falciparum isolates and anti-rosette activity of sera from Gambians with cerebral or uncomplicated malaria. Am J Trop Med Hyg. 1992;46:503-510.

7. Hourcade D, Miesner DR, Bee C, Zeldes W, Atkinson JP. Duplication and divergence of the amino-terminal coding region of the complement receptor 1 (CR1) gene. J Biol Chem. 1990;265: 974-980.

8. Cohen JHM, Atkinson JP, Klickstein LB, Oudin S, Subramanian B, Moulds JM. The C3b/C4b receptor (CR1, CD35) on erythrocytes: methods for study of the polymorphisms. Mol Immunol. 1999; 36:819-825.

9. Moulds JM, Moulds JJ, Brown M, Atkinson JP. Antiglobulin testing for CR1-related (Knops/McCoy/Swain-Langley/York) blood group antigens: negative and weak reactions are caused by variable expression of CR1. Vox Sang. 1992;62:230-235.

10. Helgeson M, Swanson J, Polesky HF. Knops-Helgeson (Kna), a high frequency erythrocyte antigen. Transfusion. 1970;10:137-138.

11. Wilson JG, Murphy EE, Wong WW, Klickstein LB, Weis JH, Fearon DT. Identification of a restriction length polymorphism by a CR1 cDNA that correlates with the number of CR1 on erythrocytes. J Exp Med. 1986;164:50-59.

12. Herrera AH, Xiang L, Martin SG, Lewis J, Wilson JG. Analysis of complement receptor type one (CR1) expression on erythrocytes and of CR1 allelic markers in Caucasian and African American populations. Clin Immunol Immunopathol. 1999;87:176-183.

13. Moulds JM, Nickells MW, Moulds JJ, Brown M, Atkinson JP. The C3b/C4b receptor is recognized by the Knops, McCoy, Swain-Langley and York blood group antisera. J Exp Med. 1991;173:1159-1163.

14. Rao N, Ferguson DJ, Lee S-F, Telen MJ. Identification of human erythrocyte blood group antigens on the C3b/C4b receptor. J Immunol. 1991;146: 3502-3507.

15. Daniels G. Human Blood Groups. Oxford, United Kingdom: Blackwell Science; 1995.

16. Moulds JM, Kassambara L, Middleton JJ, et al. Identification of complement receptor one (CR1) polymorphisms in West Africa. Genes Immunity. 2000;1:325-329.

17. Zimmerman PA, Dadzie KY, DeSole G, Remme J, Alley ES, Unnasch TR. Onchocerca volvulus DNA probe classification correlates with epidemiologic patterns of blindness. J Infect Dis. 1992; 65:964-968.

18. Walker RH, ed. Technical Manual. 10th ed. Arlington, Va: American Association of Blood Banks; 1990.

19. Quadri RA, Schifferli JA. Over-estimation of the number of complement receptor type 1 (CR1) on erythrocytes. Scand J Immunol. 1992;36:125-130.

20. Nickells M, Hauhart R, Krych M, et al. Mapping epitopes for 20 monoclonal antibodies to CR1. Clin Exp Immunol. 1998;112:27-33.

21. Moulds JM, Brai M, Cohen J, et al. Reference typing report for complement receptor one (CR1). Exp Clin Immunogenet. 1998;15:291-294.

22. Krych-Goldberg M, Hauhart EL, Subramanian BV, et al. Decay accelerating activity of complement receptor type one (CD35): two active sites are required for dissociating C5 convertases. J Biol Chem. 1999;274:31160-31168.

23. Moulds JM, Rowe K. Neutralization of Knops system antibodies using soluble complement receptor one. Transfusion. 1996;36:517-520.

24. Weir BS. Genetic Data Analysis. Sunderland, MA: Sinauer; 1996.

25. Molthan L. Expansion of the York, Cost, McCoy, Knops blood group system: the new McCoy antigens McCc and McCd. Med Lab Sci. 1983;40: 113-121.

26. Xiang L, Rundles JR, Hamilton DR, Wilson JG. Quantitative alleles of CR1: coding sequence analysis and comparison of haplotypes in two ethnic groups. J Immunol. 2000;163:4939-4945.

27. Ross GD, Yount WJ, Walport MJ, et al. Disease-associated loss of erythrocyte complement receptors (CR1, C3b receptors) in patients with systemic lupus erythematosus and other diseases involving autoantibodies and/or complement activation. J Immunol. 1985;135:2005-2014.

28. Corvetta A, Pomponio G, Bencivenga R, et al. Low number of complement C3b/C4b receptors (CR1) on erythrocytes from patients with essential mixed cryoglobulinemia, systemic lupus erythematosus and rheumatoid arthritis: relationship with disease activity, anticardiolipin antibodies, complement activation and therapy. J Rheumatol. 1991;18:1021-1025.

29. Van Dyne S, Holers VM, Lublin DM, Atkinson JP. The polymorphism of the C3b/C4b receptor in the normal population and in patients with systemic lupus erythematosus. Clin Exp Immunol. 1987; 68:570-579.

30. Yu CY, Campbell RD, Porter RR. A structural model for the location of the Rodgers and the Chido antigenic determinants and their correlation with the human complement component C4A/C4B isotypes. Immunogenetics. 1988;27: 399-405.