

Erythrocyte complement receptor 1 (CR1) expression level is not associated with polymorphisms in the promoter or 3' untranslated regions of the CR1 gene

I. A. Cockburn* & J. A. Rowe

Summary

Complement receptor 1 (CR1) expression level on erythrocytes is genetically determined and is associated with high (H) and low (L) expression alleles identified by a *Hind*III restriction fragment-length polymorphism (RFLP) in intron 27 of the *CR1* gene. The L allele confers protection against severe malaria in Papua New Guinea, probably because erythrocytes with low *CR1* expression, are less able to form pathogenic rosettes with *Plasmodium falciparum*-infected erythrocytes. Despite the biological importance of erythrocyte CR1, the genetic mutation controlling *CR1* expression level remains unknown. We investigated the possibility that mutations in the upstream or 3' untranslated regions of the *CR1* gene could control erythrocyte CR1 level. We identified several novel polymorphisms; however, the mutations did not segregate with erythrocyte CR1 expression level or the H and L alleles. Therefore, high and low erythrocyte CR1 levels cannot be explained by polymorphisms in transcriptional control elements in the upstream or 3' untranslated regions of the *CR1* gene.

Introduction

Complement receptor 1 (CR1) is a large (≈ 200 kDa) immune-regulatory protein expressed on the surface of erythrocytes, monocytes, neutrophils, leukocytes and glomerular

Abbreviations: CR1, complement receptor one; H, high expression allele; L, low expression allele; RFLP, restriction fragment length polymorphism; UTR, untranslated region.

* Current address: Department of Molecular Microbiology and Immunology, Malaria Research Institute, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA.

Institute of Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

Received 7 July 2005; revised 5 August 2005; accepted 5 September 2005

Correspondence: J. Alexandra Rowe, Institute of Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, Ashworth Laboratories, West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131650 5492; Fax: +44 131650 6564; E-mail: alex.rowe@ed.ac.uk

podocytes (Birmingham & Hebert, 2001). CR1 on the surface of erythrocytes binds to immune complexes and pathogenic microorganisms and facilitates their clearance (Birmingham & Hebert, 2001); however, the malaria parasite, *Plasmodium falciparum*, exploits CR1 as an adhesion ligand in the process of rosetting — the pathogenic clumping of infected erythrocytes to uninfected red cells (Rowe *et al.*, 1997). The normal level of erythrocyte CR1 in Europeans is between 50 and 1200 molecules per cell (Wilson *et al.*, 1986). In Papua New Guinea, low erythrocyte CR1 levels (< 100 molecules per cell) are extremely common (Cockburn *et al.*, 2004) and are thought to have been selected for because such low levels of erythrocyte CR1 reduce *P. falciparum* rosetting (Rowe *et al.*, 1997), and confer protection against severe malaria (Cockburn *et al.*, 2004). Low erythrocyte CR1 levels in Europeans (Wilson *et al.*, 1986) and Papua New Guineans (Cockburn *et al.*, 2004) are genetically determined and are associated with a *Hind*III restriction fragment-length polymorphism (RFLP) in intron 27 of the *CR1* gene. Subsequent work has shown that this RFLP is part of a haplotype called the L allele, which consists of a number of synonymous and non-synonymous mutations in the *CR1* gene (Xiang *et al.*, 1999). It is not known which, if any, of these coding sequence polymorphisms are functionally important in determining erythrocyte CR1 expression level. It has been suggested that the amino acid substitutions in the L allele render CR1 more susceptible to proteolytic cleavage, and this could lead to low erythrocyte CR1 levels (Herrera *et al.*, 1998). Direct evidence for this hypothesis is lacking, and other factors do not support this suggestion. The L allele only affects the expression of *CR1* on erythrocytes, but not cells of other lineages (Wilson *et al.*, 1986), whereas in Africa, the H and L alleles exist but are not associated with *CR1* expression (Rowe *et al.*, 2002). This is most simply explained if the mutation causing CR1 deficiency has not arisen in African populations, or it is present but not linked to the H and L alleles.

The possibility that polymorphisms in the transcriptional control elements of the *CR1* gene could determine high and low expression on erythrocytes has not yet been examined. An analogous situation has been described in relation to the Duffy blood group antigen. The Duffy antigen is the receptor for erythrocyte invasion by *Plasmodium*

vivax. Individuals with Duffy-negative erythrocytes are extremely common in West Africa, and they are completely resistant to vivax malaria. The Duffy-negative phenotype occurs because of a mutation in a GATA1 binding site at nucleotide -46 of the Duffy antigen/chemokine receptor gene (Tournamille *et al.*, 1995). As GATA1 is an erythroid-specific transcription factor, this mutation affects expression of the Duffy antigen only on erythroid lineage cells. We hypothesized that a similar mutation at an erythroid-specific transcription factor site in the CR1 promoter could be responsible for low erythrocyte CR1 expression. The exact promoter region of CR1 has not been clearly defined; however, a 38 base-pair (bp) region between -79 and -41 bp upstream of the transcription start site is known to be important, and the promoter region could stretch more than 1800 kb upstream of the start site (Funkhouser & Vik, 2000). The closest potential GATA1 binding site to the transcription start site in the CR1 gene is at position -1009. Binding sites for other transcription factors such as Ets and AML1 have been characterized at -49 and -43, respectively (Kim *et al.*, 1999).

The aim of this study was to sequence the upstream and 3' untranslated region (UTR) of the CR1 gene to determine if they contain any mutations linked to the L allele that could control erythrocyte CR1 expression level.

Materials and methods

Samples

Blood samples were collected and DNA extracted as described previously (Rowe *et al.*, 2002; Cockburn *et al.*, 2004).

CR1 expression level

Erythrocyte CR1 levels of the samples from Papua New Guinea and the United Kingdom had been determined previously by flow cytometry (Cockburn *et al.*, 2004). The erythrocyte CR1 levels of the samples from Mali were determined previously by ELISA (Rowe *et al.*, 2002). The flow cytometry assay is described in full by Cockburn *et al.* (2002), and the ELISA assay by Moulds *et al.* (1992). Briefly, for the flow cytometry assay, the CR1 level was determined using the anti-CR1 monoclonal antibody J3D3 (Nickells *et al.*, 1998), with comparison to a standard curve derived from a set of reference erythrocytes with known CR1 levels. The reference erythrocytes used to establish the assay had their CR1 expression initially determined by the use of ¹²⁵I-labeled antibody and Scatchard analysis to determine the number of antigenic sites per cell, controlling for the number of antibody-binding sites per CR1 molecule. For the ELISA assay, erythrocyte 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 molecules per erythrocyte was determined using J3D3 for capture and E11 as the detection antibody. A standard curve was prepared for each assay from which test values were interpolated. A

single donor previously studied as part of the VIIIth International Complement Genetics Workshop was used for the standard.

PCR and sequencing

The upstream region of the gene was amplified as three overlapping polymerase chain reaction (PCR) products. The regions were from -788 to +251 using the primers CR1pro1F-CCCCACCCCACAACAG and CR1pro1R-ACGCCCCCGCCCTCTCAC; from -1507 to -313 using the primers CR1pro2F-GAGATTGGTTAAAGGGC ACAGGAT and CR1pro2R-AAGGGCAATTGCAGGC GATGTGTA and a region from -673 to -2050 using the primers CR1pro3F-GAAAAGAGTTGGATTGGTTGGT-TAG and CR1pro3R-AGAGTTAGCCCCCTTTGCGC TAGT. All sequence positions are based on the cDNA sequence published previously (Vik & Wong, 1993). The entire exon 39, including the 3'UTR (to +7590), was amplified as a 1526-bp fragment using the primers CR13utrF1-AGCCTGTAAATTCTGGGTAGTTCA and CR13utrR1-CTGTCAAAGTAAATGTGGCAAAT. The PCR products were sequenced with internal primers using the BigDye terminator kit (Applied Biosystems, Warrington, UK).

Single nucleotide polymorphism (SNP) analysis

Several polymorphisms were identified and analysed in more detail. A mutation at position -873 was analysed by digestion of the CR1pro2 PCR product with EcoT221 (Amersham Pharmacia, Little Chalfont, UK), using the buffer supplied by the manufacturer for 2 h at 37 °C. With a T at position -873 the 1195-bp product remains uncut, while with the C-873 genotype it is cut into fragments of 636 bp and 559 bp. The mutation at position -159 was analysed by sequencing of the CR1pro1 PCR product with the primer CR1seq1F2-TCGGCAAAGCTCCCCCT GAC, while polymorphisms at +6994 and +7119 were analysed by sequencing of 3' UTR PCR product using the primer CR13utrF3-TGGCGTAATCTCGGCTCACT.

Data analysis

Genetic linkage between polymorphisms was determined using a permutations method in the program GENETIX (University of Montpellier, France), whereas statistical analysis was performed using STATVIEW (SAS Institute, Cary, NC, USA).

Results and discussion

Initially, we sequenced 2 kb of the upstream region of the CR1 gene in six individuals, three homozygous for the L allele and three homozygous for the H allele. Four of the individuals (two HH and two LL) had identical upstream sequences, suggesting that that erythrocyte CR1 expression level is not associated with a promoter polymorphism. The other two individuals, one HH and one LL,

Table 1. Gene frequencies of polymorphisms in the upstream and 3' untranslated regions of the *CR1* gene

	-873	-159	H/L alleles <i>Hind</i> III RFLP ^a	+6994	+7119
New Ireland (Papua New Guinea)	C – 0.08 T – 0.92 (88) ^b	A – 0.01 G – 0.99 (88)	L – 0.73 H – 0.27 (94)	T – 0.25 C – 0.75 (28)	T – 0.25 C – 0.75 (28)
Mali	C – 0.41 T – 0.59 (76)	A – 0.23 G – 0.76 T – 0.01 (76)	L – 0.15 H – 0.85 (298)	T – 0.21 C – 0.79 (24)	T – 0.21 C – 0.79 (24)
Edinburgh (UK)	C – 0.31 T – 0.69 (58)	A – 0.19 G – 0.81 (58)	L – 0.23 H – 0.77 (60)	T – 0.46 C – 0.54 (24)	T – 0.46 C – 0.54 (24)

^a Determined previously (Rowe *et al.*, 2002; Cockburn *et al.*, 2004).

^b Number of chromosomes tested.

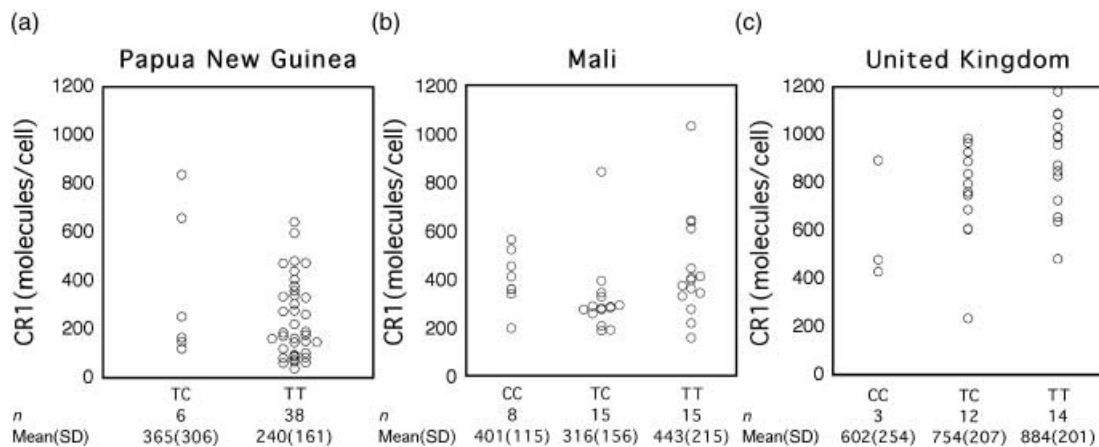


Figure 1. CR1 level on erythrocytes and the T-873C CR1 promoter polymorphism in (a) New Ireland, Papua New Guinea; (b) Mali; and (c) Edinburgh, UK. Each point represents the mean erythrocyte CR1 level of a single individual in molecules per cell. SD, standard deviation.

were heterozygous at three sites: T-873C, C-391T and G-159A (most common allele first). Subsequent sequencing did not find T-391 in any other individuals, but the other two polymorphisms occurred commonly and were analysed in more detail. We did not detect any polymorphisms in the previously identified transcription factor binding sites in the promoter region.

To determine whether the polymorphisms in the upstream region are associated with erythrocyte CR1 expression level, we genotyped 44 individuals from New Ireland (Papua New Guinea), 38 individuals from Mali (from various locations) and 29 individuals from Edinburgh (UK) at the -873 and -159 positions (Table 1), using RFLP analysis and direct sequencing, respectively. All individuals in these populations had been characterized previously for their erythrocyte CR1 level and H and L genotypes (Rowe *et al.*, 2002; Cockburn *et al.*, 2004). In the three populations studied, the -873 and -159 polymorphisms were tightly linked ($P < 0.01$), with the A-159 mutation being rarer than C-873, perhaps having arisen later (Table 1). In Papua New Guinea and Mali, we found that the polymorphisms were not linked to either the H or L allele, nor were they associated with any decrease or increase in CR1

expression ($P > 0.2$ by two-way ANOVA, including *Hind*III genotype in the model; Fig. 1a & b — data for the -873 polymorphism only is shown). In the European population, both the -873 and -159 polymorphisms were associated with CR1 expression level ($P < 0.05$ one-way ANOVA; Fig. 1c — data for the -873 polymorphism only is shown). However, linkage analysis showed that these polymorphisms are genetically linked to the *Hind*III polymorphism ($P = 0.04$), and that once this was included in the statistical model, the promoter polymorphisms had no effect on CR1 expression ($P > 0.10$ by two-way ANOVA). It is perhaps not surprising that there is some linkage between the promoter and the *Hind*III polymorphism in intron 27 (≈ 80 kb apart) in Europeans but not in Africans, as linkage disequilibrium extends across much larger distances in non-Africans because they have recently passed through a small population bottleneck (Reich *et al.*, 2001).

A more limited analysis was performed to examine if 3'UTR polymorphisms were associated with the *Hind*III H and L alleles and erythrocyte CR1 expression levels. In the original six donors, we identified linked polymorphisms at positions C6994T and C7119T that appeared initially to segregate with the L and H alleles. However, subsequent

sequencing of this region in at least 12 individuals from each population showed no linkage between these polymorphisms and either CR1 expression or the *Hind*III polymorphism (Table 1).

These data suggest that erythrocyte CR1 expression levels are not controlled by mutations in the upstream or in the 3' UTRs of the *CR1* gene. The element determining erythrocyte CR1 expression level must therefore be in the internal regions of the *CR1* gene. It has been suggested that coding sequence polymorphisms linked to the *Hind*III polymorphism might make the CR1 molecule more amenable to cleavage (Herrera *et al.*, 1998). However, the fact that the L allele exists in Africa without affecting erythrocyte *CR1* expression means this explanation requires another factor acting *in trans*, for example, a protease, to vary between the two populations, and this has not been shown. The precise cause of the variation in erythrocyte CR1 expression levels associated with the H and L genotypes remains unknown. One possibility that remains to be explored is that the messenger RNA encoded by the L allele may be less stable or less efficiently spliced than that of the H allele.

In the current study, we have looked at the upstream and 3'UTR polymorphisms with respect to erythrocyte CR1 expression; however, *CR1* is also expressed on myeloid and lymphoid lineage cell lines. The promoter and 3'UTR polymorphisms described here might have been selected if they affect *CR1* expression in these cell types. Alternatively, the polymorphisms may be non-functional and could have arisen as a result of genetic drift. In summary, we have described a number of novel upstream and 3'UTR polymorphisms in the *CR1* gene. These polymorphisms are not associated with erythrocyte CR1 expression level. These results suggest that the level of CR1 on erythrocytes is under the control of an unusual means of regulation that remains to be elucidated.

Acknowledgements

We thank all the volunteers in Edinburgh, New Ireland and Mali who donated blood for the study, and Dr Elizabeth MacKinnon and Dr Joann Moulds for providing samples from New Ireland and Mali. This work was supported by a Wellcome Trust Senior Research Fellowship to JAR (Grant number 067431) and a United Kingdom Medical Research Council Studentship to IAC.

References

Birmingham, D.J. & Hebert, L.A. (2001) CR1 and CR1-like: the primate immune adherence receptors. *Immunological Reviews*, **180**, 100.

- Cockburn, I.A., Donvito, B., Cohen, J.H. & Rowe, J.A. (2002) A simple method for accurate quantification of complement receptor 1 on erythrocytes preserved by fixing or freezing. *Journal of Immunological Methods*, **271**, 59.
- Cockburn, I.A., Mackinnon, M., O'Donnell, A., Allen, S.J., Moulds, J.M., Baisor, M., Bockarie, M., Reeder, J.C. & Rowe, J.A. (2004) A human complement receptor 1 polymorphism that reduces *Plasmodium falciparum* rosetting confers protection against severe malaria. *Proceedings of the National Academy of Sciences of the USA*, **101**, 272.
- Funkhouser, T. & Vik, D.P. (2000) Promoter activity of the 5' flanking region of the complement receptor type 1 (CR1) gene: basal and induced transcription. *Biochimica et Biophysica Acta*, **1490**, 99.
- Herrera, A.H., Xiang, L., Martin, S.G., Lewis, J. & Wilson, J.G. (1998) Analysis of complement receptor type 1 (CR1) expression on erythrocytes and of CR1 allelic markers in Caucasian and African American populations. *Clinical Immunology and Immunopathology*, **87**, 176.
- Kim, J.H., Lee, S., Rho, J.K. & Choe, S.Y. (1999) AML1, the target of chromosomal rearrangements in human leukemia, regulates the expression of human complement receptor type 1 (CR1) gene. *International Journal of Biochemistry and Cell Biology*, **31**, 933.
- Moulds, J.M., Moulds, J.J., Brown, M. & Atkinson, J.P. (1992) 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 Sanguinis*, **62**, 230.
- Nickells, M., Hauhart, R., Krych, M., Subramanian, V.B., Geoghegan-Barek, K., Marsh, H.C. Jr & Atkinson, J.P. (1998) Mapping epitopes for 20 monoclonal antibodies to CR1. *Clinical and Experimental Immunology*, **112**, 27.
- Reich, D.E., Cargill, M., Bolck, S., Ireland, J., Sabeti, P.C., Richter, D.J. *et al.* (2001) Linkage disequilibrium in the human genome. *Nature*, **411**, 199.
- Rowe, J.A., Moulds, J.M., Newbold, C.I. & Miller, L.H. (1997) *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature*, **388**, 292.
- Rowe, J.A., Raza, A., Diallo, D.A., Baby, M., Poudiougo, B., Coulibaly, D. *et al.* (2002) Erythrocyte CR1 expression level does not correlate with a *Hind*III restriction fragment length polymorphism in Africans; implications for studies on malaria susceptibility. *Genes and Immunity*, **3**, 497.
- Tournamille, C., Colin, Y., Cartron, J.P. & Le Van Kim, C. (1995) Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genetics*, **10**, 224.
- Vik, D.P. & Wong, W.W. (1993) Structure of the gene for the F allele of complement receptor type 1 and sequence of the coding region unique to the S allele. *Journal of Immunology*, **151**, 6214.
- Wilson, J.G., Murphy, E.E., Wong, W.W., Klickstein, L.B., Weis, J.H. & Fearon, D.T. (1986) Identification of a restriction fragment length polymorphism by a CR1 cDNA that correlates with the number of CR1 on erythrocytes. *Journal of Experimental Medicine*, **164**, 50.
- Xiang, L., Rundles, J.R., Hamilton, D.R. & Wilson, J.G. (1999) Quantitative alleles of CR1: coding sequence analysis and comparison of haplotypes in two ethnic groups. *Journal of Immunology*, **163**, 4939.