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

Citation for published version:

Digital Object Identifier (DOI):
10.1111/j.1744-313X.2005.00552.x

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
International journal of immunogenetics

Publisher Rights Statement:
RoMEO yellow

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
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* and J. A. Rowe
Institute of Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

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 HindIII 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 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 HindIII 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 VIIth 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- 
CCCCCACCACCAACAG and CR1pro1RACGCCGCCCCGCTTCAC; from −1507 to 
−313 using the primers CR1pro2F-GAGATTTGAATGGAAGGTAAT and 
CR1pro2R-AAGGGCTATCGTGGAGATCTGTA and a region from −673 to −2050 
using the primers CR1pro3F-GAAAAGAGTTGGATTAGTGGTGGTTAG and CR1pro3R- 
AGAGTTAGCCCCCTTTTGCAGC 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- 
AGCTGTAAAATCTCGTGGTGGTTAGTGAATATT and CR13utrR1- 
CTGTCAAGATAGTGCGGAAAAAT. 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- 
TCGGCAAAAGCTCCCCTGAC, while polymorphisms at +6994 and +7119 were 
analysed by sequencing of 3′ UTR PCR product using the primer CR13utrF3- 
TGGCGTAATCTCGTGGCTCCT.

**Data analysis**

Genetic linkage between polymorphisms was determined using a permutations method in 
the program **GENETIX** (University of Montpelier, 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 
hozygous 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, 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 HindIII 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 HindIII 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 HindIII 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 HindIII 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 HindIII 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 HindIII 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.

Acknowledgments

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.

**Abbreviations**

- CR1: complement receptor one
- H: high expression allele
- L: low expression allele
- RFLP: restriction fragment length polymorphism
- UTR: untranslated region

**References**


Vik DP, Wong WW. 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. 1993; 151:6214.

*Int J Immunogenet.* Author manuscript; available in PMC 2010 May 13.

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.
Table 1

Gene frequencies of polymorphisms in the upstream and 3′ untranslated regions of the \textit{CR1} gene

<table>
<thead>
<tr>
<th></th>
<th>-873</th>
<th>-159</th>
<th>H/L alleles</th>
<th>HindIII RFLP$^a$</th>
<th>+6994</th>
<th>+7119</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New Ireland (Papua New Guinea)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.08</td>
<td>A</td>
<td>0.01</td>
<td>L</td>
<td>0.73</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>0.92</td>
<td>G</td>
<td>0.99</td>
<td>H</td>
<td>0.27</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>(88)</td>
<td>(88)</td>
<td>(94)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mali</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.41</td>
<td>A</td>
<td>0.23</td>
<td>L</td>
<td>0.15</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>0.59</td>
<td>G</td>
<td>0.76</td>
<td>H</td>
<td>0.85</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>(76)</td>
<td>(298)</td>
<td></td>
<td></td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td><strong>Edinburgh (UK)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.31</td>
<td>A</td>
<td>0.19</td>
<td>L</td>
<td>0.23</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>0.69</td>
<td>G</td>
<td>0.81</td>
<td>H</td>
<td>0.77</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>(58)</td>
<td>(58)</td>
<td>(60)</td>
<td></td>
<td>24</td>
<td>24</td>
</tr>
</tbody>
</table>

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

$^b$ Number of chromosomes tested.