## *Brief Definitive Report*

# A PHYSICAL MAP OF THE HUMAN REGULATOR OF COMPLEMENT ACTIVATION GENE CLUSTER LINKING THE COMPLEMENT GENES *CR1*, *CR2*, *DAF*, AND *C4BP*

BY JAVIER REY-CAMPOS,* PABLO RUBINSTEIN,* AND SANTIAGO RODRIGUEZ DE CORDOBA*‡

*From the *Department of Immunogenetics, The Lindsley F. Kimball Research Institute of The New York Blood Center, New York 10021; and the ‡Centro de Investigaciones Biologicas, Consejo Superior de Investigaciones Cientificas (CSIC), Madrid, Spain*

The human genes encoding the regulatory complement components C4-binding protein (*C4BP*), the C3b/C4b receptor (*CR1*), the decay-accelerating factor (*DAF*), and factor H (*H*) are linked and define the regulator of complement activation (RCA) gene cluster (1–3), which maps to band q32 of chromosome 1 (4, 5). The same chromosomal location has been reported for the human gene encoding the C3dg receptor (*CR2*) (4), suggesting that *CR2* also belongs to this linkage group. Since the RCA gene cluster encodes the proteins involved in the control of the C3-convertases (reviewed in reference 6), it represents the regulatory counterpart of the class III gene cluster of the MHC that encodes the structural components of the C3-convertases C2, B, and C4 (reviewed in reference 7). In spite of their nonsyntenic chromosomal location, the functionally related MHC–class III and RCA gene clusters may share a common evolutionary history. They encode proteins that bind to C3b and/or C4b and share a particular structural organization of repeats of 60 amino acids characterized by a framework of highly conserved residues (reviewed in reference 8). Early attempts to determine the genetic organization of the RCA cluster have separated the *H* gene from the *CR1*, *C4BP*, and *DAF* genes (9), but the lack of recombinations between the latter has not allowed a more detailed map to emerge from formal genetic analyses. This difficulty has now been overcome through the analysis of very large genomic DNA fragments using pulsed field gel electrophoresis (PFGE) and we have aligned the *CR1*, *CR2*, *DAF*, and *C4BP* genes on an 800-kb DNA segment.

## Materials and Methods

*PFGE Analysis.* Human genomic DNA was prepared from an EBV-transformed B cell line as described in reference 10. Briefly, cells were collected by pelleting, washed twice with PBS, and resuspended at $30 \times 10^6$ cells/ml in PBS. An equal volume of melted low-gelling temperature agarose (SeaPlaque; FMC Corp., Rockland, ME), 1.6% in PBS at 50°C, was mixed with the cell suspension and immediately poured into ice-cooled 20 × 6 × 10-mm molds. When set, the agarose blocks were cut into 1-mm-thick slices that were

This work was supported in part by grant DK-19631-11 from the National Institutes of Health, and by a grant from the Juvenile Diabetes Foundation.

incubated in 0.5 M EDTA, pH 8.0, 1% SDS, and 2 mg/ml Proteinase K at 55°C for 2 d. The agarose slices were washed twice in TE (10 mM tris-HCl, pH 8.0, 1 mM EDTA) at room temperature for 30 min, twice in TE with 0.05 mg/ml PMSF at 55°C for 30 min, and then two more times with TE at room temperature. Plugs of 6 × 3 × 1 mm (containing ~1.5 μg DNA) were cut from the slices and subjected to restriction enzyme digestions. These were done in 200-μl vol for 4–8 h with 12–30 U of enzyme per digestion. In the double digestion experiments, when enzymes were not buffer compatible, the agarose plugs were washed twice with 1 ml of $H_2O$ before the second enzyme digestion. PFGE was done in 13 × 13-cm 1.5% agarose gels in 0.25× TBE (1× TBE: 89 mM Tris borate, 89 mM boric acid, 2 mM EDTA) using a 120° angle alternating homogeneous electrical field apparatus constructed according to Chu et al. (11). The gels were usually run at 185 V (voltage gradient of ~6 V/cm) for 42 h with pulse lengths of 95 s for the first 21 h and 35 s during the last 21 h. After running, the gels were stained with 0.5 μg/ml ethidium bromide and then UV irradiated for 6 min before alkaline transfer to Biotrace-RP nylon membranes (Gelman Sciences, Inc., Ann Arbor, MI). The membranes were prehybridized in 5× SSC, 1% SDS, 2% skim milk, 200 μg/ml denatured salmon testes DNA at 68°C for 3–8 h and then hybridized for 24–40 h under the same conditions with probes $^{32}P$ labeled by the random oligopriming method (12). After high stringency washes (0.2× SSC, 1% SDS at 68°C) the membranes were exposed for 1–3 d with Kodak RX-Omat films at −70°C. Before reprobing, the membranes were stripped by washing in 0.4 M NaOH at 45°C for 30 min and then in 0.2 M Tris-HCl, pH7.5, 0.1× SSC, 1% SDS at 45°C for 30 min. Both *Saccharomyces cerevisiae* (strain 344-12A) chromosomes and λ-cIts857 DNA concatemers were used as molecular weight markers. Yeast chromosomes were obtained as described (11) and λ DNA concatemers were prepared by autoannealing of the cos ends in solution (unit size, 48.5 kb). The CR1 (CR1-1), DAF, and H (NZ-2B12) cDNA probes used here have been previously described (13–15) and were generous gifts of Dr. D. T. Fearon (Department of Medicine, Harvard Medical School, Boston, MA), Dr. I. Caras (Genentech, Inc., South San Francisco, CA), and Dr. B. F. Tack (Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, CA), respectively. The C4BP probe is an 810-bp cDNA clone corresponding to the 3′ end of the human C4BP mRNA. This clone was isolated in our laboratory by antibody screening of a human liver cDNA λgt-11 expression library. The identity of this clone as human C4BP was confirmed by comparison of its nucleotide sequence with that reported by Chung et al. (16). The CR2-cDNA clone CR2-1.6 was obtained from the American Type Culture Collection, (Rockville, MD), and is described in reference 17.

## Results and Discussion

To determine the physical map of the chromosomal region including the *CR1*, *CR2*, *DAF*, *C4BP*, and *H* genes, we have generated large fragments of human genomic DNA with infrequently cutting restriction enzymes, separated them by PFGE, and sequentially probed Southern blots of these gels with locus-specific cDNA probes. The identity of fragments apparently hybridizing with two or more probes has been confirmed with double-digestion experiments that have also allowed the elaboration of the long range restriction map of the region.

Fig. 1 shows the hybridization of the different probes with Sal I, Sac II, Mlu I, Not I, and Sfi I digests. CR1, CR2, DAF, and C4BP probes hybridize with a 640-kb Sal I restriction fragment, while the C4BP probe also recognizes two additional Sal I fragments of 75 kb and >2,500 kb, respectively. This, therefore, locates the *C4BP* gene at one end of this group of genes.

A 475-kb Sfi I fragment hybridizes with both the C4BP and DAF but not with the CR1 and CR2 probes locating the *DAF* gene between *C4BP* on one side and *CR1* and *CR2* on the other. These relationships are confirmed by an 800-kb Mlu

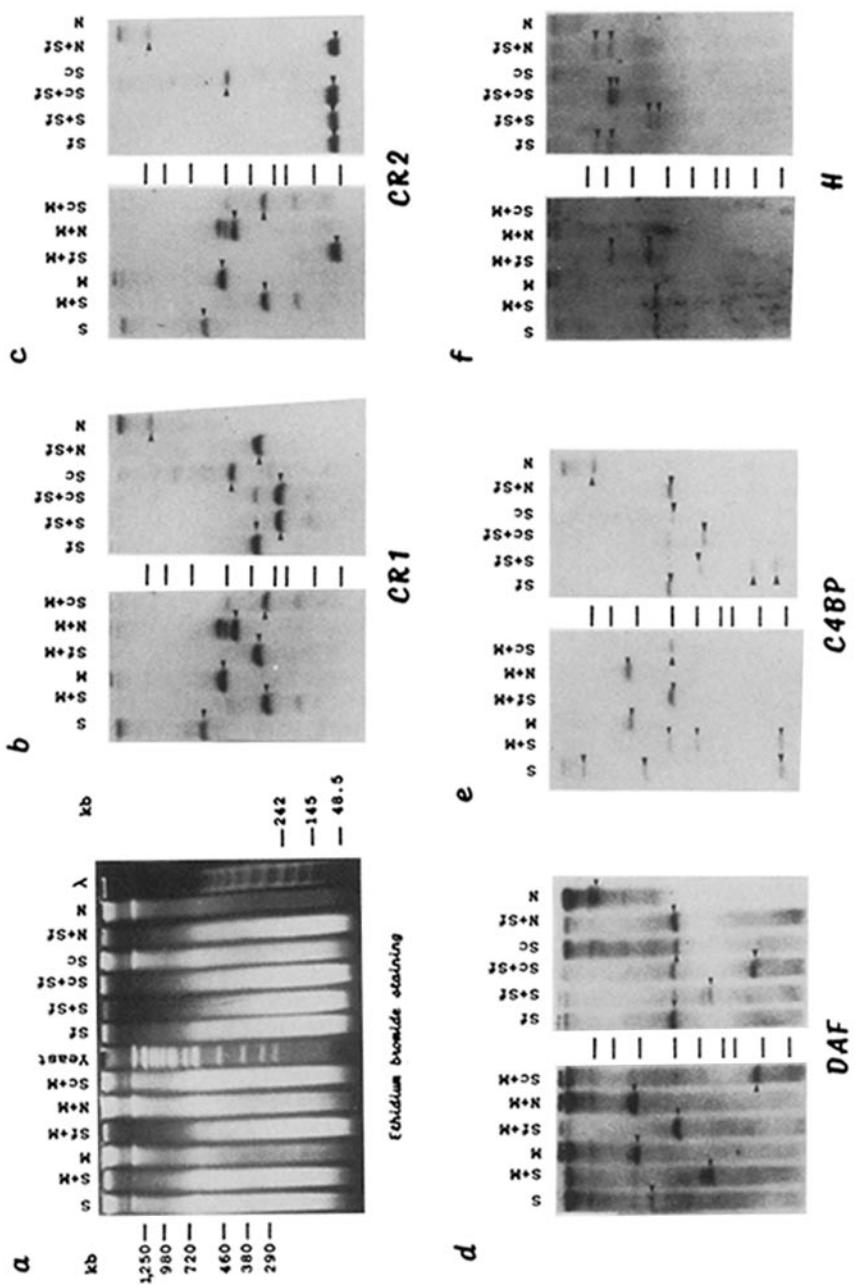


FIGURE 1. Pulsed field gel analysis of human genomic DNA using CR1-, CR2-, DAF-, C4BP-, and H-specific cDNA probes. The figure shows the ethidium bromide staining of a gel (*a*) and the subsequent hybridization of its Southern blot with $^{32}$P-labeled cDNA probes specific for *CR1* (*b*), *CR2* (*c*), *DAF* (*d*), *C4BP* (*e*), and *H* (*f*). The code for the restriction enzymes used is as follows: *S*, Sal I; *Sc*, Sac I; *Sf*, Sfi I; *M*, Mlu I; *N*, Not I. λ phage concatemers and yeast (*S. cerevisiae*) chromosomes were used as molecular markers. Their positions are indicated on both sides of (*a*) and in the center of all the others. Arrows indicate the relevant bands.

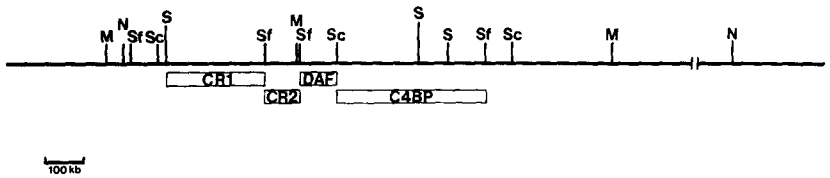


FIGURE 2. Restriction map of the human genomic region containing the *CR1*, *CR2*, *DAF*, and *C4BP* genes. The code for the restriction enzymes is included in the legend to Fig. 1.

I fragment that hybridizes to both the C4BP and DAF but to neither the CR1 or CR2 probes.

The CR1 and CR2 probes, and none of the others, identify a 480-kb Mlu I fragment that upon digestion with Sal I, results in a 320-kb fragment also common to both of them. The CR1 and CR2 probes hybridize with different Sfi I restriction fragments of 340 and 85 kb, respectively. Further digestion of these Sfi I fragments with Sal I decreases the size of the fragment hybridizing with the CR1 probe to 255 kb, but not that of the fragment that hybridizes with the CR2 probe. Since the CR1, CR2, and DAF probes hybridize with different Sfi I fragments, but with a common Sal I restriction fragment, the decrease in the size of the fragments hybridizing with the CR1 and the DAF probes, respectively, produced by the double-digestion Sfi I/Sal I, indicates that the *CR1* and *DAF* genes flank *CR2*.

Taken together, these data are only compatible with the order *CR1-CR2-DAF-C4BP* for these genes within the RCA gene cluster. The corresponding restriction map for this region of the human genome is illustrated in Fig. 2. It is worth noting that the *CR2* gene, which was previously assigned to the same chromosomal band (1q32) by in situ hybridization experiments (4), maps, in fact, in the close vicinity of the *CR1* gene, supporting the suggestion that these two genes originated from a common ancestor (17).

The H probe did not hybridize to any of the fragments recognized by the CR1, CR2, DAF, or C4BP probes (Fig. 1) and, therefore, a physical link between the gene encoding H and those for CR1, CR2, DAF, and C4BP is still lacking. This does not allow us to determine the orientation of the *CR1-CR2-DAF-C4BP* gene cluster in relation to *H* or to estimate the length of the region between them and, hence, the size of the RCA gene cluster. Based on the data presented here, however, RCA exceeds 1.0 megabases (Mb) of DNA in length and given the recombination data (9) it might be as long as 7 Mb, assuming 1 cM to correspond roughly to 1 Mb.

Preliminary evidence has been obtained of the existence of an RCA cluster in the mouse (18), which suggests that, as in the case of the MHC–class III genes, those in the RCA cluster may persist as a linked group in different species. Acquisition of mechanisms (C3-convertases) to amplify the C3b deposition on target surfaces and of parallel mechanisms to efficiently control these C3-convertases and to prevent host tissue damage may have favored the generation of the MHC–class III region and the RCA gene cluster, respectively. The nature of the selective advantage conferred by the maintenance of the linked configuration, if any, is at the moment unclear.

Interestingly, both systems, the MHC–class III region and the RCA gene

cluster, show extreme linkage disequilibrium between alleles of the loci involved. Although genetic variation of RCA genes has been much less explored than that of the MHC–class III genes, we have previously reported that the *C4BP*2* allele (gene frequency 0.01) is found only in haplotypes also carrying *CR1*B* (gene frequency 0.137; reference 9). Since the region between these genes, including *CR2* and *DAF*, measures several hundred kilobases, this complete linkage disequilibrium probably portends absence of crossing over. In the case of the MHC–class III gene cluster, linkage disequilibrium might be a predictable consequence of its small size and of the short physical distances between its genes. It might be significant, however, that both gene clusters show differences among haplotypes in either the number of genes (i.e., deletions or duplications of the C4A, C4B regions [19]) or the size of the genes (i.e., *CR1* [20]). If reduced frequencies of recombination were consequences of these differences in the genetic organization of the different haplotypes, one of the possible mechanisms could be an interference with the pairing of the homologous chromatids in these regions.

## Summary

We report the organization of the human genes encoding the complement components C4-binding protein (C4BP), C3b/C4b receptor (CR1), decay accelerating factor (DAF), and C3dg receptor (CR2) within the regulator of complement activation (RCA) gene cluster. Using pulsed field gel electrophoresis analysis these genes have been physically linked and aligned as *CR1-CR2-DAF-C4BP* in an 800-kb DNA segment. The very tight linkage between the *CR1* and the *C4BP* loci, contrasted with the relative long DNA distance between these genes, suggests the existence of mechanisms interfering with recombination within the RCA gene cluster.

We thank Dr. I. Caras, Dr. B. F. Tack, and Dr. D. T. Fearon for providing us with the DAF, H, and CR1 human cDNA probes, respectively; Ms. C. Carrier and Mr. N. Mollen for their help with the cell lines; Ms. T. Huima for the photographic work; and Ms. A. Varipapa for her help with the manuscript.

*Received for publication 14 October 1987.*

## References

1. Rodriguez de Cordoba, S., T. R. Dykman, F. Ginsberg-Fellner, G. Ercilla, M. Aqua, J. P. Atkinson, and P. Rubinstein. 1984. Evidence for linkage between the loci coding for the binding protein for the fourth component of the human complement (C4BP) and for the C3b/C4b receptor. *Proc. Natl. Acad. Sci. USA.* 81:7890.
2. Rodriguez de Cordoba, S., D. M. Lublin, P. Rubinstein, and J. P. Atkinson. 1985. Human genes for three complement components that regulate the activation of C3 are tightly linked. *J. Exp. Med.* 161:1189.
3. Rey-Campos, J., P. Rubinstein, and S. Rodriguez de Cordoba. 1987. Decay accelerating factor: genetic polymorphism and linkage to the RCA (regulator of complement activation) gene cluster in humans. *J. Exp. Med.* 166:246.
4. Weis, J. H., C. C. Morton, G. A. P. Bruns, J. J. Weis, L. B. Klickstein, W. W. Wong, and D. T. Fearon. 1987. A complement receptor locus: genes encoding C3b/C4b receptor and C3d/Epstein-Barr virus receptor map to 1q32. *J. Immunol.* 138:312.
5. Lublin, D. M., R. S. Lemons, M. M. Le Beau, V. M. Holers, M. L. Tykocinski, M. E.

Medof, and J. P. Atkinson. 1987. The gene encoding decay accelerating factor (DAF) is located in the complement-regulatory locus on the long arm of chromosome 1. *J. Exp. Med.* 165:1731.

6. Holers, V. M., J. L. Cole, D. M. Lublin, T. Seya, and J. P. Atkinson. 1985. Human C3b-regulatory and C4b-regulatory proteins. A new multi-gene family. *Immunol. Today.* 6:188.
7. Porter, R. R. 1985. The complement components coded in the major histocompatibility complexes and their biological activities. *Immunol. Rev.* 87:7.
8. Kristensen, T., P. D'Eustachio, R. T. Ogata, L. P. Chung, K. B. M. Reid, and B. F. Tack. 1987. The superfamily of C3b/C4b-binding proteins. *Fed. Proc.* 46:2463.
9. Rodriguez de Cordoba, S., and P. Rubinstein. 1987. New alleles of C4-binding protein and Factor H and further linkage data in the regulator of complement activation (RCA) gene cluster in humans. *Immunogenetics.* 25:267.
10. Kenwrick, S., M. Patterson, A. Speer, K. Fischbeck, and K. Davies. 1987. Molecular analysis of the Duchenne Muscular Dystrophy Region using Pulsed field gel electrophoresis. *Cell* 48:351.
11. Chu, G., D. Vollrath, and R. W. Davis. 1986. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science (Wash. DC).* 234:1582.
12. Feinberg, A. I., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6.
13. 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.
14. Caras, I. W., M. A. Davitz, L. Rhee, G. Weddell, D. W. Martin, Jr., and V. Nussenzweig. 1987. cDNA cloning of decay accelerating factor indicates novel use of splicing to generate two protein forms. *Nature (Lond.).* 325:545.
15. Kristensen, T., R. A. Wetsel, and B. F. Tack. 1986. Structural analysis of human complement protein H: homology with C4b-binding protein, β2-glycoprotein I and the Ba fragment of B. *J. Immunol.* 136:3407.
16. Chung, L. P., D. R. Bentley, and K. B. M. Reid. 1985. Molecular cloning and characterization of the cDNA coding for C4b-binding protein, a regulatory protein of the classical pathway of the human complement system. *Biochem. J.* 230:133.
17. Weiss, J. J., D. T. Fearon, L. B. Klickstein, W. W. Wong, S. A. Richards, A. D. Bruyn-Kops, J. A. Smith, and J. H. Weis. 1986. Identification of a partial cDNA clone for the C3d/Epstein-Barr virus receptor of human β lymphocytes: homology with the receptor for fragments C3b and C4b of the third and fourth components of complement. *Proc. Natl. Acad. Sci. USA.* 83:5639.
18. Kristensen, T., R. T. Ogata, L. P. Chung, K. B. M. Reid, and B. F. Tack. 1987. cDNA structure of murine C4b-binding protein, a regulatory component of the serum complement system. *Biochemistry.* 26:4668.
19. Carroll, M. C., K. T. Belt, A. Palsdottir, and Y. Yu. 1985. Molecular genetics of the fourth component of human complement and steroid 21-hydroxylase. *Immunol. Rev.* 83:39.
20. Wong, W. W., C. A. Kennedy, E. T. Bonaccio, J. G. Wilson, L. B. Klickstein, J. H. Weis, and D. T. Fearon. 1986. Analysis of multiple restriction fragment length polymorphism of the gene for the human complement receptor type I. Duplication of genomic sequences occurs in association with a high molecular weight receptor allotype. *J. Exp. Med.* 164:1531.