Molecular Analysis of the Murine C4b-Binding Protein Gene. Chromosome Assignment and Partial Gene Organization†

Scott R. Barnum,‡,§ Torsten Kristensen,∥ David D. Chaplin,⊥ Michael F. Seldin,# and Brian F. Tack*,‡

Department of Immunology, Research Institute of Scripps Clinic, La Jolla, California 92037, Department of Molecular Biology and Plant Physiology, University of Aarhus, Denmark, Department of Internal Medicine and Howard Hughes Medical Institute, Washington University School of Medicine, St. Louis, Missouri 63110, and Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710

Received April 7, 1989; Revised Manuscript Received June 20, 1989

ABSTRACT: Murine C4b-binding protein (C4BP) is a regulatory molecule in the classical pathway of complement. C4BP is composed predominantly of short consensus repeats (SCRs) approximately 60 amino acids in length, which contain a framework of conserved residues. The SCRs are found in many complement molecules and a growing number of noncomplement molecules as well and are a major structural feature of some of these molecules. To characterize the structure of the murine C4BP gene, a cosmid library constructed from Balb/c liver DNA was screened. Several nearly identical, overlapping clones were identified; however, none of the clones, alone or in combination, covered the entire C4BP gene. One clone (D26) was chosen for detailed analysis and found to contain all but the leader region, the first SCR, and the first half of the second SCR. The SCRs three through six were each encoded by single exons. Only the latter half of the second SCR was present on the clone, and it was encoded by a single exon, demonstrating that murine C4BP has a split SCR at the genomic level. Structural mapping of this portion of the gene demonstrates that the region extending from the second half of the second SCR through the the nonrepeat and untranslated region spans ~12 kb; however, genomic Southern blot analysis suggests that the gene is between 20 and 30 kb in length. Analysis of the 3′ genomic sequence demonstrates that this region of the gene has homology with SV-40 late (class II) RNA sequences. These sequences may play a role in 3′ cleavage of the precursor RNA prior to polyadenylation. The C4BP gene was localized to chromosome 1, ~2.1 centiMorgans centromeric of the renin genes by typing interspecific back-cross mice [(C3H/HeJ-*gld*/*gld* × *Mus spretus*)F1 × C3H/HeJ-*gld*/*gld*] and identifying informative C4BP restriction length polymorphisms, thus allowing linkage analysis with other genes previously localized to mouse chromosome 1.

Mouse C4b-binding protein (C4BP) is a serum glycoprotein with a molecular weight of ~1-1.5 million by gel filtration and is composed of seven identical noncovalently associated polypeptide chains of M_r 60 000-80 000 (Ferreira et al., 1977, 1978; Kaidoh et al., 1981; Rodriguez de Cordoba et al., 1985). Each polypeptide chain is composed of 6 short consensus repeats (SCRs) in tandem followed by a nonrepeat region of 54 amino acids (Kristensen et al., 1987a). The SCRs found in C4BP are approximately 60 amino acids in length and have the typical consensus framework of conserved residues including 4 half-cysteines, 2 prolines, 1 tryptophan residue, and several other partially conserved glycine and hydrophobic residues. The presence of the SCRs is widespread in complement proteins and a growing number of noncomplement molecules as well [Lublin et al., 1988; Kotwal et al., 1988; for reviews, see Reid et al. (1986) and Kristensen et al. (1987b)]. The recently derived amino acid sequence demonstrates that each polypeptide chain of murine C4BP is 413 residues in length with a molecular weight of 45 281 (Kristensen et al., 1987a). The cDNA sequence reveals 2 in-frame translational start signals which would yield signal peptides of 13 or 56 amino acids. In the 3′ untranslated region, there are a total of 7 potential polyadenylation signal sequences with 2 overlapping signals 19 bases upstream from the poly(A) tail. Human C4BP differs from mouse in that it has two additional SCRs and two additional cysteine residues in the nonrepeat region. These additional cysteine residues are thought to be involved in the disulfide linkage between the polypeptide chains in human C4BP (Chung et al., 1985).

C4BP functions as an essential cofactor in regulating the activation of the classical pathway of complement. C4BP mediates its regulatory role by accelerating the decay of the C3 convertase (C4b2a) through binding to C4b, displacing C2a and acting as a cofactor in the the complement protein I mediated cleavage of C4b to C4d and C4c (Gigli et al., 1979; Burge et al., 1980; Nagasawa et al., 1980; Fujita et al., 1978; Fujita & Nussenzweig, 1979). In addition, C4BP can bind to newly generated C4b in the fluid phase, preventing formation of the C3 convertase. The binding specificity of mouse C4BP is limited to mouse and human C4 and C4b (Ferreira et al., 1977, 1978; Kaidoh et al., 1981; Kai et al., 1980); however, human C4BP has a low affinity for C3b and can serve as a cofactor in the I-mediated cleavage of C3b to C3bi (Ziccardi et al., 1984).

In humans, C4BP is part of a large linkage group on chromosome 1 that includes CR1, CR2, decay-accelerating

† Publication No. 5828-IMM from the Research Institute of Scripps Clinic. This work was supported by USPHS Grant AI17354 to B.F.T. and Individual National Research Service Award AI07648 to S.R.B. D.D.C. is supported by the Howard Hughes Medical Institute. M.F.S. is supported by awards from the Charles E. Culpeper and Arthritis Foundations.

* Correspondence should be addressed to this author.

‡ Research Institute of Scripps Clinic.

§ Present address: Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294.

∥ University of Aarhus.

⊥ Washington University School of Medicine.

Duke University Medical Center.

0006-2960/89/0428-8312$01.50/0

factor (DAF), membrane cofactor protein (MCP), and factor H (Carroll et al., 1988; Rey-Campos et al., 1988; Lubin et al., 1988). With the exception of factor H, these genes are all contained within 900 kb of DNA (Carroll et al., 1988; Rey-Campos et al., 1988; Bora et al., 1989).

In this paper, we report the isolation and characterization of a murine C4BP cosmid clone which encompasses most of the coding sequence on the structural gene. From this clone, we were able to determine that the majority of the SCRs in murine C4BP were, with one exception, encoded by single exons. The latter half of the second SCR is encoded by a single exon demonstrating that C4BP, like many other SCR-containing molecules, has a split SCR at the genomic level. In addition, analysis of interspecies back-cross mice allowed the unambiguous assignment of the C4BP structural gene to chromosome 1.

Experimental Procedures

Construction and Screening of the Cosmid Library. A cosmid library was constructed in the pTCF vector (Grosveld et al., 1982) as previously described (Chaplin et al., 1983) using genomic DNA from a Balb/c liver. Filters were prehybridized and hybridized at 42 °C in 50% formamide buffer (50% formamide, 5× SSC, 1× Denhardt's solution, 5 mM Tris-HCl, pH 7.5, 0.5% SDS, and 50 μg/mL salmon sperm DNA). A full-length C4BP cDNA insert (pMBP.15; Kristensen et al., 1987a) was used as a probe after labeling with [α-^{32}P]dCTP by nick translation using a kit from Boehringer Mannheim. Positive clones were colony-purified, and cosmid DNA was isolated (Maniatis et al., 1982).

Southern Hybridizations. One to two micrograms of cosmid DNA was digested to completion with the appropriate restriction endonuclease (Boehringer Mannhein, Indianapolis, IN) and electrophoresed in a 1.0% agarose gel. The DNA was transferred to nylon filter paper; the filters were prehybridized (3–6 h) and then hybridized in 50% formamide buffer as described above at 42 °C. The probes used were ^{32}P-labeled M13 subclones derived from sequencing the intron/exon junctions of the gene. The phage subclones were labeled by using standard sequencing reaction conditions (Bankier & Barrell, 1983) in the presence of [α-^{32}P]dCTP and in the absence of dideoxynucleotides. After overnight hybridization, the filters were stringently washed (0.2X SSC/0.1% SDS at 65 °C for 30 min) and exposed to Fuji film with an intensifying screen for the appropriate time at −70 °C. For the chromosome assignment, DNA was isolated from mouse organs by standard techniques and digested with restriction endonuclease enzymes, and 10 μg of each sample was electrophoresed in 0.9% agarose gels. DNA was then transferred to Nytran membranes (Schleicher & Schuell, Keene, NH), hybridized at 65 °C, and washed under stringent conditions, all as previously described (Huppi et al., 1985).

Sequencing of Intron/Exon Junctions. Cosmid DNA was sequenced according to the method of Bankier and Barrell (1983). Fragments of 300–600 bp in length, randomly generated by sonication, were ligated into the *Sma*I site of M13mp8 after end-repairing with T4 DNA polymerase and removal of any remaining overhanging ends with mung bean nuclease. Exon-containing phage were identified by transferring a portion of the phage to MSI nylon filters, in duplicate, and probing the filters with the appropriate ^{32}P-labeled C4BP cDNA fragments. The fragments were labeled with [α-^{32}P]dCTP by the random prime labeling method of Feinberg and Vogelstein (1983) using a kit from Boehringer Mannheim. Filters were prehybridized and hybridized at 42 °C in 50% formamide buffer. Phage that screened positive were then sequenced by the dideoxy chain termination method (Sanger et al., 1977) using [α-^{35}S]thio-labeled deoxyadenosine 5′-thiotriphosphate (Biggin et al., 1983). Sequence results were aligned and compiled by using the DB programs of Staden (1980, 1982). Sequence comparison analyses were performed by using the BESTFIT program of the University of Wisconsin Genetics Computer Group. Intron/exon junction sequences were identified manually by comparison to the murine C4BP cDNA sequence and the consensus sequence of acceptor and donor splice sites (Mount, 1982).

Molecular Probes for Linkage Analysis. For the chromosome assignment, all probes were labeled as described above with [α-^{32}P]dCTP (3000 Ci/mmol; Amersham, Arlington Heights, IL) using a Pharmacia (Piscataway, NJ) oligolabeling kit and protocol. *C4bp* polymorphisms were identified by using a 1.8-kb *Bam*HI fragment from the mouse cDNA clone pMBP.15 (Kristensen et al., 1987a). *Renin* (*Ren-1,2*) was detected with a 1.4-kb *Pst*I insert from the mouse cDNA clone Id-2 (Field et al., 1984). *Ly-5* polymorphisms were detected by using a 2.4-kb *Bam*HI fragment isolated from the cDNA clone pLy5-68 (Shen et al., 1985). *Ctla-4* polymorphisms were detected by using a 1.9-kb *Eco*RI, *Hin*dIII insert from the mouse cDNA clone F41F4 (Bruner et al., 1987).

The distance (in centiMorgans) calculated between each locus equals the number of cross-over events divided by the total number of miotic events multiplied by 100. The upper and lower confidence limits (data not shown) were calculated by the binomial distribution method of Silver and Buckler (1986).

Results and Discussion

Isolation and Characterization of DNA Encoding the Mouse C4BP Gene. The murine C4BP cDNA was used as a probe to screen approximately 300 000 recombinants of a Balb/c genomic cosmid library. Five C4BP-specific clones were identified, all of which contained the latter half of the second SCR and extended through the 3′ end of the gene as determined by Southern blotting and probing with the appropriate 5′ and 3′ M13 templates derived from sequencing the C4BP cDNA. Restriction analysis demonstrated similar and unique bands for all of the clones, indicating that they represent overlapping but unique portions of the C4BP gene. Comparison of Southern blots of Balb/c genomic DNA and the cosmid clones also demonstrates the presence of identical bands and, in addition, indicates that approximately an additional 5–10 kb of genomic DNA is needed to complete the gene structure. Numerous attempts to identify a clone containing the 5′ end of the gene have been unsuccessful. This may suggest the possibility of multiple or hypersensitive *Mbo*I sites in this portion of the gene. One clone, D26, was selected for analysis of intron/exon junctions and for mapping the distances between the exons. The D26 clone contains six exons of the mouse C4BP gene, extending from the second half of the second repeat through the 3′ nonrepeat and nontranslated region of the molecule (Table I). The first exon contains 86 bp and codes for the latter half of the second repeat, while exons two through five each encode repeats three through six, respectively (Table I). The last exon encodes the 3′ nonrepeat and untranslated sequence. The average exon size is 205 bp and thus would code for 68.3 amino acids. The exons encoding repeat units (i.e., internal exons) are on average 158 bp in length and would code for 52.6 amino acids, which is larger than the reported average of 44.6 amino acids for internal exons (Traut, 1988). Also shown in Table I are the splice donor and splice acceptor sequences demarcating the 5′ and 3′ ends of the five introns found in this portion of the C4BP

Table I: Intron/Exon Junction Sequences and Size of the Exons and Introns of the Murine C4BP Gene

exon	repeat no.	length (bp)	intron/exon[a] junction sequence	intron	length (kb)
1	2b	86	tctcagA TTT...GTA Agtaagt	1	2.7
2	3	191	tatcagTT GTA...GAA Agtaagt	2	4.3
3	4	179	ccatagAG ATC...TCC Ggtaagt	3	2.1
4	5	164	atcaagAC TGT...AAA Ggtaatg	4	1.8
5	6	175	ctgtagCT CTA...CAGgtaagt	5	0.5
6	NR and UT[b]	442[c]	ttcaagGAG		

[a] Exon base pairs are capitalized and grouped in codons. [b] NR and UT, nonrepeat and untranslated regions. [c] Ending at the poly(A) addition site.

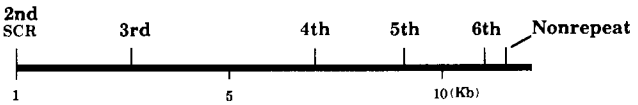

FIGURE 1: Partial physical map of the C4BP gene structure. A partial structural map of the murine C4BP gene is shown. The numbering underneath the line refers to the distance in kilobases (kb) while that above the line refers to the position of the exons as determined by Southern blot analysis and restriction mapping.

gene. The consensus of the donor sequences is gtaagt with a frequency of 100%, 100%, 100%, 100%, 80%, and 80%, respectively. This consensus is in perfect agreement with that reported in the literature (Mount, 1982; Green, 1986; Padgett et al., 1986). The 3′ intron consensus sequence was ttt/ccag with a frequency of 50%, 50%, 33%, 50%, 100%, and 100%, respectively (Table I).

A partial structural map of the C4BP gene has been completed, and a map of the region is shown in Figure 1. The intron sizes, estimated by restriction mapping and Southern blot analysis, range in size from 0.5 to 4.3 kb with an average of 2.28 kb. The introns are predominantly type 1 introns (80%) while the remaining introns are type 0 (10%) and type 2 (10%). Type 0 introns interrupt between codons while type 1 and 2 introns interrupt after the first and second nucleotides of a codon, respectively. This distribution deviates from the recently reported average composition of 54%, 27%, and 18% for type 0, 1, and 2 introns, respectively (Traut, 1988). The region from the second half of the second repeat (exon 1) to the untranslated region (exon 6) is ~12.6 kb. Genomic Southern blot analysis of DNA isolated from Balb/c kidney suggests that the minimum gene size ranges from 11.5 to 24 kb (data not shown). This would suggest that the murine C4BP gene is smaller than the human gene which has been estimated to be 30 kb (Lintin & Reid, 1986).

Variation in Exons Encoding SCRs. The first exon identified in the D26 clone codes for only half of the second SCR found in the murine C4BP gene. Initially, this was a unique finding with only a few SCR-containing genes displaying this split (Barnum et al., 1987; Vik et al., 1988; Maeda, 1984). However, as more SCR-containing genes have been examined, the split appears to be a more common structural feature. For example, human CR1 (Wong et al., 1989), CR2 (Fujisaku et al., 1989), MCP (Post & Atkinson, 1989), DAF (Lubin et al., 1989), and a CR1-like gene (Hourcade et al., 1989) all contain split SCRs. Interestingly, the location of the split is near the middle of the SCR in all molecules whose gene structure is available for detailed analysis. Comparison of three genes containing split SCRs (Figure 2) indicates that the split occurs in all cases after the third conserved glycine residue (immediately 3′ to the second conserved cysteine residue). Whether the split SCRs represent primordial subunits or are the result of a more recent transpositional event remains unclear. The presence of exons encoding two SCRs (Hourcade et al., 1989; Fujisaku et al., 1989) and the fact that the majority of SCRs are encoded by single exons perhaps suggest that split SCRs are not primordial subunits.

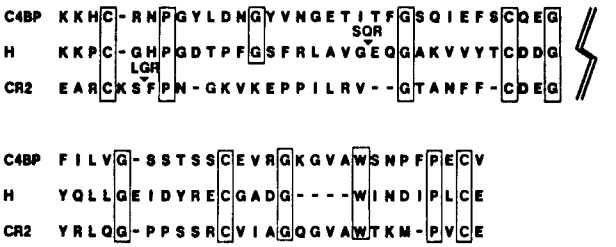

FIGURE 2: Alignment of the amino acid sequence of proteins containing SCRs encoded by two exons. The amino acid sequences of the second SCR of murine C4BP and factor H and the fourth SCR of human CR2 are aligned to show the identical position of the split in the SCRs of all three molecules. The double line represents the position of the split at the genomic level. In each case, the split occurs after the conserved glycine residue on the carboxy side of the second conserved cysteine residue. Conserved framework residues in the SCRs are boxed. Gaps or insertions have been introduced to maximize alignment of the sequence.

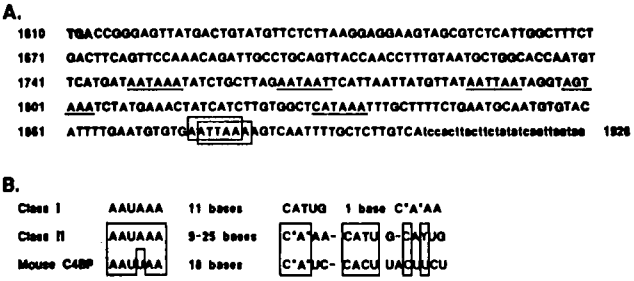

FIGURE 3: Genomic sequence of the 3′ end of the murine C4BP gene. (A) The sequence of the 3′ nontranslated region of murine C4BP along with the genomic sequence derived from the D26 clone is shown. The sequence shown in capital letters represents the previously determined cDNA sequence while that in lower case letters represents the genomic sequence. The boldface letters represent the stop codon. The underlined sequence and the boxed sequence represent the seven putative polyadenylation signal sequences. (B) The consensus sequences of SV-40 early (class I) or late (class II) RNA sequences are compared to the deduced precursor RNA sequence of mouse C4BP. The homology between the SV-40 class II sequences and mouse C4BP sequences is boxed. The asterisks indicate possible sites of cleavage and polyadenylation.

Murine C4BP 3′ Genomic Structure. Limited sequence was obtained from the 3′ end of the murine C4BP gene during the determination of the intron/exon structure of the gene. Figure 3A shows the previously determined untranslated 3′ sequence (Kristensen et al., 1987a) containing seven putative polyadenylation sites followed by an additional 3′ genomic sequence obtained from the D26 clone. The additional 3′ sequence contains no poly(T) or G/T clusters often seen in the 3′ end of genes [for a review, see Birnsteil et al. (1985)]. However, there is fairly strong sequence homology in this region of the gene with the sequence recognition element CAYUG described by Benoist and colleagues (Benoist et al., 1980). This element is found in both early (class I) and late (class II) SV-40 RNAs. The two classes are distinguished by their number and position relative to the polyadenylation site. In class I RNAs, the CAYUG is present once and located one base immediately 5′ to the polyadenylation site (Figure 3B), while the class II

Table II: Analysis of Restriction Fragment Length Polymorphism Typing of [(C3H/HeJ-*gld*/*gld* × *M. spretus*)F1 × C3H/HeJ-*gld*/*gld*] Back-Cross Mice

gene	genotype of [(C3H/HeJ-*gld*/*gld* × *M. spretus*)F1 × C3H/HeJ-*gld*/*gld*] mice[a]							
Ctla-4	CC	SC	SC ×	CC ×	CC	SC	CC	SC
C4BP	CC	SC	CC	SC	CC	SC	CC ×	SC ×
Ren-1,2	CC	SC	CC	SC	CC ×	SC ×	SC	CC
Ly-5	CC	SC	CC	SC	SC	CC	SC	CC
no. of mice	36	27	12	11	4	3	1	1

[a] CC = C3H type; SC = F1 genotype. Each column represents a pair of chromosomes that was observed in the back-cross mice and was determined by restriction fragment length polymorphism analysis at each locus. The number of chromosomes with each cross-over combination is noted at the bottom of each column. Presumed sites of recombination are denoted with an "×". Assuming the loci order given, no chromosomes with two or more recombinations were found.

RNAs have the sequence element as many as 3 times in tandem and are located immediately 3′ to the polyadenylation site (Figure 3B). The murine C4BP gene contains class II CAYUG sequence elements as shown in Figure 3B. Following the polyadenylation signal sequence, there is a gap of 18 bp in the mouse C4BP gene which is in the range of the 9–25 bp gap seen in the class II RNA sequences. Moving downstream, there is 80% homology with the first element and lesser homology (40%) further downstream. No additional CAYUG sequence elements were identified 3′ to the shown sequence (data not shown). The exact function of this sequence element remains unclear; however, it has been suggested that it may play a role in cleavage of precursor RNA by orienting the precursor in the proper conformation for cleavage. This role would theoretically be mediated by complementary base pairing between U4 snRNA sequences and the CAYUG and AAUAAA polyadenylation signal elements (Berget, 1984). More recent studies have indicated that U4 (along with U1, U2, and U6) is required for in vitro splicing but not polyadenylation (Berget & Robberson, 1986; Sharp, 1987), lending support to the suggestion that these sequence elements may function in determining the specificity of precursor RNA cleavage.

Chromosomal Assignment. In order to precisely map *C4bp*, we utilized a series of interspecific back-cross mice. This allowed the easy recognition of informative *C4bp* restriction-length polymorphisms (RFLPs) and linkage analysis with other genes (*Ctla-4*, *Ren-1,2*, and *Ly-5*) localized to mouse chromosome 1. *Ren-1,2* and *Ly-5* have been localized to the middle of mouse chromosome 1 using RFLP analysis in recombinant inbred and back-cross mice (Seldin et al., 1988a,b). *Ctla-4* (which codes for a cytotoxic T lymphocyte associated molecule) has been localized to a more proximal region of mouse chromosome 1 by in situ hybridization (Bruner et al., 1987).

Ninety-five [(C3H/HeJ-*gld*/*gld* × *Mus spretus*)F1 × C3H/HeJ-*gld*/*gld*] back-cross mice were typed for each of these genes shown in Figure 4. The results of analysis of RFLP typing are shown in Table II. At each locus, mice displayed either the homozygous C3H pattern (CC) or the heterozygous F1 pattern (SC). The gene order was established by minimization of chromosome cross-over events and is unambiguous since any other order would necessitate multiple double-cross-over events. The data indicate that *C4bp* is located on mouse chromosome 1, 2.1 centiMorgans centromeric of the renin genes.

The localization of *C4bp* on mouse chromosome 1 is consistent with previous observations indicating a large conserved linkage group common to mouse and human chromosome 1. In the human genome, a linkage group for regulation of complement activity (RCA) including *MCP*, *CR1*, *CR2*, *DAF*,

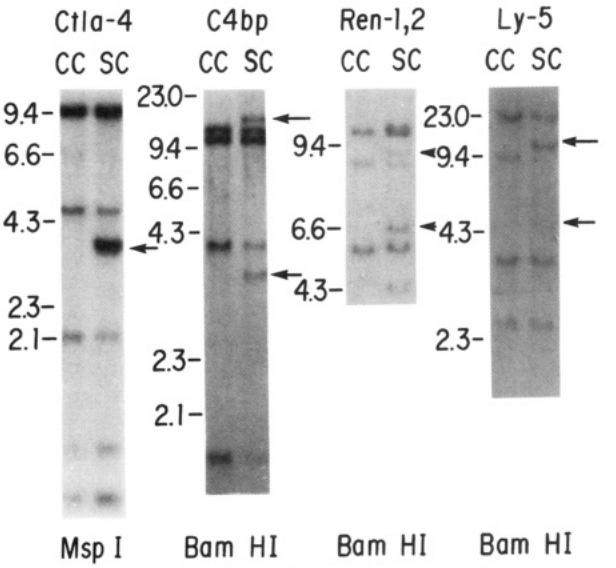

FIGURE 4: Southern blot identification for unique *Mus spretus* polymorphisms with *C4bp*, *Ctla-4*, *Ren-1,2*, and *Ly-5*. The gene probes are indicated at the top of the figure with the restriction endonucleases indicated at the bottom. Arrows signify polymorphisms present in (C3H-*gld*/*gld* × *Mus spretus*)F1 (SC) but not in homozygous C3H-*gld*/*gld* (CC) mice. Molecular size standards are shown at the left of each panel.

and *C4BP* is encoded within a region of approximately 900 kb (Rey-Campos et al., 1988; Carroll et al., 1988; Bora et al., 1989). *CR1* and *CR2* have been localized by using in situ hybridization to human chromosome 1q32 (Weis et al., 1987). Both *REN* and *CD45* (T200), the human homologues of *Ren-1,2* and *Ly-5*, have similarly been localized to this region of human chromosome 1q (McGill et al., 1987; Ralph et al., 1987). Thus, *C4bp* is another member of the mouse chromosome 1/human chromosome 1q conserved linkage group which includes *Ren-1,2*, *Ly-2* (*CD45*), *Lamb-2* (laminin B-2), *At-3* (antithrombin-3), *Apoa-2* (apolipoprotein A-2), *Sapr* (*ATPC*), and *Spna-1* (*SPTA*) (Seldin et al., 1988a,b). This suggests that other members of the human linkage group may be members of a mouse RCA linkage group. However, preliminary results suggest that other members of the RCA linkage group while present on mouse chromosome 1 are not closely linked with *C4bp* (Kingsmore et al., 1989). In addition, while *Cr1*/*Cr2*-related sequences, also known as Mcry and Mcr2, and *Cfh* have been localized to mouse chromosome 1, no murine homologues for DAF or MCP have been identified at the protein or gene level. These findings highlight how differently the murine RCA region is organized from its human counterpart (see Figure 5), despite the conservation of linkage of other genes in the region. These data have led to

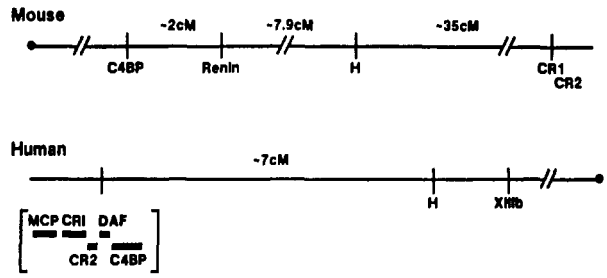

FIGURE 5: Map of the regulator of complement activation (RCA) region of mouse and humans on chromosome 1. In the murine RCA region, the C4BP gene is located approximately 10 centiMorgans (cM) centromeric to the factor H gene and 45 cM centromeric of the murine equivalent of CR1 and CR2. The human RCA region is significantly more clustered than the murine region. The unit cM (centiMorgan) corresponds to approximately 1 million bases.

the suggestion that Mcry and Mcr2 were translocated or that a gene inversion event occurred during murine gene evolution moving these genes outside the linkage group described in humans (see Figure 5) (Kingsmore et al., 1989). Ongoing studies should clarify the linkage and molecular relationships of these genes in the mouse, and studies in other species should allow further analysis of the genomic organization and functional divergence of this linkage group.

ACKNOWLEDGMENTS

We thank Jill Keeney for preparation and use of the cosmid library and Deborah Noack and Lori Graham for excellent technical assistance. We also thank Bonnie Towle for preparation of the manuscript.

REFERENCES

Bankier, A. T., & Barrell, B. G. (1983) in *Techniques in Nucleic Acid Biochemistry* (Flavell, R. A., Ed.) Vol. B5-08, pp 1-34, Elsevier/North-Holland Scientific, Limerick, Ireland.

Barnum, S. R., Keeney, J., Kristensen, T., Noack, D., Seldin, M., D'Eustachio, P., Chaplin, D., & Tack, B. F. (1987) *Complement 4*, 131.

Benoist, C., O'Hare, K., Breathnach, R., & Chambon, P. (1980) *Nucleic Acids Res. 8*, 127-142.

Berget, S. M. (1984) *Nature 309*, 179-182.

Berget, S. M., & Robberson, B. L. (1986) *Cell 46*, 691-696.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963-3965.

Birnstiel, M. L., Busslinger, M., & Strub, K. (1985) *Cell 41*, 349-359.

Bora, N. S., Lublin, D. M., Kumar, B. V., Hockett, R. D., Holers, V. M., & Atkinson, J. P. (1989) *J. Exp. Med. 169*, 597-602.

Bruner, J.-F., Denizot, F., Luciani, M.-F., Roux-Dosseto, M. R., Suzan, M., Mattei, M.-G., & Golstein, P. (1987) *Nature 328*, 267-270.

Burge, J., Nicholson-Weller, A., & Austen, K. F. (1981) *J. Immunol. 126*, 232-235.

Carroll, M. C., Alicot, E. M., Katzman, P. J., Klickstein, L. B., Smith, J. A., & Fearon, D. T. (1988) *J. Exp. Med. 167*, 1271-1280.

Chaplin, D. D., Woods, D. E., Whitehead, A. S., Goldberger, G., Colten, H. R., & Seidman, J. G. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6947-6951.

Chung, L. P., Bentley, D. R., & Reid, K. B. M. (1985) *Biochem. J. 230*, 133-141.

Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6-13.

Ferreira, A., Takahashi, M., & Nussenzweig, V. (1977) *J. Exp. Med. 146*, 1001-1018.

Ferreira, A., Weisz-Carrington, P., Gigli, I., & Nussenzweig, V. (1978) *J. Immunol. 120*, 1772.

Field, L. J., McGowan, R. A., Dickinson, D. P., & Gross, K. W. (1984) *Hypertension 6*, 597-603.

Fujisaku, A., Harley, J. B., Frank, M. B., Gruner, B. A., Frazier, B., & Holers, V. M. (1989) *J. Biol. Chem. 264*, 2118-2125.

Fujita, T., & Nussenzweig, V. (1979) *J. Exp. Med. 150*, 267-276.

Fujita, T., Gigli, I., & Nussenzweig, V. (1978) *J. Exp. Med. 148*, 1044-1051.

Gigli, I., Fujita, T., & Nussenzweig, V. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6596-6660.

Green, M. R. (1986) *Annu. Rev. Genet. 20*, 671-708.

Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Pahl, H. M., & Flavell, R. A. (1982) *Nucleic Acids Res. 10*, 6715-6732.

Hourcade, D., Miesner, D. R., Bee, C., Zeldes, W., & Atkinson, J. P. (1989) *FASEB J. 3*, 789.

Huppi, K., Jouvin-Marche, E., Scott, C., Potter, M., & Weigert, M. (1985) *Immunogenetics 21*, 445-457.

Kai, S., Fujita, T., Gigli, I., & Nussenzweig, V. (1980) *J. Immunol. 125*, 2409-2415.

Kaidoh, T., Natsuume-Sakai, S., & Takahashi, M. (1981) *J. Immunol. 126*, 463-467.

Kingsmore, F. F., Vik, D. P., Kurtz, C. B., Leroy, P., Tack, B. F., Weis, J. H., & Seldin, M. F. (1989) *J. Exp. Med. 169*, 1479-1484.

Kotwal, G. J., & Moss, B. (1988) *Nature 335*, 176-178.

Kristensen, T., Ogata, R. T., Chung, L. P., Reid, K. B. M., & Tack, B. F. (1987a) *Biochemistry 26*, 4668-4674.

Kristensen, T., D'Eustachio, P., Ogata, R. T., Chung, L. P., Reid, K. B. M., & Tack, B. F. (1987b) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 46*, 2463-2469.

Lintin, S. J., & Reid, K. B. M. (1986) *FEBS Lett. 204*, 77-81.

Lubin, D. M., Liszewski, M. K., Post, T. W., Arce, M. A., Le Beau, M. M., Rebentisch, M. B., Lemons, R. S., Seya, T., & Atkinson, J. P. (1988) *J. Exp. Med. 168*, 181-194.

Lubin, D. M., Post, T. W., Arce, M. A., Liszewski, M. K., & Atkinson, J. P. (1989) *FASEB J. 3*, 801.

Maeda, N., Yang, F., Barnett, D. R., Bowman, B. H., & Smithies, O. (1984) *Nature 309*, 131-135.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

McGill, J. R., Chirgwin, J. M., Moore, C. M., & McCombs, J. L. (1987) *Cytogenet. Cell Genet. 45*, 55-57.

Mount, S. M. (1982) *Nucleic Acids Res. 10*, 459-472.

Nagasawa, S., Ischihara, C., & Stroud, R. M. (1980) *J. Immunol. 125*, 578-582.

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., & Sharp, P. A. (1986) *Annu. Rev. Biochem. 55*, 119-1150.

Post, T. W., & Atkinson, J. P. (1989) *FASEB J. 3*, 368.

Ralph, S. J., Thomas, M L., Morton, C. C., & Trowbridge, I. S. (1987) *EMBO J. 6*, 1251-1257.

Reid, K. B. M., Bentley, D. R., Campbell, D. R., Chung, L. P., Sim, R. B., Kristensen, T., & Tack, B. F. (1986) *Immunol. Today 7*, 230-233.

Rey-Campos, J., Rubinstein, P., & Rodriquez de Cordoba, S. (1988) *J. Exp. Med. 167*, 664-669.

Rodriguez de Cordoba, S. R., Ferreira, A., & Rubinstein, P. (1985) *Immunogenetics 21*, 257-265.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Seldin, M. F., Morse, H. C., III, Reeves, J. P., Scribner, C. L., LeBoeuf, R. C., & Steinberg, A. D. (1988a) *J. Exp. Med. 167*, 688-693.
Seldin, M. F., Morse, H. C., LeBoeuf, R. C., & Steinberg, A. D. (1988b) *Genomics 2*, 48-56.
Sharp, P. A. (1987) *Science 235*, 766-771.
Shen, F.-W., Saga, Y., Litman, G., Freeman, G., Tung, J.-S., Cantor, H., & Boyse, E. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7360-7363.
Silver, J., & Buckler, C. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1423-1427.
Staden, R. (1980) *Nucleic Acids Res. 8*, 3673-3694.
Staden, R. (1982) *Nucleic Acids Res. 10*, 4731-4751.
Traut, T. W. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2944-2948.
Vik, D. P., Keeney, J. B., Munoz-Canoves, P., Chaplin, D. D., & Tack, B. F. (1988) *J. Biol. Chem. 263*, 16720-16724.
Weis, J. H., Morton, C. C., Bruns, G. A. P., Weis, J. J., Klickstein, L. B., Wong, W. W., & Fearon, D. T. (1987) *J. Immunol. 138*, 312-315.
Wong, W. W., Cahill, J. M., Rosen, M. D., Kennedy, C. A., Bonaccio, E. T., Morris, M. J., Wilson, J. G., Klickstein, L. B., & Fearon, D. T. (1989) *J. Exp. Med. 169*, 847-863.
Ziccardi, R. J., Dahlback, B., & Muller-Eberhard, H. J. (1984) *J. Biol. Chem. 259*, 13674-13679.

Purification of Botrocetin from *Bothrops jararaca* Venom. Analysis of the Botrocetin-Mediated Interaction between von Willebrand Factor and the Human Platelet Membrane Glycoprotein Ib-IX Complex†

Robert K. Andrews,‡ William J. Booth,‡ Jeffrey J. Gorman,§ Peter A. Castaldi,‡ and Michael C. Berndt*,‡

Research Centre for Thrombosis and Cardiovascular Disease, Department of Medicine, University of Sydney, Westmead Hospital, Westmead, NSW 2145, Australia, and Commonwealth Scientific and Industrial Research Organisation, Australian Animal Health Laboratory, Geelong, VIC 3220, Australia

Received February 14, 1989; Revised Manuscript Received May 26, 1989

ABSTRACT: Interaction of von Willebrand factor (vWF) with its platelet receptor only occurs in vitro in the presence of a modulator such as ristocetin. We have recently confirmed that the human platelet membrane glycoprotein (GP) Ib-IX complex is the receptor involved in the ristocetin-dependent binding of vWF by reconstitution with the purified components [Berndt, M. C., Du, X., & Booth, W. J. (1988) *Biochemistry 27*, 633-640]. We have now developed a similar solid-phase reconstitution assay using an alternate modulator, botrocetin, for the competitive analysis of functional domains in both vWF and the GP Ib-IX complex. Botrocetin was purified from *Bothrops jararaca* venom by ammonium sulfate fractionation and subsequent DEAE-cellulose and hydroxylapatite chromatography. The purified protein was a 25-kilodalton (kDa) disulfide-linked dimer with apparent subunit molecular weights of 14 000 and 14 500. Binding studies with immobilized botrocetin demonstrated that botrocetin bound to vWF and to a 52/48-kDa region of vWF that contains the GP Ib binding domain, but not to glycocalicin, a proteolytic fragment of GP Ib that contains the vWF binding site. Binding of ^{125}I-labeled vWF to GP Ib-IX complex coated beads and to platelets was strictly botrocetin-dependent with half-maximal binding at a botrocetin concentration of $\simeq$0.27 μM. Botrocetin-dependent binding of vWF was specific, saturable, and comparable to that observed with ristocetin. An anti-vWF monoclonal antibody, 3F8, inhibited ristocetin- but not botrocetin-dependent binding of vWF, suggesting the presence of distinct ristocetin and botrocetin modulator sites on vWF. The botrocetin reconstitution assay was at least an order of magnitude more sensitive than the corresponding ristocetin assay for the competitive analysis of functional domains on both vWF and the GP Ib-IX complex and has confirmed the localization of the vWF-binding domain to the 45-kDa N-terminal region of GP Ib.

One human platelet membrane receptor, the glycoprotein (GP)[1] Ib-IX complex, appears to play a key role in several aspects of the hemostatic process. Binding of the adhesive glycoprotein, von Willebrand factor (vWF), to the GP Ib-IX complex is crucial for the initial contact adhesion of platelets to the exposed subendothelium at high shear flow [for a review, see Girma et al. (1987)] and for the binding of platelets to fibrin (Loscalzo et al., 1986; Parker & Gralnick, 1987).

vWF has a subunit molecular weight of 275 000 consisting of 2050 amino acid residues (Titani et al., 1986) and circulates in plasma as a series of disulfide-linked multimers ranging in molecular weight from 1×10^6 to $>10 \times 10^6$ (Girma et al., 1987). The GP Ib-IX complex binding domain of vWF resides in a 52/48-kDa region, as revealed by the biological activity of a reduced and alkylated tryptic fragment extending

† This investigation was supported by a grant (6K14443) from the National Health and Medical Research Council of Australia. M.C.B. is the recipient of a Wellcome Australian Senior Research Fellowship.
‡ Westmead Hospital.
§ Australian Animal Health Laboratory.

[1] Abbreviations: Da, dalton; DEAE, diethylaminoethyl; EDTA, ethylenediaminetetraacetic acid; GP, glycoprotein; PTH, phenylthiohydantoin; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)-aminomethane; vWF, von Willebrand factor.

0006-2960/89/0428-8317$01.50/0