

Membrane Cofactor Protein of the Complement System: Alternative Splicing of Serine/Threonine/Proline-rich Exons and Cytoplasmic Tails Produces Multiple Isoforms that Correlate with Protein Phenotype

By Theodore W. Post, M. Kathryn Liszewski, Elizabeth M. Adams, Indriati Tedja, Elizabeth A. Miller, and John P. Atkinson

From the Howard Hughes Medical Institute Laboratories and Department of Medicine, Division of Rheumatology, Washington University School of Medicine, St. Louis, Missouri 63110

Summary

Membrane cofactor protein (MCP) is a complement regulatory protein that is expressed on human cells and cell lines as two relatively broad species with M_r of 58,000–68,000 and 48,000–56,000. The structure of a previously reported cDNA clone indicated that MCP was a type 1 membrane glycoprotein and a member of the regulators of complement activation gene/protein cluster. However, it did not provide an explanation for the unusual phenotypic pattern of MCP. Therefore, in parallel with an analysis of the gene, additional cDNAs were cloned and characterized. Six different MCP cDNA classes were identified. All encode the same 5' untranslated signal peptide, four SCRs, transmembrane domain, and basic amino acid anchor. However, they differ in the length and composition of an extracellular serine/threonine/proline (STP)-rich area, a site of heavy O-glycosylation, and cytoplasmic tail. Analysis of the MCP gene demonstrated that the variation in cDNA structure was a result of alternative splicing. Peripheral blood cells and cell lines predominantly expressed four of the six isoforms. These varied by the presence or absence of an STP-rich segment of 15 amino acids (STP^B) and by the use of one of two cytoplasmic domains. Analysis by polymerase chain reaction, Northern blots, and transfection indicated that the predominance of MCP cDNA isoforms with STP^B correlated with the high molecular weight protein phenotype, while the predominance of isoforms without STP^B correlated with the lower molecular weight phenotype. The expression in a single cell of four distinct protein species with variable STP-rich regions and cytoplasmic tails represents an interesting example of the use of alternative splicing to provide variability in a mammalian protein.

Membrane cofactor protein (MCP;¹ CD46) is a widely distributed regulatory protein of the complement system that binds to the complement activation products C3b and C4b and serves as a cofactor for their proteolytic inactivation by factor I (1–3). Except for erythrocytes, MCP has been found on nearly every cell and tissue examined: cells of epithelial, endothelial, and fibroblast lineages (4, 5), PBL (6–8), platelets (5, 7, 9), sperm (10), trophoblast tissues (11–13), and in serum, saliva, and seminal fluid (5). Because of this tissue distribution and functional profile, it has been postulated that MCP helps prevent complement activation on host cells (reviewed in references 14 and 15). Recently, this hy-

pothesis has been confirmed through two separate demonstrations that MCP can protect cells from complement-mediated lysis (16, 17).

MCP migrates on SDS-PAGE as a broad doublet with M_r of the two species of 58,000–68,000 and 48,000–56,000 (6, 18). Although the larger protein form contains more sialic acid (18, 19) and, therefore, presumably more O-linked sugars, the two species are similar functionally, structurally, and antigenically. The relative quantity expressed of each MCP species is inherited in an autosomal codominant fashion (20, 21).

One MCP cDNA has been isolated and characterized (22). Beginning at its NH₂ terminus (see Fig. 1), MCP is composed of four of the ~60 amino acid, cysteine-rich repeating motifs known as short consensus repeats (SCRs). Such units are the building blocks of the regulators of complement activation (RCA) protein/gene cluster found at 1q32 (reviewed

¹ Abbreviations used in this paper: MCP, membrane cofactor protein; nt, nucleotide; RCA, regulators of complement activation; SCR, short consensus repeat; STP, serine/threonine/proline.

in references 1 and 23), at which the structural gene for MCP also maps (24–26). Immediately 3' of the SCRs is a serine/threonine/proline (STP)-rich region, a likely area for extensive O-glycosylation. This is followed by a short segment of unknown significance, a transmembrane domain, basic amino acid anchor, and a cytoplasmic tail.

We now report six additional cDNA clones, each encoding a distinct protein class of MCP. Using PCR of mRNA, Northern blots, and transfections of cDNAs, we demonstrate that four of these isoforms are predominantly expressed by peripheral blood cells and cell lines. Through the cloning and characterization of the MCP gene, we show that alternative splicing produces these distinct classes of MCP and that the presence or absence of an STP-rich exon determines the protein phenotype.

Materials and Methods

Construction and Screening of cDNA and Genomic Libraries. A cDNA library was prepared from 10 µg of HeLa poly(A)⁺ RNA (Clontech, Palo Alto, CA). cDNA inserts were size selected to be >3 kb by agarose gel electrophoresis, ligated into λ Zap^R II, packaged in GigapackTM Gold, and plated on XL1-Blue *Escherichia coli* (all from Stratagene, La Jolla, CA). This library was probed with MCP-1 cDNA (22), which was labeled by random hexanucleotide priming (27).

Two human genomic bacteriophage libraries, obtained from the American Type Culture Collection (ATCC; Rockville, MD) and Stratagene, were constructed by partial EcoRI or Sau3A1 digestion of human genomic DNA, followed by insertion into the vectors Charon 4A and λDASH, respectively. They were screened by standard techniques (27) with MCP cDNA or MCP genomic DNA labeled by random hexanucleotide priming. A chromosome 1-specific library (from ATCC) was also screened with a 3' UT specific oligonucleotide labeled at its 5' end by T4 kinase. This library was constructed by complete EcoRI digestion followed by insertion into the vector Charon 21A.

DNA Isolation and Subcloning. Positive clones from the Zap^R II HeLa library were plaque purified followed by *in vivo* excision of the phagemid per manufacturer's protocol (Stratagene). Positive clones from the genomic library were plaque purified, and DNA was isolated using a small-scale bacteriophage preparation (28). Bacteriophage insert DNA was subcloned into the EcoRI site of pUC 19 by standard techniques (27).

RNA Isolation. Total RNA was isolated from cell lines or peripheral blood cells by either the guanidium isothiocyanate/CsCl₁ (29) or acid phenol (30) methods. PBMC were isolated using standard Ficoll-Hypaque separation methods (6, 31). Isolated RNA was quantitated by either optical density or ethidium bromide staining of agarose gels (27). Previously studied human cell lines used in this investigation were U-937 (monocyte/macrophage), K-562 (erythroleukemia), HSB-2 (T lymphocyte), HeLa (carcinoma of the cervix), and HEp-2 (carcinoma of the larynx) (4, 6).

cDNA Synthesis and PCR of Full-length cDNA. 20 µg of total RNA, obtained from K-562 and HEp-2 cell lines, as described above, was used for each first-strand synthesis reaction. RNA was incubated at 65°C for 3 min and then quick-chilled at 4°C. First-strand cDNA synthesis was then performed using the Copy Kit (Invitrogen, San Diego, CA); however, a 3' UT MCP-specific antisense primer was used in place of oligo(dT) or random primers. This primer was AAGCCACATTGCAATATTAGCTAAGCCACAGT. First strand cDNA was then ethanol precipitated and the product was used

for each subsequent PCR. Except for primers, reagents used for the PCR were from the GeneAmpTM kit (Perkin Elmer Cetus, Norwalk, CT). The MCP-specific primers were as follows: the 5' primer possessed an EcoRI site added to the 5' UT MCP sequence of ATTGTTGCGTCCCATATCTGGACCCAGAAGGG; the antisense MCP primer was the same as used for first-strand synthesis, except an EcoRI restriction site was added. Cycling conditions were an initial denaturation of 94°C for 5 min followed by 39 cycles consisting of a denaturation temperature of 94°C for 1 min, an annealing temperature of 55°C for 1 min, and an extension temperature of 72°C for 2 min. At the completion of the cycling reactions, a final extension was performed for 10 min at 72°C.

The PCR products were ethanol precipitated, digested with EcoRI, and gel isolated. The product was then subcloned into the EcoRI site of pUC 19. Distinct clones were sequenced in their entirety.

DNA Sequence Determination and Mapping. Double-stranded DNA sequencing was performed using the Sequenase Kit (United States Biochemical Corp., Cleveland, OH). Sequencing primers included forward and reverse M13 primers and synthesized oligonucleotides corresponding to known regions of MCP cDNA and MCP intron-specific sequences. Bacteriophage inserts of genomic clones were aligned by restriction mapping, hybridization to MCP exon- and intron-specific probes, and sequencing.

Northern Blots. Total RNA was isolated (as above). RNA samples, along with RNA markers (0.24–9.5 kb; Bethesda Research Laboratories, Gaithersburg, MD) were run on a 1% agarose gel by standard methods (27), except that the concentration of formaldehyde in the gel was decreased to 0.22 M. Gels were blotted onto Nitroplus 2000 membrane (Micron Separations, Inc., Westborough, MA). Blots were hybridized and washed at high stringency (27).

cDNA Synthesis and PCR for "Short Form" of cDNA. 2 µg of U-937, 10 µg of K-562, 12 µg of HSB-2, 9 µg of HeLa, 0.2 µg of HEp-2 total RNA, and 5–20 µg of six individuals' peripheral blood cells total RNA were used for each first-strand synthesis reaction in a total volume of 20 µl (Copy Kit; Invitrogen). Reagent concentrations were as per manufacturer's protocol, except that both oligo(dT) and random primers were used. Two first-strand controls were used. One was the same amount of cell line RNA and all reaction reagents but without reverse transcriptase. The other was 20 µg of yeast tRNA with all first-strand reaction reagents including reverse transcriptase. First-strand reaction mixtures were incubated for 2 h at 42°C and then at 72°C for 10 min before being ethanol precipitated and resuspended in 30 µl of water.

PCR amplification was performed as described earlier using GeneAmpTM Kit, except that cycling numbers were varied to provide linearity of the reaction. The sequence of the 5' MCP-specific primer was TGTCGACTGGAATTCCTGGTCAAATGTCGATT-TCCAGTAGTCGAAAAT, and of the 3' primer was GATATTCA-CCTGCAGGAAGCCACATTGCAATATTAGCTAAGCCACAGT. The PCR product was analyzed on a 3% agarose gel. PCR experiments using plasmids containing known MCP cDNAs were performed as above. The PCR products from cell lines and plasmids were analyzed by digestion with either SalI or StuI.

Transfection of MCP Isoforms. The cDNA isoforms were subcloned into the EcoRI site of the expression vector SFFV.neo (32). Plasmid DNA was subsequently prepared using the pZ523 kit from 5 Prime 3 Prime, Inc. (West Chester, PA). Chinese Hamster Ovary cells (CHO-K1; ATCC), were then transfected with the prepared DNA using LipofectinTM Reagent (Bethesda Research Laboratories) following the manufacturer's suggestion for preparation of stable transfectants. CHO cells were grown in Ham's F-12 with 10% FCS, penicillin (100 U/ml), and streptomycin (100 µg/ml).

2 d after transfection, cells were grown in the same medium supplemented with Geneticin® (0.5 mg/ml activity concentration) (Gibco Laboratories, Grand Island, NY). Transfected cells were sorted on an EPIC 753 (Coulter Corp., Hialeah, FL) one to four times depending on expression levels.

Labeling and Immunoprecipitation of Cells. Unless otherwise indicated, reagents were obtained from Sigma Chemical Co. (St. Louis, MO). Cell lines and stably transfected cells, the latter grown to 70–80% confluency, were washed and “starved” for 1 h at 37°C in cysteine-free medium containing 10% dialyzed FCS. After this, ³⁵S-cysteine was added (100 μCi/ml) (1,000 Ci/mmol; New England Nuclear, Boston, MA), and the incubation continued for 4 h. PBMC were surface labeled with ¹²⁵I (6).

To immunoprecipitate MCP, cells were solubilized in lysis buffer (1% NP-40 in 10 mM TRIS, pH 7.4, 150 mM NaCl, 0.05% SDS, 2 mM PMSF, and 2.5 mM EDTA). After incubating at 4°C for 20 min, cell debris was pelleted at 10,000 g for 10 min. Cell lysate supernatants were “pre-cleared” for 30 min at 4°C by adding 5 μl normal rabbit serum, 10 μl (10 μg) UPC-10 (a mouse monoclonal IgG2a of unknown specificity), and 50 μl packed protein A-agarose (Boehringer Mannheim Biochemicals, Indianapolis, IN). The “cleared” supernatant was then incubated for 1 h at 4°C with 5 μl of ascites fluid containing E4.3, an IgG2a mAb to MCP (33–35). Subsequently, 15 μl of protein A-agarose was incubated for 20 min. After washing the pelleted sample with cell lysis buffer containing 0.5 M NaCl, gel buffer was added and the sample heated to 100°C for 3 min. Gel buffer contained 10% (vol/vol) glycerol, 0.2% SDS, 62 mM TRIS-HCl (pH 6.8). The supernatant was fractionated on a 10% SDS polyacrylamide gel using the Laemmli system (36). Gels were fixed, stained, treated with Amplify (Amersham Corp., Arlington Heights, IL), dried, and exposed to Kodak XAR film.

Results

Multiple MCP cDNAs. The initial MCP cDNA clone (MCP-9) did not provide an explanation for the two bands on SDS-PAGE and was nearly 2 kb smaller than the major transcript (22). Therefore, additional MCP cDNA clones were sought by screening a cDNA library and by PCR. Since HeLa cells express a relatively large quantity of both MCP species (4), a size-selected (>3.0 kb) cDNA library was prepared and two informative clones were characterized. Clone HL-1 possessed sequence identical to MCP-9 (22) and therefore encoded for the same protein. However, HL-1 added 102 nucleotides (nt) of 5' UT sequence and 1,580 nt of 3' UT sequence for a total length of 3,212 nt (All clone sequences have been

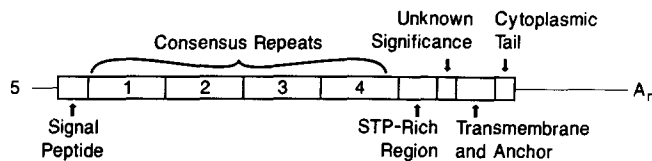


Figure 1. A schematic diagram of MCP cDNA. Data are derived from a previously reported cDNA clone, MCP-1 (22). Coding regions are shown by boxes, and 5' and 3' UT regions by lines. The protein consists of 384 amino acids and, beginning at its NH₂ terminus, has the following structure: signal peptide, 34 amino acids; consensus repeats 1–4, 251 amino acids; STP-rich region, 29 amino acids; unknown region, 13 amino acids; transmembrane domain and cytoplasmic anchor, 34 amino acids, cytoplasmic tail 23 amino acids.

submitted to the GenBank™/EMBL Data Bank). A Northern blot, using total RNA from K-562, U-937, HeLa, and HEp-2 cell lines, was hybridized at high stringency with a 0.6-kb EcoR1 fragment unique to this additional 3' UT sequence. The blot indicated that this sequence was transcribed and was part of the major transcript of all four cell lines (data not shown). Additionally, HL-1 possessed the more common polyadenylation signal AATAAA, as opposed to the atypical signal of MCP-9 (AATATA or AATGAA) (22).

Another clone from the HeLa Library (HL-27) was identical to HL-1, except for a 93-bp sequence located immediately after the transmembrane domain. Because of an inframe stop codon, this segment coded for an alternate cytoplasmic tail of 16 amino acids followed by 45 nt of 3' UT sequence (Fig. 2 A). The remaining 3' UT sequence in HL-27 was identical to the cDNA sequence of HL-1. Thus, the sequence for the cytoplasmic tail identified in the original MCP clone (MCP-9) (22) became part of the 3' UT area in clone HL-27. Due to the ordering of these segments, this alternate tail of clone HL-27 was designated CYT¹, and the initially characterized MCP cytoplasmic tail was designated CYT². A Northern blot, hybridized at high stringency with an antisense oligonucleotide to this 93-bp insert, demonstrated that all four of the cell lines examined possessed transcripts bearing this insert but that the quantity was variable (Fig. 2 B).

PCR was also used to generate cDNAs. MCP-specific primers from the 5' UT and 3' UT regions were used to produce clones that span the coding region of MCP. From both K-562 and HEp-2 total RNA, these primers produced PCR products of ~1.4 kb. Subclones, sequenced in their entirety, revealed four additional forms of MCP cDNA.

Clones K5-9 and K5-22, derived from cell line K-562, possessed an additional segment immediately after SCR-4 and before the previously defined STP region of MCP. This segment was termed STP^A (Figs. 3 and 4) because, if translated, 9 of 15 amino acids were S, T, or P. The cDNA bearing this 45-nt insert appeared with CYT¹ in clone K5-9 and with CYT² in clone K5-22 (Fig. 4). A Northern blot, hybridized with an antisense probe in STP^A, gave a weak signal (data not shown), suggesting that transcripts bearing this insert represented rare species in the four cell lines examined (also, see below).

PCR of HEp-2 total RNA yielded two other clones, H2-14 and H2-15, in which a portion of the STP region, termed STP^B, was deleted. H2-14 and H2-15 were identical except that H2-14 contained CYT¹ and H2-15 contained CYT². Fig. 4 summarizes the clones characterized. Of note, the addition of CYT¹ and STP^A, or deletion of STP^B, did not disrupt the reading frame.

PCR Analysis of the Variable Regions within MCP mRNAs. We next asked which of the six transcripts were commonly expressed. To answer this, two primers spanning the area of variability, from SCR-4 to the 3' UT region, were used in the PCR.

The PCR products derived from the total RNA of the cell lines were found to consist of variable quantities of four bands (Fig. 5). To identify these products, the same primers were used to amplify this area from cloned full-length MCP

A

1 CACATACCTAACTGATGAGACCCACAGAGAAGTAAATTTACTTCTCTGAGAGGAGAGATGAGAGAAAGGTTTGGCTTTATCATTAAAG
 70
 T Y L T D E T H R E V K F T S L U

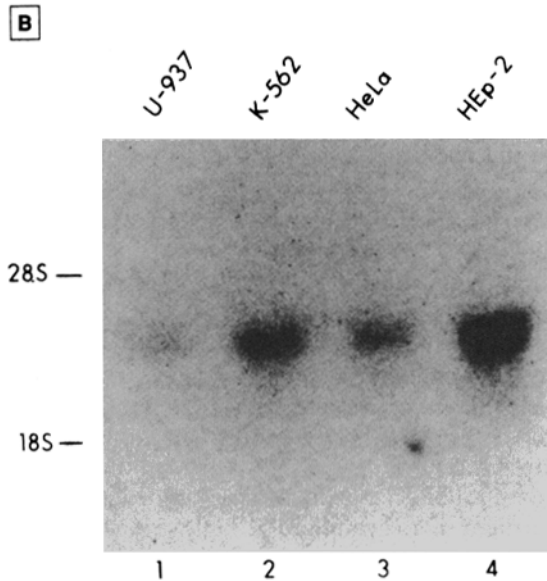


Figure 2. Cytoplasmic tail one (CYT¹). (A) nt and derived amino acid sequence. The encoded amino acids are centered under the second nt of each triplet codon. (B) Northern blot probed with an antisense oligonucleotide to CYT¹. Total RNA from each of the four cell lines was probed with an antisense oligonucleotide to CYT¹ (the sequence of this probe is antisense to nt 49-2 of A). The four cell lines show a positive signal at ~3.2 kb.

GTGCTGCCTCCATCTAGTACAAAACCTCCAGCTTTGAGTCATTCA
 V L P P S S T K P P A L S H S

Figure 3. nt and derived amino acid sequence of STP^A. This 45-nt insert codes for 15 amino acids, of which nine are S, T, or P. The insert is positioned immediately after SCR-4 and before the previously described STP region (see Fig. 1).

cDNAs bearing known isoforms (Fig. 6). The resultant product was compared with those obtained from the HeLa cell line. The known PCR products of Fig. 6, lanes 1, 2, 5, and 6 align with the four bands from HeLa cells, and also have the expected size for their respective mRNAs.

To further verify the identity of the four PCR products, a restriction enzyme analysis was undertaken. A Sal1 site is located in the sequence of STP^B, and a Stu1 site is in STP^C

(22). Only the PCR products from clones bearing STP^B were cut by Sal1 into the anticipated fragments (Fig. 7, lanes 3 and 5). All four PCR products were cut by Stu1 into fragments of the expected size (lanes 6-9), indicating they are derived from cDNA inserts with STP^C. If the PCR products of each cell line are the same as the products from the known cDNA clones, they should be digested in a similar fashion. Consequently, PCR products derived from HeLa were digested with Sal1 and compared with the undigested products (Fig. 8). Bands 1 and 3 were not cut, while bands 2 and 4 were digested into the fragments of the expected size. In addition, all four bands were digested with Stu1, and the correct fragments were produced (Figs. 7 and 8). A similar result was observed for the K-562 PCR products, as Sal1 cut bands 1 and 3 (lane 5) while Stu1 cut all four bands (Fig.

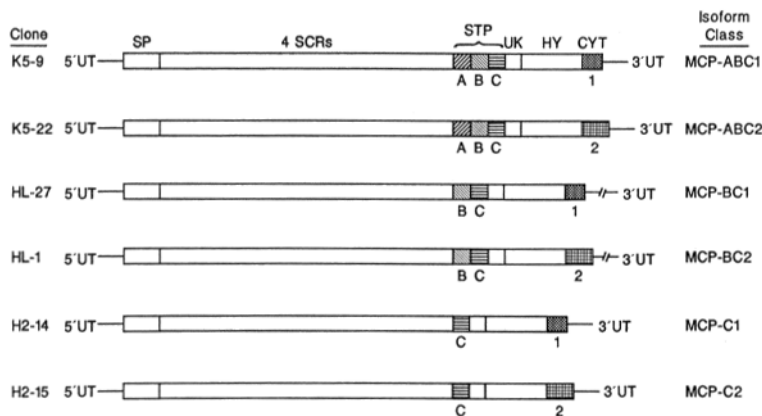


Figure 4. Schematic diagram of the six classes of MCP cDNA clones obtained by both cDNA cloning and PCR analysis of mRNA. Abbreviations are SP, signal peptide; SCRs, short consensus repeats; STP, three domains enriched in serines, threonines, and prolines; UK, area of unknown significance; HY, hydrophobic transmembrane domain with anchor; and CYT, cytoplasmic tail. Clones designated by K were derived from cell line K-562; HL, from HeLa; and H2, from HEP-2. Isoform class denotes content of STP-rich areas using capital letters, and denotes cytoplasmic tails by the number 1 or 2.

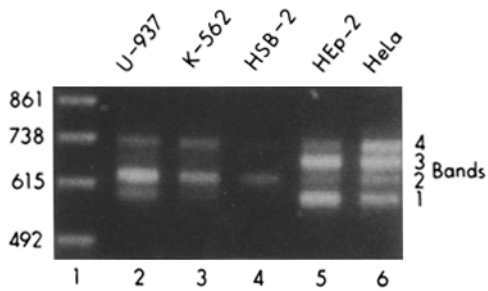


Figure 5. PCR products derived from RNA of five human cell lines. The oligonucleotides used as primers were from SCR-4 and the 3' UT region of MCP (see Materials and Methods). Lane 1, molecular weight markers in nt; lanes 2-6, PCR products of the cell lines noted. The bands marked 1-4 in this and subsequent figures refer to the four PCR products.

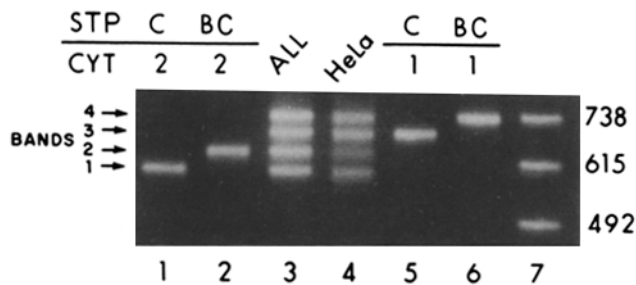


Figure 6. PCR products derived from MCP cDNAs compared with PCR products derived from HeLa cell RNA. Lanes 1, 2, 5, and 6 are the PCR products derived from plasmids with known MCP cDNA inserts. The protein domains contained within these cDNAs are shown above each lane. Lane 3 shows all four PCR products together, and lane 4 shows the PCR products derived from HeLa cell RNA. Lane 7, molecular weight markers in nt. Abbreviations are as per Fig. 4. Primers are the same as Fig. 5.

8, lane 7). This same pattern of enzyme digestion was found for U-937, HEP-2, and HSB-2 cell lines (data not shown).

Therefore, the PCR products amplified in the cell lines are the same as known MCP RNA isoforms (Fig. 6). Bands 1 and 3 represent mRNA species that contain STP^C with either CYT¹ (band 3) or CYT² (band 1), while bands 2 and 4 represent those that contain STP^B and STP^C with either CYT¹ (band 4) or CYT² (band 2).

Correlation of Protein Phenotype to mRNA Isoforms. PCR, Northern blot, and transfection analyses were undertaken to determine if there was a correlation between the mRNA species and protein phenotype.

PCR Analysis. Total RNA, derived from cell lines and from peripheral blood cells of individuals possessing the three protein phenotypes, was examined by PCR. The identity of the PCR products was verified by restriction enzyme analysis (data not shown). The relative amounts of bands 2 (with STP^B and CYT²) and 4 (with STP^B and CYT¹) were increased in individuals expressing predominantly the larger protein phenotype of MCP (Fig. 9). In contrast, the relative amounts of bands 1 (without STP^B and with CYT²) and 3 (without STP^B and with CYT¹) were increased in individ-

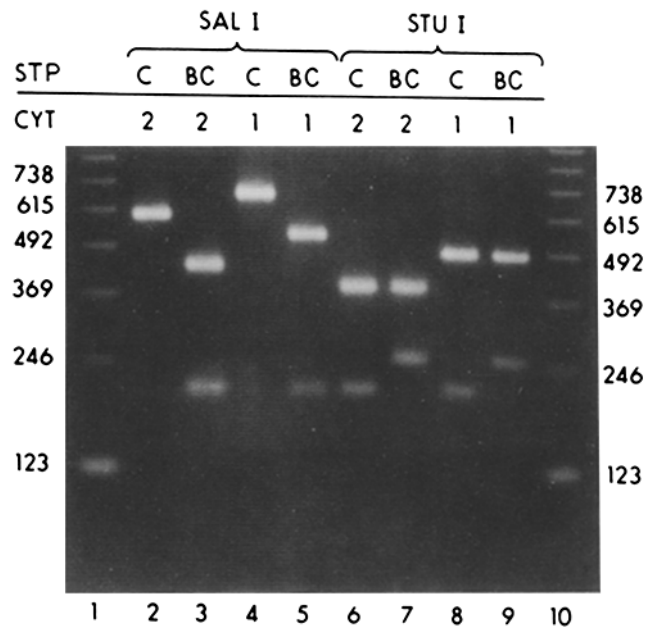


Figure 7. Restriction enzyme analysis of the PCR products. The starting (uncut) PCR products were the same as those in Fig. 6. Lanes 1 and 10, molecular weight markers in nt; lanes 2-5, SalI digests of the PCR products derived from plasmids with the protein domains as noted. Lanes 6-9, StuI digests of the same PCR products as in lanes 2-5. Primers are the same as in Fig. 5.

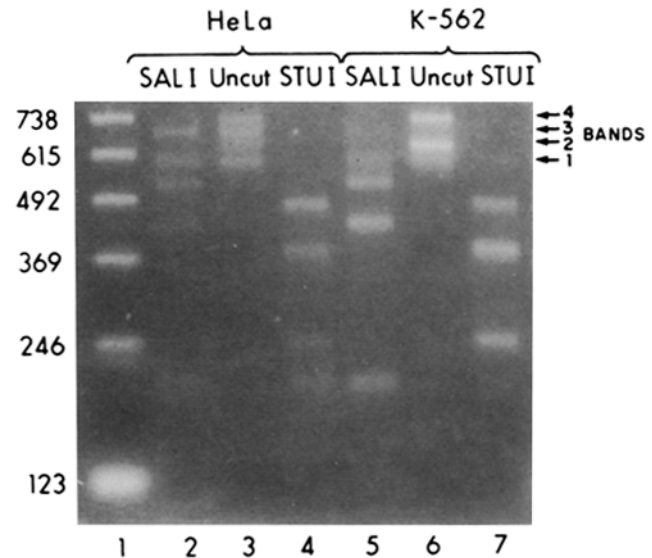


Figure 8. Restriction enzyme analysis of HeLa and K-562 PCR products. The starting (uncut) PCR products are the same as those in Fig. 5. Lane 1, molecular weight markers in nt; lanes 2 and 5, SalI digests; lanes 3 and 6, uncut PCR products; lanes 4 and 7, StuI digests. Primers are the same as in Fig. 5.

uals expressing predominantly the lower protein phenotype, and comparable quantities of bands 1 and 3 vs. 2 and 4 were present in individuals expressing equal quantities of the lower and upper protein forms. Also, this same correlation was ob-

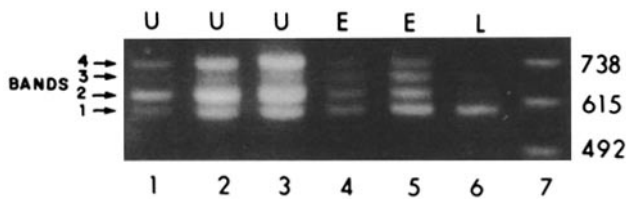


Figure 9. PCR analysis of total RNA derived from PBMC. Lanes 1–3, individuals with the upper (*U*) protein phenotype; lanes 4 and 5, from equal (*E*) protein phenotypes; and lane 6, from a lower (*L*) band predominant protein phenotype. Lane 7, molecular weight markers in nt. Primers are the same as in Fig. 5.

served in the cell lines (Fig. 5); thus, mRNA of cell lines expressing mostly the higher molecular weight protein phenotype (U-937, K-562, and HSB-2) produced predominantly PCR bands 2 and 4; the lower molecular weight protein phenotype (HEp-2), bands 1 and 3; and the equal protein phenotype (HeLa), comparable quantities of bands 1 and 3 vs. 2 and 4.

Northern Analysis. A Northern blot containing total RNA from several cell lines was probed with two antisense oligonucleotide probes (Fig. 10). One probe was antisense to SCR-4, while the other was antisense to all of STP^B. A plasmid containing both the SCR-4 and STP^B was used as a positive control. Two cell lines (U-937 and K-562) expressing predominantly the larger protein phenotype demonstrated comparable hybridization signals to both probes (Fig. 10). However, a cell line (HeLa) expressing about equal quantities and the other one (HEp-2) expressing primarily the lower protein form demonstrate a diminished signal to the probe to STP^B, as compared with the signal from SCR-4. Therefore, this Northern blot analysis supported the results obtained by PCR and indicated that the presence or absence of STP^B determines the protein phenotype.

Transfection. CHO cells were transfected with the cDNAs and the expressed proteins were precipitated and analyzed by SDS-PAGE followed by autoradiography (Fig. 11). Cells trans-

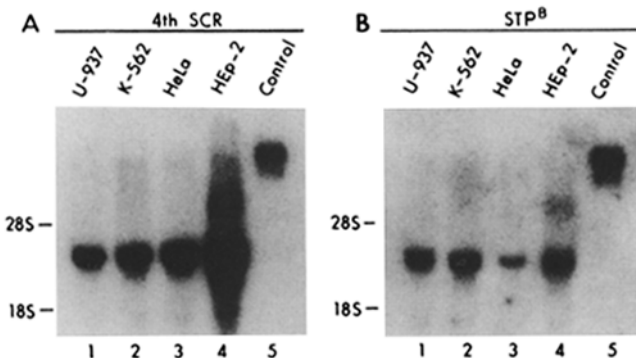


Figure 10. Northern blots of four cell lines probed with antisense oligonucleotides to SCR-4 and STP^B. (A) 20 μ g of total RNA from each of four cell lines was probed with an antisense oligonucleotide to SCR-4; (B) an equivalent blot probed with an antisense oligonucleotide to STP^B. Lane 5 of A and B, same quantity of a plasmid containing both SCR-4 and STP^B.

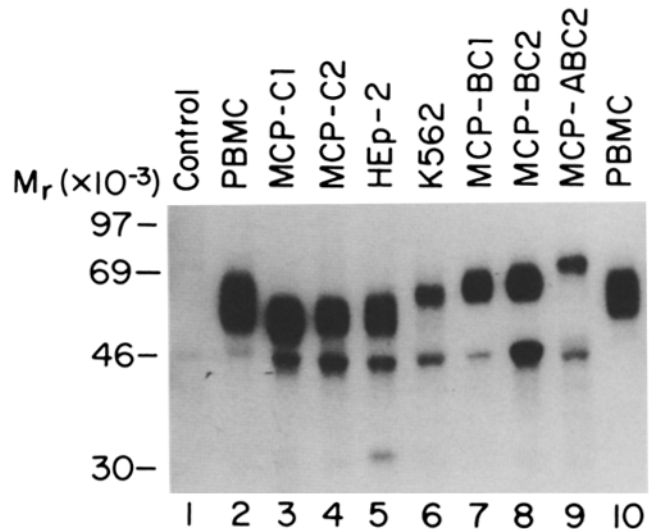


Figure 11. A comparison of MCP of transfectants and cells and cell lines. SDS-PAGE (10% reducing) analysis of MCP immunoprecipitates of cDNA transfectants, PBMC, and two human cell lines. The transfectants and cell lines were biosynthetically labeled with ³⁵S-cysteine. Blood mononuclear cells of an individual expressing about equal quantities of the two forms of MCP were labeled with ¹²⁵I. The control is CHO cells transfected with an MCP clone in reverse orientation. HEp-2 is a predominantly lower (*L*) band protein phenotype, while K-562 is a predominantly upper (*U*) band protein phenotype. Transfected MCP isoforms are identified on the basis of their STP regions and cytoplasmic tails (see Fig. 4). The bands in lanes 3–9 with M_r of 46,000–48,000 represent intracellular precursors of MCP.

fected with cDNAs containing STP^C migrated with an M_r of 45,000–55,000 (Fig. 11, lanes 3 and 4). In contrast, MCP expressed in CHO cells with STP^B migrated with an M_r of 55,000–65,000 (Fig. 11, lanes 7 and 8). The M_r of the expressed proteins were similar to those of MCP of cell lines (lanes 5 and 6) and of PBMC (lanes 2 and 10). The type of cytoplasmic tail had no detectable effect on the M_r . Thus, the presence of STP^B defines the larger protein phenotype and its absence produces the lower protein phenotype. These results are entirely consistent with the PCR and Northern analyses.

The MCP cDNA containing STP^{ABC} was also expressed and had an M_r of 65,000–75,000 (Fig. 11, lane 9). Such species have not been detected on human peripheral blood cells or cell lines (4, 6, 7, 18, 19) (Fig. 11, lanes 2, 5, 6, and 10). Therefore, these results are consistent with those derived from Northern blots (Fig. 3) and PCR (Figs. 5 and 6), and indicate that STP^{ABC} bearing mRNAs are rare in these cell populations.

Organization of the MCP Gene. To determine how the isoforms arise, three genomic bacteriophage libraries were screened and genomic fragments obtained that contained 14 exons and 13 introns of the MCP gene (Fig. 12). These exons, spread over a length of \sim 43 kb, contained one area that did not overlap. All exons and surrounding intronic areas were sequenced. The exonic sequences matched those of MCP cDNAs and PCR clones. All exon/intron boundaries conformed to the GT/AG consensus sequence rule (37). Exon

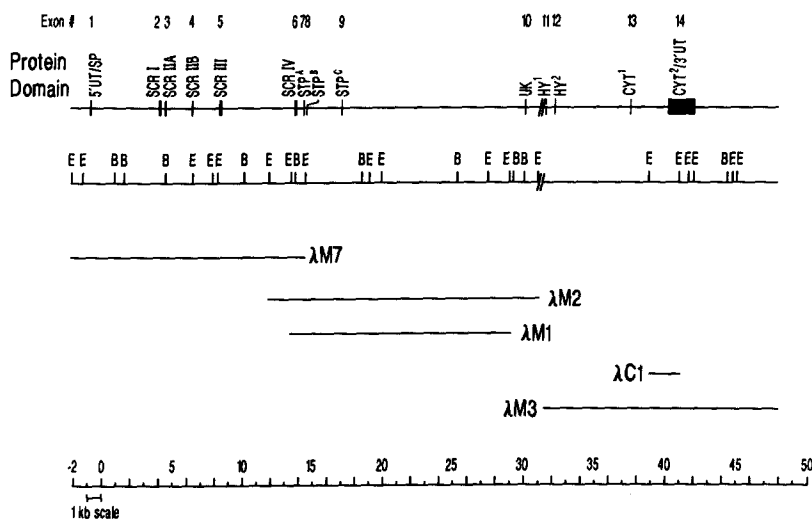


Figure 12. Structure of the human MCP gene. The exon number and corresponding encoded protein domain are shown above each exon. Exon and intron sizes are drawn to scale. As depicted by the slashed horizontal lines in the first line, the intron between exons 10 and 11 has not been cloned. The second line shows an EcoRI and BamHI restriction map of the MCP gene. Five bacteriophage clones (λ M7, λ M2, λ M1, λ M3, and λ C1) contained all 14 exons. Abbreviations of protein domains are as per Fig. 4.

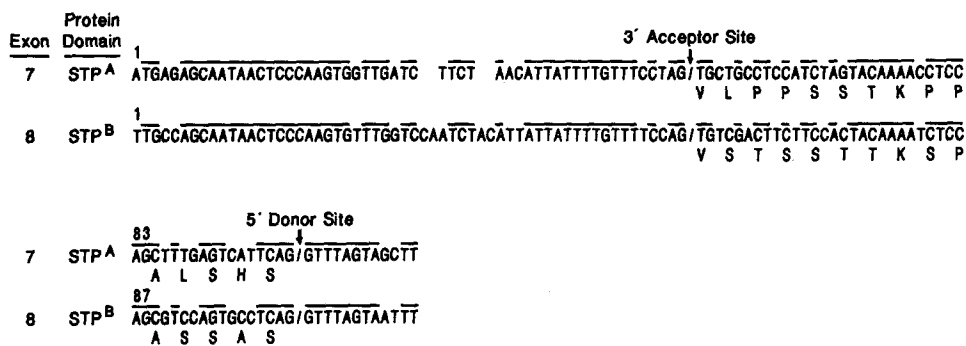


Figure 13. Homology between STP^A and STP^B and surrounding intronic areas. The nt sequences of the 3' acceptor intron, exon, and 5' donor intron of exons 7 and 8 are shown. The encoded amino acids of exons 7 and 8 are centered under the second nt of each triplet codon. Homologous nt between the two sequences are overlined. Spacing of the 3' acceptor intron of exon 7 was aligned to maximize homology.

sizes ranged from 36 to 1,985 nt, while the sizes of the characterized introns ranged from 127 nt to ~13 kb.

The first exon encodes the 5' UT/signal peptide region. The next five exons encode the four SCRs of MCP. SCR-1, SCR-3, and SCR-4 are encoded by a single exon, while SCR-2 is encoded by two equally sized exons. Interestingly, this "split" exon occurs after the second nt of glycine 34. This pattern is identical to the split exons of other complement regulatory proteins (1). Exons 7, 8, and 9 encode the three STP-rich areas. Since the 5' donor splice site of exons 7 and 8 occurs after the first nt of the triplet codon, the addition or deletion of these exons does not disrupt the reading frame. Interestingly, exons 7 and 8 have an overall homology of ~70%. Moreover, there is extensive homology in the 3' acceptor and 5' donor introns of both of these exons (Fig. 13), indicating that these exons arose by intragenic duplication.

Exon 10 encodes for a 13 amino acid stretch of unknown significance. Exons 11 and 12 encode for the hydrophobic transmembrane domain and the basic amino acid anchor. Exon 13 encodes for the 93-nt segment CYT¹. If exon 13 is present in a transcript, this nt sequence codes for a cytoplasmic tail of 16 amino acids. Since this exon also codes for an in-frame stop codon, the presence of exon 13 converts exon 14 into the 3' UT region of MCP. If exon 13 is not present,

then exon 14 encodes the cytoplasmic tail of 23 amino acids, named CYT², and the 3' UT region of MCP.

Fig. 14 illustrates how alternative splicing of the STP and CYT exons produces the multiple isoforms of MCP.

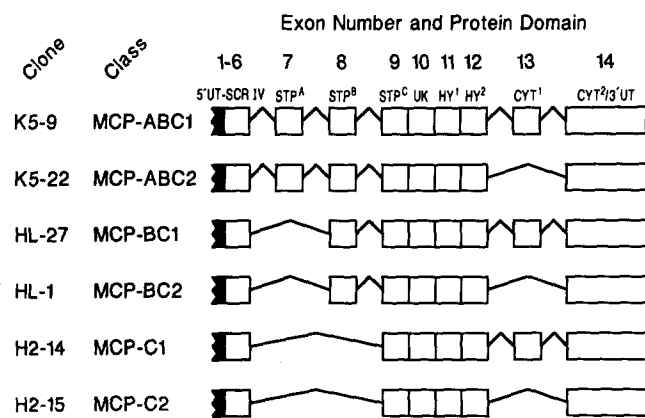


Figure 14. Multiple MCP cDNAs are generated by alternative splicing. The top illustrates the clone, isoform class, exon number, and the protein domains of MCP cDNAs. Abbreviations are as per Fig. 4.

Discussion

An unusual aspect of MCP is the variable expression of two relatively broad protein species (2, 3, 6). The goal of the present investigation was to determine the molecular basis of this phenotypic observation. The approach taken was to identify and characterize additional cDNA clones. As this was progressing, the organization of the MCP gene was also being determined. Six unique mRNA species were characterized. Each mRNA encoded for an identical 5' UT region, signal peptide, four SCRs, transmembrane domain, and basic amino acid anchor. Due to variable splicing, the cDNAs differed in the length of the STP-rich area and in the type of cytoplasmic tail. Alternative splicing produces three different STP-rich regions and two distinct cytoplasmic tails. The alternative splicing event that gives rise to the two tails occurs with each of the splicing events involving the STP exons; that is, each variation in the STP region was found with each tail.

To relate these molecular observations to those previously made at the protein level, three types of mRNA analyses were performed. PCR analysis, using primers encompassing the area of variability within the mRNA, demonstrated that variable quantities of four primary MCP mRNA species were expressed in cell lines and in peripheral blood cells. Two of these mRNA species possessed and two lacked STP^B. Northern blots indicated that the relative predominance of isoforms with STP^B correlated with the higher molecular weight protein phenotype, while the predominance of isoforms without STP^B correlated with the lower molecular weight protein phenotype. This finding was supported by transfection analysis, as cDNAs with STP^B produced proteins with an M_r of the larger species, while those without STP^B had an M_r similar to the smaller protein species. Taken together, these data establish that a cell's MCP phenotype depends on the content of the STP area of the predominant mRNA isoforms.

A role for STP-rich regions (and therefore O-glycosylation) in defining the two major forms of MCP was in part anticipated, since the larger species possesses more sialic acid than the smaller one (19). Although STP^B is only 15 amino acids long, 11 residues are S, T, or P. Thus, the attachment of O-linked sugars could account for the molecular weight difference between the protein species. Further, the broadness of

the forms of MCP is largely accounted for by post-translational modifications, at least in part due to the addition of carbohydrate (19). It is likely that variability in the number of O-linked sugars attached to the STP-rich areas explains this heterogeneity. That MCP transfected into CHO cells produced mature proteins exhibiting broad bands on gels supports this suggestion. The seven amino acid difference between the two tails may also account for some of the variability in size, although transfectants identical except for their cytoplasmic tails were not separable on SDS-PAGE.

The transfectants bearing the different forms of MCP will be useful in future studies. For example, up to four forms of MCP are expressed on a single cell. The biologic significance of such diversity for a regulatory membrane protein is unknown and would be interesting to examine. Also, the role of STP-rich domains adjacent to a transmembrane domain is not well understood. Several receptor and membrane-bound glycoproteins, such as the low density lipoprotein receptor (38), the IL-2 receptor (39-41), and another RCA protein, DAF (42, 43), have STP-rich areas immediately adjacent to their transmembrane domains. While the function of the STP domain is unknown, some evidence suggests that O-glycosylation may protect the protein from proteolysis (44, 45). Alternative splicing of the STP-rich exons in these genes has not been described, and MCP is unique in this regard. Interestingly, the leukocyte common antigen family, whose structural gene is also located at 1q32 (46), resembles MCP in that variable splicing among STP-rich exons is used to produce multiple cDNA species. In this gene, though, alternative splicing may be cell specific, and this region, located at the NH₂ terminus of the protein, is postulated to be a lectin binding domain. The variation in the STP region of MCP isoforms is a model system to analyze the function of such domains in mammalian proteins.

Finally, the biologic significance of the two cytoplasmic tails will be of interest to evaluate. Approximately equal quantities of MCP forms bearing each tail are expressed on human cells. Previous studies demonstrated that the processing of two precursor forms of MCP varied fourfold (19). Preliminary studies indicate that the cytoplasmic tail determines this difference, and further analysis of this observation should provide insights into the signals required for the processing and compartmentalization of mammalian precursor proteins.

We thank Lorraine Whiteley, Pat Parvin, and Anita Zinna for efficient manuscript preparation; Lisa Westfield for oligonucleotide synthesis; Catherine Dirnbeck and Jan Bragdon for skillful cell sorting on the EPIC analyzer; and Dennis Hourcade, V. Michael Holers, Douglas Lublin, V. Kumar, and Matthew Thomas for helpful discussions.

Portions of this work have appeared in abstract form in *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 3:A368, 1989; *Complement Inflammation*. 6:389, 1989; and *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 4:A2187, 1990.

Address correspondence to John P. Atkinson, Division of Rheumatology, Washington University School of Medicine, 4566 Scott Avenue, Box 8045, St. Louis, MO 63110.

Received for publication 19 March 1991.

References

- Hourcade, D., V.M. Holers, and J.P. Atkinson. 1989. The regulators of complement activation (RCA) gene cluster. *Adv. Immunol.* 45:381.
- Lublin, D.M., and J.P. Atkinson. 1989. Decay-accelerating factor and membrane cofactor protein. *Curr. Top. Microbiol. Immunol.* 153:123.
- Liszewski, M.K., T.W. Post, and J.P. Atkinson. 1991. Membrane cofactor protein (MCP or CD46): newest member of the regulators of complement activation gene cluster. *Annu. Rev. Immunol.* 9:431.
- McNearney, T., L. Ballard, T. Seya, and J.P. Atkinson. 1989. Membrane cofactor protein of complement is present on human fibroblast, epithelial and endothelial cells. *J. Clin. Invest.* 84:538.
- Purcell, D.F.J., G.J. Clark, M.A. Brown, I.F.C. McKenzie, M.S. Sandrin, and J.J. Deacon. 1990. HuLy-m5, an antigen sharing epitopes with envelope gp70 molecules of primate retroviruses, and a structural relationship with complement regulatory molecules. In *Leukocyte Typing IV. White Cell Differentiation Antigens*. W. Knapp, editor. Oxford University Press, Oxford. 653-656.
- Cole, J., G.A. Housley, Jr., T.R. Dykman, R.P. MacDermott, and J.P. Atkinson. 1985. Identification of an additional class of C3-binding membrane proteins of human peripheral blood leukocytes and cell lines. *Proc. Natl. Acad. Sci. USA.* 82:859.
- Seya, T., L.L. Ballard, N.S. Bora, V. Kumar, W. Cui, and J.P. Atkinson. 1988. Distribution of membrane cofactor protein (MCP) of complement on human peripheral blood cells. *Eur. J. Immunol.* 18:1289.
- Sparrow, R.L., and I.F.C. McKenzie. 1983. HuLy-m5: a unique antigen physically associated with HLA molecules. *Hum. Immunol.* 7:1.
- Yu, G.H., V.M. Holers, T. Seya, L. Ballard, and J.P. Atkinson. 1986. Identification of a third component of complement-binding glycoprotein of human platelets. *J. Clin. Invest.* 78:494.
- Fenichel, P., B.-L. Hsi, D. Farahifar, M. Donzeau, D. Barrier-Delpech, and C.J.G. Yeh. 1989. Evaluation of the human sperm acrosome reaction using a monoclonal antibody, GB24, and fluorescence-activated cell sorter. *J. Reprod. Fert.* 87:699.
- Stern, P.L., N. Beresford, S. Thompson, P.M. Johnson, P.D. Webb, and N. Hole. 1986. Characterization of the human trophoblast-leukocyte antigenic molecules defined by a monoclonal antibody. *J. Immunol.* 137:1604.
- Hsi, B.-L., J.S. Hunt, and J.P. Atkinson. 1991. Differential expression of complement regulatory proteins on subpopulations of human trophoblast cells. *Am. J. Reprod. Immunol.* In press.
- Hunt, J.S., and B.-L. Hsi. 1990. Evasive strategies of trophoblast cells: Selective expression of membrane antigens. *Am. J. Reprod. Immunol.* 23:57.
- Atkinson, J.P., and T. Farries. 1987. Separation of self from non-self in the complement system. *Immunol. Today.* 8:212.
- Seya, T., and J.P. Atkinson. 1989. Functional properties of membrane cofactor protein of complement. *Biochem. J.* 264:581.
- Seya, T., T. Hara, M. Matsumoto, Y. Sugita, and H. Akedo. 1990. Complement-mediated tumor cell damage induced by antibodies against membrane cofactor protein (MCP, CD46). *J. Exp. Med.* 172:1673.
- Oglesby, T.J., D. White, I. Tedja, K. Liszewski, L. Wright, J. Van der Bogarde, and J.P. Atkinson. 1991. Protection of mammalian cells from complement (C)-mediated lysis by transfection of human membrane cofactor protein (MCP) and decay accelerating factor (DAF). *Clin. Res.* 39:251A.
- Seya, T., J. Turner, and J.P. Atkinson. 1986. Purification and characterization of a membrane protein (gp45-70) that is a cofactor for cleavage of C3b and C4b. *J. Exp. Med.* 163:837.
- Ballard, L.L., N.S. Bora, G.H. Yu, and J.P. Atkinson. 1988. Biochemical characterization of membrane cofactor protein of the complement system. *J. Immunol.* 141:3923.
- Ballard, L., T. Seya, J. Teckman, D.M. Lublin, and J.P. Atkinson. 1987. A polymorphism of the complement regulatory protein MCP (membrane cofactor protein or gp 45-70). *J. Immunol.* 138:3850.
- Bora, N.S., T.W. Post, and J.P. Atkinson. 1991. Membrane cofactor protein (MCP) of the complement system: a Hind III RFLP that correlates with the expression polymorphism. *J. Immunol.* 146:2821.
- Lublin, D.M., M.K. Liszewski, T.W. Post, M.A. Arce, M.M. LeBeau, M.B. Rebentisch, R.S. Lemons, T. Seya, and J.P. Atkinson. 1988. Molecular cloning and chromosomal localization of human membrane cofactor protein (MCP): evidence for inclusion in the multi-gene family of complement-regulatory proteins. *J. Exp. Med.* 168:181.
- Campbell, R.D., S.K.A. Law, K.B.M. Reid, and R.B. Sim. 1988. Structure, organization and regulation of the complement genes. *Annu. Rev. Immunol.* 6:161.
- Carroll, M.C., E.A. Alicot, P.J. Katzman, L.B. Klickstein, J.A. Smith, and D.T. Fearon. 1988. Organization of the gene encoding complement receptors type 1 and 2, decay accelerating factor and C4-binding protein in the RCA locus on human chromosome 1. *J. Exp. Med.* 167:1271.
- Rey-Campos, J., P. Rubinstein, and S. Rodriguez De Cordoba. 1988. A physical map of the human regulator of complement activation gene cluster linking the complement genes CR1, CR2, DAF, and C4bp. *J. Exp. Med.* 167:664.
- Bora, N.S., D.M. Lublin, B.V. Kumar, R.D. Hockett, V.M. Holers, and J.P. Atkinson. 1989. Structural gene for MCP maps to within 100 kb of the 3' end of the C3b/C4b receptor (CR1) gene. *J. Exp. Med.* 169:597.
- Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1987. *Current Protocols in Molecular Biology*. John Wiley and Sons, Inc., New York. pg. 1.13.5.
- Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
- Chomezynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156.
- Boyum, A. 1966. Isolation of mononuclear cells and granulocytes from human blood. *Scand. J. Clin. Lab. Invest.* 21(Suppl. 97):77.
- Fuhlbrigge, R.C., S.M. Fine, E.F. Unanue, and D.D. Chaplin. 1988. Expression of membrane interleukin 1 by fibroblasts transfected with murine pro-interleukin 1 α cDNA. *Proc. Natl. Acad. Sci. USA.* 85:5649.
- Purcell, D.F.J., I.F.C. McKenzie, D.M. Lublin, P.M. Johnson, J.P. Atkinson, T.J. Oglesby, and N.J. Deacon. 1990. The human cell-surface glycoproteins HuLy-m5, membrane cofactor protein (MCP) of the complement system, and trophoblast-leukocyte common (TLX) antigen are CD46. *Immunology.* 70:155.

34. Nickells, M.W., and J.P. Atkinson. 1990. Characterization of CR1- and MCP-like proteins of two primates. *J. Immunol.* 144:4262.
35. Cho, S.-W., T.J. Oglesby, B.-L. Hsi, E.M. Adams, and J.P. Atkinson. 1991. Characterization of three monoclonal antibodies to membrane cofactor protein (MCP) of the complement system and quantitation of MCP by radioassay. *Clin. Exp. Immunol.* 83:257.
36. Laemmli, U. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)* 227:680.
37. Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50:349.
38. Yamamoto, T., C.G. Davis, M.S. Brown, W.J. Schneider, M.L. Casey, J.L. Goldstein, and D.W. Russell. 1984. The human LDL receptor: a cysteine rich protein with multiple Alu sequences in its mRNA. *Cell.* 39:27.
39. Leonard, W.J., J.M. Depper, G.R. Crabtree, S. Rudikoff, J. Pumphrey, R.J. Robb, M. Kronke, P.B. Svetlik, N.J. Peffer, T.A. Waldmann, and W.C. Greene. 1984. Molecular cloning and expression of cDNAs for the human interleukin-2 receptor. *Nature (Lond.)* 311:626.
40. Nikaido, T., A. Shimizu, N. Ishida, H. Sabe, K. Teshigawara, M. Maeda, T. Uchiyama, J. Yodoi, and T. Honjo. 1984. Molecular cloning of cDNA encoding human interleukin-2 receptor. *Nature (Lond.)* 311:631.
41. Cosman, D., D.P. Cerretti, A. Larsen, L. Park, C. March, S. Dower, S. Gillis, and D. Urdal. 1984. Cloning, sequence and expression of human interleukin-2 receptor. *Nature (Lond.)* 312:768.
42. Caras, I.W., M.A. Davitz, L. Rhee, G. Weddell, D.W. Martin, Jr., and V. Nussenzweig. 1987. Cloning of decay-accelerating factor suggests novel use of splicing to generate two proteins. *Nature (Lond.)* 325:545.
43. Medof, M.E., D.M. Lublin, V.M. Holers, D.J. Ayers, R.R. Getty, J.F. Leykam, J.P. Atkinson, and M.L. Tykocinski. 1987. Cloning and characterization of cDNAs encoding the complete sequence of decay accelerating factor of human complement. *Proc. Natl. Acad. Sci. USA.* 84:2007.
44. Reddy, P., I. Caras, and M. Krieger. 1989. Effects of O-linked glycosylation on the cell surface expression and stability of decay-accelerating factor, a glycopospholipid-anchored membrane protein. *J. Biol. Chem.* 264:17329.
45. Kozarsky, K., D. Kingsley, and M. Krieger. 1988. Use of a mutant cell line to study the kinetics and function of O-linked glycosylation of low density lipoprotein receptors. *Proc. Natl. Acad. Sci. USA.* 85:4335.
46. Thomas, M.L. 1987. The leukocyte common antigen family. *Annu. Rev. Immunol.* 7:339.