*Proc. Natl. Acad. Sci. USA*
Vol. 80, pp. 264–267, January 1983
Immunology

# Cloning of a human complement component *C4* gene

(HLA class III antigen/cosmid library/synthetic oligonucleotide/cDNA library)

MICHAEL C. CARROLL AND RODNEY R. PORTER

Medical Research Council Immunochemistry Unit, Biochemistry Department, Oxford University, Oxford, United Kingdom

*Contributed by R. R. Porter, October 4, 1982*

**ABSTRACT** Six overlapping cosmid clones having an average insert size of 40 kilobase pairs were identified and isolated from a human genomic library by using a cDNA probe, Alu-7, specific for the amino acid sequence of C4d, a known region of the fourth component of human complement. Analysis of these genomic clones by restriction digestion and Southern blotting shows that all six probably contain the same complete *C4* gene. Nucleotide sequence comparison of the genomic clone Cos-A and the cDNA clone Alu-7 shows an identical sequence except for the presence of a 1,500-base-pair intron in the genomic sequence. The amino acid sequence predicted from the nucleotide sequence agrees with the known C4d region amino acid sequence with one exception.

In the major histocompatibility locus of man, HLA, there are genes coding for three classes of antigen. Class I HLA-A, -B, and -C are transplantation antigens, class II HLA-D are the human analogues of the mouse *I*-region-associated antigens, and class III are the complement components C2 and C4 and factor B. Class I and class II antigens show an exceptionally high degree of polymorphism (1); class III shows much less but many allelic forms of the complement components have been recognized, mainly by differences in charge. In C4, there is evidence from population studies of two loci with several alleles at each locus (2).

The genetics of C4 was clarified when O'Neill *et al.* (3, 4) showed that the erythrocyte antigens Chido and Rogers are fragments of C4 and these were identified as C4d (5). They are bound to the erythrocyte surface and their antigenic specificities correlate with electrophoretic variants of C4. Rogers correlates with C4A (or C4F) and Chido correlates with C4B (or C4S). The pattern of inheritance suggests that there are separate loci for C4A and C4B and that each have multiple allelic forms.

C4 is a protein of $M_r$ about 200,000 with three peptide chains, $\alpha$ ($M_r$, 95,000), $\beta$ ($M_r$, 75,000), and $\gamma$ ($M_r$, 30,000). C4d is a degradation fragment of $M_r$ about 40,000 coming from the middle of the $\alpha$ chain (6). C4 is, however, synthesized as a single peptide chain in mouse (7–9) and guinea pig (10, 11), and presumably in other species, that undergoes proteolytic cleavage on secretion to give the $\beta$, $\alpha$, and $\gamma$ chains in that order (12, 13). The C4d fragment corresponds therefore to the center section of pro-C4 and much of the polymorphism of C4 appears to derive from structural changes in this part of the molecule.

The molecular organization of *C4* genes has been investigated by isolation of the cDNA coding for sections of C4d and using this to identify a *C4* gene in a human cosmid library provided by F. G. Grosveld (National Institute for Medical Research, London).

## MATERIALS AND METHODS

**Isolation of RNA.** Fresh postmortem human liver was sliced and dropped into liquid nitrogen. RNA was extracted from homogenized tissue in 4 M guanidine thiocyanate/14% 2-mercaptoethanol (14) and precipitated by addition of 0.5 vol of ethanol and 0.025 vol of 1 M acetic acid with stirring for 20 min in a salt water/ice bath. After two extractions with 6 M guanidine·HCl, the precipitates were combined and dissolved in 10 mM Tris·HCl/1 mM EDTA/0.1% $NaDodSO_4$ and 0.1 vol of dimethyl sulfoxide was added. The mixture was heated to 65°C for 5 min and then the RNA was fractionated on a 15–30% sucrose gradient for 16 hr at 25,000 rpm in a Beckman SW28 rotor at 20°C. One-milliliter fractions were translated in a rabbit reticulocyte cell-free system (Amersham) with $[^{35}S]$methionine (600 Ci/mmol; 1 Ci = 37 GBq).

Translation of pro-C4 was determined by immunoprecipitation with C4-specific rabbit antiserum. Cell-free translation reaction mixtures (15) of 200 $\mu$l were diluted to a final vol of 500 $\mu$l with detergent buffer to a final concentration of 25 mM Tris·HCl, pH 7.5/150 mM NaCl/0.5% Nonidet P-40/methionine (1 mg/ml)/1 mM EDTA/10 mM diisopropyl fluorophosphate. Excess antibody and 5 $\mu$g of carrier C4 were added and mixtures were incubated at 4°C for 2 hr.

**Synthesis of DNA *In Vitro*.** Approximately 100 $\mu$g of 28S RNA was primed with 6 $\mu$g of oligo$(dT)_{12-18}$ and transcribed with 48 units of avian myeloma virus reverse transcriptase (J. W. Beard) in a 60-$\mu$l reaction mixture at 42°C for 90 min. The reaction was terminated by addition of 20 mM EDTA, and the mixture was extracted with phenol and then desalted on a 1-ml Sephadex G-100 column. RNA was degraded by dissolving the ethanol precipitate from the pooled peak fractions in 0.1 M NaOH/1 mM EDTA at 70°C for 20 min. The mixture was neutralized with 0.1 M HCl/1 M Tris·HCl, pH 8.0, and the RNA was precipitated with 2.5 vol of ethanol.

Double-stranded DNA was synthesized using conditions modified from Wickens *et al.* (16). After brief denaturing of single-stranded cDNA at 70°C for 3 min and quick cooling in ice water, 8 units of the Klenow subfragment of *Escherichia coli* DNA polymerase I (Boehringer Mannheim) was added (final vol, 40 $\mu$l) and the mixture was incubated at 25°C for 2 hr. The reaction was terminated by addition of 20 mM EDTA, the mixture was extracted with phenol, and RNA was precipitated with ethanol.

**Preparation of the Recombinant Library.** A human cDNA library containing $5 \times 10^5$ recombinants was prepared as follows. Approximately 0.8 $\mu$g of double-stranded DNA was digested with 10 units of *Alu* I at 37°C for 2 hr. After termination of the reaction with 20 mM EDTA and extraction of the mixture with phenol, the material was fractionated on a 1-ml Sephacryl 300 column in 200 mM NaCl/10 mM Tris·HCl/1 mM EDTA. The peak fractions were pooled and concentrated by ethanol precipitation.

The resultant fragments were ligated by using T4 DNA ligase to the plasmid vector pAT-153-PVU II-8 that had previously been cleaved at its unique *Pvu* II site and treated with calf intestinal phosphatase to minimize self-ligation. This ligation mixture was then used to transform competent *E. coli* K-12 strain

Abbreviations: bp, base pair(s); kbp, kilobase pair(s); C2, C4, etc., second, fourth, etc., components of human complement.

MC 1061 (17). An aliquot of the transformed cells was removed for analysis and the remainder was amplified in 2 liters of Luria broth containing ampicillin at 100 $\mu$g/ml for 8 hr. Transformants were then centrifuged and suspended in fresh Luria broth/ampicillin/15% glycerol, and aliquots were stored at −70°C.

**Synthesis of the Oligonucleotide Probe.** A mixture of 16 different 14-nucleotide-long oligonucleotides was synthesized (18) based on the known C4d amino acid sequence Met-Ala-Gln-Glu-Thr, using the solid-phase phosphotriester technique. The oligonucleotide chosen

$$5'\text{-G-T-}\begin{matrix}\text{T}\\\text{C}\end{matrix}\text{-T-C-}\begin{matrix}\text{T}\\\text{C}\end{matrix}\text{-T-G-}\begin{matrix}\text{A}\\\text{T}\\\text{C}\\\text{G}\end{matrix}\text{-G-C-C-A-T-3}'$$

was complementary to the mRNA.

**Screening the cDNA Library.** Approximately 10,000 colonies were plated on ampicillin-containing bacteriological plates, transferred to Whatman 541 filter paper by blotting, and processed for hybridization as described by Gergen *et al.* (19). Filters were incubated at 55°C in 0.9 M NaCl/90 mM Tris·HCl, pH 7.5/6 mM EDTA/0.5% Nonidet P-40 containing boiled sonicated salmon sperm DNA at 100 $\mu$g/ml for 3 to 4 hr and then hybridized for 16 hr at 32°C in the same solution with about $2 \times 10^5$ dpm (0.5 ng/ml) of 5′-$^{32}$P-labeled oligonucleotide probe (20). Filters were washed with 0.9 M NaCl/0.09 M sodium citrate at 20°C for 1 hr and then four times at 34°C over a 2-hr period, air dried, and autoradiographed at −70°C.

**Isolation of the C4 Genomic Clones.** Approximately 200,000 colonies of the human genomic library constructed with the pTM cosmid vector and partially *Mbo* I-digested placental DNA by F. G. Grosveld were plated onto 16 nitrocellulose filters (14-cm Millipore HAWP, 0.45 $\mu$M) on nutrient agar plates with ampicillin. Replicas were prepared and processed for colony hybridization according to Grosveld *et al.* (21). Prehybridization incubation and hybridization were carried out at 42°C in 50% formamide solution (22).

The 301-base-pair (bp) C4 cDNA probe, Alu-7, used for screening was labeled to a specific activity of approximately $10^8$ dpm/$\mu$g of DNA by using the Amersham nick-translation kit (23). The hybridization mixture contained 2–5 × $10^5$ dpm/ml.

After hybridization, the filters were washed with four changes of 0.3 M NaCl/0.03 M sodium citrate at room temperature, with four changes of 0.15 M NaCl/0.015 M sodium citrate at 65°C over 2 hr, and finally with two changes of 0.03 M NaCl/3 mM sodium citrate at 65°C for 30 min each. Filters were autoradiographed at −70°C.

**Isolation of DNA.** Cosmid and plasmid DNAs were extracted from bacterial colonies by using the alkaline NaDodSO$_4$ method (24).

**Partial Characterization of a C4 Cosmid Clone.** The cosmid Cos-A containing the *C4* gene was partially mapped by a series of single and double restriction enzyme digestions, under manufacturers conditions, and separation of the fragments by electrophoresis on agarose gels in 90 mM Tris borate, pH 8.3/1 mM EDTA. DNA fragments were blotted onto nitrocellulose filters (25) and hybridized with either the C4 cDNA probe Alu-7 or the pTM vector.

## RESULTS

**Isolation of C4 mRNA from Human Liver.** Approximately 2 mg of total RNA per gram of frozen human liver was recovered after a single extraction with 4 M guanidine thiocyanate followed by two extractions with 6 M guanidine·HCl.

Cell-free translation in a rabbit reticulocyte lysate using RNA fractionated on sucrose gradients followed by immunoprecipitation and analysis on NaDodSO$_4$ polyacrylamide gels (Fig. 1) showed that the 28S RNA fraction was enriched in C4 mRNA. The minor bands seen in track 1 are probably breakdown products of C4. No bands are seen when the irrelevant ovalbumin–antiovalbumin is precipitated in the translation mixture. As also occurs in the mouse (7–9) and guinea pig (10–12), pro-C4 was translated as a single polypeptide chain of $M_r$ approximately 180,000. The intensity of the $M_r$ 180,000 band in the total RNA translation mixture may be accounted for by the observation that immunoprecipitation with antisera to complement proteins C3 (Fig. 1, track 2) and C5 (not shown) gave similar sized products.

**Identification and Isolation of C4 cDNA Clones.** The C4d region amino acid sequence Met-Ala-Gln-Glu-Thr (D. N. Chakravarti, R. D. Campbell, and J. Gagnon, personal communication) was chosen for preparing a mixture of 16 synthetic 14-base-long oligonucleotides that were complementary to the mRNA sequence. This mixture was end labeled with [$\gamma$-$^{32}$P]ATP and used to screen approximately 10,000 cDNA clones

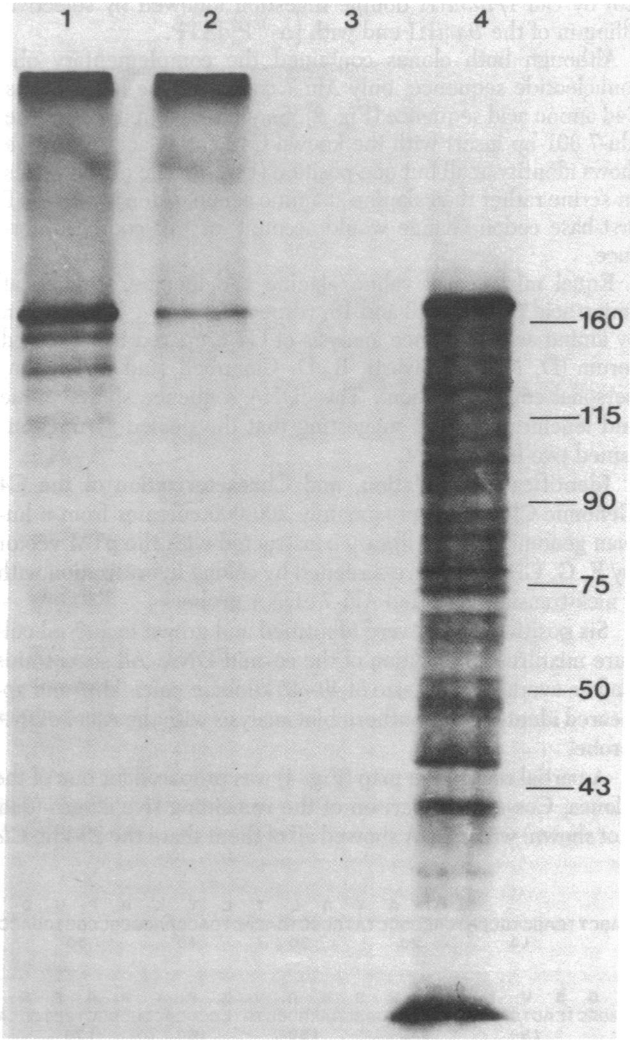

FIG. 1. The 28S fraction of human liver RNA contains the message for C3 and C4. [$^{35}$S]Methionine-labeled proteins were translated in a cell-free rabbit reticulocyte lysate from the 28S fraction of human liver RNA, precipitated with specific antisera, and analyzed on NaDodSO$_4$/6% polyacrylamide gels. Lanes: 1, immunoprecipitation with rabbit anti-human C4; 2, immunoprecipitation with rabbit anti-human C3; 3, immunoprecipitation in the presence of 5 $\mu$g of ovalbumin and an equivalent amount of antiovalbumin; 4, total translated protein. Numbers on the right represent $M_r \times 10^{-3}$.

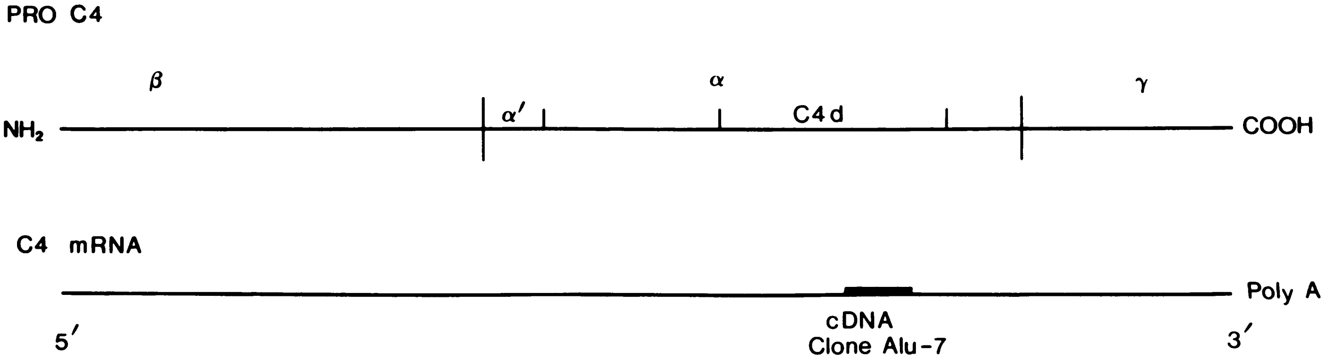

FIG. 2. Comparison of the pro-C4 molecule, the proposed C4 mRNA, and Alu-7 cDNA. The 301-bp insert cDNA of the Alu-7 clone isolated from a human liver cDNA library is complementary to a region of the C4 mRNA that codes for the amino acid sequence of a known C4d region.

by colony hybridization. Two positive clones, Alu-7 and -16, were identified over background after a final wash at 41°C. The cloned cDNA inserts were excised for Maxam–Gilbert analysis (26) by *Cla* I/*Bam*HI double digestion followed by selective filling-in of the *Bam*HI end with [$\alpha$-$^{32}$P]dATP.

Although both clones contained the complementary oligonucleotide sequence, only Alu-7 coded for the homologous C4d amino acid sequence (Fig. 2). Sequence comparison of the Alu-7 301-bp insert with the known C4d amino acid sequence shows identity at all but one position (Fig. 3). The cDNA codes for serine rather than alanine at amino acid position 92. A G→T first-base codon change would account for this coding difference.

Equal mixtures of valine/alanine and leucine/arginine at amino acid positions 13 and 16, respectively, have been shown by amino acid sequence analysis of C4 prepared from pooled serum (D. N. Chakravarti, R. D. Campbell, and J. Gagnon, personal communication). The cDNA sequence shows valine and leucine together, suggesting that the pooled serum contained two forms of C4.

**Identification, Isolation, and Characterization of the C4 Genomic Clones.** Approximately 200,000 colonies from a human genomic cosmid library constructed with the pTM vector by F. G. Grosveld were screened by colony hybridization with a nick-translated labeled Alu-7 cDNA probe.

Six positive clones were identified and grown in 100-ml culture mixtures for isolation of the cosmid DNA. All six cosmids had an average insert size of 40–45 kilobase pairs (kbp) and appeared identical by Southern blot analysis with the Alu-7 cDNA probe.

A partial restriction map (Fig. 4) was prepared for one of the clones, Cos-A. Comparison of the remaining five clones (data not shown) with Cos-A showed all of them share the 25-kbp *Cla* I fragment that hybridizes to the cDNA probe but some extend as much as 8 kbp further on the 5′ end.

Southern blot analysis of Cos-A genomic DNA with the Alu-7 cDNA probe showed that there are additional restriction sites not seen in the cDNA sequence. Further digestion with *Alu* I suggested the presence of at least one intron of approximately 1,500 bp, occurring after nucleotide position 100 (Fig. 4). To compare the genomic sequence in this region with the cDNA, DNA fragments on both the 5′ and the 3′ ends of the intron were subcloned and the nucleotide sequences were determined.

The nucleotide sequence of the genomic DNA was identical to that of the cDNA up to position 107 (Fig. 3), where the consensus dinucleotide splice signal G-T appeared, and there was also a stop codon several nucleotides downstream. The nucleotide sequence on the 3′ end of the intron showed the consensus splice signal A-G followed by the cDNA coding sequence beginning at nucleotide 108 and continuing with the identical cDNA sequence.

## DISCUSSION

A cDNA library was prepared from human liver mRNA that had been enriched for higher molecular weight fractions. One cDNA clone, Alu-7, was found to hybridize with an oligonucleotide probe synthesized to correspond to a known pentapeptide sequence in C4d. The 301-bp nucleotide sequence of Alu-7 coded for a 100-amino acid sequence identical to the known peptide sequence except in one position, where a serine replaced an alanine. Polymorphism had already been recognized in the amino acid sequence, where both alanine and valine were found at position 13 and arginine and leucine were found at position 15. The cDNA sequence coded for valine and leucine in these positions. It is most likely therefore that the Ser→ Ala (position 92) change represents an allelic or locus variation that

```
              I  T  A  Y  A  L  T  L  T  K  A  P  V  D  L  L  G  V  A  H  N  N  L  M  A  M  A  Q  E  T  G  D  N  L  Y  W
AGCTTCAGCTGCCATCACGGCCTATGCCCTGACACTGACCAAGGCGCCGGTGGACCTGCTCGGTGTTGCCCACAACAACCTCATGGCAATGGCCCAGGAGACTGGAGATAACCTGTACTG
         10        20        30        40        50        60        70        80        90       100       110       120

 G  S  V  T  G  S  Q  S  N  A  V  S  P  T  P  A  P  R  N  P  S  D  P  M  P  Q  A  P  A  L  W  I  E  T  T  A  Y  A  L  L
GGGCTCAGTCACTGGTTCTCAGAGCAACGCCGTGTCGCCCACCCCGGCTCCTCGCAACCCATCCGACCCCATGCCCCAGGCCCCAGCCCTGTGGATTGAAACCACAGCCTACGCCCTGCT
         130       140       150       160       170       180       190       200       210       220       230       240

 H  L  L  L  H  E  G  K  A  E  M  A  D  Q  A  S  A  W  L  T  R  Q  G  S
GCACCTCCTGCTTCACGAGGGCAAAGCAGAGATGGCAGACCAGGCTTCGGCCTGGCTCACCCGTCAGGGCAGCTGCA
         250       260       270       280       290       300       310
```

FIG. 3. Nucleotide sequence of the Alu-7 cDNA insert. Insert cDNA was excised from the plasmid vector pAT-153-PVU II-8 by digestion with *Bam*HI/*Cla* I and the 3′ end was selectively labeled with [$\alpha$-$^{32}$P]dATP. Labeled cDNA fragments were isolated from 4% native polyacrylamide gels and the nucleotide sequence was determined by the Maxam–Gilbert technique (26). The sequence complementary to the oligonucleotide probe is underlined. Amino acids: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; D, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

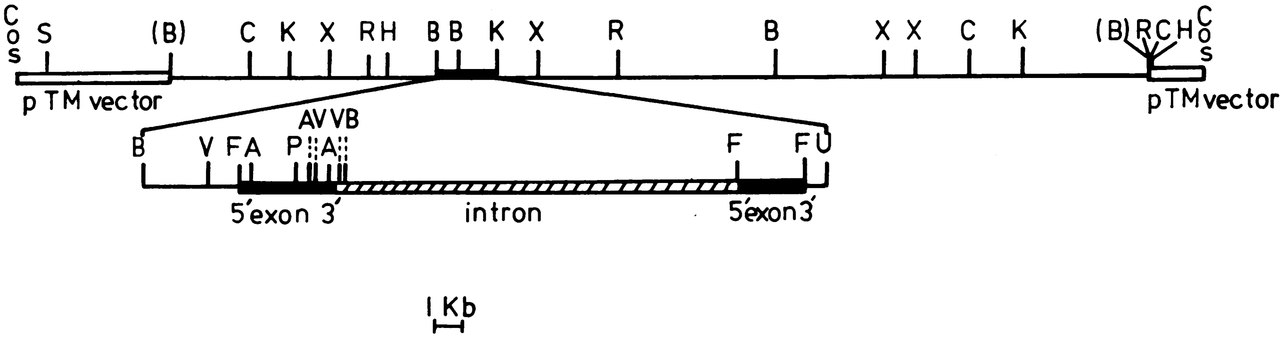

FIG. 4. Partial restriction map of cosmid clone Cos-A. A partial restriction map was prepared by single and double digestion of Cos-A DNA followed by separation of the DNA fragments on 0.6% and 0.8% agarose gels. For Southern blot analysis, DNA fragments were transferred to nitrocellulose and hybridized with either nick-translated [$\alpha$-$^{32}$P]dATP-labeled Alu-7 cDNA or the pTM vector probe. The C4d region coding sequence is approximately 10 and 25 kbp from the 5′ and 3′ fragment ends, respectively. The expanded region of the map shows the 1,500-bp intron with the C4d coding sequence on the 5′ and 3′ ends. Restriction enzyme sites used: A, *Hae* II; B, *Bam*HI; C, *Cla* I; F, *Hin*fI; H, *Hin*dIII; K, *Kpn* I; P, *Pvu* II; R, *Eco*RI; S, *Sal* I; U, *Alu* I; V, *Ava* II; X, *Xho* I.

is not detected in the amino acid sequence because of the relatively poor yield of the phenylthiohydantoin derivative of serine in the Edman degradation.

Using the Alu-7 cDNA as a probe, we found that six overlapping cosmids hybridized in a human genomic library constructed by F. G. Grosveld. All six have been investigated by Southern blot analysis and found to contain an identical 25-kbp *Cla* I fragment. Fifteen different restriction enzymes were used and this suggests that, if more than one form of the *C4* gene is present, their sequences are very similar.

One of the cosmids, Cos-A, probably contains the whole *C4* gene; mapping showed that the section hybridizing with the probe is not less than 10 and 25 kbp from the 5′ and 3′ vector ends, respectively. The probe corresponds to a central part of the precursor C4 molecule with approximately 1,100 and 700 amino acid residues on the amino- and carboxyl-terminal sides, respectively. Even if several introns are present, it is likely that this cosmid contains the complete gene. However, further characterization will be required to prove this.

The nucleotide sequences of the subcloned Cos-A restriction fragments show identity with the cDNA sequence, including positions corresponding to the amino acids serine, valine, and alanine at positions 92, 13, and 15, respectively. As mentioned above, this sequence is likely to be that of one genetic variant of C4. A noncoding nucleotide sequence was, however, found beginning at base 108 and was approximately 1,500 bp long with the consensus dinucleotide splice signals G-T and A-G at the 5′ and 3′ ends, suggesting that it is an intron.

Genetic evidence suggests the two *C4* loci are tightly linked and may be tandemly arranged on the chromosome (2). An end fragment of a second *C4* gene could be present in Cos-A but undetected because the cDNA probe recognizes only a region near the middle of the gene. Further analysis with 5′- and 3′-end probes will be necessary to detect the presence of an additional *C4* gene and more restriction enzyme digestion should identify the genes coding for the different variants of C4.

We thank Professor G. G. Brownlee for his exceptional help in providing laboratory facilities and essential advice throughout this work. We also thank Mr. Keith Gould for technical assistance in synthesis of oligonucleotide probes; Dr. R. D. Campbell for providing the rabbit anti-human C4 serum; Dr. R. Sim for rabbit anti-human C4 and C5 sera; and Drs. F. E. Baralle, K. H. Choo, D. H. Bishop, I. Jones, and W. C. Tan and Mr. J. G. Rees for advice and helpful discussion. M.C.C. is a recipient of an American Arthritis Foundation Fellowship.

1. Barnstaple, C. J., Jones, E. A. & Bodmer, W. F. (1979) in *Defense and Recognition IIB, Cellular Aspects*, MTP Int. Rev. Sci. Ser. Biochem., ed. Lennox, E. S. (University Park Press, Baltimore), Vol. 22, pp. 151–224.
2. Raum, D., Donaldson, V. H., Rosen, F. S. & Alper, C. A. (1980) *Curr. Top. Haematol.* 3, 111–174.
3. O'Neil, G. T., Yang, S. Y., Tegoli, J., Berger, R. & Dupont, B. (1978) *Nature (London)* 273, 668–670.
4. O'Neil, G. J., Yang, S. Y. & Dupont, B. (1978) *Proc. Natl. Acad. Sci. USA* 75, 5165–5169.
5. Tilley, C. A., Romans, D. G. & Crookson, M. C. (1978) *Nature (London)* 276, 713–715.
6. Reid, K. B. M. & Porter, R. R. (1981) *Annu. Rev. Biochem.* 50, 433–464.
7. Roos, M. H., Atkinson, T. P. & Shreffler, D. C. (1978) *J. Immunol.* 121, 1106–1115.
8. Parker, K. L., Roos, M. H. & Shreffler, D. C. (1979) *Proc. Natl. Acad. Sci. USA* 76, 5853–5857.
9. Fey, G., Odink, K. & Chapuis, R. M. (1980) *Eur. J. Immunol.* 10, 75–82.
10. Hall, R. E. & Colten, H. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 1707–1710.
11. Goldberger, G., Abraham, G. N., Williams, J. & Colten, H. R. (1980) *J. Biol. Chem.* 255, 7071–7074.
12. Goldberger, G. & Colten, H. R. (1980) *Nature (London)* 286, 514–516.
13. Karp, D. R., Parker, K. L., Shreffler, D. C. & Capra, J. D. (1981) *J. Immunol.* 126, 2060–2061.
14. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* 18, 5294–5299.
15. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* 67, 247–256.
16. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* 253, 2483–2495.
17. Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* 138, 179–207.
18. Duckworth, M. L., Gait, M. J., Goelet, P., Hong, G. F., Singh, M. & Titmas, R. C. (1981) *Nucleic Acids Res.* 9, 1691–1706.
19. Gergen, J. P., Stern, R. H. & Websink, P. C. (1979) *Nucleic Acids Res.* 7, 2115–2136.
20. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* 9, 879–894.
21. Grosveld, F. G., Dahl, H. M., Boer, E. D. & Flavell, R. A. (1981) *Gene* 13, 227–237.
22. Bernards, R. & Flavell, R. A. (1980) *Nucleic Acids Res.* 8, 1521–1534.
23. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* 113, 237–251.
24. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* 7, 1521–1534.
25. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503–517.
26. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560–564.