Biochem. J. (1985) 231, 729–735 (Printed in Great Britain)

Molecular cloning and characterization of the complementary DNA and gene coding for the B-chain of subcomponent C1q of the human complement system

Kenneth B. M. REID
M. R. C. Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Plasmid clones containing cDNA coding for the B-chain of human C1q were isolated from a liver cDNA library. The longest cDNA insert isolated contained all the coding sequence for amino acid residues B1 to B226 plus a 3′ non-translated region of 264 nucleotides that extended into the poly(A) tail, thus accounting for 950 nucleotides of the mRNA. The B-chain mRNA was estimated by Northern-blot analysis to be 1.46 kb (kilobases) long, which indicated that approx. 500 bases were not accounted for in the cDNA clone. A cosmid clone containing the C1q-B chain gene was isolated from a human genomic DNA library. The precise 5′ limit of gene was not established, but from the data available it appears that the gene is approx. 2.6 kb long. The coding sequence for residues B1 to B226 in the gene is interrupted by one intron, of 1.1 kb, which is located within the codon coding for glycine at position B36. This glycine residue is located in the middle of the triple-helical regions found in C1q at exactly the position where there is an unusual structural feature, i.e. a bend in each of the helical regions brought about by the interruption of the Gly-Xaa-Yaa repeating triplet sequences in the A- and C-chains and the presence of an 'extra' triplet in the B-chain. Nucleotide sequencing of the 5′ end of the gene indicates the presence of a predominantly hydrophobic stretch of 29 amino acids, immediately before residue B1, which could serve as a signal peptide.

INTRODUCTION

Normal human serum C1q, of M_r 460000, contains 18 polypeptide chains (6A, 6B, 6C), each 226 amino acids long and each containing a collagen-like region (of 81 amino acids), located near the *N*-terminus, and a *C*-terminal globular region (of approx. 136 amino acids) [for review see Reid (1983)]. The role of C1q is to interact with the Fc regions of aggregated immunoglobulin, or some other activator, via its six globular 'heads' (each 'head' being formed from the *C*-terminal regions of an A-, a B- and a C-chain). This interaction brings about the activation of the proenzyme C1r to $\overline{C1r}$ in the C1q–$C1r_2$–$C1s_2$ complex, thus allowing the $\overline{C1r}$ to activate proenzyme C1s. The $C1r_2$–$C1s_2$ complex is considered to be located on the triple-helical 'connecting strands' and 'stalk' of C1q, which are formed from the *N*-terminal collagen-like regions present in the chains of C1q. After activation the $\overline{C1}$ complex comes under the control of $\overline{C1}$-inhibitor, which rapidly removes $\overline{C1r}$ and $\overline{C1s}$, leaving the collagen-like regions of C1q free to interact with cell-surface receptors and other molecules such as fibronectin. There appear to be several different forms of C1q, since a variety of molecules that are antigenically related to normal serum C1q, some having haemolytic activity, have been reported, but these molecules differ with respect to their M_r on comparison with each other and normal C1q. The different C1q-like molecules are synthesized in fibroblasts (Reid & Solomon, 1977) and epithelial cells (Morris *et al.*, 1978), and are present in normal colostrom (Yonemasu *et al.*, 1979) and the sera of patients suffering from immune-complex-related renal disease as a consequence of C1q dysfunction (Thompson *et al.*, 1980; Chapius *et al.*, 1982; Hannema *et al.*, 1982; Reid & Thompson, 1983). Macrophages also synthesize C1q, and it has been suggested that C1q may serve as an Fc receptor on guinea-pig peritoneal macrophages (Loos, 1983). Therefore although the major site of sythesis of human serum C1q is considered to be the columnar cells of the small intestine (Colten *et al.*, 1968; Colten, 1976), it is clear that other cells can synthesize C1q or C1q-like molecules. A study, at the DNA level, of these apparent structural variants of C1q has been initiated by the isolation and preliminary characterization of a cDNA clone for the B-chain of normal human C1q (Reid *et al.*, 1984). The present paper gives the cDNA sequence corresponding to residues 1–226 of the B-chain plus the entire 3′ non-translated region, and also describes the isolation and characterization of a cosmid clone containing the B-chain gene.

MATERIALS AND METHODS

Synthesis and radiolabelling of oligonucleotide probes, preparation of the recombinant cDNA libraries and isolation of cDNA clones

All these procedures were performed as described in Reid *et al.* (1984).

Isolation of a genomic clone

Approx. 200000 colonies of a cosmid library were plated on to nitrocellulose filters, replicas prepared and

Abbreviation used: kb, kilobase.

processed for colony hybridization as described by Grosveld *et al.* (1981). The cosmid library was constructed by Dr. A. Palsdottir of this Unit by using human white-cell DNA, partially digested with endonuclease *Mbo*1, the cosmid pTCF and employing the procedures described by Grosveld *et al.* (1982). The filters were prehybridized and hybridized at 42 °C in a buffer containing 50% (v/v) formamide (Bernards & Flavell, 1980). An 84 base-pair insert (coding for residues B199–B226) carried in the plasmid vector pAT 153/*Pvu*II/8 was cut out by digestion with *Bam*H1/*Cla*1 and purified by electrophoresis on a 4%-(w/v)-polyacrylamide gel (Reid *et al.*, 1984). This 84-base pair C1q cDNA probe was used for the screening after it had been labelled to a specific radioactivity of approx. 10^8 c. p. m./μg of DNA by nick translation (Rigby *et al.*, 1977). After hybridization for 20 h, the filters were washed with two changes of 0.3 M-NaCl/0.03 M-sodium citrate at room temperature, with two changes of 0.15 M-NaCl/0.015 M-sodium citrate at 65 °C over 1 h, and finally with two changes of 0.03 M-NaCl/3 mM-sodium citrate at 65 °C over 1 h. Filters were autoradiographed at − 70 °C for 40 h.

Preparation and analysis of cloned DNA

Plasmid DNA (Birnboim & Doty, 1979) and cosmid DNA (Grosveld *et al.*, 1981) were extracted from bacterial colonies by standard procedures. Large-scale preparations were further purified by isopycnic centrifugation (Radloff *et al.*, 1967; Maniatis *et al.*, 1982). Partial characterization of the cosmid was achieved by a series of single and double digests with restriction enzymes and by separation of the fragments on agarose gels. The DNA fragments were blotted on to nitrocellulose filters as described by Wahl *et al.* (1979) and hybridized with either a radioactively labelled section of C1q B chain cDNA or the pTCF cosmid vector as a probe.

Restriction fragments from the cosmid DNA found to hybridize with 5′ or 3′ fragments of the cDNA probe were subcloned into the pAT 153/*Pvu*II/8 plasmid vector (Anson *et al.*, 1984) as follows. Restriction-enzyme digests of the cosmid clone (using *Ava*II, *Hin*f1, *Stu*1, or *Taq*1) were each ligated into the *Pvu*II site of the vector (after repair of the digested DNA with DNA polymerase 1 Klenow fragment in the case of the *Ava*II, *Hin*f1 and *Taq*1 digests) and clones corresponding to 5′ or 3′ ends of the C1q gene detected by colony hybridization with the cDNA probes. DNA sequence analysis was carried out as described by Maxam & Gilbert (1980), and preparation of DNA fragments for sequencing or nick translation was carried out as recommended in Maniatis *et al.* (1982). The cloned DNA inserts could be conveniently excised from the pAT 153/*Pvu*II/8 plasmid by digestion with *Bam*H1/*Hind*III, *Bam*H1/*Cla*1 or *Bam*H1/*Eco*R1.

Northern-blot analysis

RNA was extracted from human liver by the guanidine thiocyanate method of Chirgwin *et al.* (1979). Human liver RNA either fractionated by sucrose-gradient centrifugation or chromatographed on an oligo(dT)–cellulose column were designated as 28S, 18S-1 (from the leading shoulder of the main 18S peak), 18S-II (from the trailing shoulder of the 18S peak) and poly(A)$^+$. Approx. 20 μg of each RNA sample was denatured in 50% (v/v) formamide/6% (v/v) formaldehyde at 60 °C for 15 min. Radioactively labelled *Hind* III fragments of bacteriophage λ were denatured in the same manner as the RNA and used as size markers. The RNA and marker DNA samples were electrophoresed in a 1% (w/v) agarose/6% (v/v) formaldehyde gel with 40 mM-Mops (pH 7.0)/10 mM-sodium acetate/1 mM-EDTA (Lehrach *et al.*, 1977). The gel was washed consecutively in water, 50 mM-NaOH, 0.1 M-Tris/HCl, pH 7.4, and 1 × SSC (0.15 M-NaCl/0.015 M-sodium citrate). The RNA was blotted overnight on to a nitrocellulose filter. After baking at 80 °C for 3 h the filter was hybridized as described for the use of the nick-translated probes in the screening of the genomic DNA library.

RESULTS AND DISCUSSION

Characterization of the C1q B-chain cDNA clones

Several cDNA clones, each containing an 84-base-pair insert, were isolated as previously described by screening a cDNA fragment library from human liver with a synthetic oligonucleotide mixture (Reid *et al.*, 1984). The cDNA insert from these clones was found to code for the *C*-terminal 28 amino acids of the B-chain (B198–B226) and the use of this cDNA allowed the isolation and partial characterization of longer cDNA clones that contained all the coding sequence for residues B1–B226 (Reid *et al.*, 1984). These clones have now been completely characterized and the sequence of 950 nucleotides of cDNA, along with the derived amino acid sequence, is shown in Fig. 1. The derived amino acid sequence is consistent with the published amino acid sequence (Reid *et al.*, 1982), except at position B73, where the derived sequence predicts a glycine residue (which is correct), whereas the published sequence, determined by standard protein-sequencing techniques, gives proline at that position. This difference has thus now been reconciled, since it was due to a misinterpretation of the protein-sequence data (Reid *et al.*, 1982) rather than to the presence of two types of C1q B-chain in human serum. It has now been determined that residue B73 is glycine by studies at the amino acid level (K. B. M. Reid, unpublished work). The cDNA sequence data show that the *N*-terminal amino acid is a glutamine residue, which is consistent with the finding of a pyroglutamate residue in this position by protein-sequencing studies (Reid & Thompson, 1978). The *N*-terminal glutamine residue at position B1 is preceded by the sequence -Gln-Ala-, as indicated from the derived amino acid sequence. Therefore, to obtain the mature serum form of the C1q B-chain, an Ala–Gln bond is split by a processing enzyme and the resulting *N*-terminal glutamine residue must then cyclize to give a 'blocked' *N*-terminal residue. There is a termination codon immediately after the codon for alanine at B226, showing that any processing of a larger form of the B-chain must take place at the *N*-terminal end of the chain. In the 3′ non-translated region there is an AATAAA sequence element at 815–820 which is not used (Fig. 1). The poly(A) tail appears to start at nucleotide 936, and it is possible that the sequence element AATGAA, rather than the more common sequence AATAAA (Proudfoot & Brownlee, 1976), may be involved in the signal for polyadenylation, although the AATGAA sequence has been reported to prevent mRNA 3′-end formation (Wickens & Stephen, 1984).

```
        Q  L  S  C  T  G  P  P  A  I  P  G  I  P  G  I  P  G  I  P    20
CCAGGCCCAGCTCAGCTGCACCGGGCCCCCAGCCATCCCTGGCATCCCGGGTATCCCTGGGACACC    66

  G  P  D  G  Q  P  G  T  P  G  I  K  G  E  K  G  L  P  G  L  A  G    42
TGGCCCCGATGGCCAACCTGGGACCCCAGGGATAAAAGGAGAGAAAGGGCTTCCAGGGCTGGCTGG   132

  D  H  G  E  F  G  E  K  G  D  P  G  I  P  G  N  P  G  K  V  G  P    64
AGACCATGGTGAGTTCGGAGAGAAGGGAGACCCAGGGATTCCTGGGAATCCAGGAAAAGTCGGCCC   198

  K  G  P  M  G  P  K  G  G  P  G  A  P  G  A  P  G  P  K  G  E  S    86
CAAGGGCCCCATGGGCCCTAAAGGTGGCCCAGGGGCCCCTGGAGCCCCAGGCCCCAAAGGTGAATC   264

  G  D  Y  K  A  T  Q  K  I  A  F  S  A  T  R  T  I  N  V  P  L  R   108
GGGAGACTACAAGGCCACCCAGAAAATCGCCTTCTCTGCCACAAGAACCATCAACGTCCCCCTGCG   330

  R  D  Q  T  I  R  F  D  H  V  I  T  N  M  N  N  N  Y  E  P  R  S   130
CCGGGACCAGACCATCCGCTTCGACCACGTGATCACCAACATGAACAACAATTATGAGCCCCGCAG   396

  G  K  F  T  C  K  V  P  G  L  Y  Y  F  T  Y  H  A  S  S  R  G  N   152
TGGCAAGTTCACCTGCAAGGTGCCCGGTCTCTACTACTTCACCTACCACGCCAGCTCTCGAGGGAA   462

  L  C  V  N  L  M  R  G  R  E  R  A  Q  K  V  V  T  F  C  D  Y  A   174
CCTGTGCGTGAACCTCATGCGTGGCCGGGAGCGTGCACAGAAGGTGGTCACCTTCTGTGACTATGC   528

  Y  N  T  F  Q  V  T  T  G  G  M  V  L  K  L  E  Q  G  E  N  V  F   196
CTACAACACCTTCCAGGTCACCACCGGTGGCATGGTCCTCAAGCTGGAGCAGGGGGAGAACGTCTT   594

  L  Q  A  T  D  K  N  S  L  L  G  M  E  G  A  N  S  I  F  S  G  F   218
CCTGCAGGCCACCGACAAGAACTCACTACTGGGCATGGAGGGTGCCAACAGCATCTTTTCCGGGTT   660

  L  L  F  P  D  M  E  A  *                                          226
CCTGCTCTTTCCAGATATGGAGGCCTGACCTGTGGGCTGCTTCACATCCACCCCGGCTCCCCCTGC   726

CAGCAACGCTCACTCTACCCCCAACACCACCCCTTGCCCAGCCAATGGACACAGTAGGGCTTGGTG   792

AATGCTGCTGAGTGAATGAGTAAATAAACTCTTCAAGGCCAAGGAACAGTGGTCTAATTCAACTCT   858

GTGTCCCAGCACTGGCACACCAGAAGTGCCATGCTCAGAAATGTTGGTTACATGAATGAATGAACC   924

ATGAATGAATGAAAAAAAAAAAAAA                                           950
```

Fig. 1. cDNA sequence coding for the B-chain of C1q

The sequence of 950 nucleotides of the cloned cDNA, which includes the coding sequence for residues 1–226 of the B-chain of human C1q, is shown. Nucleotide 8 is the first base of the codon for the *N*-terminal amino acid of the B-chain [i.e. a glutamine residue which is found as pyroglutamate in the mature protein (Reid & Thompson, 1978)]. The poly(A) tail appears to start at nucleotide 936, which implies that the sequence AATGAA may be involved in the signal for polyadenylation. The amino acid sequence derived from the cDNA sequence is identical with that obtained by protein sequencing. It is estimated that the C1q B-chain mRNA is approx. 500 bases longer than the cDNA sequence shown in this Figure.

By Northern-blot analysis the C1q B-chain mRNA was estimated to be 1.46 kb long, and most of the message was found in the 18S-II fraction (Fig. 2). The weak signal from the poly(A)$^+$ track indicates that the C1q mRNA may be at a low level in the liver, since hybridization of the same filter with radioactive cDNA probes for factor B or C4b-binding protein gave strong signals in the poly(A)$^+$ track. The molar concentrations of all three proteins in serum are similar, therefore the Northern-blot analysis shows that whereas the liver may be the major site of synthesis of factor B and C4b-binding protein, this is probably not the case for the B-chain of C1q. Knowing that the mRNA for the B-chain is 1.46 kb allows an estimate to be made of the length of the 5′ nucleotide sequence before the codon for B1. If it is assumed that the average length of a poly(A) tail is 100–150 bases (Bawerman, 1976), then this leaves approx. 360 nucleotides to account for the 5′ region before the codon for B1. Since the genomic sequence indicates that the residue B1 may be preceded by a predominantly hydrophobic string of up to 29 amino acids, most of which could act as a signal peptide (see below and Fig. 3), then it may be that most of the 5′ region of the mRNA will be 5′ non-translated nucleotide sequence. However, in view of the finding of C1q-like chains having a significantly higher M_r than normal in studies *in vitro* on fibroblasts

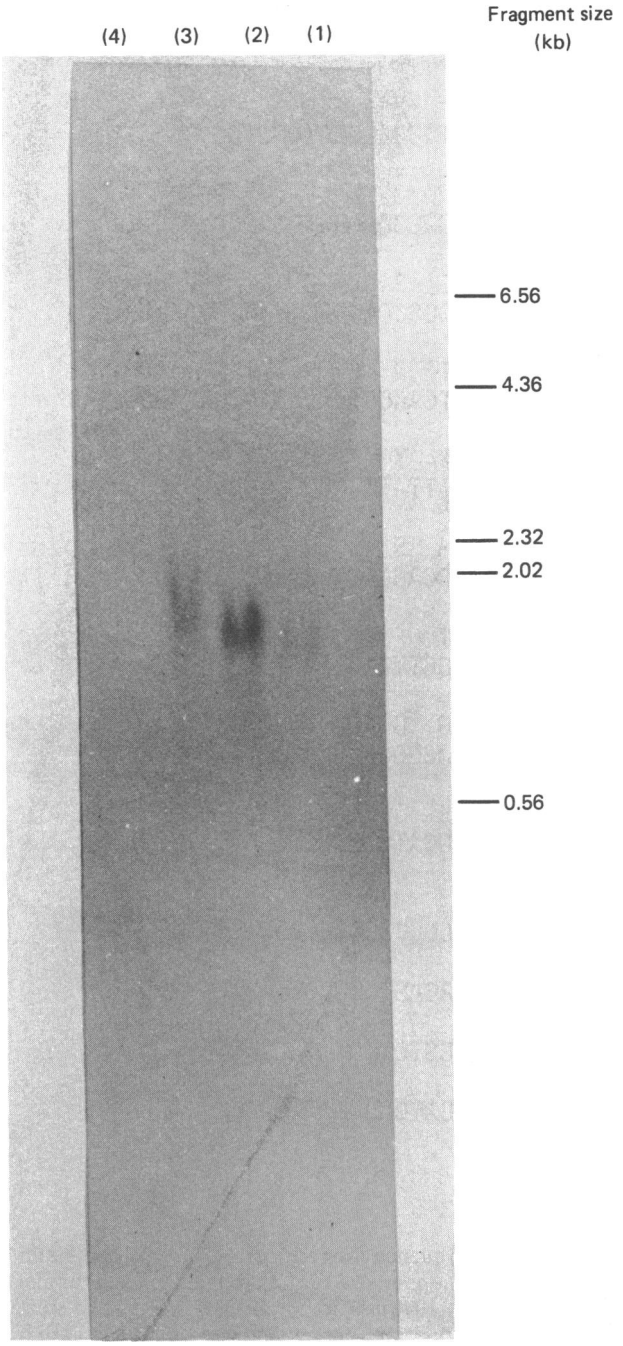

Fig. 2. Northern-blot analysis of the C1q B-chain mRNA

Four human liver RNA samples [(1) total poly(A)$^+$; (2) 18S-II; (3) 18S-1; and (4) 28S] were electrophoresed in an agarose gel, transferred to a nitrocellulose filter and hybridized with a nick-translated 84-base-pair probe as described in the text. Marker DNA fragments were derived from a *Hin*dIII digest of bacteriophage λ DNA and treated in the same way as the RNA samples.

(apparent M_r 46000 as against 27000; Reid & Solomon, 1977), the possibility that there may be a precursor, or pro-form, of the B-chain cannot be excluded until the 5′ end of the cDNA, corresponding to the liver mRNA, has been completely sequenced.

Isolation and characterization of a genomic clone for the B-chain of C1q

Screening of approx. 200000 clones from a human genomic DNA cosmid library with a nick-translated 84-base-pair cDNA probe (coding for residues B198–B226) allowed the isolation of one positive clone. The total length of genomic DNA in this cosmid was approx. 45 kb, and Southern-blot analysis showed that the C1q B-chain cDNA probes hybridized to restriction fragments occupying a central portion in the cosmid insert. The Southern-blot analysis indicated the presence of only one C1q B-chain gene. In order to study the gene in greater detail, 1.33 kb *Ava*II, 1.47 kb *Taq*1, 0.63 kb *Stu*1 and 0.78 kb *Hin*f1 genomic fragments were subcloned in the pAT/153 vector and isolated by using 5′ and 3′ cDNA fragments as probes (Figs. 3 and 4). The results of the nucleotide sequence analyses of these subclones are shown in Fig. 3. The nucleotide sequence coding for residues B1–B226 and for the entire 3′ non-translated region is identical with that obtained for the cDNA sequence. The genomic nucleotide sequence before the codon for the *N*-terminal residue B1 indicates the presence of a predominantly hydrophobic stretch of 29 amino acids (Fig. 3), most of which, as discussed above, could serve as a signal peptide. It is possible that there is no large pro-form of the B-chain of C1q; however, three proteins, transferrin (Schneider *et al.*, 1984), Ia-associated invariant polypeptide (Claessum *et al.*, 1983) and ovalbumin (Braell & Lodish, 1982), have hydrophobic stretches of approx. 25 amino acids located near the *N*-terminal end of the chain that are not cleaved off and could serve as an internal signal sequence. In the case of transferrin it is considered possible that the 25-amino-acid-long hydrophobic stretch, as well as acting as a signal sequence, may also play a role in anchoring the protein in the membrane.

From the studies shown in Figs. 3 and 4, approx. 2040 base-pairs have been identified as being present in the C1q B-chain gene and, since this does not account for approx. 500 base-pairs of the mRNA, it can be calculated that the gene is at least 2.6 kb long. Only one intron of 1.1 kb was found in the region coding for residues B1–B226 (Figs. 3 and 4). This intron is located within the codon for glycine-36 and obeys the GT–AG splicing rule (Fig. 3). It is of interest that this is exactly the point where the triple-helical chains of C1q appear to bend when viewed in the electron microscope (Fig. 4). The bending of the triple helices is considered to be brought about by an interruption of the Gly-Xaa-Yaa repeating triplet sequences in the A- (by insertion of a threonine at position A39 between two triplets) and C- (by the presence of an alanine residue at C36, where a glycine residue would be expected) chains and the apparent insertion of an extra triplet in the B-chain (Reid, 1983). If the A- and C-chain genes also have introns in positions corresponding to the points of disruption of the triple-helical chains, then this could indicate some role for the introns in the conservation of an unusual structural feature in the C1q molecule, especially as this feature is probably functionally important in the binding of aggregated IgG and/or the activation of proenzymes C1r and C1s in the $C1r_2$–$C1s_2$ complex.

Unlike the human and chick fibrillar collagens, the amino acid sequence within the collagen-like region of the

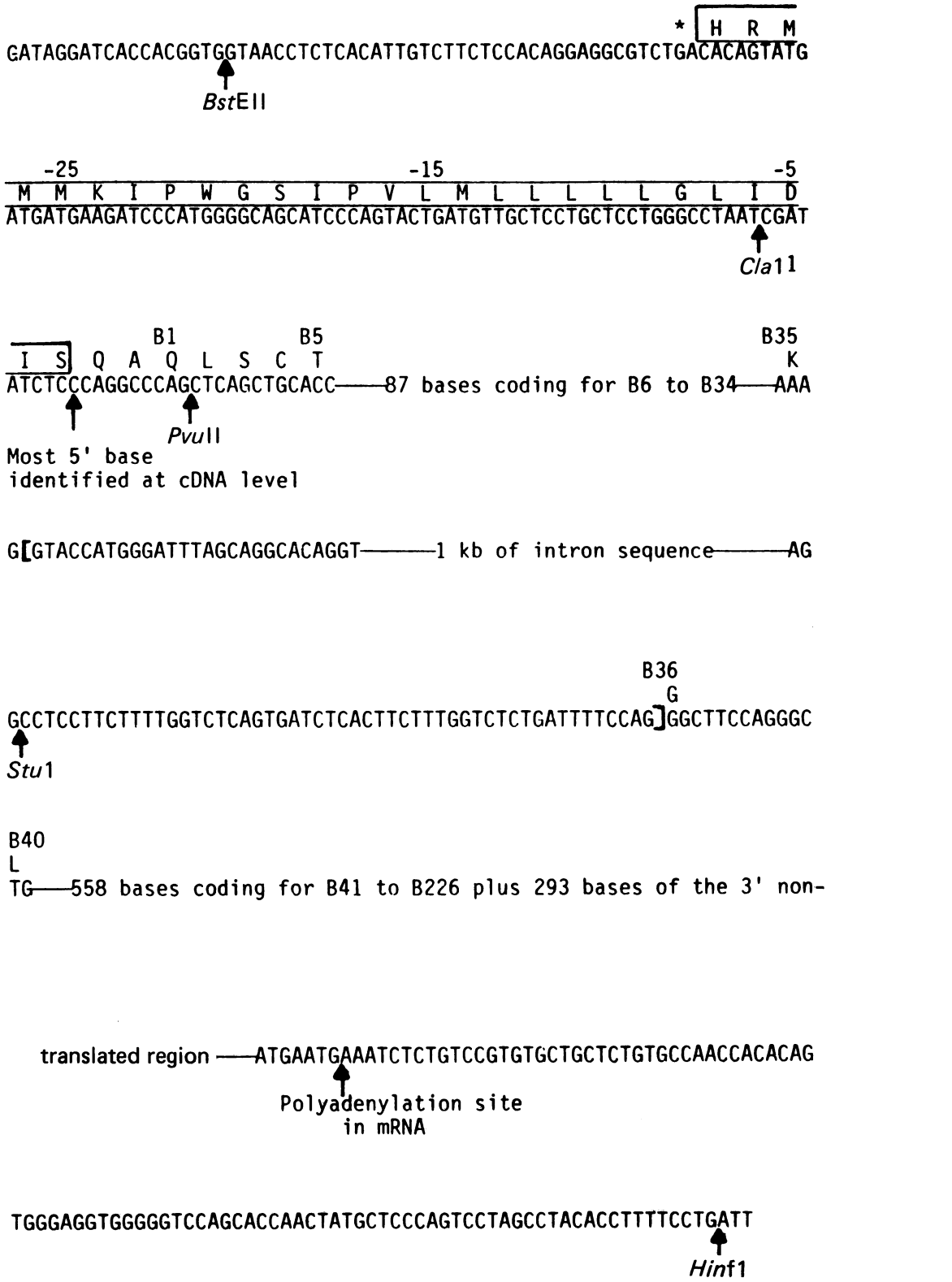

Fig. 3. DNA sequence of the C1q B-chain gene

The coding of the C1q B-chain gene is shown. The nucleotide sequence found in the cDNA clone starts at the codon for glutamine at position −2, therefore the nucleotide sequence coding for the derived amino acid sequence designated as '−3 to −29' has not yet been shown to be present in the mRNA for the B-chain and is enclosed in a box. Only certain sections of the nucleotide sequence coding for amino acid residues B1–B226 and for the 3′ non-translated region of the mRNA are shown, since these portions of the gene sequence are identical with that shown in Fig. 1 for the cDNA sequence. The square brackets denote the start and finish of the intron within the codon for glycine at B36.

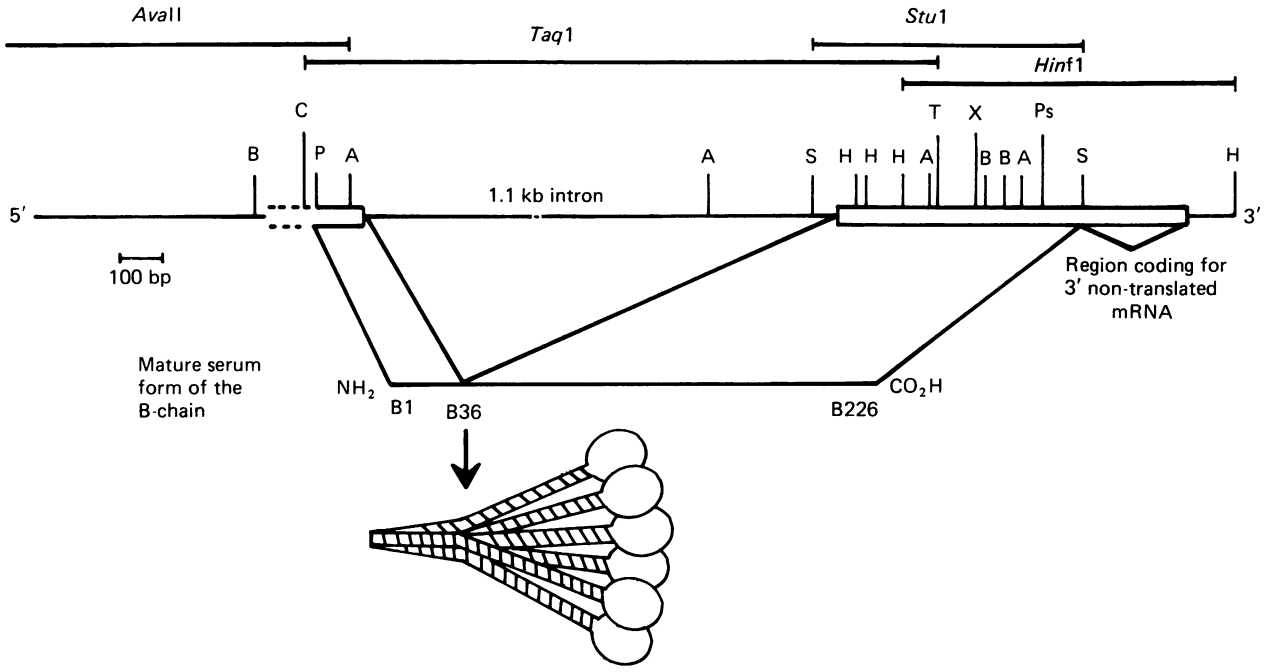

Fig. 4. Restriction map of the C1q B-chain gene

A restriction map of the region of the C1q B-chain gene was characterized by isolation of subclones prepared from *Ava*II, *Taq*1, *Stu*1 and *Hin*f1 restriction-enzyme digests of the cosmid clone as indicated at the top of the Figure. The continuous-line boxed regions show definitely established exon sequence coding for the B-chain mRNA; the broken-line box shows possible exon sequence coding for predominantly hydrophobic amino acid sequence. The precise extent of the 5′ exon (or exons) has not yet been defined. Restriction-enzyme sites that were useful in preparation of subclones and sequencing are shown: A, *Ava*II; B, *Bst*EII; C, *Cla*1; H, *Hin*f1; P, *Pvu*II; Ps, *Pst*1; S, *Stu*1; T, *Taq*1; X, *Xho*1. The position of the glycine residue at B36 in each of the six B-chains present in one molecule of C1q is indicated by the arrow (↓).

C1q B-chain does not follow the general rule of being encoded by exons of 45, 54, 99, 108 or 162 base-pairs in which the Gly-Xaa-Yaa triplet pattern is precisely conserved (Chu *et al.*, 1984; Yamada *et al.*, 1984). In some respects the C1q structure seems more similar to that of the non-fibrous collagen of two invertebrates (Monson *et al.*, 1982; Kramer *et al.*, 1982); for example, the *Drosophila* non-fibrous collagen has discontinuities in its Gly-Xaa-Yaa pattern and contains intron splice sites in a similar position to that found in the C1q B-chain gene, i.e. within a glycine codon (Monson *et al.*, 1982).

The availability of cDNA and genomic probes for the B-chain of C1q has allowed the examination, by Southern-blot analysis, of restriction-enzyme digests of genomic DNA from patients suffering from lack of C1q function, i.e. due to either the complete lack of C1q or the presence of a dysfunctional molecule in the serum. However, initial studies have shown no difference between normal subjects and deficient patients (R. McAdam & K. B. M. Reid, unpublished work). Further studies with DNA probes for the A- and C-chains or involving examination of the processing of the whole molecule are required to provide a proper understanding of these deficiency states. Use of the B-chain cDNA probe has allowed, by analysis of human–mouse somatic-cell hybrids, the assignment of the human C1q B-chain to chromosome 1p (Solomon *et al.*, 1985).

I thank Dr. A. K. Bentley (Sir William Dunn School of Pathology, Oxford, U.K.) and Dr. D. R. Bentley (of this Unit) for providing the cDNA libraries and Ms. J. Parsons for excellent technical assistance.

REFERENCES

Anson, D. S., Choo, K. H., Rees, D. J. G., Giannelli, F., Gould, K., Huddleston, J. A. & Brownlee, G. G. (1984) EMBO J. **3**, 1053–1060

Bawerman, G. (1976) Prog. Nucleic Acid Res. Mol. Biol. **17**, 1513–1523

Bernards, R. & Flavell, R. A. (1980) Nucleic Acids Res. **8**, 1521–1534

Birnboim, H. C. & Doty, J. (1979) Nucleic Acids Res. **7**, 1513–1523

Braell, W. A. & Lodish, H. F. (1982) J. Biol. Chem. **257**, 4578–4582

Chapius, R. M., Hauptmann, G., Grosshams, E. & Isliker, H. (1982) J. Immunol. **129**, 1509–1512

Chirgwin, J. M., Przbyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) Biochemistry **18**, 5294–5299

Chu, M.-L, deWet, W., Bernard, M., Ding, J.-F., Morabito, M., Myers, J., Williams, C. & Ramirez, F. (1984) Nature (London) **310**, 337–340

Claessum, L., Larhammar, D., Rask, L. & Peterson, P. A. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 7395–7399

Colten, H. R. (1976) Adv. Immunol. **22**, 67–118

Colten, H. R., Gordon, J. M., Borsos, T. & Rapp, J. H. (1968) J. Exp. Med. **128**, 595–604

Grossveld, F. G., Dahl, H. H. M., de Boer, E. & Flavell, R. A. (1981) Gene **13**, 227–237

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

Hannema, A. J., Kluin-Nelemans, H. C., Hack, C. E., Mallee, C., van Helden, H. P. T. & Erenber-Belmer, A. J. M. (1982) Mol. Immunol. **19**, 1373
Kramer, J. M., Cox, G. N. & Hirsch, D. (1982) Cell **30**, 599–606
Lehrach, H., Diamond, D., Wezney, J. M. & Boedtker, H. (1977) Biochemistry **16**, 4743–4751
Loos, M. (1983) Curr. Top. Microbiol. Immunol. **102**, 1–56
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor
Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. **65**, 499–560
Monson, J. M., Natzle, J., Friedman, J. & McCarthy, B. J. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 1761–1765
Morris, K. M., Colten, H. R. & Bing, D. H. (1978) J. Exp. Med. **148**, 1007–1019
Proudfoot, N. J. & Brownlee, G. G. (1976) Nature (London) **263**, 211–214
Radloff, R., Bauer, W. & Vinograd, J. (1967) Proc. Natl. Acad. Sci. U.S.A. **57**, 1514–1521
Reid, K. B. M. (1983) Biochem. Soc. Trans. **11**, 1–12
Reid, K. B. M. & Solomon, E. (1977) Biochem. J. **167**, 647–660
Reid, K. B. M. & Thompson, E. O. P. (1978) Biochem. J. **173**, 863–868
Reid, K. B. M. & Thompson, R. A. (1983) Mol. Immunol. **20**, 1117–1125
Reid, K. B. M., Gagnon, J. & Frampton, J. (1982) Biochem. J. **203**, 559–569
Reid, K. B. M., Bentley, D. R. & Wood, K. J. (1984) Philos. Trans. R. Soc. London Ser. B **306**, 345–354
Rigby, P. W., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. Mol. Biol. **113**, 237–251
Schneider, C., Owen, M. J., Banville, D. & Williams, J. G. (1984) Nature (London) **311**, 675–678
Solomon, E., Skok, J., Griffin, J. & Reid, K. B. M. (1985) Cytogenet. Cell Genet. (Int. Workshop Human Gene Mapping 8th) in the press
Thompson, R. A., Haeney, M., Reid, K. B. M., Davies, J. G., White, R. H. R. & Cameron, A. H. (1980) New Engl. J. Med. **303**, 2–24
Wahl, G. M., Stern, M. & Stark, G. R. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 3683–3687
Wickens, A. & Stephen, C. (1984) Science **226**, 1045–1047
Yamada, Y., Liau, G., Mudryj, M., Obici, S. & de Crombrigghe, B. (1984) Nature (London) **310**, 333–337
Yonemasu, K., Kitajuma, H., Tanabe, S., Ochi, T. & Shinkai, H. (1979) Immunology **35**, 523–530

Received 26 February 1985/Accepted 9 July 1985