

# Complementary DNA and Derived Amino Acid Sequence of the $\beta$ Subunit of Human Complement Protein C8: Identification of a Close Structural and Ancestral Relationship to the $\alpha$ Subunit and C9<sup>†</sup>

O. M. Zack Howard, A. Gururaj Rao, and James M. Sodetz\*

Department of Chemistry and School of Medicine, University of South Carolina, Columbia, South Carolina 29208

Received January 30, 1987; Revised Manuscript Received March 5, 1987

**ABSTRACT:** A cDNA clone encoding the  $\beta$  subunit ( $M_r$  64 000) of the eighth component of complement (C8) has been isolated from a human liver cDNA library. This clone has a cDNA insert of 1.95 kilobases (kb) and contains the entire  $\beta$  sequence [1608 base pairs (bp)]. Analysis of total cellular RNA isolated from the hepatoma cell line HepG2 revealed the mRNA for  $\beta$  to be  $\sim$ 2.5 kb. This is similar to the message size for the  $\alpha$  subunit of C8 and confirms the existence of different mRNAs for  $\alpha$  and  $\beta$ . This finding supports genetic evidence that  $\alpha$  and  $\beta$  are encoded at different loci. Analysis of the derived amino acid sequence revealed several membrane surface seeking segments that may facilitate  $\beta$  interaction with target membranes during complement-mediated cytolysis. Determination of the carbohydrate composition indicated 1 or 2 asparagine-linked but no O-linked oligosaccharide chains. Comparison of the  $\beta$  sequence to that reported for  $\alpha$  in the preceding paper [Rao, A. G., Howard, O. M. Z., Ng, S. C., Whitehead, A. S., Colten, H. R. & Sodetz, J. M. (1987) *Biochemistry* (preceding paper in this issue)] and to that of human C9 revealed a striking homology between all three proteins. For  $\beta$  and  $\alpha$ , the overall homology is 33% on the basis of identity and 53% when conserved substitutions are allowed. For  $\beta$  and C9, the values are 26% and 47%, respectively. All three have a large internal domain that is nearly cysteine free and N- and C-termini that are cysteine-rich and homologous to the low-density lipoprotein receptor repeat and epidermal growth factor type sequences, respectively. The overall homology and similarities in size and structural organization are indicative of a close ancestral relationship. It is concluded that  $\alpha$ ,  $\beta$ , and C9 are members of a family of structurally related proteins that are capable of interacting to produce a hydrophilic to amphiphilic transition and membrane association.

**H**uman C8 participates in complement-mediated cell lysis through interaction with the intermediate C5b-7 complex to form C5b-8 (Müller-Eberhard, 1986). Although capable of promoting slow membrane lysis itself, the major role of C5b-8 is to serve as a receptor for C9. This role involves a single, specific site on the  $\alpha$  subunit of C8 that mediates C9 binding to yield C5b-9, the terminal cytolytic complex of complement.

Human C8 consists of three nonidentical subunits ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) that are arranged asymmetrically as a disulfide-linked  $\alpha$ - $\gamma$  dimer and a noncovalently associated  $\beta$  chain (Steckel et al., 1980). Several well-defined functional domains have been identified from studies of the individual subunits. The  $\alpha$  subunit has separate domains involved in interactions with  $\beta$  (Brickner & Sodetz, 1984),  $\gamma$  (Brickner & Sodetz, 1985), C9 (Stewart & Sodetz, 1985), and the target membrane bilayer (Steckel et al., 1983). The primary role of  $\beta$  is to direct C8 into the nascent complex through interaction with a specific binding site on C5b-7 (Monahan & Sodetz, 1981). The nature of this site is unknown; however, studies of interactions between purified C5, C6, C7, and C8 (Kolb et al., 1973) or  $\beta$  (Stewart et al., 1987) suggest it resides on C5b. A role for  $\gamma$  has not been identified, and evidence indicates it is not essential for C8 activity (Brickner & Sodetz, 1984).

In the preceding paper, we described the characterization of cDNA clones that yielded the entire amino acid sequence of the  $\alpha$  subunit (Rao et al., 1987). That study also yielded evidence for the existence of a separate mRNA for  $\alpha$ , a finding

in support of genetic studies which concluded that  $\alpha$ - $\gamma$  and  $\beta$  are encoded at different loci (Raum et al., 1979; Alper et al., 1983; Rittner et al., 1983; Rogde et al., 1986). In this report, we present the entire amino acid sequence of  $\beta$  as derived from a cDNA clone encoding this subunit. This extends previous structural data on  $\beta$ , which was limited to its amino acid composition and amino-terminal sequence (Steckel et al., 1980). Our results also confirm the existence of a separate mRNA for  $\beta$  and reveal significant homologies to both the  $\alpha$  subunit and C9.

## EXPERIMENTAL PROCEDURES

**Purification of Proteins and Amino Acid Sequencing.** The  $\beta$  subunit was purified from C8 and converted to its *S*-(pyridylethyl) derivative as described previously (Steckel et al., 1980; Rao & Sodetz, 1984). *S*-(Pyridylethyl)- $\beta$  was digested by CNBr or by *Staphylococcus aureus* protease (Miles Laboratories) under conditions favoring cleavage at glutamic acid (50 mM ammonium acetate, pH 4.0). Peptides from CNBr digests were fractionated by reverse-phase high-performance liquid chromatography (HPLC) as described (Rao et al., 1987). *S. aureus* protease peptides were fractionated similarly but with a trifluoroacetic acid/isopropanol solvent system. Amino-terminal amino acid sequencing was performed by automated Edman degradation on a Beckman 890C sequencer. Carboxy-terminal sequencing was performed by digestion with carboxypeptidase P (Rao et al., 1987).

**Screening of the cDNA Library.** On the basis of the amino acid sequence of peptides derived from  $\beta$ , two mixed-sequence oligonucleotide probes were synthesized and used to screen an adult human liver cDNA library (Belt et al., 1984). Ap-

<sup>†</sup> This research was supported by NIH Grants A116856 and BRSG S07 RR07160 and American Heart Association Established Investigator Award 82-121 to J.M.S.

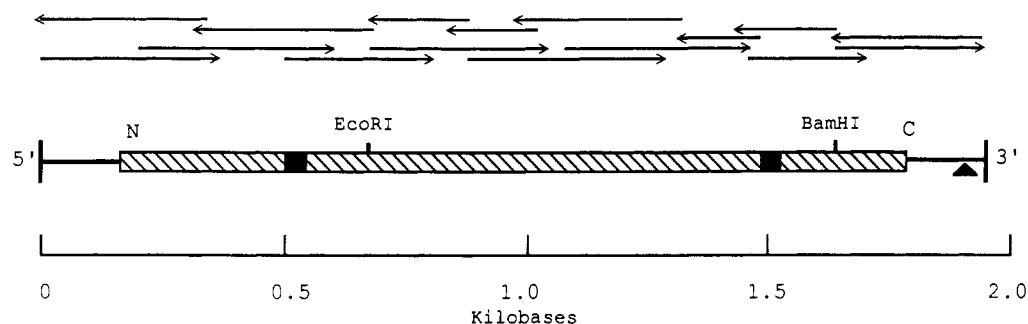


FIGURE 1: Map and sequencing strategy for  $\beta$  cDNA. Shown is the  $\beta$  cDNA insert and internal restriction sites used to generate fragments for sequencing. Arrows identify increments sequenced either by the rapid-deletion subcloning method or by using synthetic primers. The polyadenylation signal is indicated by ( $\blacktriangle$ ). The hatched area shows the  $\beta$  coding region along with the oligonucleotide probe sites ( $\blacksquare$ ).

proximately 40 000 recombinant clones were screened as described earlier (Rao et al., 1987). Hybridizations were performed at 40 °C with washing at 42 °C for the 17-mer probe and 44 °C for the 20-mer probe. Washing was in 6  $\times$  SSC (0.9 M NaCl/90 mM sodium citrate) containing 0.05% sodium pyrophosphate.

**DNA Sequencing.** Fragments to be sequenced were isolated by electroelution from acrylamide gels, subcloned into M13 mp18 or mp19 and sequenced by a modified dideoxy chain-termination method (Rao et al., 1987). Fragments of extended length and restriction site overlaps were sequenced either by the rapid-deletion M13 subcloning approach (Dale et al., 1985) or by using synthetic primers prepared by the University of South Carolina Oligonucleotide Synthesis Facility.

**Analysis of RNA Blots.** Blots containing total cellular RNA from the human hepatoma cell line HepG2 were prepared and probed as described (Rao et al., 1987). A labeled single-stranded probe was prepared by primed synthesis on an M13 template containing the middle 0.95-kb fragment from an *EcoRI*–*Bam*HI digest of the  $\beta$  cDNA insert.

**Data Analysis.** Sequence analysis and secondary structure predictions were performed as described previously (Rao et al., 1987).

## RESULTS

**Protein Sequencing and Carbohydrate Composition.** Four peptides from CNBr digests and six from *S. aureus* protease digests were partially sequenced by conventional methods to yield a total of 160 residues of  $\beta$  sequence. From this, the sequence Asp-Gln-Tyr-Trp-Gly-Ile near the N-terminus was selected to make the following 17-mer probe:

GAC/T-CAG/A-TAC/T-TGG-GGC/T/G/A-AT

Another sequence of Glu-Glu-Phe-Gln-Lys-Glu-Val near the C-terminus was used to design the following 20-mer probe:

GAG/A-GAG/A-TTT/C-CAA/G-AAG/A-GAG/A-GT

Both probes were found to be highly selective for  $\beta$  cDNA clones under stringency conditions described above.

To determine the C-terminal sequence, *S*-(pyridylethyl)- $\beta$  was subjected to carboxypeptidase P digestion. Yields (mol/mol, in parentheses) of amino acids released at 30 min were Ser (0.72), Leu (0.70), Asp (0.67), and *S*-(pyridylethyl)cysteine (0.50). Yields after 60 min were Ser (1.2), Leu (0.97), Asp (0.91), and *S*-(pyridylethyl)cysteine (0.73). The time course of release for all four residues was such that a sequence could not be deduced. However, one of the peptides isolated from *S. aureus* protease digests was found to have Thr-Leu-Asp-Cys-Ser as its complete sequence. In view of results from carboxypeptidase digestions, it was assumed this

peptide corresponded to the C-terminus of  $\beta$ . This was subsequently confirmed by the cDNA sequence.

Carbohydrate analyses of  $\beta$  yielded the following composition (mol/mol,  $\pm 10\%$ ): glucosamine (GlcN) (3.5), galactosamine (GalN) (0), Man (2.7), Gal (2.5), *N*-acetylneuraminic acid (NeuAc) (1.7), Fuc (0.1), Glc (4.4). This composition is nearly identical with that for  $\alpha$  with the exception of glucose, which is variable and is assumed to be a contaminant. As with  $\alpha$ , the composition indicates one but no more than two asparagine-linked oligosaccharide chains and no O-linked sugars.

**Sequence of  $\beta$  cDNA.** After the cDNA library was screened, the clone with the largest insert that was positive to both oligonucleotide probes was selected for further characterization. Results in Figure 1 show that this clone has an insert [1944 base pairs (bp)] that includes the entire coding region for mature  $\beta$ . The 3' extension has a single polyadenylation signal but no poly(A) sequence.

The complete nucleotide sequence and derived amino acid sequence are shown in Figure 2. Translation of the 5' region yields a continuous 54-residue sequence with no in-frame termination codons. One methionine occurs at position -2; however, the flanking nucleotide sequences is one that is rarely found at a translation initiation methionine (Kozak, 1981). Moreover, the arginine at -1 is not a common residue recognized by signal peptidase (Heijne, 1984). Considering the absence of other methionines and the potential for 5' loop-back artifacts referred to earlier (Rao et al., 1987), no conclusion can be made regarding a signal peptide or propeptide for  $\beta$ .

The coding region for mature  $\beta$  is 1608 bp in length and contains 536 amino acids for a total  $M_r$  of 60 742. This compares well to 553 amino acids predicted from the amino acid composition and a  $M_r$  of 64 000 for the glycosylated protein (Steckel et al., 1980). There are three asparagines that are candidates for glycosylation. In  $\alpha$ , it was proposed that only one of two candidates was glycosylated on the basis of behavior of the corresponding peptides in sequencing reactions and the overall carbohydrate composition. It is noteworthy that  $\beta$  has nearly an identical carbohydrate composition and that one of its three possible sites contains the same Ser-Asn-Trp-Ser sequence found at the proposed glycosylation site in  $\alpha$ . Therefore, this seems like a most probable site for glycosylation in  $\beta$ . The  $\beta$  subunit also resembles  $\alpha$  in that its cysteine residues are clustered near the N- and C-termini. Of a total of 28, there are none within the middle 178 residues and only 4 in the middle 347.

**RNA Blot Analysis.** The cDNA for  $\beta$  contains a polyadenylation signal in the 3' untranslated region but no poly(A) sequence. Likewise, there is no clear indication whether the 5' extension is complete. To determine if this cDNA is nearly full length, total RNA from HepG2 cells was analyzed on

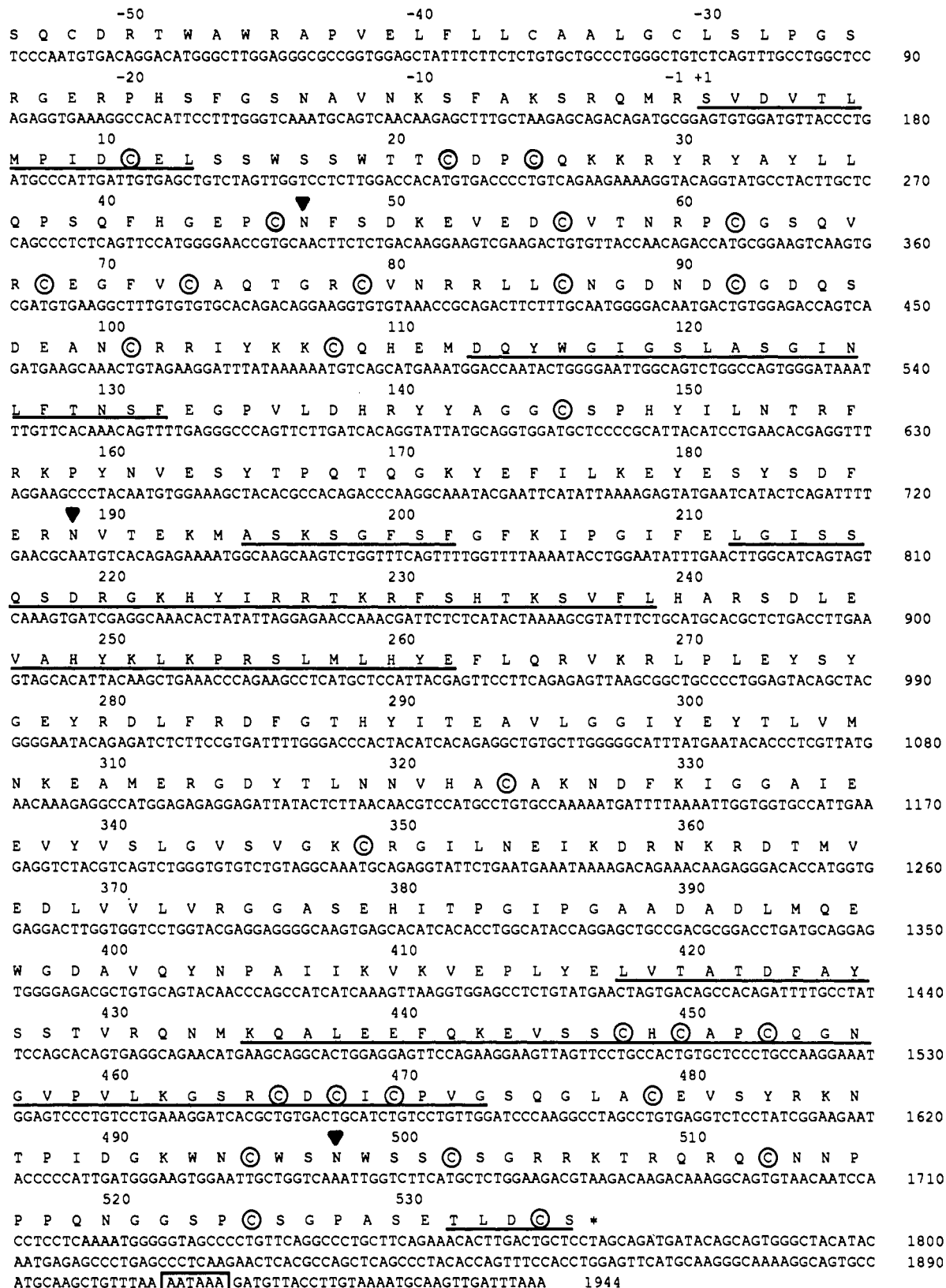


FIGURE 2: cDNA and amino acid sequence of  $\beta$ . Numbering of amino acids begins with the N-terminus of mature  $\beta$  as +1. Cysteines are highlighted by circles, and potential asparagine-linked glycosylation sites are identified by (▼). Segments that were confirmed by amino acid sequencing are underlined. The polyadenylation signal sequence is boxed, and the translation stop codon is identified by (\*).

RNA blots by using a  $\beta$  cDNA probe. Results in Figure 3 show the prominent mRNA detected is ~2.5 kb in length. This indicates the  $\beta$  message has 5' and/or 3' sequences extending beyond those found in the cDNA. A second, less abundant mRNA of ~1.5 kb can also be detected. A second mRNA of the same size was also observed with the  $\alpha$  cDNA

probe and was attributed to either an alternatively processed or a highly homologous mRNA (Rao et al., 1987). The former possibility was considered because an internal AATAAA polyadenylation sequence occurs at 1298 bp from the 5' end of the  $\alpha$  cDNA. Either explanation could apply again here because the  $\beta$  cDNA has the same sequence at 1226 bp from

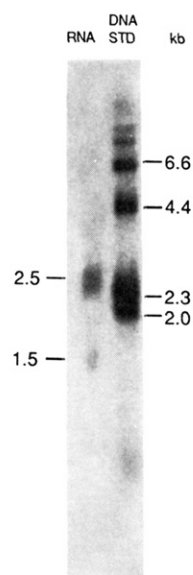


FIGURE 3: RNA blot analysis of  $\beta$  mRNA. Approximately 10  $\mu$ g of total cellular RNA from HepG2 cells was electrophoresed and blotted as described in the text. The blot was probed with a single-stranded probe corresponding to the middle 0.95-kb fragment of the  $\beta$  cDNA. Molecular weight standards consist of a *Hind*III digest of  $^{32}$ P-labeled  $\lambda$  phage DNA.

its 5' end. Importantly, detection of both the 2.5- and 1.5-kb mRNAs with either probe is not a result of shared specificity. This was established by extensive cross-hybridization experiments with  $\alpha$  and  $\beta$  cDNA fragments (data not shown). It may also be that these smaller mRNAs are unique to HepG2 cells or are a function of using total RNA. When poly(A) RNA from fresh baboon liver was blotted and probed with either an  $\alpha$  or  $\beta$  cDNA probe, only a single mRNA species of  $\sim 2.5$  kb was detected.<sup>1</sup>

**Secondary Structure Predictions.** The hydropathic profile in Figure 4 indicates that  $\beta$  is largely hydrophilic. Analysis by the method of Eisenberg et al. revealed no potential transmembrane segments; however, seven segments were identified as capable of forming amphipathic  $\beta$ -structures that are membrane surface seeking (Eisenberg et al., 1984). These segments correspond to residues 65–75, 147–162, 196–216, 236–249, 320–347, 405–418, and 489–499.

**Sequence Homologies.** Results in the preceding paper (Rao et al., 1987) described two homologies between  $\alpha$ ,  $\beta$ , and C9. One located near the N-terminus included the cysteine-rich low-density lipoprotein (LDL) receptor repeat, and the other at the C-terminus included the cysteine-rich growth factor type domain. Figure 5 shows that several other significant homologies exist, including one segment that is unique to  $\alpha$  and  $\beta$  and another to  $\beta$  and C9. Overall homology between  $\alpha$  and  $\beta$  is 33% on the basis of identity and 53% when conserved substitutions are included. For  $\beta$  and C9, the values are 26% and 47%, respectively.

## DISCUSSION

Results in this study indicate  $\beta$  is encoded in a mRNA that is similar in size yet distinct from the  $\alpha$  subunit mRNA. This observation is significant because it supports genetic studies which concluded that  $\alpha$ - $\gamma$  and  $\beta$  are encoded at different loci. As noted in the preceding paper in this issue, indirect evidence suggests that  $\gamma$  may also be encoded in a separate message, and if so, this would mean separate loci exist for  $\alpha$ ,  $\beta$ , and  $\gamma$ ,

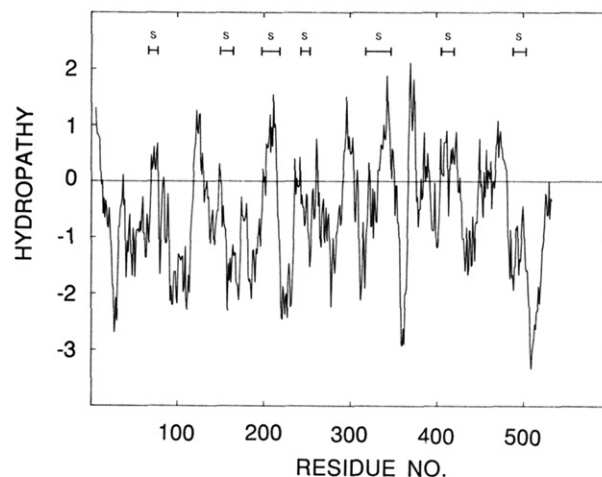


FIGURE 4: Hydropathy profile of  $\beta$ . Sequence regions with positive hydropathy values are hydrophobic, and those with negative values are hydrophilic. Brackets in the inset identify potential membrane surface seeking segments (S).

a possibility that is compatible with genetic data on C8 (Rao et al., 1987). More detailed studies of the number of loci and their linkage should now be possible with the availability of  $\alpha$  and  $\beta$  cDNA probes.

Analysis of the  $\beta$  sequence reveals several segments with potential to form amphipathic  $\beta$ -structures that are capable of interacting with membrane surfaces. These segments may serve two functions. One is that they may assist in stabilizing interactions of C8 with C5b-7 by direct association with the membrane surface. Since  $\beta$  alone can bind with high affinity to this complex, stabilization of C8 binding is probably mediated in part through interactions involving  $\beta$  (Monahan & Sodetz, 1981). Second, these segments may contribute directly to the lytic activity of C5b-8 and C5b-9 through membrane perturbation. However, this contribution is likely to be minor because, unlike  $\alpha$ , the  $\beta$  subunit in membrane-bound C5b-8 and C5b-9 is only marginally labeled by photosensitive, membrane-restricted probes (Steckel et al., 1983). Marginal photolabeling indicates limited interaction with the bilayer, a result consistent with the above structural predictions.

The  $\beta$  subunit exhibits a striking overall homology to  $\alpha$  and human C9. As summarized in Figure 6, this includes not only the LDL receptor and epidermal growth factor type domains but other large regions as well. Each homologous region occurs at the same location in all three proteins, thus revealing a remarkable similarity in structural organization. This is further indicated by the similar number of amino acids, the localization of cysteines at the N- and C-termini, and, in the case of  $\alpha$  and  $\beta$ , the nearly identical carbohydrate content. Considering these proteins are hydrophilic but can display amphiphilic characteristics, it is also significant that each contains segments capable of interacting with membranes. Most of these occur in a relatively cysteine-free region where conformational constraints are minimal. Such a design is not likely a result of chance. As it exists independently in serum, each protein must be refractory to membranes, yet still be able to expose membrane-interacting segments upon binding to C5b-7 or C5b-8. Having maximum flexibility in this region would permit exposure of such segments to be modulated by conformational changes, which in turn could be influenced by protein-protein interactions during assembly of the complement complexes. Thus, it seems one reason why each protein may have this structural arrangement is that it facilitates a function they have in common, a hydrophilic to amphiphilic transition and subsequent membrane association.

<sup>1</sup> S. C. Ng and J. M. Sodetz, unpublished results.

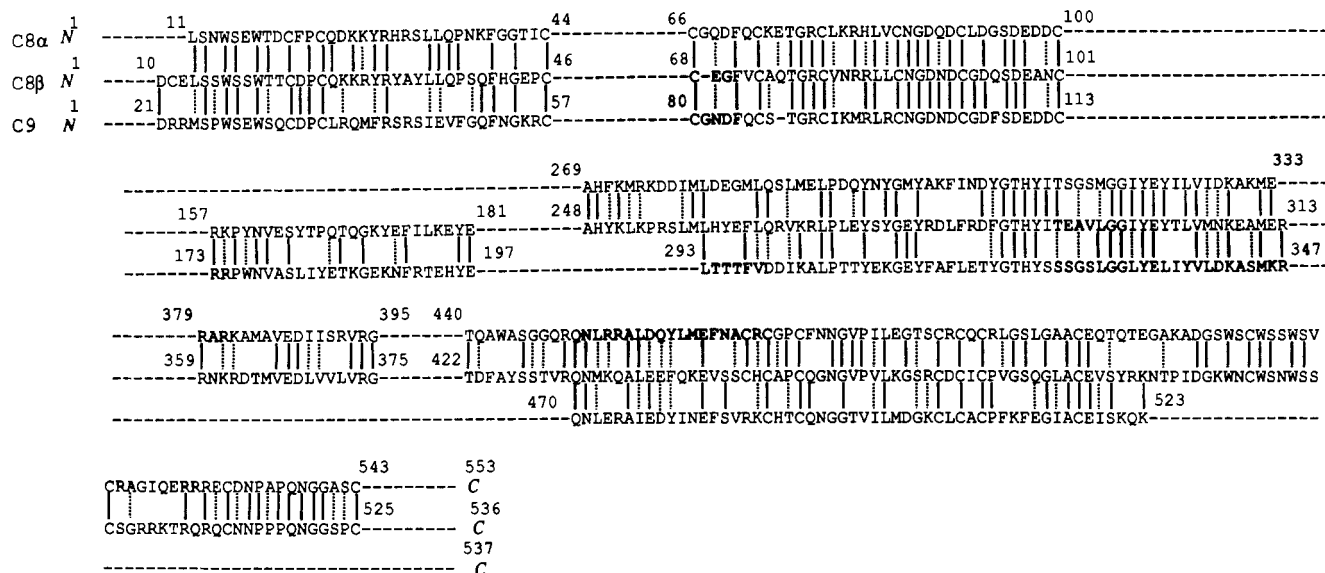


FIGURE 5: Homologous sequences in  $\alpha$ ,  $\beta$ , and C9. Segments of  $\beta$  that were identified as homologous to  $\alpha$  and/or human C9 are optimally aligned. The sequence of  $\alpha$  is from the preceding paper (Rao et al., 1987) and that of C9 from Di Scipio et al. (1984). All segments shown exhibit  $\geq 39\%$  homology on the basis of identity and  $\geq 54\%$  homology if conserved substitutions are included.

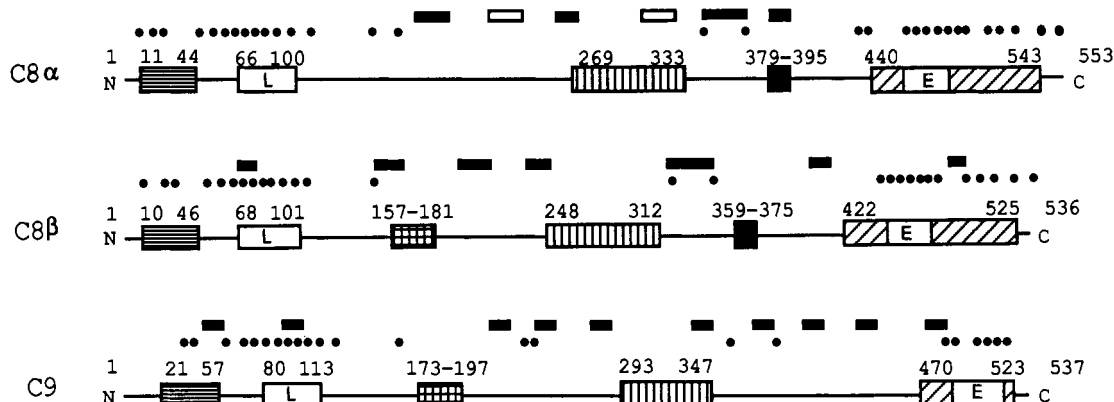


FIGURE 6: Summary of structural similarities between  $\alpha$ ,  $\beta$ , and C9. Segments of  $\alpha$ ,  $\beta$ , and C9 that are related by sequence homology are identified by residue numbers and common markings. Segments containing the LDL receptor homology are identified by L, and those with growth factor type homology by E. Cysteine residues are identified by (●). Solid bars above each protein identify the potential membrane surface seeking segments while open bars identify potential transmembrane segments. Data for  $\alpha$  is from the preceding paper (Rao et al., 1987). The sequence (Di Scipio et al., 1984) and locations of predicted membrane surface seeking segments (Shiver et al., 1986) for C9 are from published reports. It is noted that the number of cysteines shown here for C9 differs by one from that reported by others (Stanley et al., 1986).

Having conserved domains in the same region of each protein may also confer stability to C8 and C9 interactions. In solution, C8 and C9 can form a complex that is mediated by a single C9 binding site on  $\alpha$  (Stewart & Sodetz, 1985). Because  $\alpha$  is simultaneously associated with  $\beta$  (in C8), this complex must be physically arranged as  $\beta$ - $\alpha$ -C9. Monomeric C9 also has properties that allow it to self-polymerize in solution or upon binding to C5b-8 (Müller-Eberhard, 1986), suggesting that a heteropolymer of  $\beta$ - $\alpha$ -C9 $_n$  can form within membrane-bound C5b-9. Specific association between C8 and C9 also occurs during assembly of SC5b-9, the soluble counterpart of membrane-bound C5b-9. Thus, stability of these interactions is not dependent solely on the presence of membranes but also on intrinsic properties of the constituent proteins. Considering this and similarities in  $\alpha$ ,  $\beta$ , and C9 structures, one can speculate that pairs of conserved domains might align to provide a repetitive structure that stabilizes the above heteropolymer. This could explain C8 and C9 interactions and the self-polymerization of C9 but not the inability of  $\beta$  to associate with C9 nor the fact that  $\alpha$ - $\gamma$  and  $\beta$  do not self-polymerize. Undoubtedly, nonconserved regions of each protein must have a role in determining the specificity of

interactions. With the availability of sequences, specific segments of  $\alpha$ ,  $\beta$ , and C9 can now be isolated and used to delineate the various domains involved in their interactions.

On the basis of similarities in their structure and function, we conclude that  $\alpha$ ,  $\beta$ , and C9 share a close ancestral relationship and are members of a family of proteins that are capable of induced conformational changes leading to membrane interaction. Evidence suggests that other proteins may belong to this family, one of which is perforin, the pore-forming protein released from cytotoxic lymphocytes. This protein facilitates lysis of target cells through self-polymerization and formation of pores that are morphologically similar to C5b-9 formed on complement-lysed membranes (Dourmashkin et al., 1980). The perforin monomer is similar in size to C9, and like C9, it can be induced to self-polymerize by metal ions (Young et al., 1986). Recent immunological evidence also indicates perforin and C9 have common antigenic determinants, at least one of which involves the cysteine-rich LDL receptor domain (Young et al., 1986; Zalman et al., 1986; Tschoop et al., 1986). These similarities suggest homology to C9 which, if extensive, will establish perforin as a fourth member of this family.

## ADDED IN PROOF

We recently isolated a  $\beta$  cDNA clone with a longer 5' extension. Compared to that in Figure 2, the derived amino acid sequence differs only from position -54 to position -51. The revised sequence MKNS encompasses a strong translation initiation signal at methionine -54, thus suggesting that  $\beta$  has a leader sequence of 54 amino acids. The sequence preceding residue -50 in Figure 2 is now considered artifactual.

## ACKNOWLEDGMENTS

We thank K. Tertia Belt and Michael Carroll for the cDNA library and Simon C. Ng, A. Steven Whitehead, and Harvey R. Colten for helpful discussions regarding this project. We also thank Regina York for expert technical assistance with the peptide sequencing, carbohydrate analysis, and the synthesis of oligonucleotides.

## REFERENCES

- Alper, C. A., Marcus, D., Raum, D., Petersen, B. H., & Spira, T. J. (1983) *J. Clin. Invest.* 72, 1526-1531.
- Belt, K. T., Carroll, M. C., & Porter, R. R. (1984) *Cell (Cambridge, Mass.)* 36, 907-914.
- Brickner, A., & Sodetz, J. M. (1984) *Biochemistry* 23, 832-837.
- Brickner, A., & Sodetz, J. M. (1985) *Biochemistry* 24, 4603-4607.
- Dale, R. M. K., McClure, B. A., & Houchins, J. B. (1985) *Plasmid* 13, 31-40.
- Di Scipio, R. G., Gehring, M. R., Podack, E. R., Kan, C. C., Hugli, T. E., & Fey, G. H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 7298-7302.
- Dourmashkin, R. R., Deteix, P., Simone, C. B., & Henkart, P. (1980) *Clin. Exp. Immunol.* 42, 554-560.
- Eisenberg, D., Schwarz, E., Komaromy, M., & Wall, R. (1984) *J. Mol. Biol.* 179, 125-142.
- Heijne, G. V. (1984) *J. Mol. Biol.* 173, 243-251.
- Kolb, W. P., Haxby, J. A., Arroyave, C. M., & Müller-Eberhard, H. J. (1973) *J. Exp. Med.* 138, 428-437.
- Kozak, M. (1981) *Nucleic Acids Res.* 9, 5233-5252.
- Monahan, J. B., & Sodetz, J. M. (1981) *J. Biol. Chem.* 256, 3258-3262.
- Müller-Eberhard, H. J. (1986) *Annu. Rev. Immunol.* 4, 503-528.
- Rao, A. G., & Sodetz, J. M. (1984) *Complement* 1, 182-186.
- Rao, A. G., Howard, O. M. Z., Ng, S. C., Whitehead, A. S., Colten, H. R., & Sodetz, J. M. (1987) *Biochemistry* (preceding paper in this issue).
- Raum, D., Spence, A., Balavitch, D., Tideman, S., Merritt, A. D., Taggart, R. T., Petersen, B. H., Day, A. K., & Alper, C. A. (1979) *J. Clin. Invest.* 64, 858-865.
- Rittner, C., Hargesheimer, W., & Mollenhauer, E. (1983) *Hum. Genet.* 67, 166-169.
- Rogde, S., Olaisen, B., Gedde-Dahl, T., Jr., & Teisberg, P. (1986) *Ann. Hum. Genet.* 50, 139-144.
- Shiver, J. W., Dankert, J. R., Donovan, J. J., & Esser, A. F. (1986) *J. Biol. Chem.* 261, 9629-9636.
- Stanley, K. K., Kocher, P., Luzio, J. P., Jackson, P., & Tschopp, J. (1985) *EMBO J.* 4, 375-381.
- Steckel, E. W., York, R. G., Monahan, J. B., & Sodetz, J. M. (1980) *J. Biol. Chem.* 255, 11997-12005.
- Steckel, E. W., Welbaum, B. E., & Sodetz, J. M. (1983) *J. Biol. Chem.* 258, 4318-4324.
- Stewart, J. L., & Sodetz, J. M. (1985) *Biochemistry* 24, 4598-4602.
- Stewart, J. L., Kolb, W. P., & Sodetz, J. M. (1987) *J. Immunol.* (in press).
- Tschopp, J., Masson, D., & Stanley, K. K. (1986) *Nature (London)* 322, 831-834.
- Young, J. D. E., Cohn, Z. A., & Podack, E. R. (1986) *Science (Washington, D.C.)* 233, 184-190.
- Zalman, L. S., Wood, L. M., & Müller-Eberhard, H. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 6975-6979.

Conformation, Stability, and Folding of Interleukin 1 $\beta$ <sup>†</sup>

Stewart Craig,<sup>\*,†</sup> Ursula Schmeissner,<sup>§</sup> Paul Wingfield,<sup>§</sup> and Roger H. Pain<sup>†</sup>

Department of Biochemistry, University of Newcastle upon Tyne, NE1 7RU England, and BIOGEN SA, F-1227 Carouge, Geneva, Switzerland

Received October 8, 1986; Revised Manuscript Received January 28, 1987

**ABSTRACT:** Recombinant human interleukin 1 $\beta$  has been studied in solution with respect to its conformation, stability, and characteristics of unfolding and refolding. It is an all- $\beta$ -type, stable globular protein with a high cooperativity under conditions where refolding is reversible. The tryptophan residue is approximately 40% exposed to solvent, and the four tyrosines are 50% exposed. The fluorescence of the single tryptophan residue is quenched at pH 7.5 but dequenched by high salt, by titration to lower pH with a pK of 6.59, and by denaturants, resulting in an unusual biphasic change in fluorescence on unfolding. Both histidine and thiol residues have been excluded as being responsible for the pH dependence of fluorescence by site-directed mutagenesis and by chemical modification, respectively. The likely candidate is an aspartate or glutamate.

**T**he family of interleukin 1 proteins have a wide range of biological effects in vivo both local at the site of production by macrophages and also in distant tissues (Duff, 1985; Feldmann, 1985). These biological effects range from pro-

taglandin E production to stimulation of the acute phase response.

Preparations of authentic human IL-1 $\beta$ <sup>1</sup> are contaminated by other cytokines so it is therefore important to be able to

<sup>†</sup> This work was supported by BIOGEN SA, Geneva.

<sup>\*</sup> Address correspondence to this author.

<sup>†</sup> University of Newcastle upon Tyne.

<sup>§</sup> BIOGEN SA.

<sup>1</sup> Abbreviations: IL-1 $\beta$ , interleukin 1 $\beta$ ; CD, circular dichroism; UV, ultraviolet; Gdn-HCl, guanidine hydrochloride; MES, 2-(N-morpholino)ethanesulfonic acid; Tris, tris(hydroxymethyl)aminomethane; NATA, N-acetyltryptophanamide;  $\Delta G'_{H_2O}$ , free energy of stabilization in the absence of denaturant; EDTA, ethylenediaminetetraacetic acid.