Huber, R., Kukla, D., Ruehlmann, A., & Steigemann, W. (1971) in *Proceedings of the International Research Conference on Proteinase Inhibitors* (Fritz, H., & Tschesche, H., Eds.) pp 56–64, de Gruyter, Berlin.
Hvidt, A., & Pedersen, E. J. (1974) *Eur. J. Biochem. 48*, 333–338.
Karplus, M., & McCammon, J. A. (1979) *Nature (London) 277*, 578.
Kassel, B. (1970) *Methods Enzymol. 19*, 844–852.
Kassel, B., & Laskowski, M. (1965) *Biochem. Biophys. Res. Commun. 20*, 463–468.
Kress, L. F., & Laskowski, M., Sr. (1967) *J. Biol. Chem. 242*, 4925–4929.
Kress, L. F., Wilson, K. A., & Laskowski, M., Sr. (1968) *J. Biol. Chem. 243*, 1758–1762.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Masson, A., & Wüthrich, K. (1973) *FEBS Lett. 31*, 114–118.
Moses, E., & Hinz, H.-J. (1983) *J. Mol. Biol. 170*, 765–776.
Pershina, L., & Hvidt, A. (1974) *Eur. J. Biochem. 48*, 339–344.
Poland, D. C., & Scheraga, H. A. (1965) *Biopolymers 3*, 379–399.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.
Privalov, P. L., Ptitsyn, O. B., & Birshtein, T. M. (1969) *Biopolymers 8*, 559–571.
Privalov, P. L., Plotnikov, V. V., & Filimonov, V. V. (1975) *J. Chem. Thermodyn.*, 41–47.
Richarz, R., Sehr, P., Wagner, G., & Wüthrich, K. (1979) *J. Mol. Biol. 130*, 19–30.
Richarz, R., Nagayama, K., & Wüthrich, K. (1980) *Biochemistry 19*, 5189–5196.
Roder, H., Wagner, G. & Wüthrich, K. (1985a) *Biochemistry 24*, 7407–7411.
Roder, H., Wagner, G., & Wüthrich, K. (1985b) *Biochemistry 24*, 7396–7407.
Schellman, J. A. (1955) *C. R. Trav. Lab. Carlsberg, Ser. Chim. 29*, 230–259.
Snyder, G. H., Rowan, R., III, & Sykes, B. D. (1976) *Biochemistry 15*, 2275–2283.
Stassinopoulou, C. I., Wagner, G., & Wüthrich, K. (1984) *Eur. J. Biochem. 145*, 423–430.
Wagner, G., & Wüthrich, K. (1979) *J. Mol. Biol. 130*, 1–18.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 160*, 343–361.
Wagner, G., Tschesche, H., & Wüthrich, K. (1979a) *Eur. J. Biochem. 95*, 239–248.
Wagner, G., Kalb (Gilboa), A. J., & Wüthrich, K. (1979b) *Eur. J. Biochem. 95*, 249–253.
Wagner, G., Stassinopoulou, C. I., & Wüthrich, K. (1984) *Eur. J. Biochem. 145*, 431–436.
Wenzel, H. R., Beckmann, J., Mehlich, A., Siekmann, J., Tschesche, H., & Schutt, H. (1985) in *Modern Methods in Protein Chemistry. Review Articles* (Tschesche, H., Ed.) Vol. 2, pp 173–184, de Gruyter, Berlin.
Wlodawer, A., Walter, J., Huber, R., & Sjölin, L. (1984) *J. Mol. Biol. 180*, 301–329.
Wüthrich, K., & Wagner, G. (1978) *Trends Biochem. Sci. (Pers. Ed.) 3*, 227–230.

Complementary DNA Cloning of Complement C8β and Its Sequence Homology to C9†

Jacques-Antoine Haefliger,‡ Jürg Tschopp,*,‡ Denise Nardelli,§ Walter Wahli,§ Hans-Peter Kocher,∥ Mario Tosi,⊥ and Keith K. Stanley#

Institut de Biochimie, Université de Lausanne, CH-1066 Epalinges, Switzerland, Institut de Biologie animale, Université de Lausanne, CH-1015, Dorigny-Lausanne, Switzerland, Hoffman-La Roche & Company AG, CH-Basel, Switzerland, Unité d'Immunogénétique, Institut Pasteur, F-75724 Paris Cedex 15, France, and European Molecular Biology Laboratory, D-6900 Heidelberg, FRG

Received December 4, 1986; Revised Manuscript Received January 29, 1987

ABSTRACT: The complete amino acid sequence of mature C8β has been derived from the DNA sequence of a cDNA clone identified by expression screening of a human liver cDNA library. Comparison with the amino acid sequence of C9 shows an overall homology with few deletions and insertions. In particular, the cysteine-rich domains and membrane-inserting regions of C9 are well conserved. These findings are discussed in relation to a possible mechanism of membrane attack complex formation.

Efficient lysis of biological membranes by complement is due to the sequential assembly of the terminal complement components into the pore-forming complex C5b·C6·C7·C8·$(C9)_n$. C5b-7 forms the first membrane inserting complex to which C8 must be added before polymerization of C9 at the site of the C5b-8 complex causes major membrane damage and cell lysis (Lachmann et al., 1970). C8 therefore plays an essential intermediary role.

Human C8 is a glycoprotein of M_r 150000 comprising three polypeptide chains organized into two subunits. The α-chain [molecular mass 64 kilodaltons (kDa)] is linked to the γ-chain (25 kDa) through a disulfide bridge, while the β-chain (64 kDa) is noncovalently bound to this complex (Kolb & Müller-Eberhard, 1976). Both α- and β-chains of C8 are essential for C8 activity, but no direct role for the γ-chain has so far been described. The β-chain mediates the binding of C8 to C5b-7 (Monahan & Sodetz, 1981; Brickner et al., 1985), while C8α appears to be more intimately integrated into the

† Supported in part by grants from the Swiss National Science Foundation.
* Author to whom correspondence should be addressed.
‡ Institut de Biochimie, Université de Lausanne.
§ Institut de Biologie animale, Université de Lausanne.
∥ Hoffman-La Roche.
⊥ Institut Pasteur.
European Molecular Biology Laboratory.

0006-2960/87/0426-3551$01.50/0

lipid bilayer since it is predominantly labeled by membrane-restricted photoaffinity probes (Steckel et al., 1983; Hu et al., 1981; Podack et al., 1981; Amiguet et al., 1985). After membrane attack complex (MAC) formation C8β can be dissociated by boiling in sodium dodecyl sulfate (SDS), but C8α-γ remains integrated in the poly(C9) tubular structure (Podack, 1984). Since C8 binding is essential for C9 assembly into the MAC, it is likely that the overall function of C8 is to facilitate the membrane insertion and unfolding of C9 (Tschopp et al., 1986a; Stewart & Sodetz, 1985).

We present here the cDNA sequence for C8β, showing that it contains both the type A [C9/low-density lipoprotein (LDL) receptor] and type B [epidermal growth factor (EGF) precursor] cysteine-rich regions (Stanley et al., 1986) as well as an extensive overall homology to C9. We discuss this finding in relation to MAC formation.

Experimental Procedures

A library of human liver cDNA (Woods et al., 1982) was subcloned into the bacterial expression vector pEX1 (Stanley & Luzio, 1984) by using an adaptor strategy (Haymerle et al., 1986). This library of about 300 000 clones was screened with two different polyclonal antisera raised against C8β by using the colony blot procedure (Stanley, 1983, 1987). One antiserum was obtained by injecting C8β eluted from SDS–polyacrylamide gels into rabbits (a kind gift of Dr. E. R. Podack, Valhalla, NY; Podack, 1984). The second antiserum was raised in rabbits against C8β purified on a FPLC Mono P column (Pharmacia Inc.) in the presence of urea as described (Tschopp et al., 1986b).

cDNA inserts obtained by expression screening were used to screen a second human liver cDNA library in pUC9 (Tosi et al., 1986) by in situ colony hybridization. The largest clone obtained in this screening was sequenced by the dideoxy chain termination method (Sanger et al., 1977), using fragments generated by restriction enzymes or by ordered deletion of the complete fragment (Henikoff, 1984) cloned into the Wayne–Barnes vector (Barnes & Bevan, 1983) as modified by Sahli et al. (1985).

Affinity purification of polyclonal antisera on bacterial fusion proteins was carried out as described (Stanley, 1987).

C8β (400 pmol) purified according to Tschopp et al. (1986b) was subjected to automated Edman degradation in an Applied Biosystems 470A gas-phase sequencer in the presence of polybrene. Phenylthiohydantoin- (PTH-) amino acids from each cycle were analyzed on-line, by using a 120A PTH-amino acid analyzer (Applied Biosystems).

Protein comparison was performed by using a matrix comparison method (Argos, 1985, 1986) in which every possible amino acid span of a given length in one protein is compared with those of a second protein. Individual spans are compared on the basis of five physical parameters that are important in determining protein folding (Argos, 1986), and also by using the Dayhoff relatedness odds matrix (Barker et al., 1978). Both scores are scaled, combined, and entered into a matrix. The numbers in this matrix are then recalculated as a number of standard deviations (σ values) above the mean, and values above a certain value are displayed on a matrix plot. By repeating this process for spans of different lengths and displaying only the spans with highest σ values, more sensitivity and a higher signal to noise ratio on the plot is obtained (Argos, 1986).

Results

A library of human liver cDNA containing 300 000 clones in the bacterial expression vector pEX1 was screened by using the colony blot procedure with two different polyclonal sera raised against purified C8β. Five clones positive with both antisera were obtained after colony purification. All of these were capable of affinity purifying anti-C8β antibodies from both sera, showing that the fusion proteins expressed genuine C8β epitopes. Since the expression library was subcloned from a library in pBR322 with the restriction enzyme *Pst*I and therefore contained few full-length clones, a fragment of 0.6 kilobases (kb) from one clone was used as a probe for in situ colony hybridization of a human liver cDNA library constructed in pUC9 (Tosi et al., 1986). Four positive clones were obtained, of which one had an insert of 1.7 kb. The DNA sequence of this clone is shown in Figure 1.

In order to confirm the authenticity of this clone, purified C8β was subjected to amino acid amino-terminal analysis. The sequence originally published by Steckel et al. (1980) was confirmed, and additional amino acids were determined. The sequence of the first 14 amino acids of mature C8β was found 24 base pairs from the 5′-end of the cDNA sequence (underlined, Figure 1). Thus, the first eight amino acids in the open reading frame do not code for the sequence of the mature C8β protein. They do not display the characteristics of a normal leader peptide either, suggesting that C8β may be generated by the cleavage of a pro-C8β protein at this point.

The open reading frame deduced from the cDNA sequence contains 537 amino acids coding for an unglycosylated protein of M_r 62 000. Three potential sites for N-linked oligosaccharide attachment are present (bold font, Figure 1), which could account for the difference from the experimentally observed molecular weight of the purified protein (M_r 64 000). None of these sites occurs in a position equivalent to the N-linked oligosaccharide attachment sites of C9. C8β also contains an R-G-D sequence which forms the central recognition sequence of a number of cell-surface receptors (Ruoslahti & Pierschbacher, 1986). In C8β this sequence is present in a predicted turn region (residues 314–316) that might well be located on the surface of the protein, but there is no evidence for C8β binding to cell surfaces.

Homology with C9. Figure 2 shows the homology between C8β and C9 determined by using a matrix comparison method (see Experimental Procedures) in which the protein sequences are compared on the basis of residue physical parameters rather than simply amino acid identity. A particularly sensitive method of detecting homologous proteins is obtained (Argos, 1986) by repeating the comparison with different window lengths and combining the most significant data points. The diagonal line in Figure 2 demonstrates an overall homology between C8β and C9 over almost the whole length of the C9 molecule. Points along the lines in Figure 2 were in the range 5–9 times the standard deviation for the whole comparison matrix, indicating an extremely significant correlation. Only at the ends of the C8β molecule and in two central regions is the homology interrupted.

By use of the sequence coordinates generated by the computer from the data in Figure 2, an alignment of the two sequences was deduced (Figure 3A). In this figure both identical and conserved types of amino acid are identified. In the amino-terminal region of C8β (residues 9–189) all 14 cysteine residues align perfectly or within one residue of those found in C9 Considerable other homology also exists, notable within the class A cysteine-rich domain (residues 67–104) which is found in C8β as well as in C9. Twenty-three amino acids of conserved type are found in this region, including the cluster of acidic residues at the carboxy-terminal end (see Table I) which are thought in the LDL receptor to mediate

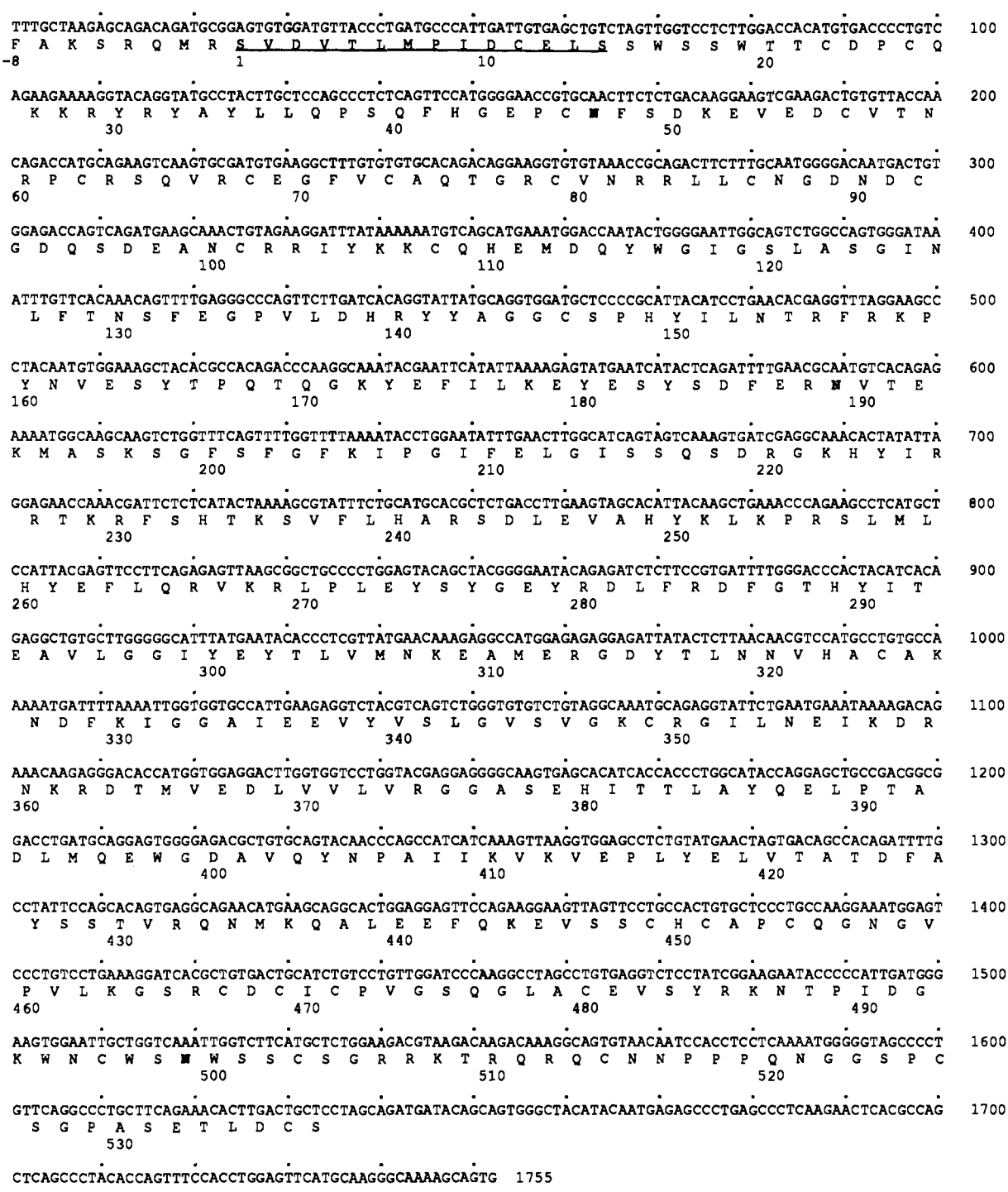

FIGURE 1: cDNA sequence and derived amino acid sequence of C8β. Underlined amino acids were obtained by amino-terminal sequencing of purified C8β protein. The numbering of amino acids refers to the amino-terminal serine in this sequence. Possible sites of N-linked oligosaccharide attachment are shown in bold characters.

apolipoprotein binding and a group of positive charges in the center of the sequences characteristics of C9.

In the central region two shifts in the alignment occur. The line in Figure 2 was used to generate the alignment shown although other alignments with smaller gaps are also possible, suggesting that this region is not highly conserved. The gap at residue 270 of C9 corresponds to the position of an intron in the C9 gene (D. Marazziti, G. Eggertsen, P. Argos, K. K. Stanley, and G. Fey, unpublished results).

In the carboxy-terminal half of C8β a further close homology is observed extending from residue 237 to residue 486 (Figure 3A). Seven out of eight of the cysteine residues in this part of C8β align with cysteine residues in the C9 sequence. Homology of many other amino acids is also observed, especially in the class B cysteine-rich region and through sequences that potentially could be involved in membrane interactions in C9 (Stanley et al., 1986). Indeed, all of the amphipathic elements of secondary structure predicted in C9, which occur in the carboxy-terminal half of the protein known to be labeled by membrane restricted photoaffinity probes, are found in regions well conserved in C8β (underlined, Figure 3A). Furthermore, the hydrophobic amino acids in these regions are mostly conserved between the two sequences. This is consistent with the observed labeling of C8β with membrane-restricted photoaffinity probes (Amiguet et al., 1985). Throughout this segment of the molecule (residues 237–486)

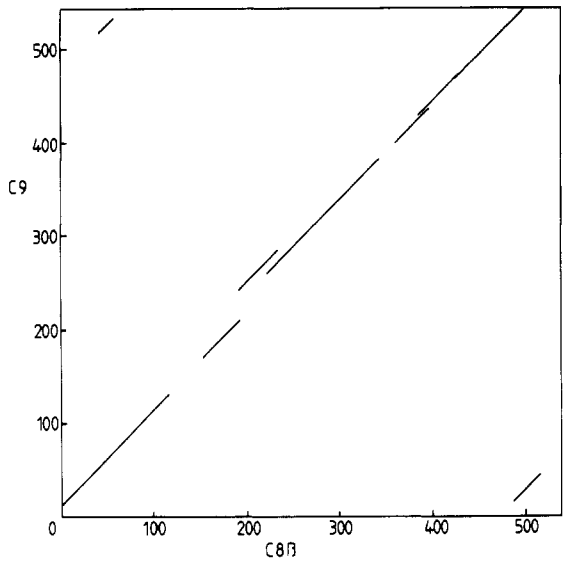

FIGURE 2: C8β–C9 homology computed on the basis of residue physical parameters. The amino acid sequence of C8β was compared with that of C9 by using 15 windows of size 7–35 amino acids and comparing on the basis of five residue physical parameters and the Dayhoff relatedness odds matrix. Points where a coefficient was 5.0σ or more above the mean of the matrix are shown.

Table I: Alignment of the Class A Cysteine-Rich Sequences Showing Charge Distribution

LDL receptor		
	4-44	DRCER-NEFQC--QDGKCISYKWVCDGSAECQDGSDESQETCLS
	45-85	VTC-KSGDFSCGGRVNRCIPQFWRCDGQVDCDNGSDEQG--CPP
	86-124	KTC-SQDEFRC--HDGKCISRQFVCDSDRDCLDGSDEAS--CPV
	125-165	LTC-GPASFQC--NSSTCIPQLWACDNDPDCEDGSDEWPQRCRG
	174-212	SPC-SAFEFHC--LSGECIHSSWRCDGGPDCKDKSDEEN--CAV
	213-251	ATC-RPDEFQC--SDGNCIHGSRQCDREYDCKDMSDEVG--CVN
	253-294	TLCEGPNKFKC--HSGECITLDKVCNMARDCRDWSDEPIKECGT
C9	78-115	DDC--GNDFQC--STGRCIKMRLRCNGDNDCGDFSDEDD--CES
C8β	67-104	VRCEG---FVCA-QTGRCVNRRLLCNGDNDCGDQSDEAN--CRR
Consensus		C F C G CI CD DC D SDE C
Peptide B		DNDCGDFSDE

to some other part of the molecule. Altogether C8β contains 28 cysteine residues, suggesting a structure with 14 disulfide bonds. In the amino-terminal half of the molecule, 12 cysteines are present, of which the latter 6 are present in the class A homology, suggesting two disulfide-bonded protein domains. Human C9 contains an extra pair of cysteines in the central nonhomologus region, but these are unlikely to be of structural or functional significance since they are also missing from mouse C9 (J. Herz, and K. K. Stanley, unpublished results). In the second half of the molecule, one extra cysteine is found in C8β at position 450, which does not have an equivalent in C9. This is therefore a candidate for disulfide bond formation to one of the five cysteines in the carboxy-terminal domain (Figure 3B). Formation of a covalent link between these two regions would create a very different structure at the carboxy

the predicted secondary structure often agrees with that of C9, reflecting the high level of homology detected by the comparison of the two sequences by using residue physical parameters.

At the carboxy terminus of C8β (residues 488–537) a duplication of the amino-terminal domain of C9 and C8β is found (Figure 3B). This stretch has only five of the six cysteines found in the corresponding region at the amino terminus, suggesting that a disulfide bond must link this extra portion

A

```
C8β   9 IDCELSSWSSWTTCDPCQKKRYRYAYLLQPSQFHGEPCNFSDKEVEDCVTNRPCRSQVR.CEGFVCAQTGRCVNRRLLCNGDNDCGDQSDEANCRRIYKK 107
C9   20 IDCRMSPWSEWSQCDPCLRQMFRSRSIEVFGQFNGKRCTDAVGDRRQCVPTEPCEDAEDDCGNDFQCSTGRCIKMRLRCNGDNDCGDFSDEDDCESEPRP 119
C8β 108 CQHEMDQYWGIGSLASGINLFTNSFEGPVLDHR..YYAGGCSPHY..ILNTRFRKPYNVESYTPQTQGKYEFILKEYESYSDFERN..............  189
C9  120 PCRDRVVEESELARTAGYGINILGMDPLSTPFDNEFYNGLCNRDRDGNTLTYYRRPWNVASLIYETKGEKNFRTEHYEEQIEAFKSIIQEKTSNFNAAIS 219
C8β 190 ..................VTEKMASKSGFSFGFKIPGIFELGISSQSDRGKHYIRRTKRFSHTKSVFLHARSDLEVAHYKLKPRSLMLHYEFLQRVKRL 270
C9  220 LKFTPTETNKAEQCCEETASSISLHGKGSFRFSYSKNETYQLFLSYSSKK.............EKMFLHVKGEIHLGRFVMRNRDVVLTTTFVDDIKAL 304
C8β 271 PLEYSYGEYRDLFRDFGTHYITEAVLGGIYEYTLVMNKEAMERGDYTLNNVHACAKNDFKIGGAIEEVYVSLGVSVGKCRGILNEIKDRNKRDTMVEDLV 370
C9  305 PTTYEKGEYFAFLETYGTHYSSSGSLGGLYELIYVLDKASMKRKGVELKDIKRCLGYHLDVSLAFSEISVGAEFNKDDCVKRGEGRAVNITSENLIDDVV 404
C8β 371 VLVRGGASEHITTLAYQELPTADLMQEWGDAVQYNPAIIKVKVE....PLYELVTATDFAYSSTVRQNMKQALEEFQKEVSSCHCAPCQGNGVPVLKGSR 466
C9  405 SLIRGGTRKYAFELKEKLLRGTVIDVTDFVNWASSINDAPVLISQKLSPIYNLVPV.KMKNAHLKKQNLERAIEDYINEFSVRKCHTCQNGGTVILMDGK 503
C8β 467 CDCICPVGSQGLACEVSYRK 486
C9  504 CLCACPFKFEGIACEISKQK 523
```

B

```
C9   18 SHIDCRMSPWSEWSQCDPCLRQMFRSRSIEVFGQFNGKRCTDAVGDRRQCV  69
C8β   7 MPIDCELSSWSSWTTCDPCQKKRYRYAYLLQPSQFHGEPCNFSDKEVEDCV  58
C8β 488 TPIDGKWNCWSNWSSCSGRRKTRQR-QCNNPPPQNGGSPCSGPASETLDCS 537
```

FIGURE 3: Aligned sequence of C8β and C9 deduced from the coordinates of the lines in Figure 2. Putative membrane spanning segments of C9 are shown underlined. Identical amino acids are shown with a solid line, and conserved amino acids according to the scheme P, G; S, T; D, E, N, Q; K, R; and AVILMCHWYF are identified with a dotted line. The valine residue at residue 293 of C9 is present in only some cDNA clones.

terminus of C8β compared with that of C9.

Discussion

We have shown that C8β and C9 have an overall sequence, cysteine pattern, and predicted structural homology with only smal differences. In particular one region, the region with homology to the LDL receptor, is very highly conserved, especially at the carboxy-terminal end, which contains five negatively charged amino acids. Antibodies raised against a peptide corresponding to this region (peptide B, Table I; Tschopp & Mollness, 1986; Tschopp et al., 1986a) cross-react with C8α and C7 but not with C8β, even though C8β, as shown here, contains this region with a high degree of conservation. Indeed, only one amino acid residue out of ten is different compared with the peptide (Table I), suggesting that the antibody binding is rather sensitive to changes in this position. When all the published class A repeats are compared (Table I), it can be seen that this residue is very variable, suggesting that it might have an exposed rather than buried location.

In the homologous region of the LDL receptor, this cluster of negative charges is preserved and has been ascribed the function of apolipoprotein binding (Yamamoto et al., 1984). Thus, modification of basic residues on the apolipoprotein (Innerarity et al., 1984) and competition with small negatively charged molecules like suramin (Schneider et al., 1982) and heparin (Goldstein et al., 1976) inhibit binding. The same reagents inhibit hemolytic activity of the membrane attack complex (Tschopp et al., 1986a; Tschopp & Masson, 1987), suggesting that this conserved cluster of negative charges plays a functional role in complement action. Further evidence from binding of a monoclonal antibody, M34, to complement component C9 suggests that regions in the amino-terminal half of this protein (where the class A homology is located) are involved in C9–C9 interaction, since the antibody prevents polymerization (Stanley et al., 1986; D. Maldonado, and K. K. Stanley, unpublished results) and its binding site is occluded after polymerization (Morgan et al., 1984). It is likely, therefore, that the effect of suramin and heparin is to block polymerization of C9 by competing with the negatively charged region on the class A homology. When a peptide corresponding to the negatively charged region of C9 was directly tested in a hemolysis assay, however, it was found to stimulate C9-dependent hemolysis although it had a powerful inhibitory effect on hemolysis with the related protein called perforin from cytotoxic T-killer cells (Tschopp et al., 1986a). A possible interpretation of this finding is that the peptide imitates the normal catalytic function of C8 in the unfolding of the C9 molecule, although the data only provide circumstantial evidence for this hypothesis, which requires testing by site-directed mutagenesis. It is interesting to note, however, that C8β does indeed contain a sequence almost identical with that found in the peptide B used in these experiments. Thus both α- and β-chains could play a role in C9 polymerization, although only C8α has been shown to form stable associations with C9 (Podack, 1984; Stewart & Sodetz, 1985). This view is supported by the observation that monoclonal antibodies raised against both α and β subunits are capable of inhibiting complement cytolysis (Abraha et al., 1986).

The dissimilarities between C9 and C8β are potentially important in describing the functional difference between the two molecules. Of particular note in this respect is the presence of a extra disulfide-bonded domain at the carboxy terminus of C8β and the deletion of a stretch of amino acids in the center of the molecule. Contained within this central region in C9 is the epitope of monoclonal antibody bC5 (Stanley et al., 1985; K. K. Stanley, unpublished results) which only recognizes C9 in its unfolded conformation (Mollness et al., 1985). This antibody also cross-reacts with C5b-7 and C5b-8 but not with the individual components in their globular form (Tschopp et al., 1986a), suggesting that the same epitope must be present on other terminal components in their unfolded form. It does not appear, however, to be located on C8β. The truncation of this region, which in C9 appears to be involved in unfolding, suggests that C8β may not be able to undergo the same conformational change as C9 It will be interesting to see if this region is present in C8α, which is apparently integrated into the poly(C9) tubule of the MAC (Podack, 1984).

Acknowledgments

We thank Dr. P. Argos for his critical appraisal of the manuscript.

Registry No. Complement C9, 80295-59-6; complement C8, 80295-58-5; complement C8 (human liver β-chain protein moiety reduced), 108007-31-4.

References

Abraha, A., Bailyes, E. M., Richardson, P. J., Campbell, A. K., & Luzio, J. P. (1986) *Biochem. Soc. Trans. 14*, 779.

Amiguet, P., Brunner, J., & Tschopp, J. (1985) *Biochemistry 24*, 7328–7334.

Argos, P. (1985) *EMBO J. 4*, 1351–1355.

Argos, P. (1987) *J. Mol. Biol. 193*, 385–396.

Barker, W. C., Ketcham, L. K., & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 359–362, NBRF, Washington, D.C.

Barnes, W. M., & Bevan, M. (1983) *Nucleic Acids Res. 11*, 349–369.

Brickner, A., Lambert, S. J., & Sodetz, J. M. (1985) *Complement 2*, 13 (Abstr.).

Haymerle, H., Herz, J., Bressan, G. M., Frank, R., & Stanley, K. K. (1986) *Nucleic Acids Res. 14*, 8615–8624.

Henikoff, S. (1984) *Gene 28*, 351–359.

Hu, V. W., Esser, A. F., Podack, E. R., & Wisnieski, B. J. (1981) *J. Immunol. 127*, 380–386.

Kolb, W. P., & Müller-Eberhard, H. J. (1976) *J. Exp. Med. 143*, 1131–1139.

Goldstein, J. L., Basu, S. K., Brunschede, G. Y., & Brown, M. S. (1976) *Cell (Cambridge, Mass.) 7*, 85–95.

Innerarity, T. L., Weisgraber, K. H., Arnold, K. S., Rall, S. C., & Mahley, R. W. (1984) *J. Biol. Chem. 259*, 7261–7267.

Lachmann, P. J., Murun, E. A., & Weissmann, G. (1970) *Immunology 19*, 983–986.

Mollnes, T. E., Lea, T., Harboe, M., & Tschopp, J. (1985) *Scand. J. Immunol. 22*, 183–195.

Monahan, J. B., & Sodetz, J. M. (1981) *J. Biol. Chem. 256*, 3258–3262.

Morgan, P. B., Luzio, J. P., & Campbell, A. K. (1984) *Biochem. Biophys. Res. Commun. 118*, 616–622.

Pfeffer, S., & Ullrich, A. (1985) *Nature (London) 313*, 184.

Podack, E. R. (1984) *J. Biol. Chem. 259*, 8641–8647.

Podack, E. R., Stoffel, W., Esser, A. F., & Müller-Eberhard, H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4544–4548.

Ruoslahti, E., & Pierschbacher, M. D. (1986) *Cell (Cambridge, Mass.) 44*, 517–518.

Sahli, R., McMaster, G. K., & Hirt, B. (1985) *Nucleic Acids Res. 13*, 3617–3633.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schneider, W. J., Beisegel, U., Goldstein, J. L., & Brown, M. S. (1982) *J. Biol. Chem. 257*, 2664-2673.
Stanley, K. K. (1983) *Nucleic Acids Res. 11*, 4077-4092.
Stanley, K. K. (1987) *Methods Mol. Biol.* (in press).
Stanley, K. K., & Luzio, J. P. (1984) *EMBO J. 3*, 1429-1434.
Stanley, K. K., Kocher, H.-P., Luzio, J. P., Jackson, P., & Tschopp, J. (1985) *EMBO J. 4*, 375-382.
Stanley, K. K., Page, M., Campbell, A. K., & Luzio, J. P. (1986) *Mol. Immunol. 23*, 451-458.
Steckel, E. W., York, R. G., Monahan, J. B., & Sodetz, J. M. (1980) *J. Biol. Chem. 255*, 11997-12005.
Steckel, E. W., Wellbaum, B. A., & Sodetz, J. M. (1983) *J. Biol. Chem. 258*, 4318-4324.
Stewart, J. L., & Sodetz, J. M. (1985) *Biochemistry 24*, 4598-4602.
Sudhof, T. C., Goldstein, J. L., Brown, M. S., & Russell, D. W. (1985) *Science (Washington, D.C.) 228*, 815-822.
Tosi, M., Duponchel, C., Bougarel, P., Colomb, M., & Meo, T. (1986) *Gene 42*, 265-272.
Tschopp, J., & Mollness, T.-E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4223-4227.
Tschopp, J., & Masson, D. (1987) *Mol. Immunol.* (in press).
Tschopp, J., Podack, E. R., & Müller-Eberhard, H. J. (1985) *J. Immunol. 134*, 495-499.
Tschopp, J., Masson, D., & Stanley, K. K. (1986a) *Nature (London) 322*, 831-834.
Tschopp, J., Penea, F., Schifferli, J., & Späth, P. (1986b) *Scand. J. Immunol. 24*, 715-720.
Woods, D. E., Markham, A. F., Ricker, A. T., Goldberger, G., & Colten, H. R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5661-5665.
Yamamodo, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., & Russell, D. W. (1984) *Cell (Cambridge, Mass.) 39*, 27-38.

Complementary DNA and Derived Amino Acid Sequence of the α Subunit of Human Complement Protein C8: Evidence for the Existence of a Separate α Subunit Messenger RNA†

A. Gururaj Rao,‡ O. M. Zack Howard,‡ Simon C. Ng,‡ Alexander S. Whitehead,§,‖ Harvey R. Colten,§,⊥ and James M. Sodetz*,‡

Department of Chemistry and School of Medicine, University of South Carolina, Columbia, South Carolina 29208, and Division of Cell Biology, Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Received January 30, 1987

ABSTRACT: The entire amino acid sequence of the α subunit (M_r 64 000) of the eighth component of complement (C8) was determined by characterizing cDNA clones isolated from a human liver cDNA library. Two clones with overlapping inserts of net length 2.44 kilobases (kb) were isolated and found to contain the entire α coding region [1659 base pairs (bp)]. The 5′ end consists of an untranslated region and a leader sequence of 30 amino acids. This sequence contains an apparent initiation Met, signal peptide, and propeptide which ends with an arginine-rich sequence that is characteristic of proteolytic processing sites found in the pro form of protein precursors. The 3′ untranslated region contains two polyadenylation signals and a poly(A) sequence. RNA blot analysis of total cellular RNA from the human hepatoma cell line HepG2 revealed a message size of ~2.5 kb. Features of the 5′ and 3′ sequences and the message size suggest that a separate mRNA codes for α and argues against the occurrence of a single-chain precursor form of the disulfide-linked α-γ subunit found in mature C8. Analysis of the derived amino acid sequence revealed several membrane surface seeking domains and a possible transmembrane domain. These occur in a cysteine-free region of the subunit and may constitute the structural basis for α interaction with target membranes. Analysis of the carbohydrate composition indicates 1 or 2 asparagine-linked but no O-linked oligosaccharide chains, a result consistent with predictions from the amino acid sequence. The α subunit contains segments homologous to the negatively charged, cysteine-rich repeat sequence found in low-density lipoprotein receptor and to the cysteine-rich epidermal growth factor type sequence found in a number of proteins. Most significantly, it exhibits a striking overall homology to human C9, with values of 24% on the basis of identity and 46% when conserved substitutions are allowed. As described in an accompanying report [Howard, O. M. Z., Rao, A. G., & Sodetz, J. M. (1987) *Biochemistry* (following paper in this issue)], this homology also extends to the β subunit of C8.

Human C8 is a serum glycoprotein constituent of C5b-9, the cytolytic complex of complement composed of C5b, C6, C7, C8, and C9 (Müller-Eberhard, 1986). Assembly of this complex is initiated by conversion of C5 to C5b and proceeds in a sequential manner:

$$\text{C5b} \xrightarrow{\text{C6}} \text{C5b-6} \xrightarrow{\text{C7}} \text{C5b-7} \xrightarrow{\text{C8}} \text{C5b-8} \xrightarrow{n\text{C9}} \text{C5b-9}_n$$

The intermediate C5b-7 complex contains a metastable lipid binding site that anchors the nascent complex to target cell membranes. Once on the membrane, C5b-7 binds C8 to form a tetramolecular complex, C5b-8. This complex is capable of slowly lysing erythrocyte membranes as well as some nu-

† This research was supported by NIH Grants AI16856 and BRSG S07 RR07160 and American Heart Association Established Investigator Award 82-121 to J.M.S. and NIH Grant AI24836 to H.R.C.
‡ University of South Carolina.
§ Children's Hospital, Harvard Medical School.
‖ Present address: Division of Immunology, Children's Hospital, Harvard Medical School, Boston, MA 02115.
⊥ Present address: Department of Pediatrics, Washington University School of Medicine, St. Louis, MO 63110.

0006-2960/87/0426-3556$01.50/0