

# The Amino Acid Sequence of *Limulus* C-reactive Protein

EVIDENCE OF POLYMORPHISM\*

(Received for publication, January 9, 1986)

Nga Yen Nguyen, Akira Suzuki, Robert A. Boykins, and Teh-Yung Liu

From the Division of Biochemistry and Biophysics, Office of Biologics Research and Review, Food and Drug Administration, Bethesda, Maryland 20892

The amino acid sequence, the positions of the disulfide bonds, and the site of glycosylation for the three subunits of *Limulus* C-reactive proteins (CRPs) 1.1, 1.4, and 3.3 have been established. The three subunits were shown to exist approximately in equimolar amount and are tightly associated. The hexagonal structure of *Limulus* CRP, as revealed by electron microscopic studies of Fernandez-Moran *et al.* (Fernandez-Moran, H., Marchalonis, J., and Edelman, G. M. (1968) *J. Mol. Biol.* 32, 467-469) might consist of two each of the subunits. The three subunits share an identical amino-terminal sequence of 44 residues and a carboxyl-terminal sequence from residues 206 to 218. Microheterogeneity exists to the extent of 10 to 11% for the entire protein. The positions of 6 half-cystines that form the three disulfide bonds and the site of glycosylation are constant in all subunits. Sequence analyses of peptides derived from enzymatic and chemical cleavages of affinity purified *Limulus* CRP indicate that subunits other than the three mentioned above exist in the hemolymph. *Limulus* CRP is therefore polymorphic. Topological analyses of *Limulus* CRPs, human CRP, rabbit CRP, human amyloid P-component, and Syrian hamster female protein indicate that the seven proteins may originate from the same ancestral gene. Using the topological data generated from the amino acid sequences of the proteins, we calculate that human and *Limulus* CRPs diverged about 500 million years ago. This figure is in general agreement with the evolutionary distance postulated by anthropological estimation of 400-500 million years.

In humans the acute phase response to inflammation, infections, or various forms of tissue injury is characterized by a significant increase in the circulating concentration of a number of plasma proteins (1). The C-reactive protein (CRP)<sup>1</sup> is one of the acute phase reactants found in humans and most mammals (2). CRP exhibits the unique property of reacting with and precipitating pneumococcal C-polysaccharide in the presence of calcium ions (3). Although CRP shares many properties with the immunoglobulins, such as the ability to

promote agglutination (4), complement fixation (5), bacterial capsular swelling (6), phagocytosis (7), and precipitation of polycationic and polyanionic compounds (8), statistically significant sequence homology does not exist between CRPs and the immunoglobulins (9, 10). Thus, CRP has been assigned to a new super family of proteins. The existence of CRP in an invertebrate, *Limulus* polyphemus, has recently been reported (11). The major difference between the *Limulus* and mammalian CRPs lies in the production and release of proteins into the circulatory system. In the mammalian system, CRP is produced only under inflammatory conditions, whereas CRP is a normal component of the invertebrate hemolymph. Because invertebrates do not appear to have immunoglobulins, it has been suggested that CRP might be serving as a primitive form of immunoglobulin in *Limulus*. For these reasons, we have undertaken an investigation of *Limulus* CRP at both the protein and DNA levels. In this paper, we report the results of our protein-based studies of *Limulus* C-reactive proteins.

## EXPERIMENTAL PROCEDURES<sup>2</sup>

### RESULTS AND DISCUSSIONS

This study was initiated to establish the primary structure of *Limulus* CRP and to compare it with its mammalian counterparts, human and rabbit CRPs. Because *Limulus* is believed to be a living fossil, such a comparison was expected to provide insight into the evolutionary origin of CRPs.

Analyses by Edman degradation of intact *Limulus* CRP and peptide fragments derived from chemical and enzymatic cleavages soon revealed, however, that *Limulus* CRP purified by affinity chromatography, unlike its mammalian counterparts, is not a single protein species. *Limulus* CRP was found to be microheterogeneous. A great deal of difficulty was encountered in aligning the peptides to establish the primary structure of *Limulus* CRPs. Molecular cloning of *Limulus* CRP genes was, therefore, simultaneously started to complement protein structural analysis. Without the prior DNA sequence data it was not clear that the multiple assignments obtained in the sequencing were due to contaminants or to *bona fide* microheterogeneity. Microheterogeneities were indicated by the unique sequences predicted from the DNA sequences of the cloned genes (17). The amino acid sequences

<sup>2</sup> Portions of this paper (including "Experimental Procedures," Figs. 1S-10S, and Tables IS-XVIIS) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 86M-74, cite the authors, and include a check or money order for \$12.80 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> The abbreviations used are: CRP, C-reactive protein; HFP, hamster female protein; SAP, serum amyloid protein; HPLC, high performance liquid chromatography; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; PTH, phenylthiohydantoin; ABI, Applied Biosystems, Inc.; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; PAMs, accepted point mutations per 100 residues.

deduced from DNA analysis of three *Limulus* CRP genes were useful and sometimes essential to align the complex data derived from peptide sequence analysis. Protein structural studies on the other hand provided information that could not have been deduced from DNA analysis of the genes. These include the positions of disulfide bonds, the site of glycosylation, and the proportion of each subunit present in the affinity purified CRPs.

**The Primary Structure of *Limulus* CRPs**—The strategy used for sequence determination of *Limulus* CRPs is shown in Fig. 2. Information essential for the establishment of the primary structure for *Limulus* CRP subunits 1.1, 1.4, and 3.3 is presented in Tables IS through XVS and Figs. 1S–9S of the Miniprint Section. These data confirm the amino acid sequence deduced from DNA analysis of the three *Limulus* CRP genes and provide evidence that all three CRP genes are being expressed. The three subunits share an identical amino-terminal sequence up to residue 44 and a carboxyl-terminal sequence from residue 206 to residue 218. In the entire protein, microheterogeneity exists to the extent of 10–11%. In addition to the three subunits, a number of peptides were isolated and sequenced (Table XVIS) which do not fit into the predicted amino acid sequence of the isolated genes. Nevertheless, these peptides are thought to be derived from *Limulus* CRP genes because of their extensive sequence homology with the established CRP sequences. Based on their yield, these peptides were estimated to represent <10% of the affinity purified *Limulus* CRP.

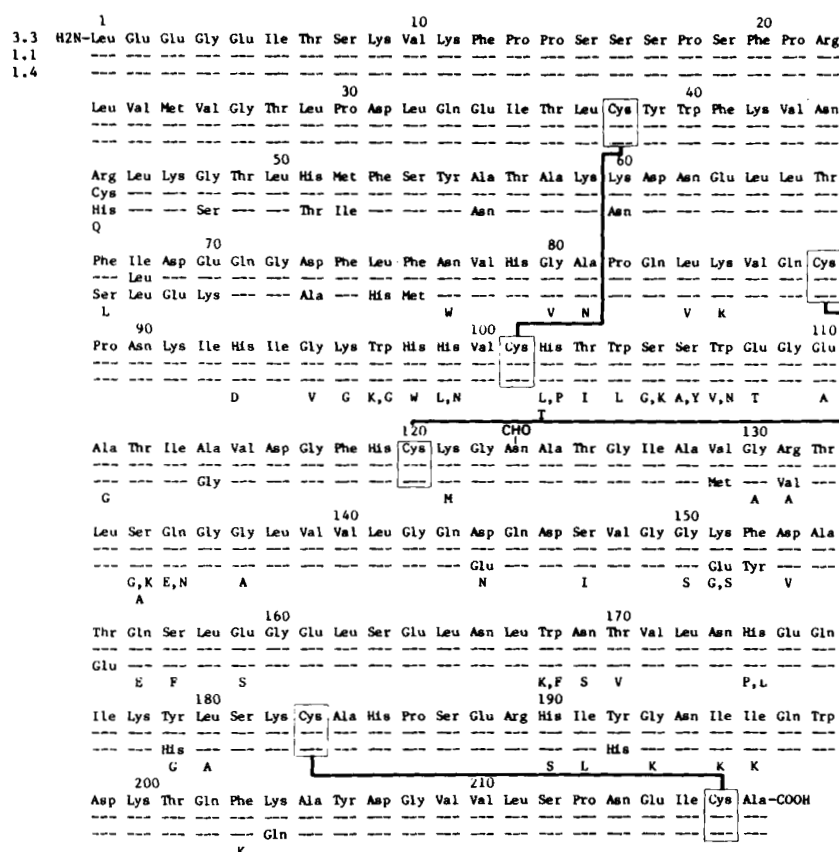
Examination of the *Limulus* CRP sequences shown in Fig. 1 revealed that subunit 1.1 has 7 half-cystines, whereas the other two subunits, 1.4 and 3.3, each possess 6 half-cystines.

The positions of these 6 half-cystines are conserved in the three subunits and they contribute to the formation of 3 intramolecular disulfide bonds (see discussion below). The 7th half-cystine found in position 45 of subunit 1.1, therefore, either must exist with a free sulfhydryl group or is involved in an intermolecular disulfide bond. Exposure of *Limulus* CRPs to 6 M guanidinium hydrochloride followed by the addition of iodoacetamide resulted in the formation of carboxymethylamido-cysteine which was quantitatively analyzed on the amino acid analyzer as carboxymethylcysteine. The yield of carboxymethylcysteine was 0.25–0.30 mol/mol of subunit CRP analyzed, indicating that 25–30% of the affinity purified CRPs had a free sulfhydryl group.

Advantage was taken of the fact that each subunit of *Limulus* CRP contains a different number of methionine residues and some at different positions in the proteins. While Met-25 is present in all three subunits, Met-52 is present in subunits 3.3 and 1.1 and is missing in subunit 1.4, and Met-76 and Met-129 are only present in subunit 1.4. Cyanogen bromide cleavage followed by Edman degradations yielded PTH-derivatives adjacent to the carboxyl termini of the methionine residues. The yields of these PTH-derivatives provided a means to estimate the amount of each subunit in the CRP preparation. The data shown in Table III suggest that the three subunits are present approximately in equimolar amounts.

An attempt was made to separate subunit 1.1 from subunits 1.4 and 3.3 by the use of a mercury-Sepharose column. Previously we were successful in separating homologous proteins with and without the free sulfhydryl group on mercury-Sepharose (12). All subunits of *Limulus* CRP bound to mercury-

FIG. 1. Primary structure of *Limulus* C-reactive protein. Three-letter coded amino acid sequences, designated as 3.3, 1.1, and 1.4 are determined by both amino acid and nucleotide sequence analyses. Single letter residues are derived from amino acid sequencing alone.



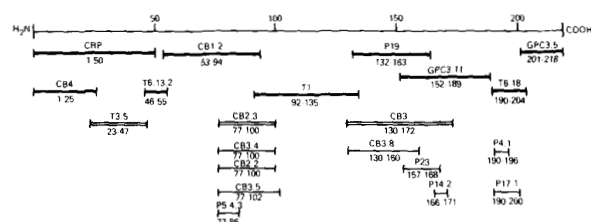


FIG. 2. Sequence strategy for *Limulus* C-reactive protein. Heavy lines, peptides used in sequence determination of subunits 3.3 and 1.1. Double lines, peptides used in sequence determination of subunit 1.4. Single lines, peptides that belong to subunits other than 1.1, 1.4, and 3.3 (see Table XVIII).

Sephacryl quantitatively. Elution of the column with buffer containing 1 M KCl failed to dissociate any subunit. The subunits coeluted when 0.5 mM  $\text{HgCl}_2$  was included in the buffer. These results suggest that the subunits aggregate perhaps to form a hexagonal structure as was revealed by electron microscopic studies (13). From the stoichiometry of the subunits present, it is suggested that each hexagonal structure generally consists of two each of the 3.3, 1.1, and 1.4 subunits. The hexagonal structure can be dissociated in 1% sodium dodecyl sulfate or 6 M guanidinium hydrochloride. Under dissociating conditions the mercury-Sephacryl column fails to function as an affinity column for sulfhydryl-proteins.

**Position of Disulfide Bonds**—The positions of the half-cysteine residues in the three subunits have been independently established by sequence analysis of various peptides which confirmed the data deduced from DNA analysis of the genes. Three disulfide-containing peptides were isolated after thermolysin cleavages of intact *Limulus* CRP. Edman degradation of these peptides established disulfide bonds Cys<sup>38</sup>-Cys<sup>101</sup>, Cys<sup>88</sup>-Cys<sup>120</sup>, and Cys<sup>183</sup>-Cys<sup>217</sup> (Tables XIII-XV and Fig. 9S). These data are consistent with a single disulfide pairing pattern but, in consideration of the existence of a free sulfhydryl group in species 1.1, the possibility of the existence of minor alternative forms cannot be excluded. The disulfide bond between Cys<sup>38</sup>-Cys<sup>101</sup> is conserved in human CRP, rabbit CRP, human SAP, and the female protein of Syrian hamster.

**Site of Glycosylation**—*Limulus* CRP is a glycoprotein containing glucosamine and mannose (11). Treatment of the protein with hydrofluoric acid or trifluoromethanesulfonic acid failed to dissociate glucosamine from the protein. The site of glycosylation was established as Asn-123 based on the isolation and characterization of peptides T1 and e4, which contain Asn-123. Of all the peptides isolated and characterized, only T1 and e4 were found to contain glucosamine. Upon Edman degradation, a blank occurred at position 123, corresponding to Asn-123 in these two peptides (Tables IV and XIII). The tripeptide sequence Asn<sup>123</sup>-Ala-Thr is consistent with the consensus amino acid sequence Asn-X-Thr/Ser as the glycosylation site in glycoproteins (14, 15). This sequence is conserved in all three subunits of *Limulus* CRP and occurs only once in the protein. Peptide sequence analysis showed that serine and threonine are not likely to be the sites of glycosylation. Galactosamine normally associated with serine/threonine at the site of glycosylation is absent from the protein.

**Topologies of CRPs and Related Proteins**—CRP, serum amyloid protein (SAP), and the female protein of the Syrian hamster (HFP) are similar, yet distinct, proteins. These proteins exhibit calcium-dependent ligand-binding activities and display a characteristic pentagonal structure under the electron microscope (16). The exception is *Limulus* CRP, for which a hexagonal structure was revealed by electron micro-

TABLE I

Alignment of CRPs, SAP, and female protein

Positions omitted from the original alignment 221 to 227 of hamster female protein.

|                          | Number of Differences |    |    |     |     |     |     |
|--------------------------|-----------------------|----|----|-----|-----|-----|-----|
|                          | 1                     | 2  | 3  | 4   | 5   | 6   | 7   |
| 1 <i>Limulus</i> CRP 3.3 |                       | 2  | 23 | 160 | 161 | 160 | 162 |
| 2 <i>Limulus</i> CRP 1.1 | 1                     |    | 22 | 160 | 161 | 160 | 162 |
| 3 <i>Limulus</i> CRP 1.4 | 11                    | 10 |    | 162 | 162 | 159 | 160 |
| 4 Rabbit CRP             | 74                    | 74 | 75 |     | 50  | 113 | 113 |
| 5 Human CRP              | 74                    | 74 | 74 | 24  |     | 101 | 104 |
| 6 Human SAP              | 73                    | 73 | 73 | 55  | 49  |     | 66  |
| 7 Hamster Female Protein | 74                    | 74 | 73 | 55  | 50  | 37  |     |

Percent differences

TABLE II

PAMs computed from per cent differences

|                        | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
|------------------------|-------|-------|-------|-------|-------|-------|-------|
| <i>Limulus</i> CRP 3.3 | 0.0   | 0.9   | 11.3  | 184.8 | 185.7 | 182.2 | 184.0 |
| <i>Limulus</i> CRP 1.1 | 0.9   | 0.0   | 10.8  | 184.8 | 185.7 | 182.2 | 184.0 |
| <i>Limulus</i> CRP 1.4 | 11.3  | 10.8  | 0.0   | 192.0 | 189.2 | 178.8 | 177.3 |
| Rabbit CRP             | 184.8 | 184.8 | 192.0 | 0.0   | 29.3  | 94.6  | 93.8  |
| Human CRP              | 185.7 | 185.7 | 189.2 | 29.3  | 0.0   | 77.1  | 80.3  |
| Human SAP              | 182.2 | 182.2 | 178.8 | 94.6  | 77.1  | 0.0   | 41.6  |
| Hamster Female Protein | 184.0 | 184.0 | 177.3 | 93.8  | 80.3  | 41.6  | 0.0   |

TABLE III

Edman degradation of mixture of CNBr fragments

2.4 nM of *Limulus* CRP purified by phosphorylcholine affinity column were desalted, lyophilized, and subjected to cyanogen bromide cleavage. The entire cleavage products were subjected to automated sequence analysis using ABI model 470A.

| Source        | Subunit | Sequence                    | PTH-derivatives |       |         |       |
|---------------|---------|-----------------------------|-----------------|-------|---------|-------|
|               |         |                             | Cycle 1         | Yield | Cycle 2 | Yield |
|               |         |                             |                 | nM    |         | nM    |
| 3.3, 1.4, 1.1 |         | Leu <sup>1</sup> -Glu       | Leu             | 0.98  | Glu     | 0.91  |
| 3.3, 1.4, 1.1 |         | Met <sup>25</sup> -Val-Gly  | Val             | 1.10  | Gly     | 0.98  |
| 3.3, 1.1      |         | Met <sup>52</sup> -Phe-Ser  | Phe             | 0.64  | Ser     | 0.18  |
| 1.4           |         | Met <sup>76</sup> -Asn-Val  | Asn             | 0.16  | Val     |       |
| 1.4           |         | Met <sup>129</sup> -Gly-Val | Gly             | 0.24  | Val     | 0.58  |

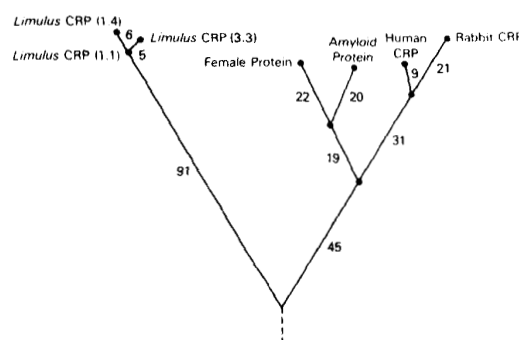


FIG. 3. Evolutionary tree of CRPs and related proteins. The method described by Dayhoff (18) was used to derive this tree. The branch lengths are proportional to the number of amino acid changes which we infer to have occurred and are shown as PAMs. The total length = 268.2 PAMs.

scopic studies (13). DNA analyses of the genes coding for these proteins, as well as protein sequence analysis, revealed extensive sequence homology (17). In order to establish the evolutionary relatedness of the seven proteins in question,

*Limulus* CRPs 1.1, 1.4, and 3.3, human CRP, rabbit CRP, human SAP, and hamster HFP, we have compared and analyzed their sequences by a computer program (18). Fig. 10S shows the primary structures of the seven proteins. Their percent differences in amino acid sequence are summarized in Table I. Based on these data a PAMs matrix (accepted point mutations/100 residues) was constructed (Table II) to generate topologies for the seven proteins using the MATTOP program (18). Five trees with the shortest absolute length were generated but only one positive tree was found. It is diagrammed in Fig. 3. The "trunk" of the tree is placed arbitrarily about in the middle. Using the relatively well-established distance between human and rabbit as 150 million years, which indicates a rate of change of 20 PAMs/100 million years, we calculate that human and *Limulus* CRPs diverged at least 500 million years ago. This figure is in general agreement with the evolutionary distance postulated by anthropological estimation of at least 500 million years (21).

The tree suggests that hamster HFP and human SAP have arisen during vertebrate evolution. The data would suggest that the branch point to their common ancestor was at least 200 million years ago. From the tree alone it would not be clear whether hamster HFP and human SAP are actually the same protein in different animals or different proteins, but from other evidence, including their mode of synthesis and ligand binding properties (19, 20), it is likely that they are different proteins. If they are different, the gene duplication would have been long before the mammalian radiation, maybe 100 million years ago. The dates mentioned could have been considerably earlier if gene conversion operated to keep the duplicated genes very similar. This appears to be a common phenomenon in multigene families. In tracing the genealogy of proteins, the relations of greatest interest are those between sequences that are about 15–25% identical. This "twilight zone" is thought to be the place where one must look for the roots of the protein family tree, to find molecules that diverged early in the course of their evolution (22). The data presented in Table I indicate that the sequence identity between *Limulus* CRPs and their mammalian counterparts is about 25%. Thus, the tree constructed based on these data would suggest strongly that human CRP, rabbit CRP, *Limulus* CRPs, human SAP, and hamster HFP share a common ancestral gene. In addition, the tree discloses the order of lineage branching and information about the time of divergence among lineages.

**Acknowledgments**—We wish to thank Drs. Derick Gates and Suzanne Epstein for a critical review of this manuscript. The assistance of Ellen Kirshbaum in preparing this manuscript is deeply appreciated.

#### REFERENCES

- Koj, A. (1974) in *Structure and Function of Plasma Proteins* (Allison, A. C., ed) Vol. 1, pp. 73–131, Plenum Press, New York
- Tillet, W. S., and Francis, T. (1930) *J. Exp. Med.* **52**, 561–571
- MacLeod, C. M., and Avery, O. T. (1941) *J. Exp. Med.* **73**, 191–200
- Gal, K., and Miltenyi (1955) *Acta Microbiol. Acad. Sci. Hung.* **3**, 41–51
- Kaplan, M. H., and Volanakis, J. E. (1974) *J. Immunol.* **112**, 2135–2147
- Löfstrom, G. (1944) *Br. J. Exp. Pathol.* **25**, 21–26
- Hokama, Y., Coleman, M. K., and Riley R. F. (1962) *J. Bacteriol.* **83**, 1017–1024
- Siegel, J. E., Rent, R., and Gewurz, H. (1974) *J. Exp. Med.* **140**, 631–647
- Oliveira, E. B., Gotschlich, E. C., and Liu, T. (1979) *J. Biol. Chem.* **254**, 489–502
- Wang, C.-M., Nguyen, N. Y., Yonaha, K., Robey, F., and Liu, T.-Y. (1982) *J. Biol. Chem.* **257**, 13610–13615
- Robey, F. A., and Liu, T.-Y. (1981) *J. Biol. Chem.* **256**, 969–975
- Kort, A. A., and Liu, T.-Y. (1973) *Biochemistry* **12**, 320–327
- Fernandez-Moran, H., Marchalonis, J., and Edelman, G. M. (1968) *J. Mol. Biol.* **32**, 467–469
- Neuberger, A., Gottschalk, A., Marshall, R. D., and Spiro, R. G. (1972) in *Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., ed) pp. 450–490, Elsevier, New York
- Spiro, R. G. (1973) *Adv. Protein Chem.* **27**, 349–467
- Osmond, A. P., Friedenson, B., Gerwurz, H., Painter, R. H., Hofman, T., and Shelton, E. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 739–743
- Nguyen, N. Y., Suzuki, A., Cheng, S. M., Zon, G., and Liu T.-Y. (1986) *J. Biol. Chem.* **261**, 10450–10455
- Dayhoff, M. O. (1979) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed) Vol. 5, Suppl. 3, pp. 1–8, National Biomedical Research Foundation, Washington, D. C.
- Mantzouranis, E. C., Dowton, S. B., Whitehead, A. S., Edge, M. D., Bruns, G. A. P., and Colten, H. R. (1985) *J. Biol. Chem.* **260**, 7752–7756
- Dowton, S. B., Woods, D. E., Mantzouranis, E. C., and Colten, H. R. (1985) *Science* **228**, 1206–1208
- Lankester, E. R. Q. *J. Microsc. Sci.* **XXI**, 504, 1881
- Doolittle, R. F. (1981) *Science* **214**, 149–159
- Roth, M. (1971) *Anal. Chem.* **43**, 880–882
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, L. E., Jr., and Weber, P. (1981) *Anal. Biochem.* **118**, 131–137
- Mort, A. J., and Lampert, D. T. A. (1977) *Anal. Biochem.* **82**, 289–309
- Simpson, R. J., Neuberger, M. R., and Liu, T.-Y. (1976) *J. Biol. Chem.* **251**, 1936–1940
- Spackman, D. H. (1963) *Fed. Proc.* **22**, 244
- Spackman, D. H., Stein, N. H., and Moore, S. (1958) *Anal. Chem.* **30**, 1190–1206
- Boykins, R. A., and Liu, T.-Y. (1980) *J. Biochem. Biophys. Methods* **2**, 71–78
- Gross, E., and Witkop, B. (1962) *J. Biol. Chem.* **237**, 1856–1860
- Dixon, H. B. F., and Perham, R. N. (1968) *Biochem. J.* **109**, 312–314
- Konigsberg, W. (1972) *Methods Enzymol.* **25**, 185–188
- Crestfield, A. M., Moore, S., and Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627
- Link, T. P., and Stark, G. R. (1968) *J. Biol. Chem.* **243**, 1082–1088
- Yonaha, K., Elliot, S. D., and Liu, T.-Y. (1982) *J. Protein Chem.* **1**, 317–333
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E., and Dryer, W. J. (1981) *J. Biol. Chem.* **256**, 7990–7997
- Tarr, G. E. (1975) *Anal. Biochem.* **63**, 361–370
- Zimmerman, C. L., Appella, E., and Pisano, J. J. (1977) *Anal. Biochem.* **77**, 569–573

Supplemental Material  
to  
Structural Polymorphism of Limulus C-reactive protein

by  
Nga Yen Nguyen, Akira Suzuki, Robert A. Boykins and Teh-Yung Liu

EXPERIMENTAL PROCEDURES

MATERIALS

1. Limulus CRP was purified from Limulus hemolymph supplied by the Associates of Cape Cod Inc., Massachusetts.
2.  $^{14}\text{C}$  labeled iodoacetamide (10–25 mCi/mmol) and methyl iodide (5–20 mCi/mmol) were obtained from New England Nuclear. Enzymes were purchased from Boehringer Mannheim and Worthington. Buffers and reagents were obtained from Burdick & Jackson, Fisher Scientific and Sigma Chemical Co. Reagents used for amino acid sequence analysis were supplied by Beckman Instruments and Applied Biosystems, Inc.

METHODS

1. Isolation and purification of Limulus CRP: Limulus CRP was purified from Limulus hemolymph by phosphorylcholine affinity chromatography as described previously (9–11). Approximately 38–50 mgs of CRP were obtained from 50 ml of hemolymph.
2. Purification of peptides: Peptides derived from chemical or enzymatic cleavage of CRP were purified by gel chromatography, ion exchange chromatography or by HPLC, alone or in combination. Detailed information related to each purification is described in the Legends to Figures. Identification of peptides by fluorimetry was performed as described (23).
3. Polyacrylamide gel electrophoresis in SDS: SDS-PAGE was carried out on 10%–12.5% polyacrylamide gels (24); slab gels (15 x 18cm) were run at a constant power of 3 mA for 16 hrs; cylindrical gels (6 mm I.D.) were run at a constant power of 1 mA for 4 hrs.
4. Immunological studies: Immunodiffusion of CRP was performed (11) on agarose plates set in Tris-saline- $\text{Ca}^{2+}$  buffer, in the presence of 0.1% sodium azide as preservative.
5. Deglycosylation: Carbohydrate removal was carried out in the presence of trifluoromethanesulfonic acid (25) or hydrofluoric acid (26). Deglycosylated proteins were used for preparation of some of the peptides used in sequence analysis.
6. Amino acid and carbohydrate analyses: Samples were hydrolyzed in *vacuo* in 4 N methanesulfonic acid (27), and their amino acid or sugar composition was determined as described either on a Beckman 121 M amino acid analyzer (28,29) or on a sugar analyzer (30). Serine, Threonine, and Tryptophan were not corrected for destruction during hydrolysis.
7. Cyanogen bromide cleavage: Cleavage at methionine residues was conducted as described by Gross and Witrop (31), and the resulting peptides were purified by gel chromatography and HPLC.
8. Reduction, alkylation and citraconylation: Citraconylation and S-carboxymethylation of CRP and its cyanogen bromide fragments were conducted as described (32–34). Carboxymethylamidation was conducted in the presence of 50  $\mu\text{Ci}$  of  $^{14}\text{C}$ -iodoacetamide. The radioactively labeled sample was subjected to trypsin digestion.
9. Radiolabeling of methionine residues: CRP was radiolabeled by  $^{14}\text{C}$ -methyl iodide (35,36) and subjected to trypsin digestion.
10. Digestion with endoproteases (enzyme:substrate = w/w): Digestions of CRP with trypsin (Worthington, TPKC-trypsin : substrate = 1:50), chymotrypsin (enzyme : substrate = 1:50) or thermolysin (enzyme : substrate = 1:20) was performed for 16 hrs at 37°C as previously described (10). Tryptic, chymotryptic and thermolysin peptides were purified by gel permeation chromatography and HPLC.
11. Digestion with carboxypeptidases: CRP, or its carboxy-terminal peptides were digested with carboxypeptidases (9,10). At specific times corresponding to 60, 180, 300 minutes and 24 hrs, 1 nmol aliquots were mixed with an equal volume of 3% TFA, pH 1.2, evaporated, dissolved in 0.2 M citrate buffer, pH 2.2 and subjected to amino acid analysis.
12. Automated Edman degradation: Amino acid sequence determination of proteins and peptides was performed either with a Beckman 890C automated sequencer or with an Applied Biosystems, Inc. microsequencer (37). The anilinothiazolinone intermediates were converted to PTH-derivatives by methanolic HCl (38) or by 25% TFA in water. Details regarding each analysis are given in the Legends to Figures.
13. Identification of PTH-amino acids: PTH-derivatives (20–60 pmoles) were identified by a modification of the method of Zimmerman et al. (39).
14. Determination of sequence homology: comparison of the amino acid sequence of Limulus CRP with those of other proteins was performed by MATTOP program (18).

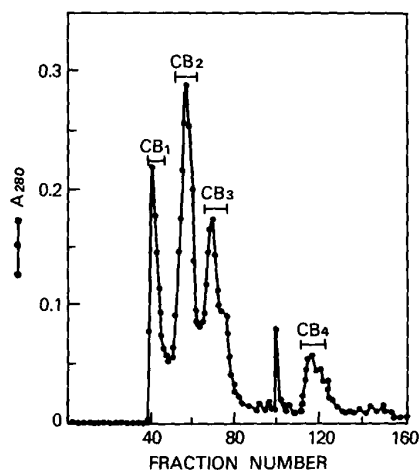


Fig. 1S. Purification of cyanogen bromide peptides. Sephadex G-100 (10 x 190 cm), 20% acetic acid; flow rate = 1 ml/min; fraction volume = 7 ml; sample = Limulus CRP (60 mg); detection = Absorbance at 280 nm. Fractions were pooled (bars) and subjected to automated Edman degradation. CB4 and CB3 yielded the amino acid sequences reported in Tables IIS and XS. Unresolved peptides CB1 and CB2 were further purified by HPLC.

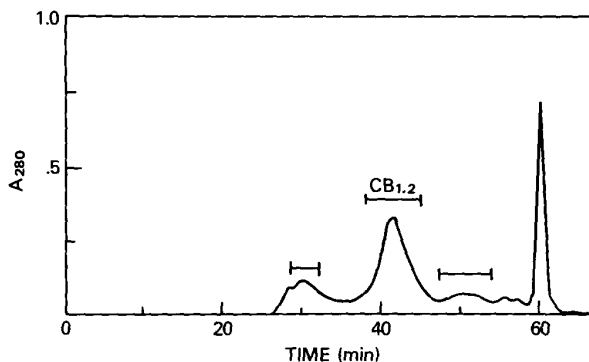


Fig. 2S. Purification of peptides derived from CB1. Sample = CB1, 11mg (Fig. 1S), reduced,  $^{14}\text{C}$ -alkylated and citraconylated, dissolved in 2 ml of 0.1 M ammonium bicarbonate; chromatography column = TSK C3000SW HPLC column (7.5 x 900 mm); mobile phase = 0.05 M phosphate buffer pH 6.5; flow rate = 0.5 ml/min; detection = Absorbance at 280; CB1.2 yielded the amino acid sequence reported in Table IIS. Peptides which generated sequence data presented in this report are given a denomination consisting of letters and numbers.

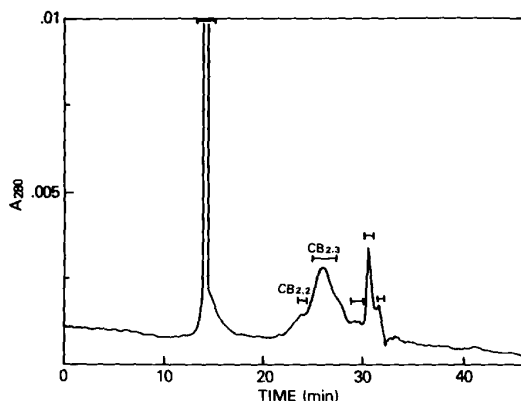


Fig. 3S. Purification of peptides derived from CB2. Sample = CB2, 10mg (Fig. 1S) reduced,  $^{14}\text{C}$ -alkylated and citraconylated, dissolved in 2 ml of 0.1 M ammonium bicarbonate; other conditions and fraction designation as in the Legends to Fig. 2S. CB2.3 and CB2.2 yielded the amino acid sequences reported in Table XIS and XVIS.

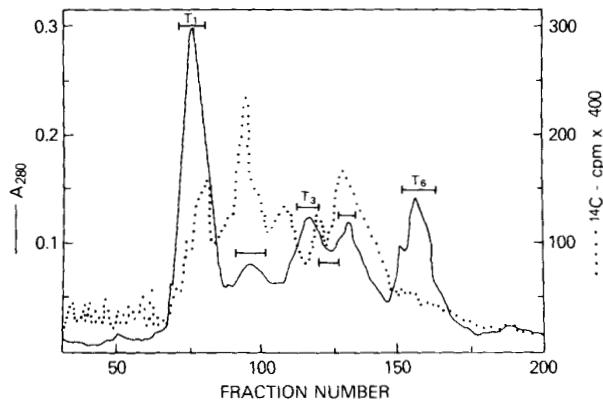


Fig. 4S. Purification of tryptic peptides derived from intact CRP.  $^{14}\text{C}$ -methyl-methionyl CRP (60 mg), reduced, alkylated and citraconylated. Sample was dissolved in 3 ml of 0.1 M ammonium bicarbonate pH 8.0; enzyme: TPCK-trypsin in 0.1 M ammonium bicarbonate; enzyme: substrate = 1:50; chromatography column = Sephadex G-50 fine (2 x 100 cm) in 0.1 M ammonium bicarbonate; flow rate = 1 ml/min; fraction volume = 7 ml; detection = absorbance at 280 nm, and radioactivity determination; fractions were pooled (bars) and subjected to automated Edman degradation. T1 yielded the amino acid sequence reported in Table IVS. Unresolved peptides (T3 and T6) were further purified by HPLC.

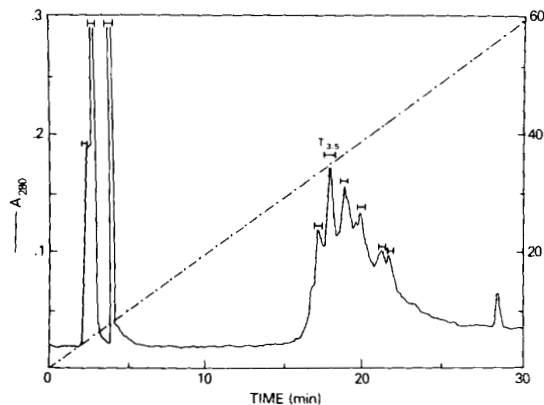


Fig. 5S. Purification of tryptic peptides derived from T3: Sample = T3, 0.2 mg (Fig. 4S); chromatography column: C18 reverse phase silica gel (4.6 x 250 mm); mobile phase: 0.1% ammonium acetate; B =  $\text{CH}_3\text{CN}$ ; flow rate = 0.5 ml/min; gradient = 10%-60% B, 1 hr. T3.5 yielded the amino acid sequence reported in Table XIIS.

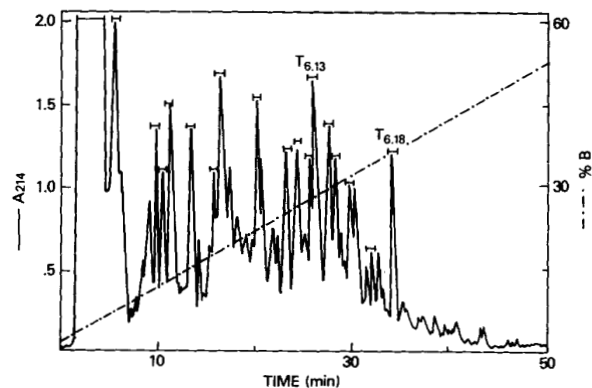


Fig. 6S. Purification of tryptic peptides derived from T6: Sample = T6, 0.2 mg (Fig. 4S); column: Vydac C18 HPLC column (4.6 x 250 mm); mobile phase: A = 0.1% TFA in water, B = 0.1% TFA in  $\text{CH}_3\text{CN}$ ; flow rate = 0.5 ml/min; gradient = 10%-50% B, 1 h. T6.13.2 derived from further purification of T6.13 under conditions identical to those listed in the legends to Fig. 2S. T6.18 was pure as shown by sequence analysis (Table VIS).

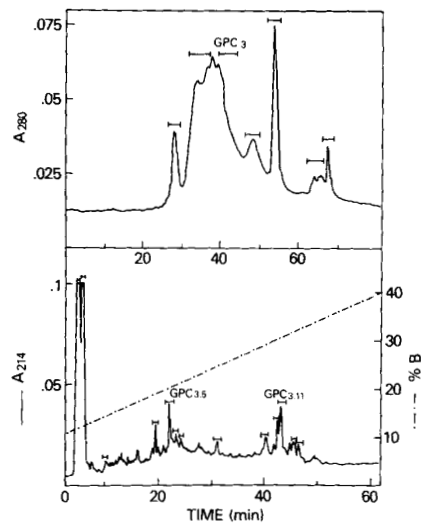


Fig. 7S. Purification of GPC3.5 and GPC3.11: Sample: CBI, 2 mg (Fig. 1S), reduced,  $^{14}\text{C}$ -alkylated, citraconylated and digested with trypsin as in the legends to Fig. 4S; upper panel = TSK G3000SW under conditions identical to those listed in Fig. 2S; lower panel: Ultrapore C3 reverse phase silica gel HPLC column (4.6 x 75 mm); mobile phase: A = 0.1% TFA in water, B = 0.1% TFA in acetonitrile; gradient = 15%-40% B, 1 h; GPC3.5 and GPC3.11 yielded the amino acid sequences reported in Tables VIIS and VIIIS.

Table IS: Partial Amino Acid Sequence of Limulus C-reactive protein

| Cycle number   | 1    | 5    | 10   | 15   |
|----------------|------|------|------|------|
| Residue number | 1    | 5    | 10   | 15   |
| Amino acid(s)  | Leu  | Glu  | Glu  | Gly  |
| nmol           | 39.3 | 37.4 | 37.4 | 29.7 |
|                | 37.3 | 35.3 | 13.8 | +    |
|                | 6.3  | 29.8 | 21.2 | 25.4 |
|                | 26.7 | 27.5 | 1.3  |      |
|                | 16   | 20   | 25   | 30   |
|                | 16   | 20   | 25   | 30   |
|                | Ser* | Ser* | Pro  | Ser* |
|                | +    | +    | 21.1 | 2.3  |
|                | 24.7 | 27.3 | 15.7 | 23.9 |
|                | 20.6 | 18.9 | 29.4 | 18.4 |
|                | 2.4  | 17.3 | 7.6  |      |
|                | 31   | 35   | 40   | 45   |
|                | 31   | 35   | 40   | 45   |
|                | Asp  | Leu  | Gln  | Glu  |
|                | 11e  | Thr* | Leu  | Cys  |
|                | Tyr  | Trp  | Phe  | Lys  |
|                | Val  | Val  | Asn  | Arg  |
|                | His  | Cys  |      |      |
|                | 6.4  | 24.4 | 17.0 | 16.3 |
|                | 13.2 | 2.6  | 12.8 | 8.2  |
|                | 3.8  | 17.8 | 2.6  | 13.2 |
|                | 12.0 | 9.0  | 5.0  |      |
|                | 46   | 50   |      |      |
|                | 46   | 50   |      |      |
|                | Leu  | Lys  | Gly  | Thr* |
|                | 10.2 | 9.3  | 8.4  | +    |
|                | 9.1  |      |      |      |

Sample amount = 50 nmol; initial yield (based on Leu<sub>1</sub>) = 78.6%; repetitive yield = 97%; cycles run = 60; cycles identified = 50; double coupling for 1st cycle, Beckman 890C sequenator.

\* = also identified as the corresponding dehydro-derivative at 313 nm

• = identified as PTH  $^{14}\text{C}$ -carboxymethylcysteine

Table IIS: Amino Acid Sequence of CB4

|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|----------------|------|------|------|------|------|------|------|------|------|------|-----|------|------|-----|------|----|
| Cycle number   | 1    |      |      |      |      | 5    |      |      |      |      | 10  |      |      |     |      | 15 |
| Residue number | 1    |      |      |      |      | 5    |      |      |      |      | 10  |      |      |     |      | 15 |
| Amino acid     | Leu  | Glu  | Glu  | Gly  | Glu  | Ile  | Thr* | Ser* | Lys  | Val  | Lys | Phe  | Pro  | Pro | Ser* |    |
| nmol           | 24.3 | 26.3 | 21.6 | 26.7 | 27.0 | 22.0 | 12.7 | 7.1  | 18.1 | 27.8 | 9.9 | 15.4 | 15.8 | 8.6 | +    |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |
|                |      |      |      |      |      |      |      |      |      |      |     |      |      |     |      |    |

Sample amount = 35 nmol; initial yield (based on Leu<sub>1</sub>) = 70.0%; repetitive yield = 94%; cycles run = 30; cycles identified = 25, Beckman 890C sequencer.

\* = also identified as the corresponding dehydro-derivative at 313 nm

X = identified as homoserine by amino acid analysis of CB4.

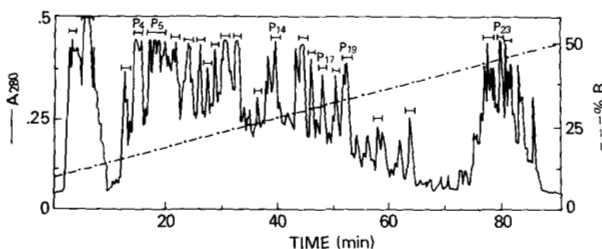


Fig. 8S: Purification of P19: Sample = CB1.2, 10mg, digested with trypsin; Vydac C18 HPLC column (4.6 x 250 mm), 1:30 h, as in the legends to Fig. 6S; P19 yielded the amino acid sequence reported in Table IXS. Other peptides, designated as P4 to P23, were further purified under conditions identical to those listed in Fig. 2S to generate the amino acid sequences reported in Table XVIS.

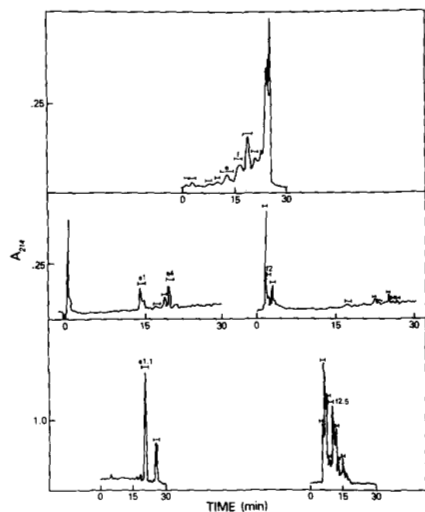


Fig. 9S: Purification of disulfide linked thermolysin peptides f2.5, el.1 and e4. Upper panel: sample = 4mg CRP dissolved in 0.16 ml of 9 M Guanidinium hydrochloride; enzyme = thermolysin in 0.5 M tris buffer pH 6.8, 0.02 M CaCl<sub>2</sub>; enzyme : substrate = 1:20; chromatography column = TSK G2000SW and TSK G3000SW; mobile phase = 0.1% TFA in 20% CH<sub>3</sub>CN; flow rate = 1 ml/min.

Center panels: rechromatography of fractions e and f; chromatography column = C<sub>3</sub> reverse phase silica gel (4 x 150 mm); mobile phase : A = 0.1% TFA in water, B = 0.1% TFA in CH<sub>3</sub>CN; gradient = 0% - 30% B, 30 min; e4 yielded the amino acid sequences reported in Table XIIS.

Lower panels: rechromatography of fractions el and f2; mobile phase : A 0.1% ammonium acetate in water, B = 0.1% ammonium acetate in CH<sub>3</sub>CN. Other conditions as in center panels; f2.5 and el.1 yielded the amino acid sequences reported in Tables XIVS and XVS.

Alignment of C-reactive proteins, SAP, and FP

|                   |                                                             |                                                             |                         |                                              |
|-------------------|-------------------------------------------------------------|-------------------------------------------------------------|-------------------------|----------------------------------------------|
|                   |                                                             | 1                                                           | 2                       | 3                                            |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | LEE                                                         | GE                                                          | ITS                     | KV - K F P P S S S S P S F P R L V M V G T L |
| 2 CRP Limulus 1.1 | LEE                                                         | GE                                                          | ITS                     | KV - K F P P S S S S P S F P R L V M V G T L |
| 3 CRP Limulus 1.4 | LEE                                                         | GE                                                          | ITS                     | KV - K F P P S S S S P S F P R L V M V G T L |
| 4 CRP Rabbit      | /A                                                          | G M H K K A F V F P K E S D T S Y V S L A Q L K K           |                         |                                              |
| 5 CRP Human       | /Q                                                          | T D M S K K A F V F P K E S D T S Y V S L A Q L K K         |                         |                                              |
| 6 SAP Human       | /H                                                          | T D L S G K V F V F P R E S V T D H V N L I T P L E K       |                         |                                              |
| 7 FP Hamster      | - -                                                         | T D L T G K V F V F P R Q S E T D Y V K L I P R L D K       |                         |                                              |
| Conserved         | LE                                                          |                                                             | T K V F V F P S S V L L | L T K                                        |
|                   |                                                             | 3                                                           | 4                       | 5                                            |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | P D L Q E I T L C                                           | - Y W F K V N R L K G T L H M F S Y A T C                   |                         |                                              |
| 2 CRP Limulus 1.1 | P D L Q E I T L C                                           | - Y W F K V N R L K G T L H M F S Y A T C                   |                         |                                              |
| 3 CRP Limulus 1.4 | P D L Q E I T L C                                           | - Y W F K V N R L K G T L H M F S Y A T C                   |                         |                                              |
| 4 CRP Rabbit      | P - L K A F I V C L Y F T I D L S M T R G Y S I F S Y A T R |                                                             |                         |                                              |
| 5 CRP Human       | P - L K A F I V C L Y F T I D L S M T R G Y S I F S Y A T R |                                                             |                         |                                              |
| 6 SAP Human       | P - L Q N F T L C F R A Y S D L S - - R A Y S L F S Y N T Q |                                                             |                         |                                              |
| 7 FP Hamster      | P - L Q N F T V C F R A Y S D L S - - R P H S L F S Y N T E |                                                             |                         |                                              |
| Conserved         | P L Q F T L C                                               | Y Y L S R S F S Y A T                                       |                         |                                              |
|                   |                                                             | 6                                                           | 7                       | 8                                            |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | K K D N E L L T F I D E Q Q D F L F N V H G A P Q L K V Q C |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | K K D N E L L T F I D E Q Q D F L F N V H G A P Q L K V Q C |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | K K D N E L L T F I D E Q Q D F L F N V H G A P Q L K V Q C |                                                             |                         |                                              |
| 4 CRP Rabbit      | R Q F N E I L L F W S K D I G Y S F S V G G D E I F I K V S |                                                             |                         |                                              |
| 5 CRP Human       | R Q D N E I L L F W S K D I G Y S F S V G G D E I F I K V S |                                                             |                         |                                              |
| 6 SAP Human       | G R D N E L L Y K E R V G E I S L Y I G R H K V T P K V I   |                                                             |                         |                                              |
| 7 FP Hamster      | Y G E N E L L Y K E R V G E I S L Y I G R H K V T P K V I   |                                                             |                         |                                              |
| Conserved         | D N E L L F                                                 | G Y F V G G L                                               |                         |                                              |
|                   |                                                             | 9                                                           | 10                      | 11                                           |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | P N K I H I G K W H H V C H T W S S W E G E A T I A V D G F |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | P N K I H I G K W H H V C H T W S S W E G E A T I A V D G F |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | P N K I H I G K W H H V C H T W S S W E G E A T I A V D G F |                                                             |                         |                                              |
| 4 CRP Rabbit      | D V P V D P T - - H L C A S W E S S T G I A E L V W D G K   |                                                             |                         |                                              |
| 5 CRP Human       | E V T V A P V - - H I C T S W E S S S G I A E F W I N G T   |                                                             |                         |                                              |
| 6 SAP Human       | E K F P A P V - - H I C V S W E S S S G I A E F W I N G T   |                                                             |                         |                                              |
| 7 FP Hamster      | E E F A S P V - - H F C T S W E S S S G I A E F W I N G K   |                                                             |                         |                                              |
| Conserved         | P                                                           | H C S W E S G I A E W V D G                                 |                         |                                              |
|                   |                                                             | 12                                                          | 13                      | 14                                           |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | H C K G N A T G I A V G R T L S Q G L V V L G Q D Q D S V   |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | H C K G N A T G I A V G R T L S Q G L V V L G Q D Q D S V   |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | H C K G N A T G I A V G R T L S Q G L V V L G Q D Q D S V   |                                                             |                         |                                              |
| 4 CRP Rabbit      | P - - N V R K S L K K G Y T V G A E A S I L L G Q E Q D S F |                                                             |                         |                                              |
| 5 CRP Human       | P - - N V R K S L K K G Y T V G A E A S I L L G Q E Q D S F |                                                             |                         |                                              |
| 6 SAP Human       | P - - L V K K G L R Q G Y F V E A Q P K I V L G Q E Q D S Y |                                                             |                         |                                              |
| 7 FP Hamster      | P - - V V K K G L R Q G Y F V E A Q P K I V L G Q E Q D N Y |                                                             |                         |                                              |
| Conserved         | P                                                           | V K G L G Y T L I V L G Q E Q D S                           |                         |                                              |
|                   |                                                             | 15                                                          | 16                      | 17                                           |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | G G K F D A T Q S L E G E L S E L N L W N T V L N H E Q I K |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | G G K F D A T Q S L E G E L S E L N L W N T V L N H E Q I K |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | G G E F D A T Q S L E G E L S E L N L W N T V L N H E Q I K |                                                             |                         |                                              |
| 4 CRP Rabbit      | G S F E K Q S L W D I G N V N M W D Y A L S P E I N         |                                                             |                         |                                              |
| 5 CRP Human       | G G N F E G S Q S L V G D I G N V N M W D Y A L S P E I N   |                                                             |                         |                                              |
| 6 SAP Human       | G G K F D R S Q S F V G E I G D L N M W D S V L T P E E I K |                                                             |                         |                                              |
| 7 FP Hamster      | G G G F D N Y Q S F V G E I G D L N M W D S V L T P E E I K |                                                             |                         |                                              |
| Conserved         | G G F D                                                     | Q S L V G E I G L N M W D V L P E I K                       |                         |                                              |
|                   |                                                             | 18                                                          | 19                      | 20                                           |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 |                         |                                              |
| 1 CRP Limulus 3.3 | Y L S K C A H P S E R H I Y G N I I Q W D K T Q F K A Y D G |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | Y L S K C A H P S E R H I Y G N I I Q W D K T Q F K A Y D G |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | H L S K C A H P S E R H I Y G N I I Q W D K T Q F K A Y D G |                                                             |                         |                                              |
| 4 CRP Rabbit      | T Y I - - - - A G G T F S P N V L D W R E L T Y Q V H G E   |                                                             |                         |                                              |
| 5 CRP Human       | T I T - - - - L G G F S P N V L D W R E L T Y Q V H G E     |                                                             |                         |                                              |
| 6 SAP Human       | S A Y - - - - Q G T P L P A N I L D W Q A L N Y E I N G Y   |                                                             |                         |                                              |
| 7 FP Hamster      | S V Y - - - - Q G V P L P A N I L D W Q A L N Y E I N G Y   |                                                             |                         |                                              |
| Conserved         | Y                                                           | G                                                           | N I L W L Y G           |                                              |
|                   |                                                             | 21                                                          | 22                      |                                              |
|                   |                                                             | 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7                           |                         |                                              |
| 1 CRP Limulus 3.3 | V V L S P N E I C A - - - - -                               |                                                             |                         |                                              |
| 2 CRP Limulus 1.1 | V V L S P N E I C A - - - - -                               |                                                             |                         |                                              |
| 3 CRP Limulus 1.4 | V V L S P N E I C A - - - - -                               |                                                             |                         |                                              |
| 4 CRP Rabbit      | V H V K P Q L W P - - - - -                                 |                                                             |                         |                                              |
| 5 CRP Human       | V F T K P Q L W P - - - - -                                 |                                                             |                         |                                              |
| 6 SAP Human       | V I I K P L V W V - - - - -                                 |                                                             |                         |                                              |
| 7 FP Hamster      | A V I R P R C V A L S S Y N K I S                           |                                                             |                         |                                              |
| Conserved         | V V                                                         | P                                                           |                         |                                              |

Fig. 10S: Comparison of the amino acid sequences of Limulus CRP (9), rabbit CRP (10), human SAP (19) and female hamster HFP (20).

Table IIIS: Partial Amino Acid Sequence of C81.2

|                |                                                                                             |                   |             |     |
|----------------|---------------------------------------------------------------------------------------------|-------------------|-------------|-----|
| Cycle number   | 1                                                                                           | 5                 | 10          | 15  |
| Residue number | 53                                                                                          | 57                | 62          | 67  |
| Amino acid(s)  | Phe Ser* <sup>†</sup> Tyr Ala Thr*Ala Lys Lys Asp Asn Glu Leu Thr*Phe Ser* <sup>†</sup> Leu | Asn               | Asn         | Leu |
| pmol           | 852 + 726 808 758 790 676 882 482 510 510 916 964 486 478 + 94                              | 68                | 159         |     |
|                | 16                                                                                          | 20                | 25          | 30  |
|                | 68                                                                                          | 72                | 77          | 82  |
|                | Leu Asp Glu Gln Gly Asp Phe Leu Phe Asn Val His Gly Ala Pro Ile Glu Lys                     | Ala               | His Met     |     |
|                | 800 332 300 432 536 426 328 728 302 272 388 232 328 214 138                                 | 71 66 62          | 103 103 225 |     |
|                | 31                                                                                          | 35                | 40          |     |
|                | 83                                                                                          | 87                | 92          |     |
|                | Gln Leu Lys Val Gln Cys Pro Asn Lys Ile His Ile                                             |                   |             |     |
|                | 244 348 198 368 216                                                                         | 106 156 186 152 + | 196         |     |

Sample amount = 1.7 nmol; initial yield (based on Phe<sub>53</sub>) = 77%; repetitive yield = 96%; cycles run = 50; cycles identified = 42, ABI 470A

\* = also identified as the corresponding dehydro-derivative at 313 nm

† = identified as PTH-<sup>14</sup>C-carboxymethylcysteine

Table IVS: Partial Amino Acid Sequence of T1

|                |                                                                                 |                                                            |             |
|----------------|---------------------------------------------------------------------------------|------------------------------------------------------------|-------------|
| Cycle number   | 5                                                                               | 10                                                         | 15          |
| Residue number | 96                                                                              | 101                                                        | 106         |
| Amino acid(s)  | Ile His Ile Gly Lys Trp His His Val Cys His Thr* Trp Ser* Ser* <sup>†</sup> Leu | Leu                                                        | Ala Lys Tyr |
| nmol           | 39.0 36.9 33.3 33.9 35.4 30.0 24.6 32.3 31.5                                    | 15.5 27.6 18.3 + 7.7<br>15.5 19.7 15.8 19.9<br>0.9 2.3 5.2 |             |
|                | 16                                                                              | 20                                                         | 25          |
|                | 107                                                                             | 111                                                        | 116         |
|                | Trp Glu Gly Glu Ala Thr* Ile Ala Val Asp Gly Phe His Cys Lys Met                |                                                            |             |
|                | 11.6 16.8 21.6 15.3 18.4 13.4 22.4 24.6 18.2 13.7 10.5 17.5 12.9                | 4.3 1.0                                                    |             |
|                | 31                                                                              | 35                                                         | 40          |
|                | 122                                                                             | 126                                                        | 131         |
|                | Gly X Ala Thr*Gly Ile Ala Val Met Gly Arg Thr*Leu Ser*Gln Gly Ala               |                                                            |             |
|                | 11.5 8.8 2.6 9.5 6.8 7.2 6.2 11.3 1.4 1.5 4.3 + 1.9                             | 3.6 0.8 9.5 3.4                                            |             |

Sample amount = 60 nmol; initial yield (based on Ile<sub>92</sub>) = 65.0%; repetitive yield = 96%; cycles run = 60; cycles identified = 44, Beckman 890C sequenator.

\* = Also identified as the corresponding dehydro-derivative at 313 nm

† = identified as PTH-<sup>14</sup>C-carboxymethylcysteine

X = Blank, proposed site of glycosylation. Deduced from DNA sequence analysis as Asn

Table VS: Partial Amino acid sequence of T6.13.2

|                |                                           |    |    |
|----------------|-------------------------------------------|----|----|
| Cycle number   | 1                                         | 5  | 10 |
| Residue number | 46                                        | 50 | 55 |
| Amino acid     | Leu Lys Gly Thr* Leu His Met Phe Ser* Tyr |    |    |
| pmol           | 160 50 100 130 70 110 80 60 60 70         |    |    |

Sample amount = 210 pmol; initial yield (based on Leu<sub>46</sub>) = 80.0%; repetitive yield = 93%; cycles run = 20; cycles identified = 10, ABI 470A

\* = also identified as the corresponding dehydro-derivative at 313 nm

Table VIS: Amino Acid Sequence of T6.18

|                |                                                                                       |     |     |     |
|----------------|---------------------------------------------------------------------------------------|-----|-----|-----|
| Cycle number   | 1                                                                                     | 5   | 10  | 15  |
| Residue number | 190                                                                                   | 194 | 199 | 204 |
| Amino acid(s)  | His Ile His Gly Asn Ile Ile Gln Trp Asp Lys Thr*Gln Phe Lys Ser* <sup>†</sup> Leu Tyr |     |     |     |
| pmol           | 109 125 73 104 86 97 100 68 11 70 30 + 39 28 25 23                                    |     |     |     |

Sample amount = 1.5 nmol; initial yield (based on Ile<sub>191</sub>) = 74.0%; repetitive yield = 97.0%; cycles run = 25; cycles identified = 15, ABI 470A

\* = identified as the corresponding dehydro-derivative at 313 nm

Table VIIS: Amino Acid Sequence of GPC3.5

|                |                                                              |     |     |     |
|----------------|--------------------------------------------------------------|-----|-----|-----|
| Cycle number   | 1                                                            | 5   | 10  | 15  |
| Residue number | 201                                                          | 205 | 210 | 215 |
| Amino acid(s)  | Thr*Gln Phe Lys Ala Tyr Asp Gly Val Val Leu Ser*Pro Asn Glu  |     |     |     |
| pmol           | 166 587 602 486 592 395 207 232 293 280 292 + 123 150 200 17 |     |     |     |
|                | 16                                                           |     |     |     |
|                | 216                                                          |     |     |     |
|                | Ile Cys Ala                                                  |     |     |     |
|                | 241                                                          | 31  |     |     |

Sample amount = 890 pmol; initial yield (based on Phe<sub>203</sub>) = 68%; repetitive yield = 93%; cycles run = 25; cycles identified = 18, ABI 470A.

† = identified as <sup>14</sup>C-carboxymethylcysteine

\* = also identified as the corresponding dehydro-derivative at 313 nm

Table VIIIS: Amino Acid Sequence of GPC3.11

|                |                                                                 |     |     |     |
|----------------|-----------------------------------------------------------------|-----|-----|-----|
| Cycle number   | 1                                                               | 5   | 10  | 15  |
| Residue number | 152                                                             | 156 | 161 | 166 |
| Amino acid(s)  | Phe Asp Ala Thr*Gln Ser*Leu Glu Gly Glu Leu Ser*Glu Leu Asn Tyr |     |     |     |
| pmol           | 208 220 190 + 240 + 150 188 118 170 100 + 110 88 165 31 41      |     |     |     |
|                | 16                                                              | 20  | 25  |     |
|                | 167                                                             | 171 | 176 |     |
|                | Leu Trp Asn Thr*Val Leu Asn His Glu Gln Ile Lys Pro Leu         |     |     |     |
|                | 100 63 58 + 75 54 40 + 34 26 17 63 33 62 20 32                  |     |     |     |

Sample amount: 0.4 nmol; initial yield (based on Asp<sub>153</sub>) = 55%; repetitive yield = 94%; cycles run = 35; cycles identified = 27, ABI 470A

\* = identified as the corresponding dehydro-derivative at 313 nm

Table IXS: Partial Amino Acid Sequence of P19

|                |                                                                 |     |     |     |
|----------------|-----------------------------------------------------------------|-----|-----|-----|
| Cycle number   | 1                                                               | 5   | 10  | 15  |
| Residue number | 132                                                             | 136 | 141 | 146 |
| Amino acid(s)  | Thr* Leu Ser* Gln Gly Leu Val Val Leu Gly Gln Asp Gln Asp       |     |     |     |
| nmol           | 0.8 2.2 0.7 1.5 2.2 2.1 1.5 2.0 2.4 1.4 1.8 1.2 1.1 1.2 0.8 0.4 |     |     |     |
|                | 16                                                              | 20  | 25  | 30  |
|                | 147                                                             | 151 | 156 | 161 |
|                | Ser* Val Gly Gly Lys Phe Asp Ala Thr* Gln Ser* Leu Glu Gly Glu  |     |     |     |
|                | 1.2 1.6 1.6 1.7 0.8 0.9 1.1 0.9 0.7 1.0 0.6 0.8 0.4 1.3 0.2     |     |     |     |
|                | 31                                                              |     |     |     |
|                | 162                                                             |     |     |     |
|                | Leu Ser*                                                        |     |     |     |
|                | 1.0 0.6                                                         |     |     |     |

Sample amount = 3.0 nmol; initial yield (based on Leu<sub>2</sub>) = 73.3%; repetitive yield = 97%; cycles run = 40; cycles identified = 32, ABI 470A

\* = also identified as the corresponding dehydro-derivative at 313 nm

## Structural Polymorphism of Limulus C-reactive Protein

Table XS: Partial Amino Acid Sequence of CB3

| Cycle number                         | 1                                | 5                                                            | 10                         | 15         |
|--------------------------------------|----------------------------------|--------------------------------------------------------------|----------------------------|------------|
| Residue number                       | 130                              | 134                                                          | 139                        | 144        |
| Amino acid(s)                        | Gly Val Thr* Leu                 | Ser* Glu Gly Gly Leu Val Val Leu Gly Gln Glu Asp             |                            |            |
| nmol                                 | 17.6 24.1 15.0 22.4 +            | 4.6 15.8 15.1 15.7 18.3 19.1 15.2 15.2 11.5 3.1 21.1 7.6 3.3 | 1.9                        | 9.4        |
| 16                                   | 20                               | 25                                                           | 30                         |            |
| 145                                  | 149                              | 154                                                          | 159                        |            |
| Gln Asp Ser* Val                     | Gly Gly Glu Tyr Ser* Gly Phe Val | Ala Glu Gln Ser* Leu Glu Ser*                                |                            |            |
| 0.99 12.9 +                          | 12.8 15.4 12                     | 1.2 2.8 1.3 11.2 17.8 1.8 +                                  | 17.6 2.3 1.3 3.8 2.6 7.6 + | 13.2 9.9 + |
| 31                                   | 35                               | 40                                                           |                            |            |
| 160                                  | 164                              | 169                                                          |                            |            |
| Gly Glu Leu Ser* Glu Leu Asn Leu Trp | Asn* Thr* Val Leu                |                                                              |                            |            |
| 10.6 12.3 8.9 +                      | 3.2 8.3 5.5 7.5 2.0              | 4.7 2.7 4.3 7.6 +                                            |                            |            |

Sample amount = 30 nmol; initial yield (based on Leu 134): 74.6%; repetitive yield = 95.6%; cycles run = 60; cycles identified = 43, Beckman 890C sequenator.  
 \* = Also identified as the corresponding dehydro-derivative at 313 nm

Table XI: Partial Amino Acid Sequence of CB2.3

| Cycle number                    | 1                                                           | 4           | 10 | 15 |
|---------------------------------|-------------------------------------------------------------|-------------|----|----|
| Residue number                  | 77                                                          | 81          | 86 | 91 |
| Amino acid(s)                   | Asn Val His Gly Ala Pro Gln Leu Lys Val Gln Cys Pro Asn Lys | Asn Val Arg |    |    |
| pmol                            | 813 1102 487 723 348 667 866 184 102 896 753 .              | 414 659 527 |    |    |
| 16                              | 20                                                          |             |    |    |
| 92                              | 96                                                          |             |    |    |
| Ile His Ile Gly Trp His His Val | Gly Lys Trp Leu                                             |             |    |    |
| 434 210 500 288 216 37 + +      | 126 248 154 32 85                                           |             |    |    |

Sample amount: 1.5 nmol; initial yield (based on Val<sub>78</sub>) = 73.3%; repetitive yield = 97%; cycles run: 30; cycles identified = 24, ABI 470A  
 \* = identified as PTH <sup>14</sup>C-carboxymethylcysteine

Table XII: Amino Acid Sequence of T3.5

| Cycle number                            | 1                                                                  | 5    | 10 | 15 |
|-----------------------------------------|--------------------------------------------------------------------|------|----|----|
| Residue number                          | 23                                                                 | 27   | 32 | 37 |
| Amino acid(s)                           | Leu Val Met Val Gly Thr* Leu Pro Asp Leu Gln Glu Ile Thr* Leu      |      |    |    |
| nmol                                    | 33.8 36.5 30.0 33.6 29.9 17.4 31.3 18.3 26.0 25.8 22.7 20.0 30.7 + | 27.8 |    |    |
| 16                                      | 20                                                                 | 25   |    |    |
| 38                                      | 42                                                                 | 47   |    |    |
| Cys Tyr Trp Phe Lys Val Asn His Leu Lys | Gln                                                                |      |    |    |
| . 15.1 4.6 7.2 9.9 8.1 6.8 3.0 1.3 0.4  | 0.3                                                                |      |    |    |

Sample amount: 45 nmol; initial yield (based on Leu<sub>23</sub>) = 74.6%; repetitive yield = 96%; cycles run = 35; cycles identified = 25, ABI 470A  
 \* = identified as the corresponding dehydro-derivative at 313 nm  
 \* = identified as <sup>14</sup>C-carboxymethylcysteine

Table XIII: Amino Acid Sequences of disulfide linked (Cys 88 - Cys 120)

|                   | peptide e4                      |                                                       |    |  |  |  |  |  |  |  |  |  |  |  |  |  |
|-------------------|---------------------------------|-------------------------------------------------------|----|--|--|--|--|--|--|--|--|--|--|--|--|--|
| Cycle number      | 1                               | 5                                                     | 10 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Residue number(s) | 84 113                          | 88 117                                                |    |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Amino acid(s)     | Leu Lys Val Gln Cys Pro Asn Lys | Ile Ala Val Asp Gly Phe His Cys Lys Gly X Ala Thr Gly |    |  |  |  |  |  |  |  |  |  |  |  |  |  |
| nmol              | 2.5 2.7 4.1 1.7 . 2.4 1.4 1.1   | 2.1 1.8 + 1.8 1.3 1.6 0.6 . 1.5 0.6 X 0.7 0.2 0.2     |    |  |  |  |  |  |  |  |  |  |  |  |  |  |

Sample amount: 3.2 nmol; initial yield (based on Leu<sub>84</sub>) = 78.1%; repetitive yield = 94%; cycles identified = 14, ABI 470A with ABI 120A  
 \* = identified as PTH-<sup>14</sup>C-carboxymethylcysteine of RCM-CB1.2, and RCM1. Analysis of performic acid oxidized e4 yielded 2.0 moles of cysteic acid per mole of disulfide-linked e4.  
 + = combined recoveries of Valine residues at cycles 86 and 115  
 X = Blank, proposed site of glycosylation. Deduced from DNA sequence analysis as Asn

Table XIV: Amino Acid Sequences of disulfide linked (Cys 38 - Cys 101)

|                   | peptide f2.5            |   |   |   |
|-------------------|-------------------------|---|---|---|
| Cycle number      | 1                       | 2 | 3 | 4 |
| Residue number(s) | 37 100                  |   |   |   |
| Amino acid(s)     | Leu Cys Val Cys His Thr |   |   |   |
| nmol              | 6.3 . 6.1 . 3.6 3.2     |   |   |   |

Sample amount: 8.0 nmol; initial yield (based on Leu<sub>37</sub>) = 76.7%; cycles identified = 4, ABI 470A with ABI 120A.  
 \* = identified as PTH <sup>14</sup>C-carboxymethylcysteine of RCM-CRP and <sup>14</sup>C-RCM1. Analysis of performic acid oxidized sample of f2.5 yielded 2.1 moles of cysteic acid per mole of disulfide-linked peptide f2.5.

Table XV: Amino Acid Sequences of disulfide linked (Cys 183 - Cys 217)

|                   | peptide e1.1                                |             |     |  |  |  |  |  |  |  |  |  |
|-------------------|---------------------------------------------|-------------|-----|--|--|--|--|--|--|--|--|--|
| Cycle number      | 1                                           | 5           | 10  |  |  |  |  |  |  |  |  |  |
| Residue number(s) | 180 216                                     | 184         | 189 |  |  |  |  |  |  |  |  |  |
| Amino acid(s)     | Leu Ser Lys Cys Ala His Pro Ser Glu Arg His | Ile Cys Ala |     |  |  |  |  |  |  |  |  |  |
| nmol              | 3.6 2.8 3.4 . 3.0 1.9 2.3 1.7 1.4 1.1 0.2   | 2.9 . 1.2   |     |  |  |  |  |  |  |  |  |  |

Sample amount: 4.6 nmol; initial yield (based on Leu<sub>180</sub>) = 78.2%; repetitive yield = 94%; cycles identified = 11, ABI 470A with ABI 120A  
 \* = identified as PTH <sup>14</sup>C-carboxymethylcysteine of RCM-GPC3.5 and GPC3.11. Analysis of performic acid oxidized sample of e1.1 yielded 1.8 moles of cysteic acid per mole of disulfide-linked peptide e1.1.

Table XVIS

| Residues<br>Identified | Amino acid sequence |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     | Peptide | Purified<br>from |     |     |                     |                     |                      |                      |                    |
|------------------------|---------------------|-----|-----|------|------------|------------|------------|------------|------------|-----|------|-----|-----|-----|-----|-----|-----|------------|-----|------------|------------|--------------|-----|-----|---------|------------------|-----|-----|---------------------|---------------------|----------------------|----------------------|--------------------|
| 77-101<br>pmol         | Asn                 | Val | His | Gly  | Ala<br>Asn | Pro        | Gln        | <u>Val</u> | Lys        | Val | Gln  | Cys | Pro | Asn | Lys | Ile | His | Ile        | Gly | <u>Lys</u> | <u>Trp</u> | His          | Val | Cys | Leu     | 22               | 30  | .   | CB3.4<br>Fig.<br>1S | CB3,<br>Fig.<br>1S  |                      |                      |                    |
|                        | 171                 | 191 | 140 | 201  | 78<br>116  | 109        | 145        | 108        | 81         | 137 | 108  | .   | 112 | 116 | 73  | 70  | 50  | 54         | 51  | 29         | 50         | 40           | 22  | 30  | .       |                  |     |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 77-101<br>pmol         | Asn                 | Val | His | Gly  | <u>Asn</u> | Pro        | Gln        | <u>Val</u> | <u>Arg</u> | Val | Gln  | Cys | Pro | Asn | Lys | Ile | His | Ile        | Gly | <u>Lys</u> | <u>Trp</u> | His          | Val | Cys | Leu     | 22               | 14  | 40  | .                   | CB3.5<br>Fig.<br>1S | CB3,<br>Fig.<br>1S   |                      |                    |
|                        | 253                 | 208 | 192 | 227  | 210        | 192        | 168        | 223        | 160        | 220 | 169  | .   | 134 | 127 | 113 | 131 | 69  | 83         | 90  | 84         | 60         | 22           | 14  | 40  | .       |                  |     |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 77-101<br>pmol         | Asn                 | Val | His | Gly  | <u>Asn</u> | Pro        | Gln        | <u>Val</u> | <u>Glu</u> | Val | Gln  | Cys | Pro | Asn | Lys | Ile | His | Ile        | Gly | <u>Lys</u> | <u>Trp</u> | <u>Asn</u>   | Val | Cys | Leu     | 240              | 340 | 420 | 520                 | .                   | CB2.2<br>Fig.<br>1S  | CB2,<br>Fig.<br>1S   |                    |
|                        | 823                 | 992 | 630 | 720  | 782        | 620        | 710        | 870        | 680        | 780 | 640  | .   | 530 | 643 | 540 | 680 |     |            | 200 | 510        | 480        | 400          | 240 | 340 | 420     | 520              | .   |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            | 320 |            | 60         |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 77-85<br>pmol          | Trp                 | Val | His | Gly  | Ala        | Pro        | Gln        | Leu        | Lys        | Val |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     | P5.4.3<br>Fig.<br>8S | P5,<br>Fig.<br>8S    |                    |
|                        | 520                 | 610 | 340 | 220  | 102        | 111        | 108        | 52         | 34         |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 130-<br>160<br>pmol    | Gly                 | Val | Thr | Leu  | <u>Lys</u> | <u>Asn</u> | Gly        | <u>Ala</u> | Leu        | Val | Val  | Leu | Gly | Cln | Asn | Gln | Asp | <u>Ile</u> | Val | Gly        | <u>Ser</u> | <u>Ser</u> * | Phe | Asp | Ala     | Thr              | Gln | Ser | Leu                 | Glu                 | Gly                  | CB3.8<br>Fig.<br>1S  | CB3,<br>Fig.<br>1S |
|                        | 299                 | 379 | 51  | 300  | 189        | 184        | 111        | 112        | 284        | 156 | 189  | 284 | 170 | 165 | 184 | 149 | 179 | 150        | 189 | 94         | 22         | +            | 86  | 79  | 67      | 40               | 51  | 20  | 77                  | 59                  | 38                   |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 157-<br>168<br>pmol    | Ser*                | Leu | Glu | Gly  | Gly        | Leu        | Ser*       | Glu        | Leu        | Asn | Leu  | Trp |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      | P23<br>Fig.<br>8S    | Fig.<br>8S         |
|                        | +                   | 141 | 177 | 157  | 170        | 113        | +          | 162        | 125        | 108 | 78   |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 166-<br>169<br>nmol    | Asn                 | Leu | Trp | Ser* |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      | P14.2<br>Fig.<br>8S  | P14,<br>Fig.<br>8S |
|                        | 1.4                 | 1.2 | 0.3 | +    |            |            |            |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 170-<br>181<br>nmol    | Val                 | Leu | Asn | His  | Glu        | Gln        | Ile        | Lys        | His        | Leu | Ser* | Lys |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      | P5.3.2<br>Fig.<br>8S | P5,<br>Fig.<br>8S  |
|                        | 2.1                 | 1.9 | 1.4 | 0.9  | 1.1        | 1.0        | 0.9        | 0.9        | 0.3        | 0.7 | +    | 0.6 |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
|                        |                     |     |     |      |            |            |            |            | 0.6        | 0.3 |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 190-<br>200<br>pmol    | Ser*                | Ile | Tyr | Gly  | Asn        | Ile        | Ile        | Gln        | Trp        | Asp | Lys  |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      | P17.1<br>Fig.<br>8S  | P17,<br>Fig.<br>8S |
|                        | +                   | 724 | 810 | 630  | 810        | 840        | 740        | 697        | 400        | 375 | 408  |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |
| 190-<br>8S<br>nmol     | His                 | Ile | Tyr | Gly  | Asn        | <u>Lys</u> | <u>Lys</u> |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      | P4.1<br>Fig.<br>8S   | Fig.<br>8S         |
|                        | 1.2                 | 1.3 | 0.9 | 0.6  | 0.7        | 0.4        | 0.1        |            |            |     |      |     |     |     |     |     |     |            |     |            |            |              |     |     |         |                  |     |     |                     |                     |                      |                      |                    |

\* = identified as the corresponding dehydro-derivative at 313 nm; + identified as PTH <sup>14</sup>C-carboxymethylcysteine

1-5 nmol of these peptides were sequenced with Applied Biosystems 470A microsequencer. Underlined residues represent amino acids that do not fit into the predicted amino acid sequence of the three isolated CRP genes (17). The yields of these peptides are estimated to be <10% of the affinity purified protein.

Table XVIIIS  
Amino Acid Composition of Peptides used for Sequence Analysis

| Amino Acid    | Peptides |          |         |        |         |         |         |         |          |         |         |         |         |         |        |        | Total  |        |        |        |
|---------------|----------|----------|---------|--------|---------|---------|---------|---------|----------|---------|---------|---------|---------|---------|--------|--------|--------|--------|--------|--------|
|               | CB3.4    | CB3.5    | CB2.2   | CB2.3  | CB2.4   | CB2.5   | CB2.6   | CB2.7   | CB2.8    | CB2.9   | CB2.10  | CB2.11  | CB2.12  | CB2.13  | CB2.14 | CB2.15 | CB2.16 | CB2.17 | CB2.18 | CB2.19 |
| Aspartic Acid | 20(18)   | 5.0(4)   |         |        | 2.7(2)  | 2.0(2)  | 4.5(5)  | 5.6(6)  | 7.8(9)   | 5.2(4)  | 1.9(2)  | 3.1(3)  |         |         |        |        |        |        |        |        |
| Threonine     | 1.1(1)   | 6.2(9)   | 3.6(4)  | 1.2(1) | 0.8(1)  | 0.5(1)  | 2.6(1)  | 2.3(3)  | 3.7(3)   | 2.4(3)  | 2.5(2)  | 0.9(1)  | 1.0(1)  |         |        |        |        |        |        |        |
| Serine        | 4.9(5)   | 8.5(10)  | 3.3(4)  | 1.5(1) |         | 1.0(1)  | 3.4(2)  | 4.5(4)  | 8.8(8)   | 2.0(2)  |         |         |         |         |        |        |        |        |        | 1.8(2) |
| Glutamic Acid | 3.3(3)   | 17.8(20) | 6.9(5)  |        | 2.9(2)  | 1.9(2)  | 6.6(6)  | 11.7(9) | 17.5(16) | 3.6(4)  | 3.1(2)  | 1.1(1)  | 1.1(1)  |         |        |        |        |        |        |        |
| Proline       | 5.1(4)   | 4.0(4)   |         |        |         | 0.8(1)  |         |         | 2.5(2)   | 2.8(2)  | 1.0(1)  | 0.7(1)  | 1.0(1)  |         |        |        |        |        |        |        |
| Glycine       | 1.2(1)   | 20(16)   | 9.1(11) | 1.6(1) | 5.1(1)  | 1.3(1)  | 1.6(1)  | 7.2(6)  | 8.9(9)   | 6.1(7)  | 1.8(1)  | 3.1(3)  |         |         |        |        |        |        |        |        |
| Alanine       |          | 11(11)   | 4.9(4)  |        |         | 1.5(2)  | 1.2(1)  | 1.7(1)  | 5.2(4)   | 3.4(4)  | 0.6(0)  | 1.6(2)  |         |         |        |        |        |        |        | 2.0(2) |
| 1/2 Cysteine  |          | 4.4(5)*  | 1.5(2)* |        |         | 0.5(1)* |         |         | 1.9(2)*  | 2.0(3)* | 0.8(1)* | 2.0(2)* | 2.1(2)* | 1.8(2)* |        |        |        |        |        |        |
| Valine        | 1.5(2)   | 9.2(11)  | 6.9(6)  |        |         | 1.2(2)  | 1.2(1)  | 3.2(4)  | 3.1(7)   | 3.0(4)  | 1.2(3)  | 2.1(2)  | 1.1(1)  |         |        |        |        |        |        |        |
| Methionine    |          |          |         |        | 0.6(1)  |         |         |         |          |         |         |         |         | 1.0(1)  |        |        |        |        |        |        |
| Isoleucine    | 0.9(1)   | 9.0(10)  | 3.9(4)  |        | 2.6(3)  | 0.8(1)  | 0.5(1)  | 0.5(1)  | 2.1(5)   | 2.9(4)  | 1.1(1)  | 0.9(1)  |         |         |        |        |        |        |        | 0.7(1) |
| Leucine       | 1.6(2)   | 13.8(14) | 3.0(3)  | 1.8(2) | 1.1(0)  | 0.9(1)  | 4.8(5)  | 6.7(8)  | 8.9(10)  | 0.6(1)  | 4.2(5)  | 0.9(1)  | 1.0(1)  | 1.1(1)  |        |        |        |        |        |        |
| Tyrosine      |          | 3.0(4)   |         | 1.3(1) | 1.0(1)  | 0.7(1)  |         |         | 1.4(2)   |         | 1.2(1)  |         |         |         |        |        |        |        |        |        |
| Phenylalanine | 1.5(2)   | 5.8(7)   | 1.4(1)  | 1.0(1) | 0.9(1)  | 1.1(1)  | 1.1(1)  | 1.5(1)  | 1.4(1)   | 1.0(1)  | 0.9(1)  | 1.2(1)  |         |         |        |        |        |        |        |        |
| Lysine        | 1.2(2)   | 7.8(10)  | 3.4(3)  | 1.0(1) | 1.1(2)  | 1.2(1)  | 1.2(1)  | 1.6(2)  | 3.0(3)   | 5.2(4)  | 1.2(2)  | 2.8(3)  |         |         |        |        |        |        |        | 1.3(1) |
| Histidine     |          | 9.0(9)   | 4.3(5)  | 1.0(1) | 1.0(1)  |         | 1.0(1)  | 1.0(1)  | 4.4(5)   | 5.3(6)  | 0.5(1)  | 1.0(1)  | 2.0(2)  |         |        |        |        |        |        |        |
| Arginine      | 0.8(1)   | 1.7(2)   | 1.0(1)  |        |         |         |         |         | 1.0(1)   | 0.5(0)  |         |         |         |         |        |        |        |        |        | 1.0(1) |
| Tryptophan    |          | 0(5)     | 0.8(3)  |        | 0.8(1)  |         | 0.6(1)  | 0.7(1)  | 1.2(2)   | 2.8(3)  | 0.8(1)  |         |         |         |        |        |        |        |        |        |
| Homoserine    | 0.9(1)   | 0.7(1)   |         |        |         |         |         |         |          |         |         |         |         | 0.9(1)  |        |        |        |        |        |        |
| Position      | 1-25     | 51-218   | 92-151  | 46-55  | 190-204 | 201-218 | 152-178 | 132-178 | 130-218  | 77-129  | 23-47   | 64-91   | 37-38   | 180-190 |        |        |        |        |        |        |
| Total         | 25       | 166      | 60      | 10     | 15      | 18      | 27      | 47      | 89       | 53      | 25      | 22      | 6       | 14      |        |        |        |        |        |        |
| Yields (%)    | 80.0     | 38.9     | 32.1    | 19.2   | 42.5    | 20.8    | 23.0    | 71.7    | 23.1     | 44.8    | 51.8    | 60.0    | 30.0    | 50.0    |        |        |        |        |        |        |

+ = determined as S-carboxymethylcysteine

\* = determined as cysteic acid

Values in the parentheses are taken from the amino acid sequences deduced from DNA analyses of clones; a=3.3, b=1.1, c=1.4

0 = detected but not quantitated