Hubbel, W. L., & McConnell, H. M. (1971) *J. Am. Chem. Soc. 93,* 314.
Israelachvili, J., Sjösten, J., Eriksson, L. E. G., Ehrström, M., Gräslund, A., & Ehrenberg, A. (1975) *Biochim. Biophys. Acta 382,* 125.
Kawato, S., Kinosita, K., Jr., & Ikegami, A. (1977) *Biochemistry 16,* 2319.
Kinosita, K., Jr., Mitaku, S., Ikegami, A., Ohbo, N., & Kunii, T. L. (1976) *J. Appl. Phys., Jpn. 15,* 2433.
Kinosita, K., Jr., Kawato, S., & Ikegami, A. (1977) *Biophys. J. 20,* 289.
Ladbrooke, B. D., Williams, R. M., & Chapman, D. (1968) *Biochim. Biophys. Acta 150,* 333.
Lakowicz, J. R., & Prendergast, F. G. (1978) *Science 200,* 1399.
Lentz, B. R., Barenholz, Y., & Thompson, T. E. (1976a) *Biochemistry 20,* 4521.
Lentz, B. R., Barenholz, Y., & Thompson, T. E. (1976b) *Biochemistry 20,* 4529.
Lippert, J. L., & Peticolas, W. L. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68,* 1572.
Oldfield, E., & Chapman, D. (1971) *Biochem. Biophys. Res. Commun. 43,* 610.
Oldfield, E., & Chapman, D. (1972) *FEBS Lett. 21,* 302.
Oldfield, E., Chapman, D., & Derbyshire, W. (1971) *FEBS Lett. 16,* 102.
Seelig, J. (1977) *Q. Rev. Biophys. 10,* 353.
Shinitzky, M., Dianovx, A. C., Gitler, C., & Weber, G. (1971) *Biochemistry 10,* 2106.
Stoffel, W., Tunggal, B. D., Zierenberg, O., Schreiber, E., & Binczek, E. (1974) *Hoppe Seyler's Z. Physiol. Chem. 355,* 1367.
Weber, G., Helgerson, S. L., Gramer, W. A., & Mitchell, G. W. (1976) *Biochemistry 15,* 4429.

# Purification, Characterization, and Amino Acid Sequence of Rat Anaphylatoxin (C3a)†

John W. Jacobs, Jeffrey S. Rubin, Tony E. Hugli, Richard A. Bogardt, Ida K. Mariz, John S. Daniels, William H. Daughaday, and Ralph A. Bradshaw*

ABSTRACT: C3a anaphylatoxin derived from the third component of complement has been isolated from rat serum and its complete amino acid sequence determined. A three-step purification procedure was employed that consisted of gel filtration on Sephadex G-100, followed by chromatography of the anaphylatoxin-containing pool on carboxymethylcellulose. A subsequent separation on DEAE-Sephadex resolved C3a from minor contaminating peptides. Biological studies have shown that purified rat anaphylatoxin is approximately twice as active as human or porcine C3a when tested for smooth-muscle contraction. In addition to the active form of rat anaphylatoxin, a serum carboxypeptidase B inactivated form of C3a ($C3a_{des\text{-}Arg}$) was purified from rat serum and utilized in subsequent structural studies. Sequence analysis of rat C3a was facilitated by a long automated Edman degradation which established the first 55 residues of the anaphylatoxin. Overlapping peptides were generated by cyanogen bromide and trypsin, and the resultant fragments were sequenced by either automated or manual Edman procedures. The primary structure of rat C3a is 70% identical to the previously determined structures of human and porcine anaphylatoxin. Antisera raised to the purified rat peptide do not cross-react immunologically by Ouchterlony analysis with either human or porcine C3a.

Studies of anaphylatoxins from a number of species have established that this peptide serves as a mediator of the local inflammatory response (Hugli and Müller-Eberhard, 1978). Micromolar concentrations of C3a induce histamine release from isolated mast cells (Dias da Silva and Lepow, 1967; Johnson et al., 1975) and produce smooth-muscle contraction (Dias da Silva et al., 1967; Cochrane and Müller-Eberhard, 1968). In addition, C3a has pronounced effects on vascular permeability (Dias da Silva and Lepow, 1967; Wuepper et al., 1972). The isolation and characterization of anaphylatoxins from both human and porcine sera revealed that the two peptides have essentially identical biological properties (Hugli et al., 1975b). Structural studies have shown a 70% homology in amino acid sequence between porcine and human C3a, suggesting that similarities in biological activities result from homologous primary structures (Hugli, 1975; Corbin and Hugli, 1976). Both peptides require a carboxy-terminal arginine residue for biological activity and are rapidly inactivated by carboxypeptidase B digestion (Hugli et al., 1975b). In spite of these chemical similarities, immunological studies demonstrate no visible cross-reactivity between porcine and human C3a by Ouchterlony analysis, suggesting that key antigenic determinants differ between the two anaphylatoxins.

This report extends the biological and structural characterization of anaphylatoxins to rat C3a. Studies on rat anaphylatoxin grew out of our initial efforts to purify rat somatomedin. The somatomedins are a family of insulin-like peptides which exhibit growth-promoting effects in a variety of tissues (Daughaday et al., 1972; Van Wyk et al., 1974). It was

† From the Departments of Biological Chemistry (J.W.J., J.S.R., and R. A. Bradshaw) and Medicine (I.K.M., J.S.D., and W.H.D.), Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110, and the Department of Molecular Immunology (T.E.H. and R. A. Bogardt), Scripps Clinic and Research Foundation, La Jolla, California 92037. *Received May 2, 1978.* Supported by research grants from the United States National Institutes of Health (AM 13362 and AM 01526, HL 20220, and program project HL 16411).

0006-2960/78/0417-5031$01.00/0 

observed in our partially purified somatomedin preparations that the serum carboxypeptidase-inactivated anaphylatoxin ($C3a_{des\text{-}Arg}$)[1] represented a significant fraction of the total protein. Separation of rat somatomedin and the inactivated anaphylatoxin by isoelectric focusing permitted the recovery of homogeneous $C3a_{des\text{-}Arg}$ suitable for structural studies.

## Materials and Methods

*Isolation of Rat C3a and $C3a_{des\text{-}Arg}$.* Rat C3a anaphylatoxin was generated by inulin activation of fresh serum (255 mL) from healthy Lewis rats by methods previously described (Hugli et al., 1975b). The initial purification step of rat C3a isolation employed gel filtration of the activated serum. Radiolabeled ($^{125}I$) human C3a was introduced to the activated serum prior to gel filtration to aid in monitoring the elution profile for C3a. Separation was performed on Sephadex G-100 (15 × 40 cm), equilibrated with 0.15 M sodium acetate (pH 3.7) as previously described (Hugli et al., 1975b). The pool containing anaphylatoxin activity and radioactivity was lyophilized and equilibrated by dialysis in 0.15 M ammonium formate (pH 4.5) to a final volume of 70 mL. The rat C3a pool was then subjected to CM-cellulose chromatography, as previously described by Fernandez and Hugli (1976), which resolved active rat C3a from the inactive $C3a_{des\text{-}Arg}$ form (Figure 1). The pool containing anaphylatoxin activity was lyophilized and dissolved in 0.02 M sodium phosphate buffer (pH 7.0) to a final concentration of 0.4 mg/mL. The final purification step of rat C3a isolation involved DEAE-Sephadex A-50 chromatography as previously described (Hugli et al., 1975b). The C3a anaphylatoxin, recovered from the void fraction of the DEAE-Sephadex elution, was judged homogeneous according to cellulose acetate and gel electrophoretic analyses and was used without further purification.

Rat $C3a_{des\text{-}Arg}$ was purified from the serum of female Wistar Furth rats (Microbiological Associates, Bethesda, Md.) bearing transplanted pituitary tumors (MStT/W15). This serum serves as an enriched source for the isolation of rat somatomedin; circulating somatomedin levels are approximately sixfold higher in tumor-bearing animals than in normal rats (Chochinov et al., 1977). The serum was chromatographed on an ascending Sephadex G-150 gel-filtration column (5 × 100 cm) equilibrated with 0.1 M ammonium bicarbonate (pH 8.6) at 4 °C. Protein elution was monitored at 280 nm. Since $C3a_{des\text{-}Arg}$ lacks anaphylatoxin activity (Hugli et al., 1975b), the peptide was monitored during purification by a radioimmunoassay employing $[^{125}I]C3a_{des\text{-}Arg}$, antisera raised in rabbits against rat $C3a_{des\text{-}Arg}$, and purified rat $C3a_{des\text{-}Arg}$ as a standard. The active pool was boiled at pH 5.5 for 20 min and then applied to a Sephadex G-50 column (2.5 × 100 cm) equilibrated in 1 N acetic acid. Fractions with immunoreactive $C3a_{des\text{-}Arg}$ activity were pooled, lyophilized, and applied directly to a 110-mL LKB isoelectric focusing column with a 5 to 50% linear sucrose gradient, 1% Ampholine (pH 3.5–10). Fractions of 3 mL were collected from the column, dialyzed against $H_2O$ to remove sucrose and ampholines, and lyophilized. Fractions were assayed for immunoreactive $C3a_{des\text{-}Arg}$, somatomedin biological activity (Daughaday et al., 1975), and somatomedin radioreceptor activity employing a placental membrane preparation and $[^{125}I]$insulin as the ligand (Marshall et al., 1974). Resolution of somatomedin and $C3a_{des\text{-}Arg}$ activity was achieved by isoelectric focusing under identical conditions as described above, except an Ampholine pH gradient of 7–9 was employed.

*Cellulose Acetate and Gel Electrophoresis.* Microzone electrophoresis on cellulose acetate strips was performed at room temperature and 250 V for 20 min in barbital buffer ($\mu$ = 0.75, pH 8.6) using the Beckman microzone apparatus, Model D101. Sample concentration ranged from 1.45 to 1.60 mg/mL. Protein standards employed in the microzone experiments included human and porcine C3a prepared as previously described (Hugli et al., 1975b; Corbin and Hugli, 1976) for comparison with rat C3a obtained from inulin-activated serum of normal rats.

Polyacrylamide disc gel electrophoresis with sodium dodecyl sulfate was performed in either 9 or 15% acrylamide under reducing conditions by a modification of the procedure described by Laemmli (1970). Protein samples were boiled for 5 min prior to electrophoresis in a buffer consisting of 1% $NaDodSO_4$, 4% $\beta$-mercaptoethanol, 4% glycerol, and 0.002% Bromophenol blue. Electrophoresis was carried out at 0.5 mA/gel for the first half hour and at 3 mA/gel for the succeeding 3–4 h. Protein standards employed in the gel-electrophoresis experiments included cytochrome *c*, ribonuclease, myoglobin, and chymotrypsinogen purchased from Sigma (St. Louis, Mo.). The gels were stained with Coomassie brilliant blue R250.

*Amino Acid Analysis.* Samples for amino acid analysis were hydrolyzed in 6 N HCl under reduced pressure at 110 °C for 24, 48, and 72 h. Analyses were performed on a Beckman 120C amino acid analyzer. Half-cystine was determined as either cysteic acid or *S*-carboxymethylcysteine. Free sulfhydryl groups were assayed by the procedure of Ellman (1959). Tryptophan was determined quantitatively according to Edelhoch (1967). Carboxypeptidase A and B hydrolyses were performed by the method of Ambler (1967) with released residues identified by amino acid analysis.

*Amino Acid Sequencing.* Automated Edman degradations were performed in a Beckman 890C sequencer employing the dilute (0.1 M) Quadrol program of Brauer et al. (1975). All Sequencer reagents and solvents were purchased from Beckman (Spinco Division). Degradations on less than 50 nmol of peptide were carried out in the presence of phenyl $[^{35}S]$isothiocyanate (Amersham-Searle) according to Jacobs and Niall (1975). Manual Edman degradations were performed by the method of Sauer et al. (1974). Amino acid phenylthiohydantoins (Pth) were identified by a combination of gas–liquid chromatography (Pisano and Bronzert, 1969) and thin-layer chromatography (Edman, 1960). Radioactive Pth derivatives were identified by thin-layer chromatography, followed by autoradiography (Jacobs and Niall, 1975).

*Cleavage of $C3a_{des\text{-}Arg}$ and Purification of Peptides.* CNBr Peptides. $C3a_{des\text{-}Arg}$ (1 $\mu$mol) used for CNBr cleavage reactions was first S-carboxymethylated with 1-iodo[$^{14}C$]acetic acid (New England Nuclear) according to Angeletti et al. (1971). Radioactive *S*-carboxymethyl-C3a was dissolved in 2 mL of 70% formic acid and reacted with a 50-fold excess of CNBr. After 24 h at 25 °C, the reaction mixture was diluted 20-fold with distilled $H_2O$ and lyophilized. The CNBr peptides were separated by gel filtration on Sephadex G-50 superfine, and selected pools from the Sephadex G-50 column were further fractionated on a SP-Sephadex ion-exchange column (0.9 × 25 cm) using a double linear gradient of pyridine acetate buffers as described by Bradshaw et al. (1969). Aliquots of 50 $\mu$L were removed from each fraction, dissolved in 10 mL of 3a70 cocktail solution (Research Products International, Elk Grove Village, Ill.), and measured for radioactivity in a Packard Model 3370 liquid scintillation spectrometer.

[1] Abbreviations used: $C3a_{des\text{-}Arg}$, carboxypeptidase inactivated anaphylatoxin; $NaDodSO_4$, sodium dodecyl sulfate; Pth, phenylthiohydantoin; CNBr, cyanogen bromide; MT, maleoylated tryptic; EACA, $\epsilon$-aminocaproic acid.

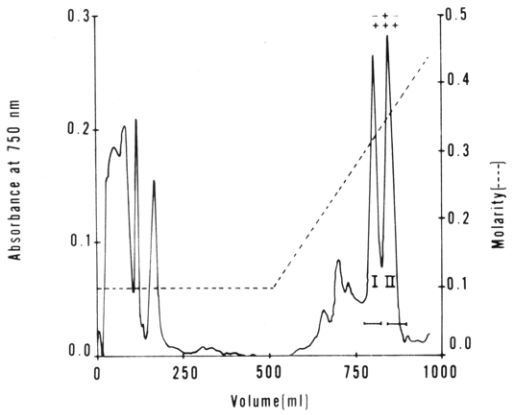


FIGURE 1: Chromatography of C3a anaphylatoxin from normal rat serum on a CM-cellulose CM 32 column (1.5 × 17 cm) equilibrated with 0.15 M ammonium formate buffer at pH 5.0. The anaphylatoxin-containing pool recovered after Sephadex G-100 gel filtration was applied to the CM-cellulose column. The column was initially developed with 510 mL of the 0.15 M buffer, and anaphylatoxin was eluted by a linear salt gradient. The gradient was composed of 250 mL of 0.15 M buffer and 250 mL of 0.5 M ammonium formate buffer at pH 5.0, and elution was performed at a flow rate of 18 mL/h. Biological activity was determined using the guinea pig ileal assay (Hugli et al., 1975b).

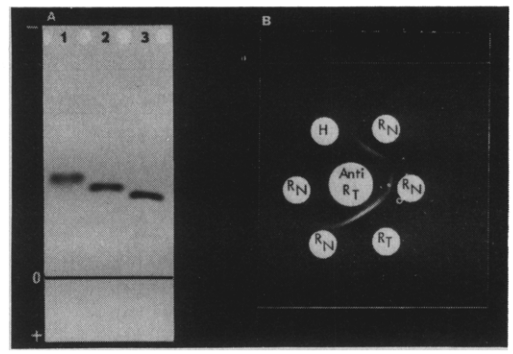


FIGURE 2: (A) Electrophoretic behavior of C3a anaphylatoxins on a microzone cellulose acetate strip. Samples were applied to the cellulose acetate strip on the line identified as the origin (O), and electrophoresis was performed for 20 min at 250 V, room temperature, and pH 8.6. The proteins were visualized after electrophoresis by staining with amido Schwarz. From left to right: (1) human C3a, (2) rat C3a from normal rats (Figure 1, pool II), and (3) rat $C3a_{des\text{-}Arg}$ from normal rats (Figure 1, pool I). (B) Double immunodiffusion analysis of C3a anaphylatoxins in agar gel. Rabbit antibody to $C3a_{des\text{-}Arg}$ obtained from tumor-bearing rats (Anti-$R_T$) was added to the center well. The C3a anaphylatoxins were placed in the outer wells in a clockwise arrangement: human C3a (H, 1.0 μg), normal rat C3a ($R_N$, 0.9 μg), normal rat C3a ($R_N$, 0.45 μg), tumor-bearing rat $C3a_{des\text{-}Arg}$ ($R_T$, 1.0 μg), normal rat C3a ($R_N$, 0.23 μg), and normal rat C3a ($R_N$, 0.11 μg).

Tryptic Peptides. Prior to tryptic digestion $C3a_{des\text{-}Arg}$ (500 nmol) was reacted with maleic anhydride in 1 mL of 8 M urea (pH 9.3). A 50-fold excess of maleic anhydride was added over free amino groups of $C3a_{des\text{-}Arg}$. The reagent was added stepwise over a 1-h period, with solid sodium carbonate added to maintain pH 9.3. The derivatized peptide was desalted by extensive dialysis and lyophilized.

Maleoylated $C3a_{des\text{-}Arg}$ was dissolved in 1 mL of 0.2 M $NH_4HCO_3$ (pH 8.2) and digested with trypsin (1%, w/w) overnight at 37 °C. The digest mixture was applied directly to a Sephadex G-25 gel-filtration column (0.9 × 200 cm) equilibrated in 0.2 M $NH_4HCO_3$ (pH 8.2). The column effluent was monitored at 280 nm for protein and at 250 nm for maleoyl groups.

## Results

*Purification of C3a and $C3a_{des\text{-}Arg}$.* Active rat anaphylatoxin, C3a, was generated in normal rat serum during complement activation by inulin. Gel filtration and CM-cellulose chromatography (Figure 1) resulted in the isolation of two anaphylatoxin pools containing approximately equal quantities of material. Pool I was found to contain essentially pure $C3a_{des\text{-}Arg}$, while pool II contained biologically active C3a. Rat C3a recovered from DEAE-Sephadex chromatography was judged homogeneous by polyacrylamide gel electrophoresis in $NaDodSO_4$ and by microzone electrophoresis (Figure 2A) with an estimated molecular weight of 9000. Ouchterlony immunodiffusion in gels was used to test cross-reactivity between purified rat C3a (active) and the antisera formed against rat $C3a_{des\text{-}Arg}$ (inactive) purified from tumor-bearing rats. $C3a_{des\text{-}Arg}$ antisera tested against human C3a, rat C3a, and rat $C3a_{des\text{-}Arg}$ (Figure 2B) showed immunochemical identity between rat anaphylatoxins prepared from normal and tumor-bearing rats. Human C3a showed no visible immunochemical cross-reactivity to any of the rat anaphylatoxins. C3a (Hugli and Müller-Eberhard, 1978) purified from normal rat serum (Figure 1, pool II) induced smooth-muscle contraction and was found to be devoid of somatomedin biological activity. The average yield of rat C3a was 15.4 mg/L, while the yield of rat $C3a_{des\text{-}Arg}$ was 13.0 mg/L when isolated from normal serum by the procedure described under Materials and Methods. The combined yield of immunoreactive C3a from complement-activated serum was then found to be 28.5 mg/L.

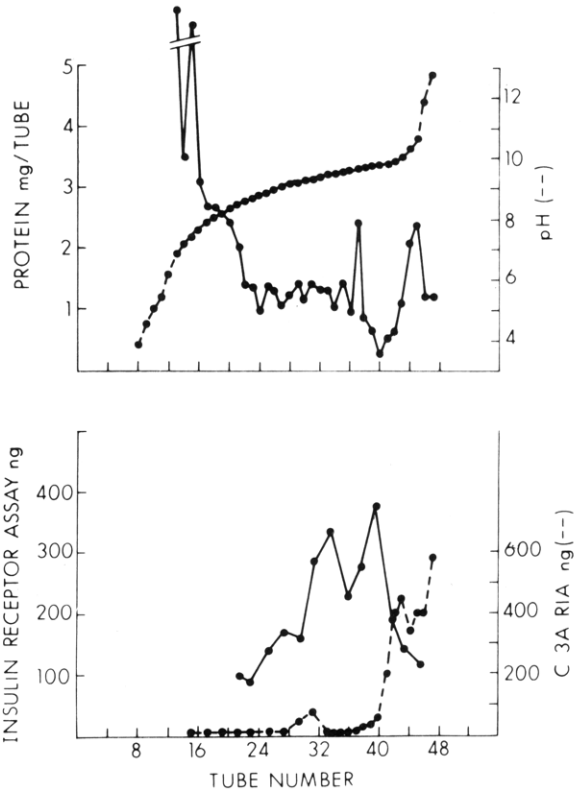


FIGURE 3: Separation of somatomedin activity from immunoreactive $C3a_{des\text{-}Arg}$ by isoelectric focusing in a linear sucrose gradient, 1% ampholine, pH range 7–9. The somatomedin/$C3a_{des\text{-}Arg}$ containing pool recovered after isoelectric focusing (pH range 3.5–10) was applied to the column. Fractions of 3 mL were collected.

$C3a_{des\text{-}Arg}$ was the major form of the anaphylatoxin purified from the serum of rats bearing transplanted pituitary tumors. Isoelectric focusing in the pH 3.5–10 range resulted in a pool containing both immunoreactive $C3a_{des\text{-}Arg}$ and somatomedin activity. Resolution of these two peptides was achieved by isoelectric focusing utilizing a narrower pH range of 7–9 (Figure 3) and resulted in the recovery of homogeneous

TABLE I: Amino Acid Analyses of C3a, $C3a_{des\text{-}Arg}$ and Peptides Isolated from Tryptic and CNBr Digestion of *S*-[$^{14}$C]Carboxymethyl-$C3a_{des\text{-}Arg}$.[a]

| | C3a | $C3a_{des\text{-}Arg}$ | CN-II | CN-I-I | CN-I-III | MT-III |
|---|---|---|---|---|---|---|
| Asp | 7.05 (7) | 7.33 (7) | 2.86 (3) | 2.54 (3) | 0.42 (0) | 0.96 (1) |
| Thr | 2.98 (3) | 3.07 (3) | 0.91 (1) | 0.78 (1) | 0.92 (1) | 0.12 (0) |
| Ser | 3.12 (3) | 2.93 (3) | (0) | (0) | 1.65 (2) | 0.88 (1) |
| Glu | 8.88 (9) | 9.20 (9) | 2.30 (2) | 1.54 (2) | 3.20 (3) | 2.16 (2) |
| Pro | 1.00 (1) | 1.27 (1) | (0) | (0) | (0) | (0) |
| Gly | 5.39 (5) | 5.20 (5) | 2.71 (3) | 1.24 (1) | 1.30 (1) | 1.26 (1) |
| Ala | 4.14 (4) | 4.07 (4) | 0.95 (1) | 0.96 (1) | 1.72 (2) | 0.89 (1) |
| ½-Cys | 5.82 (6)[b] | 5.61 (6)[c] | 1.56 (2) | 1.51 (2)[c] | 1.57 (2)[c] | (0) |
| Val | 1.95 (2) | 1.90 (2) | (0) | 0.59 (1) | (0) | 2.21 (2) |
| Met | 4.71 (5)[d] | 4.93 (5) | 0.38[e] (1) | (0) | 0.21 (1)[e] | 1.14 (1) |
| Ile | 2.90 (3) | 3.07 (3) | (0) | 0.72 (1) | 0.79 (1) | (0) |
| Leu | 6.99 (7) | 7.30 (7) | 1.00 (1) | 2.86 (3) | 1.76 (2) | 3.16 (3) |
| Tyr | 2.84 (3) | 2.93 (3) | 0.92 (1) | 0.88 (1) | 0.81 (1) | (0) |
| Phe | 0.86 (1) | 1.00 (1) | (0) | (0) | 0.92 (1) | (0) |
| Lys | 6.08 (6) | 6.13 (6) | 2.96 (3) | 1.10 (1) | 1.91 (2) | 0.31 (0) |
| His | 2.06 (2) | 1.87 (2) | (0) | 1.81 (2) | (0) | 0.86 (1) |
| Arg | 11.00 (11) | 9.97 (10) | 1.26 (1) | 3.26 (3) | 3.27 (3) | 0.91 (1) |
| total | 78 | 77 | 19 | 22 | 22 | 14 |
| res no. | 1–78 | 1–77 | 10–28 | 56–77 | 34–55 | 71–77; 1–7 |

[a] Values are given in residues/mole from 24-h hydrolyses on duplicate samples. Assumed integral values are in parentheses. [b] Determined as cysteic acid. [c] Determined as *S*-carboxymethylcysteine. [d] As methionine sulfone. [e] As homoserine.

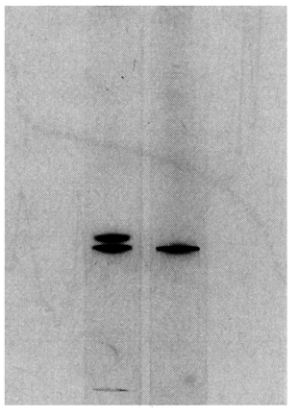

FIGURE 4: $NaDodSO_4$–polyacrylamide gel electrophoresis of 20 μg of rat $C3a_{des\text{-}Arg}$ in 15% acrylamide, 0.1% $NaDodSO_4$, and 0.5 mM 2-mercaptoethanol. Left-hand channel contains $C3a_{des\text{-}Arg}$ and cytochrome *c*, mol wt 12 400. Right-hand channel contains 20 μg of $C3a_{des\text{-}Arg}$ obtained after isoelectric focusing. Proteins were visualized after electrophoresis by staining with Coomassie blue.

$C3a_{des\text{-}Arg}$. The $C3a_{des\text{-}Arg}$ was judged pure on the basis of $NaDodSO_4$ gel electrophoresis experiments (Figure 4) and N-terminal sequence analyses as discussed below. Purified rat $C3a_{des\text{-}Arg}$ was devoid of somatomedin biological activity and was approximately 1/20 as potent as rat anaphylatoxin for stimulating the contraction of guinea pig ileum. The low level of anaphylatoxin activity associated with $C3a_{des\text{-}Arg}$ is probably due to a trace of intact C3a that was copurified. A total of 2 L of rat serum was processed with the average yield of $C3a_{des\text{-}Arg}$ being 25 mg/L serum.

*Amino Acid Composition of C3a and $C3a_{des\text{-}Arg}$.* The amino acid composition of rat anaphylatoxin and $C3a_{des\text{-}Arg}$ are shown in Table I. The two peptides differ in composition by only one arginine residue. Tryptophan was not detected. The protein was devoid of free sulfhydryl groups, suggesting that all six half-cystine residues in C3a are present in disulfide linkages.

*Sequencer Analysis of Intact Rat C3a and $C3a_{des\text{-}Arg}$.* High-sensitivity sequence analysis on 50 nmol of native rat $C3a_{des\text{-}Arg}$ established the purity of the material in addition to the first five residues. Only one polypeptide chain was observed utilizing sequencing methods capable of detecting 100 pmol of free amino group, although contaminants with a blocked α-amino group would not be seen. A second Sequencer analysis with reduced and *S*-[$^{14}$C]carboxymethyl-$C3a_{des\text{-}Arg}$ (450 nmol) permitted positive identifications of the first 55 residues (Table II). Residues 57 and 58 were established as half-cystines in this degradation on the basis of the $^{14}$C isotope released at the corresponding cycle. The average repetitive yield for each step through the first 55 cycles was 94%. Automated sequencer analyses on functionally active rat C3a anaphylatoxin (20 cycles) gave identical amino acid assignments as those determined for $C3a_{des\text{-}Arg}$ purified from tumor-bearing rats.

*CNBr Peptides.* Peptides obtained by CNBr cleavage of carboxymethylated $C3a_{des\text{-}Arg}$ were purified on a Sephadex G-50 column as shown in Figure 5. Pool II was rechromatographed on the same column and analyzed without further purification. Peptide CN-II consisted of 19 amino acids (Table I). Automated sequence analyses permitted positive identifications of the first 8 residues of CN-II (Figure 6), confirming that this peptide occupied residues 10–28 of C3a as established by the N-terminal sequence analysis. The first pool (CN-I) (Figure 5) was further fractionated on a SP-Sephadex ion-exchange column utilizing a pyridine–acetate gradient. The elution profile of this column is shown in Figure 7. Peptide CN-I-I consisting of 22 amino acids (Table I) was degraded in the automated sequencer with continuous identifications of Pth residues possible for 15 cycles (Figure 6). This peptide lacked homoserine, suggesting it occupied the C-terminal portion of the molecule. Identification of cysteine residues at positions 2 and 3 of the degradation permitted placement of peptide CN-I-I immediately following the methionine residue located at residue 55 (Figure 6). Peptide CN-I-III consisted of 22 amino acids (Table I) and was degraded successfully for eight cycles in the sequencer, as shown in Figure 6. This peptide overlapped with the previously determined sequence of residues 34–54 (Figure 6). Peptide CN-I-II, as determined by amino acid analysis, represented a small fraction of peptide CN-II

TABLE II: Automated Amino Terminal Degradation of *S*-[$^{14}$C]-Carboxymethyl-C3a$_{des\text{-}Arg}$.[a]

| cycle | res | yield (nmol)[b] | cycle | res | yield (nmol)[b] |
|---|---|---|---|---|---|
| 1 | Ser | 151.1 | 30 | Asp | 61.0 |
| 2 | Val | 396.6 | 31 | Ile | 62.3 |
| 3 | Glu | 250.7 | 32 | Pro | 25.1 |
| 4 | Leu | 354.1 | 33 | Met | 48.9 |
| 5 | Met | 326.8 | 34 | Lys | *d* |
| 6 | Glu | 291.3 | 35 | Tyr | 40.1 |
| 7 | Arg | *c* | 36 | Ser | 10.2 |
| 8 | Arg | *c* | 37 | Cys | *e* |
| 9 | Met | 253.9 | 38 | Glu | 29.8 |
| 10 | Asp | 245.6 | 39 | Arg | *c* |
| 11 | Lys | *d* | 40 | Arg | *c* |
| 12 | Ala | 231.3 | 41 | Ala | 32.0 |
| 13 | Gly | 142.1 | 42 | Arg | *c* |
| 14 | Glu | 153.2 | 43 | Leu | 28.6 |
| 15 | Tyr | 186.2 | 44 | Ile | 27.1 |
| 16 | Thr | 98.1 | 45 | Thr | 9.3 |
| 17 | Asp | 153.2 | 46 | Glu | 15.4 |
| 18 | Lys | *d* | 47 | Gly | 9.8 |
| 19 | Gly | 101.3 | 48 | Glu | 19.7 |
| 20 | Leu | 138.0 | 49 | Ser | 3.2 |
| 21 | Arg | *c* | 50 | Cys | *e* |
| 22 | Lys | *d* | 51 | Lys | *d* |
| 23 | Cys | *e* | 52 | Leu | 11.3 |
| 24 | Cys | *e* | 53 | Ala | 9.8 |
| 25 | Glu | 89.3 | 54 | Phe | 5.6 |
| 26 | Asp | 81.6 | 55 | Met | 3.5 |
| 27 | Gly | 32.1 | 56 | nd | nd[f] |
| 28 | Met | 59.6 | 57 | Cys | *e* |
| 29 | Arg | *c* | 58 | Cys | *e* |

[a] Degradation on 450 nmol of peptide employing the dilute Quadrol program of Brauer et al. (1975). Phenylthiohydantoins were identified and quantitated by a combination of gas–liquid and thin-layer chromatography according to Jacobs and Niall (1975). [b] The yield for Pth residues was determined by gas–liquid chromatography. [c] Pth-Arg was identified by the phenanthrenequinone reaction. [d] Pth-Lys was identified by thin-layer chromatography. [e] Identified as radioactivity migrating with carboxymethylcysteine on thin-layer chromatography. [f] Not determined.

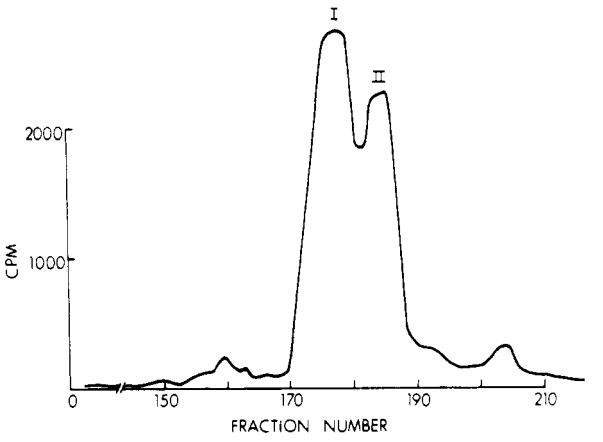


FIGURE 5: Elution profile of CNBr peptides of *S*-[$^{14}$C]carboxymethyl-C3a$_{des\text{-}Arg}$ on a 2.5 × 100 cm column of Sephadex G-50 equilibrated with 1 N acetic acid at 4 °C. Fractions of 2.0 mL were collected and radioactivity was counted in 50 μL aliquots taken from each fraction. Peaks I and II were further fractionated as detailed under Materials and Methods.

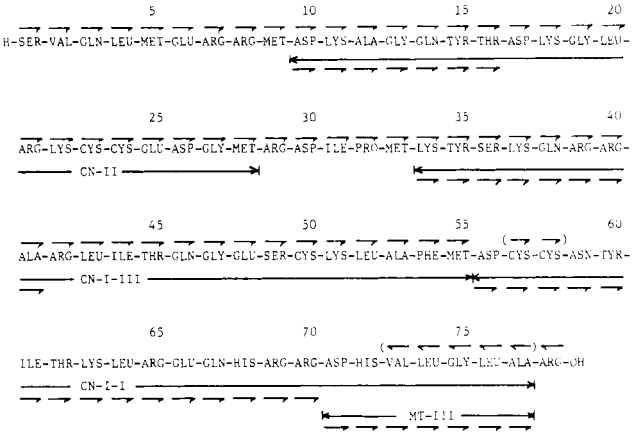


FIGURE 6: The amino acid sequence of rat anaphylatoxin (C3a). Residues identified by sequencer analyses (—) and carboxypeptidase A and/or B hydrolysis (---) are so indicated. Abbreviations used are: CN-, cyanogen bromide; MT-, maleoylated tryptic. See text for description of peptide purification.

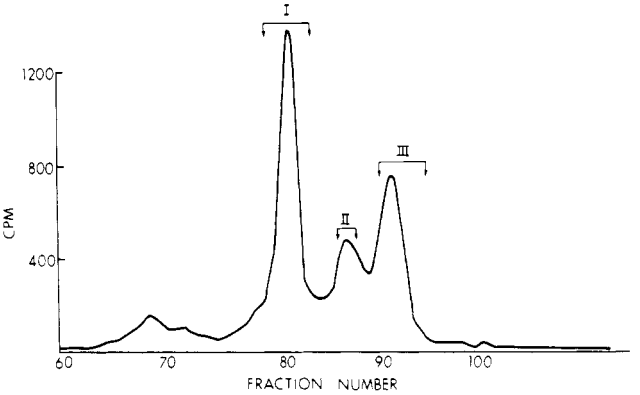


FIGURE 7: Elution profile of CNBr peptides of *S*-[$^{14}$C]carboxymethyl-C3a$_{des\text{-}Arg}$. Peak I (Figure 5) was loaded onto a 0.9 × 25 cm column of SP-Sephadex at 55 °C. The column was developed with a double linear gradient of pyridine acetate according to Bradshaw et al. (1969). Fractions of 1.5 mL were collected and radioactivity was counted from 50-μL aliquots of alternate fractions. Fractions were pooled as indicated by arrows and Roman numerals.

that was present in pool I (Figure 5) and was not further analyzed.

*Maleoylated Tryptic Peptides.* Native C3a$_{des\text{-}Arg}$, maleoylated and digested with trypsin, was applied directly to a Sephadex G-25 gel-filtration column with fractions monitored for absorbance at 250 and 280 nm (Figure 8). Amino acid analyses revealed that pool MT-I represented peptides held together by the three disulfide linkages. Pool MT-II was not analyzed but pool MT-III was found to contain 14 amino acids by amino acid analysis (Table I). From previous sequence work it was deduced that this pool contained two peptides: the N-terminal segment (seven residues) in addition to a C-terminal heptapeptide. Since the α-amino group of the N-terminal peptide was blocked with a maleoyl group, manual degradations were performed on pool MT-III without separating the two peptides. As shown in Table III, the C-terminal peptide could be unambiguously sequenced for six cycles with the C-terminal alanine residue being identified by amino acid analysis.

*C-Terminal Analyses.* Treatment of native C3a$_{des\text{-}Arg}$ with carboxypeptidase B alone for 5 min did not release any amino acids. Addition of carboxypeptidase A followed by a 1-h digestion released alanine, leucine, glycine, and valine at a rate of 1.0:1.5:0.4:0.3, respectively. Although the above experiments could not unambiguously define the C-terminal sequence, they did add confirmation to the C-terminal sequence Val-Leu-Gly-Leu-Ala determined by manual degradations on pool MT-III (Table III). Treatment of native anaphylatoxin with carboxypeptidase B (5 min) followed by carboxypeptidase A

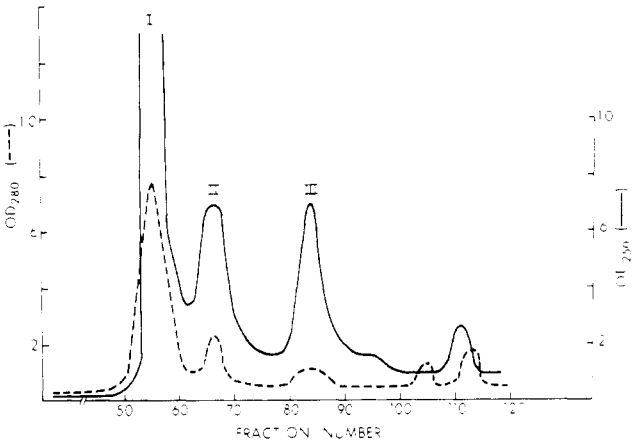


FIGURE 8: Sephadex G-25 column chromatography of maleoylated $C3a_{des\text{-}Arg}$ digested with trypsin (see Materials and Methods). Fractions (2 mL) were monitored for absorbance at 280 and 250 nm. Peak III was used for sequence analysis and contained two components, the C-terminal heptapeptide (residues 71-77) and a blocked N-terminal fragment consisting of residues 1-7.

TABLE III: Manual Edman Degradation of Fragment MT-III of Maleoylated $C3a_{des\text{-}Arg}$.[a]

| cycle | res | yield (nmol)[b] |
|---|---|---|
| 1 | Asp | 145.1 |
| 2 | His | *c* |
| 3 | Val | 124.2 |
| 4 | Leu | 101.1 |
| 5 | Gly | 32.1 |
| 6 | Leu | 48.3 |
| 7 | Ala | *d* |

[a] Manual degradation of 200 nmol of fragment MT-III employing the procedure of Sauer et al. (1974). [b] Yields of Pth residues determined by gas-liquid chromatography. [c] The phenylthiohydantoin derivative of histidine was identified by the Pauly reaction. [d] Alanine was determined by amino acid analysis.

resulted in the initial release of arginine followed by the release of amino acids in approximately the same proportions as noted above with $C3a_{des\text{-}Arg}$.

The complete amino acid sequence of rat anaphylatoxin is presented in Figure 6, indicating the placement of tryptic and CNBr peptides used in the structural analysis. A single N-terminal sequencer analysis allowed overlap of all six CNBr peptides theoretically obtainable, greatly facilitating the primary structural determination of rat C3a.

## Discussion

Determination of the complete primary structure of rat anaphylatoxin permits a comparison of this molecule with the previously determined amino acid sequences of human and porcine C3a (Hugli, 1975; Corbin and Hugli, 1976) as shown in Figure 9. It appears that anaphylatoxins from these selected species share a close genetic relatedness, as evidenced by an approximate 70% sequence identify that exists between the three polypeptides. Mutations which do occur from species to species fall, for the most part, within the two most commonly observed classes of allowed mutation as defined by Dayhoff (1972). It is not surprising, therefore, to find that rat C3a exhibits certain structure-activity relationships observed earlier with the procine and human anaphylatoxins (Hugli et al., 1975b). The rat peptide, like human and porcine C3a, has a strict requirement for a C-terminal arginyl residue to maintain full biological activity. Inactivation by either commercially available carboxypeptidase B or a carboxypeptidase B like serum enzyme results in a total loss of C3a activity, as measured by smooth-muscle contraction. Aditionally, anaphylatoxins from the three species share identical placement of half-cystine residues, including two distinctive Cys-Cys sequences. Although the disulfide pairings have yet to be rigorously determined in any of the species, owing in large part to the difficulty in effecting a cleavage between the half-cystinyl pairs, it is extremely likely that the peptides share identical disulfide linkages. This arrangement would confer, in all probability, a similar higher ordered structure among the three peptides. In this regard, previous studies have shown a high degree of similarity in the secondary structures of human and porcine C3a as assessed by circular dichroism spectra (Hugli et al., 1975a).

The biological activity of rat C3a for inducing contraction of the guinea pig ileal smooth muscle suggests that this molecule is approximately twice as active as either human or porcine C3a. It required only $5 \times 10^{-9}$ M rat C3a to induce contraction, while 1.1 to $1.4 \times 10^{-8}$ M porcine or human C3a gave a comparable response. The functional difference between these three C3a structures may be related to substitutions located near the essential carboxy terminus (Hugli and Erickson, 1976). More than half of the residue replacements that exist between the human, porcine, and rat sequences occur at common sites indicating that these sites are apparently nonessential to the preferred conformation of the molecule. Often the residues at these variable sites differ for all three sequences, suggesting a means by which these molecules may each express individual immunological behavior. If these variable sites define the individual antigenic determinants of the molecule, this would explain the lack of cross-reactivity toward the heterologous antibody.

It was observed at the inception of these studies that rat $C3a_{des\text{-}Arg}$ could be purified in high yield from the serum of rats bearing growth hormone secreting pituitary tumors. This serum provides an enriched source for the isolation of the growth factor, somatomedin, whose biosynthesis is under the control of growth hormone (Salmon and Daughaday, 1957). Although a highly enriched somatomedin preparation can be obtained from this serum (Chochinov et al., 1977), it became apparent from structural studies that this preparation was predominantly $C3a_{des\text{-}Arg}$. This possibility was addressed by Chochinov et al. (1977) who noted a relatively low potency of this initial somatomedin preparation. The authors suggested that the isolation procedure could have selected a nonsomatomedin protein with the same size and charge characteristics of rat somatomedin and that somatomedin was only a minor component of the preparation. Subsequent studies have shown that both $C3a_{des\text{-}Arg}$ and rat somatomedin are highly cationic and copurify after isoelectric focusing, utilizing a wide-range pH gradient, 3.5-10. Successful resolution of rat $C3a_{des\text{-}Arg}$ and somatomedin was achieved by an alternative isoelectric focusing step, utilizing a narrow-range (pH 7-9) gradient. The slightly more basic $C3a_{des\text{-}Arg}$ molecule was then obtained in homogeneous form suitable for the structural studies presently described. Somatomedin activity, as measured by the ability of the peptide to stimulate radiolabeled sulfate uptake into hypophysectomized rat cartilage, was recovered free of immunoreactive $C3a_{des\text{-}Arg}$ (Daughaday, W. H., unpublished results). It is interesting to note that serum from rats bearing transplanted pituitary tumors also provided an enriched starting source for the isolation of rat $C3a_{des\text{-}Arg}$. Serum C3a levels, as measured by radioimmunoassay, showed a marked

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RAT | SER | VAL | GLN | LEU | MET | GLU | ARG | ARG | MET | ASP | LYS | ALA | GLY | GLN | TYR | THR | ASP | LYS | GLY | LEU | ARG | LYS | CYS | CYS | GLU | ASP |
| PORCINE[a] | ... | ... | ... | ... | ... | ... | LYS | ... | ... | ASN | ... | LEU | ... | ... | ... | SER | — | ... | GLU | ... | ... | ARG | ... | ... | ... | HIS |
| HUMAN[b] | ... | ... | ... | ... | THR | ... | LYS | ... | ... | ASN | ... | VAL | ... | LYS | ... | PRO | — | ... | GLU | ... | ... | ... | ... | ... | ... | ... |

| | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RAT | GLY | MET | ARG | ASP | ILE | PRO | MET | LYS | TYR | SER | CYS | GLN | ARG | ARG | ALA | ARG | LEU | ILE | THR | GLN | GLY | GLU | SER | CYS | LYS | LEU |
| PORCINE | ... | ... | ... | GLN | ASN | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | GLN | ... | HIS | ... | ... | ASN | ALA | ... | VAL | LYS |
| HUMAN | ... | ... | ... | GLN | ASN | ... | ... | ARG | PHE | ... | ... | ... | ... | ... | THR | ... | PHE | ... | SER | LEU | ... | ... | ALA | ... | ... | LYS |

| | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RAT | ALA | PHE | MET | ASP | CYS | CYS | ASN | TYR | ILE | THR | LYS | LEU | ARG | GLU | GLN | HIS | ARG | ARG | ASP | HIS | VAL | LEU | GLY | LEU | ALA | ARG |
| PORCINE | ... | ... | LEU | ASN | ... | ... | GLU | ... | ... | ALA | ... | ... | ... | GLN | ... | ... | SER | ... | ASN | LYS | PRO | ... | ... | ... | ... | ... |
| HUMAN | VAL | ... | LEU | ... | ... | ... | ... | ... | ... | ... | GLU | ... | ... | ARG | ... | ... | ALA | ... | ALA | SER | HIS | ... | ... | ... | ... | ... |

FIGURE 9: Comparison of the amino acid sequence of rat anaphylatoxin (C3a) with the published sequences of human and porcine C3a. Dots indicate identical residues found in either human or porcine C3a when compared to the rat C3a sequence. The deletions at position 17 of the porcine and human proteins, indicated by dashes, were arbitrarily inserted to maximize the homology with the rat protein: (a) porcine C3a sequences as determined by Corbin and Hugli (1976); (b) human C3a sequence as determined by Hugli (1975).

increase beginning 4 days after tumor transplant. At the time of sacrifice, circulating C3a levels (approximately 200 $\mu$g/mL) were five to ten times above those levels found in normal rats (Daniels, J. S., unpublished results). The fact that the third component of complement (C3) is highly sensitive to proteolytic attack could readily explain the presence of large quantities of $C3a_{des\text{-}Arg}$ in the serum of tumor-bearing rats. Extensive conversion of C3 may have occurred either as a direct result of the tumor or from the manipulations required to free somatomedin from its carrier protein (Chochinov et al., 1977). A variety of activatable, tryptic-like enzymes of serum are capable of converting C3 to C3a and C3b (Budzko et al., 1971). Once formed, the C3a is itself readily converted to $C3a_{des\text{-}Arg}$, a serum stable and inactive form of the anaphylatoxin by serum carboxypeptidase B (EC 3.4.12.7). Component C3 is a major protein in mammalian serum and up to 60 $\mu$g/mL quantities of $C3a_{des\text{-}Arg}$ may normally be formed in extensively activated serum.

To determine if the peptide isolated from the tumor-bearing animals was indeed an inactivated derivative of rat anaphylatoxin, the present studies were initiated to obtain rat C3a from normal rat serum. This was performed by a procedure originally developed by Fernandez and Hugli (1976) for purifying both human anaphylatoxins C3a and C5a. Inulin activation of complement in normal rat serum containing the competitive carboxypeptidase inhibitor $\epsilon$-aminocaproic acid (EACA) produces levels of C3a comparable to those generated in human and porcine sera (Hugli et al., 1975b). A mixture of both rat C3a and $C3a_{des\text{-}Arg}$ was obtained from serum activated even in the presence of 1 M EACA. This mixture was presumably obtained due to higher levels of carboxypeptidase activity in rat than in human serum. Chromatography on CM-cellulose resolved rat C3a and $C3a_{des\text{-}Arg}$ and they each exhibited immunological identity to one another as well as to the material isolated from the serum of tumor-bearing rats. Confirmation that the immunoreactive material isolated from tumor-bearing rats corresponds to $C3a_{des\text{-}Arg}$ obtained from complement-activated rat serum was supplied by electrophoretic comparisons on cellulose acetate, chemical compositional analyses, immunological cross-reactivity, and partial N-terminal sequence analyses of each polypeptide chain.

In summary, elucidation of the linear amino acid sequence of rat C3a has established that a close genetic relatedness exists between rat, porcine, and human C3a. Probably this same degree of relatedness will be observed when comparing the parent C3 structures from all three species. Except for a residue (aspartic acid) deletion at position 17 in porcine and human C3a, no remarkable differences occur between the three polypeptides. In fact, conservation of certain features is perhaps the most revealing result of this comparison. Most notably conserved are the six half-cystine positions and the COOH-terminal pentapeptide sequence Leu-Gly-Leu-Ala-Arg. Caporale et al. (1977) have synthesized a series of peptides that mimic the COOH-terminal structure of human C3a and that exhibit activity characteristic of the C3a anaphylatoxin; however, synthetic peptides shorter than the pentapeptide Leu-Gly-Leu-Ala-Arg expressed progressively lower activity, emphasizing the functional importance of this particular region of the C3a polypeptide. Quantitative activity differences between intact C3a from various species are small but rat C3a is apparently more potent than the C3a obtained from the other species studied to date. Finally, it should be mentioned that identification and removal of $C3a_{des\text{-}Arg}$ as a principal contaminant of the somatomedin preparation provide a major purification step in the isolation of rat somatomedin.

## References


Ambler, R. P. (1967), *Methods Enzymol. 11,* 155–166.

Angeletti, R. H., Bradshaw, R. A., and Wade, R. D. (1971), *Biochemistry 10,* 463–470.

Bradshaw, R. A., Babin, D. R., Nomoto, M., Srinivasin, N. G., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1969), *Biochemistry 8,* 3859–3867.

Brauer, A. W., Margolies, M. N., and Haber, E. (1975), *Biochemistry 14,* 3025–3029.

Budzko, D. B., Bokisch, V. A., and Müller-Eberhard, H. J. (1971), *Biochemistry 10,* 1166–1172.

Caporale, L. H., Erickson, B. W., and Hugli, T. E. (1977), *Pept., Proc. Am. Pept. Symp. 5th, 1977,* 225–227.

Chochinov, R. H., Mariz, I. K., and Daughaday, W. H. (1977), *Endocrinology 100,* 549–556.

Cochrane, C. G., and Müller-Eberhard, H. J. (1968), *J. Exp. Med. 127,* 371–386.

Corbin, N. C., and Hugli, T. E. (1976), *J. Immunol. 117,*

990-995.
Dayhoff, M. E., Ed. (1972), Atlas of Protein Sequence and Structure, Vol. 5, Silver Springs, Md., National Biomedical Research Foundation.
Daughaday, W. H., Hall, K., Raben, M. S., Salmon, W. D., Jr., Van den Brande, J. L., and Van Wyk, J. J. (1972), *Nature (London) 235*, 107-109.
Daughaday, W. H., Phillips, L. S., and Herrington, A. C. (1975), *Methods Enzymol. 37B*, 93-96.
Dias da Silva, W., and Lepow, I. H. (1967), *J. Exp. Med. 125*, 921-926.
Dias da Silva, W., Eisele, J. W., and Lepow, I. H. (1967), *J. Exp. Med. 126*, 1027-1048.
Edelhoch, H. (1967), *Biochemistry 6*, 1948-1954.
Edman, P. (1960), *Ann. N.Y. Acad. Sci. 88*, 602-610.
Ellmann, G. L. (1959), *Arch. Biochem. Biophys. 82*, 70-77.
Fernandez, H. N., and Hugli, T. E. (1976), *J. Immunol. 117*, 1688-1694.
Hugli, T. E. (1975), *J. Biol. Chem. 250*, 8293-8301.
Hugli, T. E., and Erickson, B. W. (1976), *Proc. Natl. Acad. Sci. U.S.A. 74*, 1826-1830.
Hugli, T. E., and Müller-Eberhard, H. J. (1978), *Adv. Immunol. 26* (in press).
Hugli, T. E., Morgan, W. T., and Müller-Eberhard, H. J. (1975a), *J. Biol. Chem. 250*, 1479-1483.
Hugli, T. E., Vallota, E. H., and Müller-Eberhard, H. J. (1975b), *J. Biol. Chem. 250*, 1472-1478.
Jacobs, J. W., and Niall, H. D. (1975), *J. Biol. Chem. 250*, 3629-3636.
Johnson, A. R., Hugli, T. E., and Müller-Eberhard, H. J. (1975), *Immunology 28*, 1067-1080.
Laemmli, V. K. (1970), *Nature (London) 227*, 680-682.
Marshall, R. N., Underwood, L. E., Voina, S. J., Foushee, D. B., and Van Wyk, J. J. (1974), *J. Clin. Endocrinol. Metab. 39*, 283-291.
Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244*, 5597-5610.
Salmon, W. D., and Daughaday, W. H. (1957), *J. Lab. Clin. Med. 49*, 825-831.
Sauer, R. T., Niall, H. D., Hogan, M. L., Keutmann, H. T., O'Riordan, J. L. H., and Potts, J. T., Jr. (1974), *Biochemistry 13*, 1994-1999.
Van Wyk, J. J., Underwood, L. E., Hintz, R. L., Viona, S. J., and Weaver, R. P. (1974), *Recent Prog. Horm. Res. 30*, 259-295.
Wuepper, K. D., Bokisch, V. A., Müller-Eberhard, H. J., and Stoughton, R. B. (1972), *Clin. Exp. Immunol. 11*, 13-20.

# Tyrosyl–Base–Phenylalanyl Intercalation in Gene 5 Protein–DNA Complexes: Proton Nuclear Magnetic Resonance of Selectively Deuterated Gene 5 Protein†

Joseph E. Coleman* and Ian M. Armitage

ABSTRACT: The interactions of oligodeoxynucleotides with the aromatic residues of gene 5 protein in complexes with $d(pA)_8$ and $d(pT)_8$ have been determined by $^1H$ NMR of the protein in which the five tyrosyl residues have been selectively deuterated either in the 2,6 or the 3,5 positions. Only the 3,5 protons of the three surface tyrosyls (26, 41, and 56) interact with the bases. The remainder of the aromatic protons undergoing base-dependent upfield ring-current shifts on complex formation are phenylalanyl protons, assigned to Phe(13) on the basis of model building. $^{19}F$ NMR of the complexes of the *m*-fluorotyrosyl-labeled protein with $d(pT)_4$ and $d(pA)_8$ confirms the presence of ring-current perturbations of nuclei at the 3,5-tyrosyl positions of the three surface tyrosyls. Differential expression of the $^{19}F\{^1H\}$ nuclear Overhauser effect confirms the presence of two buried and three surface tyrosyl residues. A new model of the DNA binding groove is presented involving Tyr(26)-base-Phe(13) intercalation.

Gene 5 protein is a single-stranded DNA-binding protein elaborated by *Escherichia coli* infected with fd bacteriophage and coded for by the phage genome (Alberts and Frey, 1970; Alberts et al., 1972; Oey and Knippers, 1972). Mutants of fd temperature sensitive in the gene 5 product show that gene 5 protein is required for the shift from double-stranded DNA synthesis of the replicative forms (RF) to single-stranded DNA synthesis of the daughter viral genomes (Salstrom and Pratt, 1971; Mazur and Model, 1973). Gene 5 protein binds to single-stranded DNA of any sequence, although binding studies with defined oligodeoxynucleotides show it to have a much greater affinity for adenine-rich regions than for thymine-rich regions (Coleman et al., 1976). A variety of titration studies using spectroscopic detection of complex formation have shown that each gene 5 monomer ($M_r$ = 9689) interacts with ca. four bases (Alberts et al., 1972; Pratt et al., 1974; Anderson et al., 1975). In the presence of low salt, gene 5 protein will melt certain double-stranded homopolymers at room temperature, e.g., poly[d(A-T)], since the equilibrium is far in favor of the gene 5 protein complex with the single strand. On interaction with circular single-stranded fd DNA, both in vivo and in vitro

† From the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510. *Received May 10, 1978.* This work was supported by National Institutes of Health Grants AM 18778 and GM 21919 and National Science Foundation Grant PCM 76-82231. The 270-MHz $^1H$ NMR work was supported in part by Grant 1-P07-PR 00798 from the National Institutes of Health Division of Research Resources.

0006-2960/78/0417-5038$01.00/0 