graciously performing some of the velocity sedimentation experiments.

References

Adkins, B. J., and Foster, J. F. (1965), *Biochemistry 4*, 634.
Benson, J. V., Jr., and Patterson, J. A. (1965), *Anal. Chem. 37*, 1108.
Broome, J. (1963), *Nature 199*, 179.
Davis, B. J. (1964), *Ann. N. Y. Acad. Sci. 121*, 404.
Dayhoff, M. O., Perlmann, G. E., and MacInnes, D. A. (1952), *J. Am. Chem. Soc. 74*, 2515.
Dintzis, H. M. (1952), Ph.D. Dissertation, Harvard University, Cambridge, Mass.
Foster, J. F. (1960), *Plasma Proteins 1*, 179.
Matsubara, H., Kasper, C. B., Brown, D. M., and Smith, E. L. (1965), *J. Biol. Chem. 240*, 1125.
Okunuki, K., Matsubara, H., Nishimura, S., and Hagihara, B. (1956), *J. Biochem. (Tokyo) 43*, 857.
Ornstein, L. (1964), *Ann. N. Y. Acad. Sci. 121*, 321.
Pederson, K. O. (1962), *Arch. Biochem. Biophys., Suppl. 1*, 157.
Porath, J., and Bennich, H. (1962), *Arch. Biochem. Biophys., Suppl 1*, 152.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.
Spahr, P. F., and Edsall, J. T. (1964), *J. Biol. Chem. 239*, 850.
Sterman, M. D. (1955), Ph.D. Dissertation, Purdue University, Lafayette, Ind.
Tristam, G. R. (1953), *Proteins 1*, 215.
Weber, G., and Young, L. B. (1964a), *J. Biol. Chem. 239*, 1415.
Weber, G., and Young, L. B. (1964b), *J. Biol. Chem. 239*, 1424.
Williams, E. J., and Foster, J. F. (1960), *J. Am. Chem. Soc. 82*, 3741.
Yphantis, D. A. (1964), *Biochemistry 3*, 297.

Primary Structure of the Cytochrome *c* from the Snapping Turtle, *Chelydra Serpentina**

S. K. Chan,† I. Tulloss, and E. Margoliash

ABSTRACT: The primary structure of the cytochrome *c* from hearts of the snapping turtle, *Chelydra serpentina*, has been determined from the complete amino acid sequences of peptides isolated from a chymotryptic digest. The positions of these peptides in the over-all sequence were assigned by homology to the structures of the other 15 cytochromes *c* of known structure.

The *C. serpentina* protein bears the characteristics of all "mammalian-type" cytochromes *c* including the clustered distribution of hydrophobic and basic residues, an acetylglycine NH_2-terminal residue, a single polypeptide chain 104 residues long, a heme prosthetic group bound to cysteinyl residues in positions 14 and 17, and the typical invariant sequence Asn-Pro-Lys-Lys-Tyr-Ile-Pro-Gly-Thr-Met at residues 70–80. Position 33 is occupied by an asparaginyl residue, as contrasted to the histidyl residue commonly occurring at this location in other cytochromes *c*. The importance of this substitution is discussed in connection with recent findings concerning the nature of the hemochrome-forming groups in cytochrome *c*.

Following the determination of the complete amino acid sequence of horse heart cytochrome *c* (Margoliash and Smith, 1961; Kreil and Tuppy, 1961; Margoliash *et al.*, 1961, 1962; Margoliash and Smith, 1962; Margoliash, 1962a; Tuppy and Kreil, 1962), the primary structures of the cytochromes *c* from a number of different species have been established through the efforts of several groups of investigators. To date, in addition to the horse heart protein, the cytochromes *c* of known structure include those from man (Matsubara and Smith, 1963), pig (Stewart and Margoliash, 1965), cow (Yasunobu *et al.*, 1963), chicken (Chan and Margoliash, 1966a), tuna (Kreil, 1963, 1965), a moth, *Samia cynthia* (Chan and Margoliash, 1966b), baker's yeast (Narita *et al.*, 1963), a rhesus monkey (Rothfus and Smith, 1965), the dog (McDowall and Smith, 1965), the rattlesnake (Bahl and Smith, 1965), the mold *Neurospora crassa* (Heller and Smith, 1965), the rabbit (Needleman and Margoliash, 1966), the great grey kangaroo, *Macropus canguru* (Nolan

* From the Biochemical Research Department, Abbott Laboratories, North Chicago, Illinois. *Received February 18, 1966; revised May 16, 1966.*

† Present address, Department of Biochemistry, University of Kentucky Medical Center, Lexington, Ky.

and Margoliash, 1966), and from the yeast, *Candida krusei* (Narita and Titani, 1966).

Comparisons of amino acid sequences within this set of homologous proteins have led to interesting conclusions concerning the relations of protein structure to the evolution of species (Margoliash *et al.*, 1963; Margoliash, 1963, 1964; Smith and Margoliash, 1964; Margoliash and Smith, 1965; see also review by Margoliash and Schejter, 1966). Such conclusions can be tested and extended only through the examination of further varieties of the protein from as wide a range of species as can be obtained. Moreover, these comparisons are yielding considerable information as to the extent of primary structure variability compatible with a unique set of functional parameters (see Margoliash and Schejter, 1966). Perhaps even more importantly, this work has made available for studies of structure–function relations in cytochrome *c* an extensive series of different polypeptide chains. Thus, for example, these amino acid sequences have served to identify the effect of chemical substitutions at each of the two common methionyl residues (Tsai and Williams, 1965a,b; Ando *et al.*, 1965) and at each of the three common histidyl residues (Horinishi *et al.*, 1965). Similarly, the use of a cytochrome *c* lacking one of these three histidines has made it possible to restrict to the remaining two the possible function of providing the hemochrome-forming side chains in cytochromes *c* in which all of the ϵ-amino groups have been suitably substituted (Hettinger and Harbury, 1964, 1965; Fanger and Harbury, 1965).

Continuing the studies of the structures of cytochromes *c* from various vertebrate classes and orders this paper presents the amino acid sequence of the protein prepared from hearts of the snapping turtle, *C. serpentina*.

Experimental Section

Snapping turtle cytochrome *c* was purified and crystallized from fresh frozen hearts by the aluminum sulfate extraction procedure (Margoliash and Walasek, 1966). The preparation was thoroughly dialyzed and lyophilized. The total amino acid composition, iron content, and dry weight were determined as previously described (Chan and Margoliash, 1966a,b).

A sample of 15 μmoles of cytochrome *c* was digested at 38° with a total of 6 mg of three-times-crystallized α-chymotrypsin (Worthington), which was added in three equal portions over a period of 9 hr. This digest was fractionated by ion-exchange column chromatography on Dowex 50-X2 resin (Bio-Rad Laboratories, Richmond, Calif.). Peptides were further purified, when necessary, by paper electrophoresis and chromatography as described by Chan and Margoliash (1966a,b). The purity of all fractions was monitored by paper electrophoresis and chromatography, with the use of the various color reactions for particular residues previously employed (see Margoliash and Smith, 1962). Ninhydrin-negative peptides were detected by the hypochlorite–starch–iodine reaction

FIGURE 1: Crystals of *C. serpentina* cytochrome *c*, grown in near-saturated $(NH_4)_2SO_4$ containing 1 M NaCl. Magnification: 100 times.

(Pan and Dutcher, 1956). Hydrazinolyses of peptides were performed according to Funatsu *et al.* (1964). Sequential degradations of peptides were carried out by the Edman procedure as modified by Hirs *et al.* (1960) and Konigsberg and Hill (1962). Amino acid compositions of peptides were determined on acid or total enzymic digests (leucine aminopeptidase) in the Beckman–Spinco automatic amino acid analyzer. Yields were calculated from the analyses of the purified peptides and represent minimal recoveries, not taking into account any losses incurred during purification. Digestions of peptides with trypsin, carboxypeptidase A, leucine aminopeptidase, elastase, and chymotrypsin (Worthington) were performed as previously described (Chan and Margoliash, 1966b). The amino acids released by leucine aminopeptidase and carboxypeptidase A were determined by amino acid analysis. One peptide (5 μmoles) was subjected to partial hydrolysis in 10 ml of 0.2 M acetic acid, under reduced pressure (5 mm), for 6 hr at 105° (see Table II).

The chymotryptic heme peptide remained bound to the resin and was not recovered from the Dowex 50 column chromatography. To prepare the heme peptide, 15 μmoles of cytochrome *c* was digested with chymotrypsin as described above. The digest was first fractionated by Sephadex gel filtration (Pharmacia, G-25, bead form; column, 0.9 × 150 cm; eluent, 50% acetic acid). Fractions containing heme were pooled, lyophilized, and dissolved in 10 ml of pyridine–acetic acid buffer, pH 6.4 (pyridine–glacial acetic acid–water, 889:31:16,000, v/v/v). Further purification was carried out by electrophoresis with the Brinkmann free-flow electrophoresis apparatus, Model F.F. (Brinkmann Instruments, N. Y.), which was operated at 1900 v and 200 ma in the same pyridine–acetic acid buffer at 4°. The protein solution was applied at a rate of 2 ml/hr with the background buffer flow at 114 ml/hr,

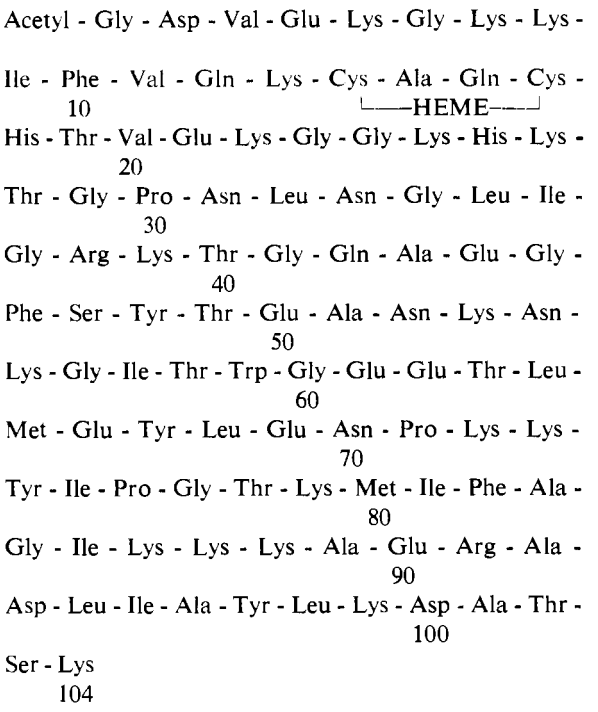

FIGURE 2: Amino acid sequence of *C. serpentina* cytochrome *c*.

the retention time between the plates being 58 min. The heme-containing fractions were pooled and lyophilized. The chymotryptic heme peptide thus prepared was shown to be homogeneous by paper electrophoresis and chromatography and its purity was ascertained by the stoichiometry of its amino acid composition (see Table III).

Results

Molecular Weight and Amino Acid Composition of Turtle Heart Cytochrome c. Figure 1 is a photograph of *C. serpentina* cytochrome *c* crystals obtained in near-saturated $(NH_4)_2SO_4$ at neutral pH. The molecular weight of the protein calculated from the iron content of 0.46% (12,195), assuming one heme per molecule of cytochrome *c*, is in good agreement with that calculated from the amino acid sequence (12,198) (see Figure 2). The amino acid composition of the protein, determined directly on acid hydrolysates, is reported in Table I and is identical with that derived from the amino acid sequence.

Separation and Purification of Peptides from the Chymotryptic Digest Nomenclature. The elution pattern of peptides obtained by column chromatography of the chymotryptic digest is given in Figure 3. The peptides were further purified, when necessary, as indicated in Tables II–X, which summarize the data used in establishing the amino acid sequences of the peptides. The purification procedures used in each case are indicated by the symbols PC for paper chromatography and PE for paper electrophoresis. The tables also list the chromatographic (*ch*) and electrophoretic (*el*) mobilities of the peptides on paper in centimeters, under the following standard conditions: electrophoresis, pH 6.4, pyridine–acetic acid buffer (Ingram, 1958), 19 v/cm, 90 min; chromatography, 1-butanol–acetic acid–water solvent (Margoliash and Smith, 1962), 16 hr. A zero (0) indicates no electrophoretic movement; a minus sign (−), movement toward the cathode, and a plus sign (+), movement toward the anode. In Tables II–X, the yield, the purification procedures, the electrophoretic–chromatographic mobilities, ninhydrin color, as well as other various color reactions for specific amino acid residues (see Chan and Margoli-

TABLE I: Amino Acid Composition of Snapping Turtle Heart Cytochrome *c*.[a]

Amino Acid	Amino Acid Residue (g/100 g of protein)	Amino Acid Residues per Molecule of Protein: From Analysis	Amino Acid Residues per Molecule of Protein: From Amino Acid Sequence
Lysine	18.55	17.9	18
Histidine	2.11	1.9	2
Arginine	2.46	2.0	2
Aspartic acid	7.55	8.2	8
Threonine	6.34	7.8	8
Serine	1.59	2.3	2
Glutamic acid	12.81	12.3	12
Proline	2.46	3.1	3
Glycine	5.99	13.0	13
Alanine	4.83	8.3	8
Half-cystine	1.68	1.4	2
Valine	2.43	3.0	3
Methionine	1.80	1.7	2
Isoleucine	6.24	6.9	7
Leucine	5.49	6.0	6
Tyrosine	4.90	3.7	4
Phenylalanine	3.58	3.0	3
Tryptophan	1.38[b]		1
Heme	5.09[b]		
Hydroxyl	0.14[b]		
Acetyl	0.22[b]		
Total	97.68		

[a] Samples of the native protein were hydrolyzed under reduced pressure in three-times glass-distilled 6 N HCl for 20 and 70 hr. Duplicate analyses were performed on each hydrolysate. The amount represented in each aliquot was calculated from the dry weight of the sample. The data reported are derived from the average or extrapolated values of all the analyses. [b] Calculated, assuming the appropriate number of residues per molecule.

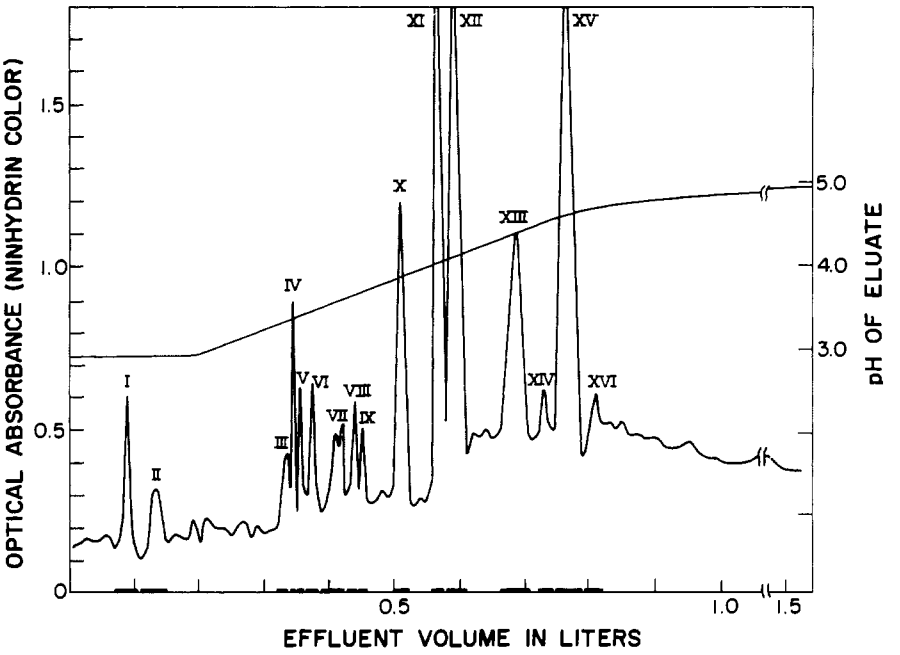

FIGURE 3: Elution pattern of peptides from a chymotryptic digest of turtle heart cytochrome *c*. The digest (15 μmoles) was chromatographed on a column of Dowex 50-X2 (0.9 × 150 cm) with a linear gradient established between pyridine–acetic acid buffer at pH 3.1 (0.2 M) and pH 4.8 (2.0 M). The column was operated at 40°. The line across the top of the pattern indicates the pH of the effluent fractions. The thick lines on the abscissa mark the fractions pooled and numbered in Roman numerals.

TABLE II: Amino Acid Sequence of Residues 1–5.

Peptide	Sequence: AcetylGly - Asp - **Val**-Glu-*Lys* ←A-3→←A-2→←—A-1—→
I	Gly, 0.97; Asp, 0.97; Val, 1.00; Glu, 0.97; Lys, 1.09 (80%; *el*, +3.8; *ch*, 14.0; blue)
Edman I	Gly, 1.00; Asp, 1.02; Val, 0.97; Glu, 1.09; Lys, 0.92
Hydrazinolysis	Free lysine, 0.90[a]
A-1	Val, 0.97; Glu, 1.00; Lys, 1.03 (PE, PC, *el*, 0; *ch*, 7.0; blue)
Edman I	**Val, 0.10;** Glu, 1.00; Lys, 0.50
A-2	Asp (PE; *el*, +11.0; *ch*, 7.5; blue)
A-3	Gly (PE, PC; *el*, +3.0; *ch*, 30.0; ninhydrin negative)

[a] This value is the molar ratio of lysine recovered to the amount of peptide subjected to hydrazinolysis.

ash, 1966b) are listed inside parentheses, in this order, immediately following the amino acid composition of the peptide. The peptides are denoted by Roman numerals according to the column chromatographic fractions in which they emerge (see Figure 3). When more than one peptide was recovered in a single chromatographic fraction, they are distinguished by lower case letters. Fragments derived from the initial chymotryptic peptides by digestion with trypsin, elastase, dilute acetic acid, and chymotrypsin are denoted by T, E, A, and C, respectively. For the digestion of peptides with carboxypeptidase A and leucine aminopeptidase, the values assigned to residues, when reported, are the molar ratios of amino acids released by the exopeptidases. Unless otherwise indicated, these ratios were calculated by taking the quantity for the residue released in the largest amount as unity. In the sequences given as headings of Tables II–X the residues in boldface type have been placed by Edman degradation, while those in italics have been placed by carboxypeptidase digestion or hydrazinolysis. In reporting Edman degradations, the residues and numbers marked in boldface type correspond to the residue removed at each step.

TABLE III: Amino Acid Sequence of Residues 11–26.

Sequence: **Val**-Gln-Lys-**Cys**-**Ala**-Gln-Cys-His-**Thr**-**Val**-Glu-Lys-**Gly**-Gly-Lys - *His*

HEME

T-1, T-2, T-3, ⟨T-4⟩; C-1, C-2, XIc

Peptide	
Chymotryptic heme peptide	Val, 1.85; Glu, 3.06; Lys, 2.90; Ala, 1.07; His, 1.80; Thr, 1.07; Gly, 2.26; Cys, not determined (80%; *el*, −3.5; *ch*, 1.5; heme, Pauly)
Edman I	**Val, 1.12;** Glu, 2.54; Lys, 1.50; Ala, 1.07; His, 2.08; Thr, 1.12; Gly, 2.07
Carboxypeptidase A	His
T-1	Val, 0.98; Glu, 1.02; Lys, 1.00 (PE; *el*, −8.0; *ch*, 7.0; blue)
Edman I	**Val, 0.10;** Glu, 1.00; Lys, 0.80
T-2	Ala, 1.17; Glu, 1.97; His, 1.02; Thr, 1.03; Val, 0.83; Lys, 0.98; Cys, not determined (PE; *el*, +2.0; *ch*, 7.0; heme; Pauly)
After oxidation	$CySO_3H$, 2.00; Ala, 1.02; Glu, 2.03; His, 0.98; Thr, 0.96; Val, 1.00; Lys, 1.01 (*el*, −3.0; *ch*, 5.0; yellow; Pauly)
Edman I	**$CySO_3H$, 1.22;** Ala, 0.82; Glu, 2.01; Thr, 0.98; Val, 0.98; Lys and His not determined
Edman II	$CySO_3H$, 1.00; **Ala, 0.31;** Glu, 1.97; Thr, 1.00; Val, 1.03; Lys and His not determined
C-1	$CySO_3H$, 1.82; Ala, 0.92; Glu, 1.18; His, 1.07 (PE; *el*, +6.2; *ch*, 4.5; Pauly)
C-2	Thr, 1.21; Val, 0.72; Glu, 0.93; Lys, 1.14 (PE; *el*, 0; *ch*, 7.0; blue)
Edman I	**Thr, 0.30;** Val, 0.93; Glu, 1.07; Lys, not determined
Edman II	Thr, 0.24; **Val, 0.25;** Glu, 1.00; Lys, not determined
T-3	Gly, 2.02; Lys, 0.98 (PE, PC; *el*, +7.5; *ch*, 3.0; blue)
T-4	His (PE, PC; *el*, −4.7; *ch*, 5.7; grey, Pauly)
XIc	Gly, 2.18; Lys, 1.00; His, 0.82 (6%; PE; *el*, −8.5; *ch*, 2.0; yellow; Pauly)
Edman I	**Gly, 1.10;** Lys, 0.10; His, 0.90
Carboxypeptidase A	His

Amino Acid Sequences of Peptides from the Chymotryptic Digest. The amino acid sequence of turtle heart cytochrome *c* is given in Figure 2. In all cases peptides from the original chymotryptic digest containing fragments of the same region of the protein were grouped together on the basis of their unique amino acid compositions. In the following descriptions these groups are considered in the order in which they appear in the over-all sequence, starting from the NH_2-terminal end.

RESIDUES 1–5. AcetylGly-Asp-Val-Glu-Lys (Table II, Peptide I). The amino acid composition of peptide I was unchanged after one step of Edman degradation, indicating a blocked NH_2 terminus. Dilute acetic acid digestion of peptide I yielded three fragments. Fragment A-3 gave no color with ninhydrin and yielded a single amino acid, glycine, after acid hydrolysis. As the electrophoretic and chromatographic mobilities of A-3 are identical with those of an authentic sample of *N*-acetylglycine, it was concluded that the NH_2-terminal residue of peptide I is acetylglycine. Fragment A-2 was characterized as free aspartic acid. From tryptic specificity and one step of Edman degradation, fragment A-1 was determined to be Val-Glu-Lys. Since hydrazinolysis of peptide I revealed the COOH-terminal residue to be lysine, the complete amino sequence of peptide I is established as given above.

RESIDUES 6–10. Gly-Lys-Lys-Ile-Phe (Peptide XIV). Peptide XIV (Gly, 1.11; Lys, 2.08; Ile, 0.88; Phe, 0.93; *el*, −9.0; *ch*, 11.5) gave a yellow ninhydrin color on paper and was recovered in a yield of 63%. One step of Edman degradation confirmed the NH_2- terminal residue as glycine (Edman I: **Gly, 0.05;** Lys, 2.00; Ile, 1.02; Phe, 0.99). Carboxypeptidase A digestion of peptide XIV released free isoleucine and phenylalanine (Ile, 1.00; Phe, 0.98) in addition to a residual peptide (Gly, 1.04; Lys, 1.96). Since chymotryptic specificity required phenylalanine to be the COOH-terminal residue, the amino acid sequence of peptide XIV is established as shown, the Lys-Lys sequence being placed by difference.

RESIDUES 11–26. Val-Gln-Lys-Cys-Ala-Gln-Cys-His-

HEME

Thr-Val-Glu-Lys-Gly-Gly-Lys-His (Table III). This heme peptide was recovered from a separate chymotryptic digest of turtle cytochrome *c* (see Experimental Section). The NH_2-terminal and COOH-terminal residues were determined to be either valine or glutamic acid (Edman degradation) and histidine (carboxypeptidase A digestion), respectively. Tryptic digestion of this heme peptide yielded four fragments. Assuming tryptic specificity, one step of Edman degradation and the electrophoretic mobility of the peptide estab-

TABLE IV: Amino Acid Sequence of Residues 27–35.

Peptide	Sequence: **Lys-Thr-Gly-Pro-Asn-Leu-**Asn-*Gly-Leu* ←——VIId——→ ←——VIIe, VIc——→
VIId	Lys, 1.11; Thr, 0.98; Gly, 1.09; Pro, 1.04; Asp, 1.92; Leu, 0.87; Ile, 0.20 (10%; PE, PC; *el*, −6.2; *ch*, 7.5; blue)
Edman I	**Lys, 0.10;** Thr, 0.89; Gly, 1.01; Pro, 1.30; Asp, 1.93; Leu, 0.88
II	Lys, <0.10; **Thr, 0.19;** Gly, 1.08; Pro, 1.04; Asp, 1.98; Leu, 0.90
III	Lys, <0.10; Thr, 0.13; **Gly, 0.36;** Pro, 1.07; Asp, 1.97; Leu, 0.97
IV	Lys, <0.10; Thr, 0.12; Gly, 0.22; **Pro, 0.28;** Asp, 1.93; Leu, 1.07
V	Lys, <0.10; Thr, <0.10; Gly, 0.21; Pro, 0.19; **Asp, 1.18;** Leu, 0.83
VI	Lys, <0.10; Thr, 0.17; Gly, 0.34; Pro, <0.10; Asp, 1.00; **Leu, 0.33**
VIIe	Lys, 1.00; Thr, 0.94; Gly, 1.98; Pro, 0.94; Asp, 2.32; Leu, 1.82 (14%; PE, PC; *el*, −5.3; *ch*, 130; blue)
Carboxypeptidase A	Gly, 0.16; Leu, 1.00
VIc	Lys, 0.85; Thr, 0.90; Gly, 2.16; Pro, 1.10; Asp, 2.00; Leu, 2.00; Ser, 0.10 (5%; PC; *el*, 0; *ch*, 7.0; blue)

TABLE V: Amino Acid Sequence of Residues 36–46.

Peptide	Sequence: **Ile-Gly-**Arg - Lys - **Thr-Gly-**Gln-*Ala*-**Glu-**Gly - *Phe* ←—T-1—→ ⟨T-2⟩ ←————T-3————→ ←——E-1——→←E-2→⟨E-3⟩
XIIIa	Ile, 1.03; Gly, 3.05; Arg, 0.86; Lys, 0.93; Thr, 1.09; Glu, 1.87; Ala, 0.99; Phe, 1.18 (35%; PC; *el*, −2.5; *ch*, 7.5; blue, Sakaguchi)
Edman I	**Ile, <0.10**; Gly, 2.95; Arg, 0.90; Lys, 0.97; Thr, 0.95; Glu, 2.03; Ala, 0.98; Phe, 1.22
II	Ile, <0.10; **Gly, 2.16;** Arg, 0.80; Lys, 0.23; Thr, 0.91; Glu, 2.09; Ala, 1.06; Phe, 0.99
Carboxypeptidase A	Phe
T-1	Ile, 0.99; Gly, 1.03; Arg, 0.98 (PE; *el*, −5.5; *ch*, 12.0; blue, Sakaguchi)
T-2	Lys (PE; *el*, −9.5; *ch*, 4.5; blue)
T-3	Thr, 0.94; Gly, 1.93; Glu, 2.21; Ala, 0.94; Phe, 0.99 (PE; *el*, +1.5; *ch*, 13.0; yellow to blue)
Edman I	**Thr, 0.13;** Gly, 1.98; Glu, 1.99; Ala, 1.10; Phe, 0.92
II	Thr, 0.09; **Gly, 1.30;** Glu, 1.96; Ala, 1.0; Phe, 0.74
E-1	Thr, 1.00; Gly, 0.80; Glu, 1.10; Ala, 1.10 (PE, PC; *el*, 0; *ch*, 6.2; blue)
Hydrazinolysis	Ala
E-2	Glu, 0.95; Gly, 1.05 (PE; *el*, +3.5; *ch*, 6.3; blue)
Edman I	**Glu, 0.15;** Gly, 1.00
E-3	Phe

lished the structure of fragment T-1 as Val·Gln·Lys, the NH_2-terminal peptide. T-2 was oxidized with performic acid, following removal of the heme by reaction with mercuric acetate (Ambler, 1963). The NH_2-terminal sequence of the oxidized product of T-2 was established as $CySO_3H$-Ala by two steps of Edman degradation. Digestion of the oxidized product of T-2 with chymotrypsin yielded two fragments. Fragment C-1 released free histidine and cysteic acid upon digestion with carboxypeptidase A. Since chymotryptic specificity indicates histidine to be the COOH-terminal residue of this fragment, and complete leucine aminopeptidase digestion of C-1 revealed no glutamic acid, the structure of C-1 was established as $CySO_3H$-Ala-Gln-$CySO_3H$-His, the glutaminyl residue being placed by difference. It may be noted that repeated Edman degradation steps after removal of the alanyl residue failed to change the analysis of the residual peptide. This indicates that the glutaminyl residue had cyclized to pyrrolidonecarboxylic acid, providing additional confirmation of the position to this residue. Because of tryptic specificity, the structure of the neutral chymotryptic fragment C-2 was established as Thr-Val-Glu-Lys by two steps of Edman degradation. The NH_2-terminal sequence of oxidized T-2 indicates that C-1 precedes C-2 in the over-all structure of

TABLE VI: Amino Acid Sequence of Residues 47–54.

Peptide	Sequence: **Ser**-Tyr-**Thr**-**Glu**-**Ala**-Asn-Lys - Asn ←——————VIIb——————→ ←IIIa→←——— Va,VIa ———→ ←——— T-2 ———→⟨T-1⟩
IIIa	Ser, 1.09; Tyr, 0.91 (48%; PE; *el*, 0; *ch*, 17.5; yellow to blue; Pauly)
Edman I	**Ser, 0.26;** Tyr, 1.00
VIa	Thr, 0.98; Glu, 1.05; Ala, 1.06; Asp, 1.94; Lys, 0.97 (28%; PC; *el*, 0; *ch*, 3.0; yellow to blue)
Edman I	**Thr, <0.10;** Glu, 0.99; Ala, 0.94; Asp, 2.08; Lys, 1.00
II	Thr, <0.10; **Glu, <0.10;** Ala, 0.97; Asp, 2.04; Lys, not determined
III	Thr, <0.10; Glu, <0.10; **Ala, <0.10;** only Asp remaining; Lys, not determined
T-1	Asn (PC; *el*, 0; *ch*, 2.0; orange)
T-2	Thr, 1.14; Glu, 0.85; Ala, 0.84; Asp, 1.14; Lys, 1.03 (PC; *el*, 0; *ch*, 7.0; yellow to blue)
Va	Thr, 0.92; Glu, 1.04; Ala, 1.00; Asp, 2.00; Lys, 1.04 (5%; PC; *el*, +2.5; *ch*, 2.5; yellow to blue)
VIIb	Ser, 0.90; Tyr, 0.98; Thr, 1.13; Glu, 1.13; Ala, 1.13; Asp, 1.78; Lys, 0.94; Leu, 0.19; Ile, 0.19 (9%; PE; *el*, −2.5; *ch*, 3.5; yellow to blue; Pauly)
Edman I	**Ser, 0.10;** Tyr, 1.10; Thr, 0.90; Glu, 1.14; Ala, 0.93; Asp, 1.93; Lys, not determined

T-2. From tryptic specificity fragment T-3 must have the structure Gly-Gly-Lys, while T-4 was characterized as free histidine. T-4, being the only tryptic fragment not affording a point of tryptic hydrolysis, represents the COOH-terminal residue of the chymotryptic heme peptide, and its NH_2-terminal valyl residue marks T-1 as the NH_2-terminal fragment. The assignment of the relative positions of T-2 and T-3 was made possible by the isolation of a minor yield peptide from the original chymotryptic digest, peptide XIc. Its structure was shown to be Gly-Gly-Lys-His by one step of Edman degradation and by carboxypeptidase A digestion. This sequence indicates that T-3 precedes the COOH-terminal fragment T-4, T-2 being therefore placed between T-1 and T-3.

RESIDUES 27–35. Lys-Thr-Gly-Pro-Asn-Leu-Asn-Gly-Leu (Table IV, Peptides VIc, VIId, VIIe). The amino acid sequence of peptide VIId was established from its electrophoretic mobility and by six steps of Edman degradation.

Except for the one extra glycyl and one leucyl residue, the composition of peptide VIIe is identical with that of peptide VIId. Since leucine and a small amount of glycine were released by carboxypeptidase A digestion of peptide VIIe, the additional sequence, Gly-Leu, must be at the COOH-terminal end. From their respective compositions and electrophoretic mobilities, it is concluded that peptide VIc is a deamidated product of peptide VIIe.

RESIDUES 36–46. Ile-Gly-Arg-Lys-Thr-Gly-Gln-Ala-Glu-Gly-Phe (Table V, Peptide XIIIa). Two steps of Edman degradation on peptide XIIIa established its NH_2-terminal sequence as Ile-Gly. Tryptic digestion of peptide XIIIa yielded three fragments. From its composition T-1 must be the NH_2-terminal fragment and have therefore the structure Ile-Gly-Arg. Tryptic fragment T-2 was shown to be free lysine. Two steps of Edman degradation of the tryptic fragment T-3 established its NH_2-terminal sequence as Thr-Gly. Elastase digestion of T-3 yielded three fragments. Alanine was the COOH-terminal residue of E-1 as shown by hydrazinolysis. From its amino acid composition, electrophoretic mobility, and the above data, E-1, which represents the NH_2-terminal segment of T-3, was found to be Thr-Gly-Gln-Ala, the glutaminyl residue being placed by difference. Fragment E-2 was shown to be the dipeptide Glu-Gly by Edman degradation, and fragment E-3 was characterized as free phenylalanine. Since the COOH-terminal residue of peptide XIIIa is the single phenylalanine in the peptide (carboxypeptidase A digestion), the elastase fragments E-1, E-2, and E-3 can be aligned in that order in tryptic fragment T-3, and the three tryptic fragments can be assigned the relative positions indicated by the over-all sequence given above.

RESIDUES 47–54. Ser-Tyr-Thr-Glu-Ala-Asn-Lys-Asn (Table VI, Peptides IIIa, Va, VIa, and VIIb). From chymotryptic specificity peptide IIIa must be Ser-Tyr. This was confirmed by one step of Edman degradation. Three steps of Edman degradation on peptide VIa established the NH_2-terminal sequence as Thr-Glu-Ala. Tryptic digestion of peptide VIa released free asparagine in addition to a neutral peptide, T-2. Thus, tryptic specificity requires the COOH-terminal sequence of peptide VIIa to be Lys-Asn and the complete structure of peptide VIa is therefore Thr-Glu-Ala-Asn-Lys-Asn, the first asparaginyl residue being placed by difference. A complete leucine aminopeptidase digest of the neutral tryptic fragment, T-2, contained free glutamic acid and asparagine, confirming the amide assignment given above.

The compositions and electrophoretic mobilities of peptides Va and VIa show that they are identical, except that the former is a deaminated product of the latter.

Peptide VIIb contains the residues in peptides IIIa

TABLE VII: Amino Acid Sequence of Residues 60–65.

Sequence: **Gly-Glu-Glu**-Thr-*Leu-Met*
←——IIb——→
←——IIa——→

Peptide	
IIa	Gly, 0.97; Gly, 2.02; Thr, 0.63; Leu, 1.01 (10%; PC; *el*, +7.0; *ch*, 18.0; yellow)
Edman I	**Gly, 0.58;** Glu, 1.91; Thr, 1.08; Leu, 1.02
II	Gly, 0.40; **Glu, 1.19;** Thr, 0.93; Leu, 0.88
III	Gly, 0.35; **Glu, 0.50;** Thr, 0.86; Leu, 1.14
IIb	Gly, 0.96; Glu, 2.07; Thr, 0.81; Leu, 1.17; Met, 1.00 (13%; PC; *el*, +6.5; *ch*, 20.0; Toennies)
Carboxypeptidase A	Leu, 0.89; Met, 1.00

TABLE VIII: Amino Acid Sequence of Residues 68–74.

Sequence: **Leu-Glu-Asn**-Pro-Lys-**Lys**-*Tyr*
←——XIIa——→
←——XIa——→
←—T-1—→←T-2→

Peptide	
XIa	Gly, 1.00; Asp, 1.12; Pro, 1.22; Lys, 1.88; Tyr, 0.78 (9%; PE; *el*, −4.2; *ch*, 4.2; blue; Pauly)
Leucine aminopeptidase digestion	Glu
T-1	Glu, 1.01; Asp, 0.98; Pro, 1.01; Lys, 1.00 (PC; *el*, 0; *ch*, 4.5; blue)
Edman I	**Glu, 0.44;** Asp, 1.07; Pro, 0.93; Lys, not determined.
II	Glu, 0.34; **Asp, 0.51;** Pro, 1.90; Lys, not determined.
T-2	Lys, 1.01; Tyr, 0.99; (PC; *el*, −5.5; *ch*, 11.5; blue, Pauly)
Edman I	**Lys, 0.10;** Tyr, 1.00
XIIa	Leu, 0.97; Glu, 1.14; Asp, 1.04; Pro, 0.95; Lys, 2.16; Tyr, 0.74; His, 0.10; Thr, 0.15 (56%; *el*, −4.5; *ch*, 6.5; blue, Pauly)
Carboxypeptidase A	Tyr
Edman I	**Leu, 0.21;** Glu, 1.05; Asp, 1.11; Pro, 0.89; Lys, 1.50; Tyr, 0.95

and VIa as evidenced by the amino acid compositions. Since one step of Edman degradation on peptide VIIb showed serine to be NH_2 terminal, the complete amino acid sequence is established as listed above.

RESIDUES 55–59. Lys-Gly-Ile-Thr-Trp (Peptide XIIIb). Peptide XIIIb (complete leucine aminopeptidase digestion: Lys, 1.11; Gly, 1.01; Ile, 0.92; Thr, 0.92; Trp, 1.04; *el*, −4.5; *ch*, 17.2) was purified by paper electrophoresis and recovered in a yield of 33%. Since chymotryptic specificity requires tryptophan to be the COOH-terminal amino acid, three steps of Edman degradation established the complete structure, the threonine residue being placed by difference; Edman I: **Lys, 0.10;** Gly, 1.14; Ile, 1.02; Thr, 0.84; Trp, not determined; Edman II: **Gly, 0.33;** Ile, 1.05; Thr, 0.96; Lys and Trp, not determined; Edman III: Gly, 0.30; **Ile, 0.36;** Thr, 1.00; Lys and Trp, not determined.

RESIDUES 60–65. Gly-Glu-Glu-Thr-Leu-Met (Table VII, Peptides IIa and IIb). Complete leucine aminopeptidase digests of peptides IIa and IIb contained no amidated residues indicating that both glutamyl residues are present as the free acids. Since the only possible points of chymotryptic hydrolysis are at the leucyl and methionyl residues of peptides IIa and IIb, respectively, the amino acid compositions show that the COOH-terminal sequence of peptide IIb is Leu-Met. This was confirmed by carboxypeptidase A digestion of peptide IIb. Three steps of Edman degradation on peptide IIa were sufficient to complete the sequence, the threonyl residue being placed by difference.

RESIDUES 66–67. Glu-Tyr (Peptide IIIb). Peptide IIIb (Glu, 1.03; Tyr, 0.97; *el*, +4.0; *ch*, 19.0) was purified by paper electrophoresis and recovered in 30% yield. The structure was characterized as Glu-Tyr by Edman degradation; Edman I: **Glu, 0.10;** Tyr, 1.00

RESIDUES 68–74. Leu-Glu-Asn-Pro-Lys-Lys-Tyr (Table VIII, Peptides XIa and XIIa). Tryptic digestion

TABLE IX: Amino Acid Sequence of Residues 83–94.

Peptide	Sequence: **Ala-Gly-Ile-**Lys-Lys-**Lys-Ala-Glu-**Arg-**Ala-Asp-***Leu*** ←——————XIV——————→ ←——IX——→ ←——————XIb——————→ ←——XIIb——→⟨T-1⟩←—T-2—→←—T-3—→
IX	Ala, 1.12; Gly, 0.97; Ile, 0.91; Lys, 1.01 (16%; *el*, −4.0; *ch*, 13.5; blue)
XIIb	Ala, 0.84; Gly, 0.96; Ile, 0.90; Lys, 2.31 (44%; PE; *el*, −7.5; *ch*, 4.7; blue)
Edman I	**Ala, <0.10;** Gly, 0.99; Ile, 0.88; Lys, 2.13
II	Ala, <0.10; **Gly, <0.10**; Ile, 1.08; Lys, 1.92
III	Only Lys remaining
XIb	Lys, 1.05; Ala, 1.96; Glu, 1.15; Arg, 0.95; Asp, 0.94; Leu, 0.96; Gly, 0.12 (50%; PE; *el*, 0; *ch*, 7.5; blue, Sakaguchi)
Carboxypeptidase A	Ala, 0.55; Asp, 0.61; Leu, 1.00
Edman I	**Lys, 0.34;** Ala, 1.85; Glu, 0.94; Arg, 1.08; Asp, 1.20; Leu, 0.94
II	Lys, 0.07; **Ala, 1.08;** Glu, 1.09; Arg, 0.94; Asp, 0.87; Leu, 1.02
III	Lys, 0.10; Ala, 0.90; **Glu, 0.16;** Arg, 1.00; Asp, 1.12; Leu, 0.98
T-1	Lys (PE; *el*, −8.0; *ch*, 5.0; blue)
T-2	Ala, 1.18; Glu, 0.88; Arg, 0.95; (PE; *el*, 0; *ch*, 6.0; blue, Sakaguchi)
Edman I	**Ala, 0.23;** Glu, 1.00; Arg, 1.00
II	Ala, <0.10; **Glu, <0.10;** Arg, 1.00
T-3	Ala, 1.02; Asp, 1.04; Leu, 0.95 (PE; *el*, +5.0; *ch*, 24.5; blue)
Edman I	**Ala, 0.12;** Asp, 1.03; Leu, 0.97
II	Ala, <0.10; **Asp, <0.10;** Leu, 1.00
XIV	Ala, 2.10; Gly, 1.07; Ile, 1.00; Lys, 2.60; Glu, 0.95; Arg, 1.16; Asp, 1.18; Leu, 0.95 (7%; *el*, −4.2; *ch*, 3.5; blue, Sakaguchi)
Carboxypeptidase A	Ala, 0.94; Asp, 0.94; Leu, 1.00

of peptide XIa yielded two fragments. Assuming the COOH-terminal residue of T-1 to be lysine because of tryptic specificity, two steps of Edman degradation were sufficient to establish the structure, the prolyl residue being placed by difference. T-2 was shown to be Lys-Tyr by Edman degradation. T-1 is electrophoretically neutral and from its composition one of the two acidic residues in it must be present as the amide. Since leucine aminopeptidase digestion of peptide XIa released free glutamic acid and no glutamine, the structure of peptide XIa is Glu-Asn-Pro-Lys-Lys-Tyr.

The composition of peptide XIIa is identical with that of peptide XIa except for one extra leucyl residue. This residue was found to be NH_2-terminal by Edman degradation, establishing the sequence of residues 68–74 as given above.

RESIDUES 75–80. Ile-Pro-Gly-Thr-Lys Met (Peptide VIIIa). Peptide VIIIa (Ile, 1.02; Pro, 1.04; Gly, 1.16; Thr, 1.08; Lys, 0.84; Met, 0.86; *el*, −4.0; *ch*, 14.5) was purified by paper chromatography and recovered in 27% yield. Both carboxypeptidase A and tryptic digestion of peptide VIIIa yielded methionine and a basic residual peptide, indicating the COOH-terminal sequence to be Lys-Met. Three steps of Edman degradation established the complete structure, the threonine residue being placed by difference; Edman I: **Ile, 0.10;** Pro, 1.21; Gly, 0.80; Thr, 1.00; Lys, 0.45; Met, 1.00; Edman II: Ile, 0.10; **Pro, 0.19;** Gly, 1.11; Thr, 1.06; Met, 0.82; Lys, not determined; Edman III: Ile, 0.10; Pro, 0.10; **Gly, 0.10;** Thr, 1.20; Met, 0.80; Lys, not determined.

RESIDUES 81–82. Ile-Phe (Peptide Vb). Peptide Vb (Ile, 0.98; Phe, 1.02; *el*, 0; *ch*, 34.0) was purified by paper chromatography and recovered in a yield of 36%. Edman degradation was used to establish the structure, as follows: Edman I: **Ile, 0.18;** Phe, 1.00.

RESIDUES 83–94. Ala-Gly-Ile-Lys-Lys-Lys-Ala-Glu-Arg-Ala-Asp-Leu (Table IX, Peptides X, IX, XIb, XIIb, and XIV). The compositions of peptides IX and XIIb and three steps of Edman degradation on peptide XIIb determine the sequence of residues 83–87. The NH_2-terminal sequence of peptide XIb was shown to be Lys-Ala-Glu by three steps of Edman degradation. Tryptic digestion of peptide XIb yielded three fragments. Fragment T-1 was free lysine. The structures of T-2 and T-3 were established by two steps of Edman degradation in each case. Since carboxypeptidase A digestion of peptide XIb releases leucine, aspartic acid and alanine, T-3 must be the COOH-terminal fragment and the complete structure of peptide XIb is shown to be Lys-Ala-Glu-Arg-Ala-Asp-Leu. T-2 is neutral and T-3 is acidic indicating both acidic residues in the over-all sequence are not amidated.

Peptide XIV is the peptide covering the entire region of residues 83–94 as evidenced by its unique amino

TABLE X: Amino Acid Sequence of Residues 98–104.

Peptide	Sequence: **Leu-Lys-Asp-Ala-Thr**-Ser-*Lys* ←T-1→←——T-2——→
X	Leu, 0.84; Lys, 2.00; Asp, 1.35; Ala, 0.88; Thr, 0.99; Ser, 0.94 (41%; *el*, −4.0; *ch*, 5.0; blue)
Edman I	**Leu, 0.05;** Asp, 1.08; Ala, 0.92; Thr, 1.01; Ser, 0.99; Lys not determined
II	Leu, <0.10; **Lys, 0.99;** Asp, 1.02; Ala, 1.07; Thr, 1.03; Ser, 0.88
III	Leu, <0.10; **Asp, 0.29;** Ala, 1.03; Thr, 0.99; Ser, 0.99; Lys not determined
T-1	Leu, 0.97; Lys, 1.03 (PC; *el*, −5.5; *ch*, 16.0; blue)
Edman I	**Leu, 0.09;** Lys, 1.00
T-2	Asp, 1.03; Ala, 1.07; Thr, 0.95; Ser, 0.95; Lys, 1.00 (PC; *el*, 0; *ch*, 4.0; blue)
Hydrazinolysis	Lys
Edman I	**Asp, 0.15;** Ala, 1.13; Thr, 1.00; Ser, 0.88; Lys, 0.50
II	Asp, 0.07; **Ala, 0.10;** Thr, 0.98; Ser, 1.02; Lys, 0.42
III	Asp, 0.03; Ala, 0.06; **Thr, 0.06;** Ser, 1.00; Lys, 0.41

acid composition. This composition shows the presence of three lysyl residues as indicated above. Carboxypeptidase A digestion of peptide XIV releases alanine, aspartic acid, and leucine, showing that peptide XIIb precedes peptide XIb and establishing the over-all sequence of residues 83–94.

RESIDUES 95–97. Ile-Ala-Tyr (Peptide IV). Peptide IV (Ile, 1.02; Ala, 1.09; Tyr, 0.89; *el*, 0; *ch*, 26.5) was recovered in 90% yield. Two steps of Edman degradation determined its structure, as follows: Edman I: **Ile, 0.10;** Ala, 0.81; Tyr, 1.20; Edman II: Ile, 0.10; **Ala, 0.10;** Tyr, 1.00.

RESIDUES 98–104. Leu-Lys-Asp-Ala-Thr-Ser-Lys (Table X, Peptide X). Three steps of Edman degradation showed the NH_2-terminal sequence to be Leu-Lys-Asp. Tryptic digestion of peptide X yielded two fragments. The structure of T-1 was found to be Leu-Lys by one step of Edman degradation. This fragment therefore represents the NH_2-terminal portion of peptide X. Hydrazinolysis of T-2 indicated lysine to be the COOH-terminal residue. Three steps of Edman degradation established the structure of T-2, the seryl residue being placed by difference. T-2 is neutral showing the aspartyl residue in the peptide to be present as the free acid. The amino acid sequence of peptide X is thus completed as given above.

Discussion

Amino Acid Sequence of C. serpentina Cytochrome c. The primary structure of snapping turtle cytochrome *c*, given in Figure 2, is derived from the complete amino acid sequences of the peptides isolated from a chymotryptic digest. It may be noted that these peptides account entirely for the amino acid composition of the protein, determined directly on acid hydrolysates (see Table I). The relative positions of the chymotryptic peptides were assigned by homology to the structures of the other 15 "mammalian-type" cytochromes *c* (Margoliash, 1962b) of known amino acid sequence, listed in the introduction. Such homologies are so extensive that there is no ambiguity in placing peptides in their correct corresponding positions in the polypeptide chain (see Margoliash and Schejter, 1966). Indeed, if one considers only cytochromes *c* from vertebrate species, identical residues in identical positions account for over 75% of the amino acid sequences, even in comparisons between the most divergent proteins. For such cytochromes *c* the independent determination of the order of peptides within complete sets of chymotryptic or of tryptic peptides would appear to be unnecessary. This may not be the case for proteins from taxonomic groups, the cytochromes *c* of which have not been extensively investigated, and which may show variations from other known structures, amounting to over 50% of the protein chain.

An unexpected tryptic-like hydrolysis at lysyl residue 5 occurred in the chymotryptic digest of snapping turtle cytochrome *c*. This led to the accumulation of peptides I and XIV in the remarkably high minimal yields of 80 and 63%, respectively. It is difficult to account for this phenomenon by assuming contamination of the chymotrypsin preparation used by trypsin, since the only other tryptic-like hydrolyses that could be detected took place at the Lys-Lys-Lys sequence, residues 86–88 (Table IX). These latter points of hydrolysis have indeed been observed in chymotryptic digests of most cytochromes *c* studied, even in the complete absence of other tryptic-like splits (see, for example, Margoliash, 1962a). Whether a proteolytic contaminant of the cytochrome *c* preparation itself was responsible for the odd hydrolysis at residue 5 cannot be assessed without further study. In this connection it may be noted that proteolytic activity does tend to be carried in cytochrome *c* preparations from some species (E. Margoliash and O. Walasek, unpublished observations).

Comparison with Other Cytochromes c and Evolutionary Relations. The amino acid sequence of the snapping turtle heart protein clearly classes it with the "mam-

malian-type" cytochromes *c*. The NH_2-terminal acetylglycine residue, the prosthetic heme group bound to cysteinyl residues in positions 14 and 17, and the single polypeptide chain 104 residues long are all characteristics common to cytochromes *c* from vertebrate species. Among the notable features the snapping turtle protein shares with all cytochromes *c* are the typical clustering of basic and hydrophobic residues, clusters which occupy positions similar to those of the corresponding groups in other cytochromes *c*; the occurrence of certain key residues in particular positions, such as the histidine in position 18, the single tryptophan residue in position 59, and the two arginyl residues in positions 38 and 91; and also the remarkable constant sequence between residues 70 and 80, Asn-Pro-Lys-Lys-Tyr-Ile-Pro-Gly-Thr-Lys-Met. This constant segment assumes a particular importance in view of the recent implication of the methionyl residue at position 80 in hemochrome formation (Harbury *et al.*, 1965).

Snapping turtle cytochrome *c* shows extensive identities of amino acid sequence with other cytochromes *c*. There are 11, 15, 14, 9, 9, 9, 9, 11, 8, 20, 18, 27, 49, 53, and 49 residue variations between the turtle protein and the cytochromes *c* from horse, man, monkey, pig, cow, dog, rabbit, kangaroo, chicken, rattlesnake, tuna, moth, baker's yeast, *Candida krusei*, and *N. crassa*. These comparisons exhibit the previously described rough proportionality between the number of amino acid substitutions and the time elapsed since the divergence of the lines of evolution leading to the species carrying these various cytochromes *c* (Margoliash, 1963; Margoliash and Smith, 1965; Margoliash and Schejter, 1966). They also support previous conclusions regarding the functional and evolutionary implications of variable and constant features of cytochrome *c* structure. These relationships have been discussed elsewhere and will not be further elaborated here (see Margoliash and Schejter, 1966).

An interesting amino acid substitution occurs in the snapping turtle protein at residue 33. In contrast to the majority of cytochromes *c* of known structure in which this position is occupied by a histidyl residue, the turtle protein carries an asparaginyl residue. This is similar to the situation in the kangaroo (Nolan and Margoliash, 1966) and the tuna (Kreil, 1963, 1965) cytochromes *c* in which asparagine and tryptophan, respectively, occur at position 33. These substitutions are important, since, taken together with the reported absence of histidyl residues at position 26 in *N. crassa* cytochrome *c* (Heller and Smith, 1965) and in iso-2-cytochrome *c* of baker's yeast (Stewart *et al.*, 1966), it appears that the only invariant histidyl residue in all cytochromes *c* is the one at position 18. Thus, if the tertiary structures of all the "mammalian-type" cytochromes *c* are identical in all essential details, it follows that no more than one histidine can be concerned with hemochrome formation. Such a conclusion is in accord with recent findings indicating that all ϵ-amino groups of lysyl residues can be alkylated without affecting the spectrum of cytochrome *c* (Fanger and Harbury, 1965) and that, in contrast to previously held views (see Lemberg and Legge, 1949; Keilin, 1960), a nonnitrogenous side chain (see Heller and Smith, 1965), such as that of a methionyl residue, can coordinate to heme, yielding spectra typical of the low-spin state characteristic of ferrocytochrome *c* (Harbury *et al.*, 1965).

The substitution of the prolyl residue at position 44 by other amino acids has been discussed previously (Chan and Margoliash, 1966a; Needleman and Margoliash, 1966) and will not be considered further.

Acknowledgments

The authors wish to thank Mr. O. F. Walasek for preparing the turtle heart cytochrome *c* used in this study and Mr. Robert Gardisky and Mr. Terrance Tanner for the amino acid analyses. They are also grateful to Miss LaVera Lazer for the free-flow electrophoreses.

References

Ambler, R. P. (1963), *Biochem. J. 89*, 341.
Ando, K., Matsubara, H., and Okunuki, K. (1965), *Proc. Japan Acad. 41*, 79.
Bahl, O. P., and Smith, E. L. (1965), *J. Biol. Chem. 240*, 3585.
Chan, S. K., and Margoliash, E. (1966a), *J. Biol. Chem. 241*, 507.
Chan, S. K., and Margoliash, E. (1966b), *J. Biol. Chem. 241*, 335.
Fanger, M. W., and Harbury, H. A. (1965), *Biochemistry 4*, 2541.
Funatsu, G., Tsugita, A., and Fraenkel-Conrat, H. (1964), *Arch. Biochem. Biophys. 105*, 25.
Harbury, H. A., Cronin, J. R., Fanger, M. W., Hettinger, T. P., Murphy, A. J., Myer, Y. P., and Vinogradov, S. N. (1965), *Proc. Natl. Acad. Sci. U. S. 54*, 1658.
Heller, J., and Smith, E. L. (1965), *Proc. Natl. Acad. Sci. U. S. 54*, 1621.
Hettinger, T. P., and Harbury, H. A. (1964), *Proc. Natl. Acad. Sci. U. S. 52*, 1469.
Hettinger, T. P., and Harbury, H. A. (1965), *Biochemistry 4*, 2546.
Hirs, C. H. W., Moore, S., and Stein, W. H. (1960), *J. Biol. Chem. 235*, 633.
Horinishi, H., Kurihara, K., and Shibata, K. (1965), *Arch. Biochem. Biophys. 111*, 520.
Ingram, V. M. (1958), *Biochim. Biophys. Acta 28*, 539.
Keilin, J. (1960), *Nature 187*, 365.
Konigsberg, W., and Hill, R. J. (1962), *J. Biol. Chem. 237*, 2547.
Kreil, G. (1963), *Z. Physiol. Chem. 334*, 154.
Kreil, G. (1965), *Z. Physiol. Chem. 340*, 86.
Kreil, G., and Tuppy, H. (1961), *Nature 192*, 1123.
Lemberg, R., and Legge, J. W. (1949), Hematin Compounds and Bile Pigments, New York, N. Y., Interscience.
Margoliash, E. (1962a), *J. Biol. Chem. 237*, 2161.

Margoliash, E. (1962b), *Brookhaven Symp. Biol. 19*, 266.
Margoliash, E. (1963), *Proc. Natl. Acad. Sci. U.S. 50*, 672.
Margoliash, E. (1964), *Can. J. Biochem. 42*, 745.
Margoliash, E., Kimmel, J. R., Hill, R. L., and Schmidt, W. R. (1962), *J. Biol. Chem. 237*, 2148.
Margoliash, E., Needleman, S. B., and Stewart, J. W. (1963), *Acta Chem. Scand. 17*, S250.
Margoliash, E., and Schejter, A. (1966), *Advan. Protein Chem. 21*, 113.
Margoliash, E., and Smith, E. L .(1961), *Nature 192*, 112.
Margoliash, E., and Smith, E. L. (1962), *J. Biol. Chem. 237*, 2151.
Margoliash, E., and Smith, E. L. (1965), *in* Evolving Genes and Proteins, Bryson, V., and Vogel, H. J., Ed., New York, N. Y., Academic, p 221.
Margoliash, E., Smith, E. L., Kreil, G., and Tuppy, H. (1961), *Nature 192*, 1125.
Margoliash, E., and Walasek, O. (1966), *Methods Enzymol. 10* (in press).
Matsubara, H., and Smith, E. L. (1963), *J. Biol. Chem. 238*, 2732.
McDowall, M. A., and Smith, E. L. (1965), *J. Biol. Chem. 240*, 4635.
Narita, K., and Titani, K. (1966), *Proc. Japan Acad. 41*, 831.
Narita, K., Titani, K., Yaoi, Y., and Murakami, H. (1963), *Biochim. Biophys. Acta 77*, 688.
Needleman, S. B., and Margoliash, E. (1966), *J. Biol. Chem. 241*, 853.
Nolan, E., and Margoliash, E. (1966), *J. Biol. Chem. 241*, 1049.
Pan, S. C., and Dutcher, J. D. (1956), *Anal. Chem. 28*, 836.
Rothfus, J. A., and Smith, E. L. (1965), *J. Biol. Chem. 240*, 4277.
Smith, E. L., and Margoliash, E. (1964), *Federation Proc. 23*, 1243.
Stewart, J. W., and Margoliash, E. (1965), *Can. J. Biochem. 43*, 1187.
Stewart, J. W., Margoliash, E., and Sherman, F. (1966), *Federation Proc. 25*, 647.
Tsai, H. J., and Williams, G. R. (1965a), *Can. J. Biochem. 43*, 1409.
Tsai, H. J., and Williams, G. R. (1965b), *Can. J. Biochem. 43*, 1995.
Tuppy, H., and Kreil, G. (1962), *Monatsh. Chem. 92*, 780.
Yasunobu, K. T., Nakashima, T., Higa, H., Matsubara, H., and Benson, H. (1963), *Biochim. Biophys. Acta 78*, 791.

N-Terminal Sequence of Actin*

R. E. Alving† and K. Laki†

ABSTRACT: The N-terminal portion of actin was isolated from a pronase digest by passage through Dowex 50 resin and subsequent fractionation on Dowex 1. This peptide was shown to be acetylated by direct determination of acetic acid by gas chromatography. The smallest residue obtained was acetyl-Asp-Glu, and larger peptides containing threonine and alanine as well were also found. Acetic acid (1 mole) was found in these peptides for each 60,000 g of actin digested, and this accounted for most of the acetic acid found in actin after hydrolysis.

It has been known for about a decade that actin does not contain an N-terminal amino acid that would react with dinitrofluorobenzene (Locker, 1954; Krans *et al.*, 1965; Carsten, 1966). On the other hand, the presence of acetic acid has been reported in the hydrolysate of actin (Krans *et al.*, 1965). These observations suggest that the N-terminal residue of actin is acetylated. We have adapted the technique of gas chromatography for the determination of acetyl groups in the muscle proteins actin, myosin and tropomyosin. Our findings show that actin and tropomyosin contain one acetyl group/mole and confirm the presence of acetyl groups in myosin (Offer, 1965). The details of these investigations will be presented elsewhere. Here we present our experiments which show that the acetyl residue in actin is bound to the N-terminal aspartic acid.

* From the Laboratory of Biophysical Chemistry, National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Maryland. *Received March 1, 1966.*

† Permanent address: National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Md.

Experimental Section

New Zealand white rabbits (closed colony since 1930) were used as a source of muscle. Ground muscle