Amino Acid Sequences of Peptides from the Tryptic Digest of *Golfingia gouldii* Hemerythrin*

W. R. Groskopf,† J. W. Holleman, E. Margoliash, and I. M. Klotz

ABSTRACT: Peptides from two separate tryptic digests of *Golfingia gouldii* hemerythrin were purified by Dowex 50W-X2 column chromatography followed by paper electrophoresis, paper chromatography, gel filtration, or various combinations of these techniques. The amino acid sequences of these peptides were completely determined, except for the relative positions of some of the residues in three peptides. These sequences account for a total of 113 residues in the entire polypeptide chain, which compares with the 112.3 residues calculated from direct analyses of the amino acid composition of the protein. The yields of these peptides indicate that they represent the predominant sequence of the *G. gouldii* hemerythrin subunits. One peptide, recovered in low yield, contained an apparent amino acid substitution which may represent one of the points of variation of the subunits in the protein.

The preceding paper (Groskopf *et al.*, 1966) gives the amino acid composition as well as the amino-terminal and the carboxyl-terminal sequences of *Golfingia gouldii* hemerythrin. The present paper will describe the amino acid sequences of all the peptides recovered from tryptic digests of the protein.

Experimental Section

Preparation of Tryptic Digests. G. gouldii hemerythrin was prepared and crystallized, and the iron-free protein was recovered, as previously described (Groskopf *et al.*, 1966). Two separate tryptic digests, one from the whole protein and the other from the iron-free protein, were utilized.

In the first case, 2.5 g (179 μmoles of subunit) of oxyhemerythrin, dissolved in 435 ml of pH 8.0, 0.02 M ammonium bicarbonate, was denatured by heating at 65° for 5 min. The resulting suspension was cooled, and to it was added 25 mg of trypsin (Worthington Biochemical Corp.) treated with dilute hydrochloric acid according to Northrop and Kunitz (1936) to destroy residual chymotryptic activity. Further additions of 25 mg of trypsin were made at 4 and 8 hr and the hydrolysis was allowed to proceed at 32° for 22 hr. The digest was lyophilized. The peptides recovered from this digest are denoted T.

For the second digestion, trypsin treated with TPCK[1] (Schoellmann and Shaw, 1962) was used to decrease the number of chymotryptic cleavages. Iron-free hemerythrin (165 μmoles of subunit) was suspended in 480 ml of pH 8.0, 0.11 M ammonium bicarbonate and hydrolyzed for 7.5 hr at 37° with TPCK-treated trypsin (3.2 μmoles). The digest was lyophilized. The peptides recovered from this digest are denoted TT.

Initial Chromatographic Separation of Peptides. The digests were fractionated on a column of Dowex 50W-X2 cation-exchange resin (J. T. Baker Chemical Co.). Preparation of the column and compositions of the pyridine–acetic acid buffers used for development were as described by Margoliash and Smith (1962). For application to the column, the lyophilized digests were dissolved in 70 ml of the pH 3.1 starting buffer. A small amount of insoluble material was removed by centrifugation. For the T digest, a linear gradient between pH 3.1 and 5.0 was used, as detailed in Figure 1.

For the TT digest, a compound gradient produced by a five-chambered gradient device (Peterson and Sober, 1959) was applied. Tank 1 of the device contained 8 l. of the 0.2 M, pH 3.1, pyridine–acetic acid starting buffer; tanks 2–4 contained 8 l. of pH 3.3, 4.4, and 3.4 buffers, respectively, made by mixing appropriate amounts of the starting and final buffers; and tank 5 contained 8 l. of the final pH 5.0, 2.0 M pyridine–acetic acid buffer. The details of this chromatography are given in Figure 2.

Purification and Characterization of Peptides. Each column chromatographic fraction was characterized

* From the Biochemistry Division, Department of Chemistry, Northwestern University, Evanston, Illinois (W. R. G., J. W. H., I. M. K.), and the Research Division, Abbott Laboratories, North Chicago, Illinois (E. M.). *Received August 29, 1966.* This investigation was supported in part by a grant (HE-08299) from the National Heart Institute, U. S. Public Health Service. It was also assisted by support made available by a U. S. Public Health Service Training Grant (5T1-GM-626) from the National Institute of General Medical Services.

† Present address: Michael Reese Research Foundation, Chicago, Ill.

[1] Abbreviations: TPCK, L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone; PTC, phenylthiocarbamyl derivative; PTH, phenylthiohydantoin derivative; nd, not determined.

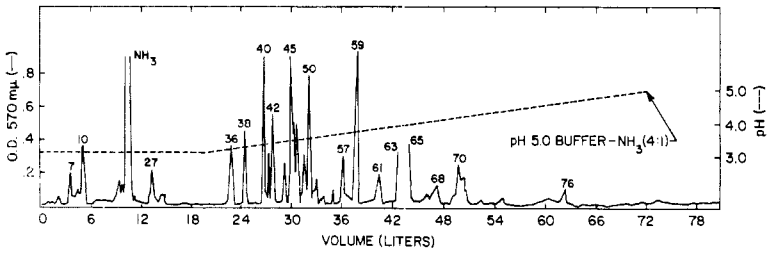


FIGURE 1: Initial column chromatographic fractionation of T digest. Resin, Dowex 50W-X2, equal-volume mixture of 120–200 and 300–325 mesh size grains; column size, 3.8 × 180 cm; flow rate, 200 ml/hr. The effluent was collected in 20-ml aliquots and samples of 0.5 ml were taken from every third one for ninhydrin monitoring (Stein and Moore, 1954). The initial temperature (40°) was raised to 60° after passage of 56 l. of eluting mixture. The linear gradient between pH 3.1 (starting buffer) and 5.0 (final buffer) was started at 20 l. A final eluting mixture of four parts of pH 5.0 buffer and one part of concentrated ammonium hydroxide (v/v) was added at the point indicated. Fractions pooled according to ninhydrin results were reduced in volume by rotary evaporation and lyophilized.

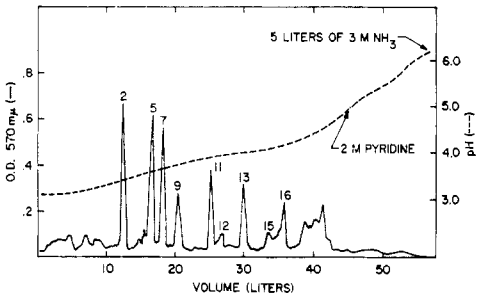


FIGURE 2: Initial column chromatographic fractionation of TT digest. Column dimensions, flow rate, collection and ninhydrin monitoring program, and fraction reduction are all as in Figure 1; resin, Dowex 50W-X2, 200–400 mesh. The initial temperature (40°) was raised to 60° after passage of 36 l. of eluting mixture. The compound gradient changed from pH 3.1 to 5.0 as indicated. Pyridine (2.0 M) and 3.0 M ammonium hydroxide were added at points indicated.

by amino acid analysis (Spackman *et al.*, 1958) following hydrolysis for 40 hr at 108° with 6 N triply glass-distilled HCl in sealed, evacuated tubes. The peptides in each fraction were also monitored by paper electrophoresis followed by chromatography (Ingram, 1958) under the following standard conditions: electrophoresis, 35 v/cm, 90 min, 5°, pH 6.4 pyridine–acetic acid–water buffer (133:4.6:1860, v/v) (Michl, 1951); chromatography, 15 hr, 1-butanol–acetic acid–water (100:30:75, v/v) or pyridine–1-butanol–acetic acid–water (60:40:12:48, v/v). Peptides were revealed with a ninhydrin reagent containing collidine (Margoliash and Smith, 1962). To detect ninhydrin-weak or ninhydrin-negative peptides, a modification of the method of Pan and Dutcher (1956) was used. The paper was sprayed with 0.25% NaOCl and after 10 min dipped in 1% formaldehyde in acetone. After drying, a light spray of aqueous starch–KI (both 0.5%) was applied, and the peptides appeared blue-black against a white background. Tryptophan (Smith, 1953a), arginine (Smith, 1953b), methionine and cysteine (Chargaff *et al.*, 1948), and tyrosine and histidine (Smith, 1960) were revealed by standard techniques.

Following the indications given by the paper electrophoresis–chromatography and the amino acid analyses, the peptide fractions were purified by paper electrophoresis, by paper chromatography, or by combinations of both techniques, under the conditions described above for paper electrophoresis–chromatography, using Whatman 3MM paper. In some cases, a preliminary fractionation was done on columns (1.2 × 120 cm) of Sephadex G-25 (Pharmacia Fine Chemicals, Inc.), using 50% acetic acid as eluent.

Determination of Amino Acid Sequences. The Edman sequential degradation procedure (Edman, 1950, 1953) was performed according to Konigsberg and Hill (1962) with the following minor changes. The formation of the PTC[1] peptide and its cyclization in anhydrous trifluoroacetic acid were carried out under nitrogen. The identity of the amino-terminal residue was determined in most cases by the subtractive method (Hirs *et al.*, 1960). In some instances, the PTH[1] derivatives released were identified by paper chromatography according to Sjöquist (1953), using solvent A.

Where necessary to complete the sequence determinations, peptides were further fragmented with chymotrypsin, pepsin, or dilute acetic acid. In a typical preliminary digestion with chymotrypsin (Worthington Biochemical Corp.), 0.2 μmole of peptide was digested in 100 μl of a 0.1 M, pH 8.0, ammonium bicarbonate buffer containing 0.1% enzyme (w/v) at room temperature. A similar procedure was used for digestion with pepsin (Worthington Biochemical Corp.), except that the reaction was run in 5% formic acid (Canfield, 1963). For partial acid hydrolysis, similar small amounts of peptide were hydrolyzed in 2 ml of 0.25 M acetic acid in sealed, evacuated tubes at 108° (Partridge and Davis, 1950; Schroeder *et al.*, 1963). Optimum reaction times were determined by subjecting samples taken at various times to paper electrophoresis–chromatography, and the digestions were then scaled up for preparative purposes.

Partial and total leucine aminopeptidase digestions (Margoliash and Smith, 1962) were employed to confirm amino-terminal sequences and for the detection of asparagine and glutamine. Two enzyme preparations were used, one prepared by Mr. Otto W. Walasek of Abbott Laboratories, North Chicago, Ill. (preparation 1), and the other a commercial sample (Worthington Biochemical Corp.) (preparation 2). In addition to aminopeptidase activity, both preparations also cleaved peptides at the carboxyl group of proline (see peptide TT-5ch4, related peptides, and Discussion).

For digestions with carboxypeptidase A (Worthington Biochemical Corp.), 0.2–0.4 μmole of peptide in 0.2–0.4 ml of 0.05 M, pH 8.5, sodium borate buffer (Olson and Kuby, 1964) was incubated at room temperature at enzyme:substrate ratios of about 1:20 (w/w) for periods of from 5 min to 6 hr. Aliquots were removed at various times and either chromatographed on paper in the solvents listed above or subjected to amino acid analysis. Digestions with carboxypeptidase B (Worthington Biochemical Corp.) were similarly performed in 0.1 M, pH 8.0, ammonium bicarbonate buffer. It should be noted that at pH 8.5 carboxypeptidase A released lysine quite readily (Olson and Kuby, 1964).

Nomenclature. Peptides derived from the series obtained by digestion of iron-containing hemerythrin with acid-treated trypsin are denoted T, and those from digestion of the iron-free protein with TPCK-treated trypsin, TT. The coding system for individual peptides within a series consists of a numeral indicating the number of the original Dowex 50W-X2 column chromatographic fraction in which the peptide appeared, followed by the notations el, ch, and G, to denote purification by paper electrophoresis, paper chromatography, or gel filtration on Sephadex G-25, respectively. The notation el is followed by A, N, or B, to denote whether the peptide was electrophoretically acidic, neutral, or basic under the conditions employed, and then by a number to distinguish between peptides of the same electrophoretic class in the case of the basic and acidic ones, counting from the origin. Peptides purified by gel filtration are numbered in the order of their elution from the column. Peptides purified by paper chromatography are numbered sequentially from the origin. The letters C, P, CA, CB, L, and AA denote digestion with chymotrypsin, pepsin, carboxypeptidase A, carboxypeptidase B, leucine aminopeptidase, and dilute acetic acid, respectively. The various operations performed to obtain each peptide appear sequentially, following the fraction number.

The tables giving the evidence presented for the proposed amino acid sequences also list the chromatographic (*ch*) and electrophoretic (*el*) mobilities of the peptides on paper in centimeters under the standard conditions given above for paper electrophoresis–chromatography. A minus sign (−) indicates movement toward the anode (acidic peptide) and a plus sign (+) movement toward the cathode (basic peptide). No movement is indicated by a zero (0). The initial colors given by ninhydrin are indicated as follows: B = blue; G = green; Gr = gray; Y = yellow; Or = orange. The presence of tryptophan, arginine, and tyrosine and histidine is indicated by the notations Ehrlich, Sakaguchi, and Pauly, respectively.

Tables II–VIII list the amino acid compositions of the purified peptides in residues per molecule of peptide, followed, within brackets, by the yield, the paper electrophoretic and chromatographic mobilities, the initial ninhydrin color, and the color reactions for specific residues, in that order. The yields of peptides listed were calculated with respect to the initial amount of digest employed, from the compositions of the pure peptides referred to the amino acid analyses of the crude Dowex 50W-X2 column chromatographic fractions in which they appeared. Residues observed in the amino acid analyses but which were in yields below 10% of the major constituents are not reported. When amino acids released by exopeptidases were identified only by paper chromatography, they were rated on a scale of +1 to +4 according to the intensity of their ninhydrin color. In all tables, and in the text, the numbers given in parentheses are the assumed stoichiometric numbers of residues per molecule of pure peptide. In reporting Edman degradations, the residues and numbers marked in boldface type correspond to the residue removed at each step.

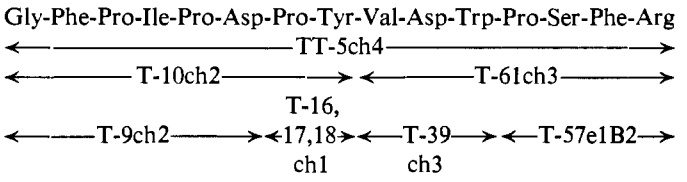
Results

Recovery of Peptides from Dowex 50W-X2 Chromatography. The elution patterns of peptides from the Dowex 50W-X2 column chromatography of the T digest (iron-containing hemerythrin, acid-treated trypsin) and the TT digest (iron-free hemerythrin, TPCK-treated trypsin) are shown in Figures 1 and 2, respectively. Most of the structural work was done on the T digest. Peptides from the TT run were used when they were more easily purified or when they provided overlapping sequences for two or more peptides from the T run. The TT run also made up for the loss of

TABLE I: Over-All Yield of Amino Acids from the TT Run.

Amino Acid	% Yield	Amino Acid	% Yield
Lys	75	Ala	81
His	63	Cys	23
Arg	62	Val	67
Asp	82	Met	87
Thr	72	Ile	55
Ser	97	Leu	69
Glu	81	Tyr	79
Pro	100	Phe	84
Gly	93	Trp	
Numerical Average 74.7%			

TABLE II



TT-5ch4[c]	Gly, 1.07 (1); Phe, 1.92 (2); Pro, 3.96 (4); Ile, 1.19 (1); Asp, 2.02 (2); Tyr, 0.45 (1); Val, 0.84 (1); Trp, (1);[a] Ser, 0.87 (1); Arg, 0.92 (1); Leu, 0.15 [27%; *el*, −2; *ch*, 20; Y to G; Pauly, Ehrlich].
Carboxypeptidase B	Phe, 0.93 (1); Arg, 1.08 (1) (4 hr, room temperature).
Leucine aminopeptidase	Gly, 1.00 (1); Tyr, 1.02 (1); Val, 1.00 (1); Asp, 0.33; Ser, 0.95 (1); Phe, 1.05 (1); Arg, 0.98 (1) (preparation 1, 18 hr, 40°). Gly, 0.80 (1) (preparation 2, 4 hr, 40°).
T-10ch2	Gly, 1.02 (1); Phe, 0.98 (1); Pro, 2.97 (3); Ile, 1.00 (1); Asp, 1.03 (1); Tyr, 0.73 (1) [43%; *el*, −6; *ch*, 28; G; Pauly].
Edman I	**Gly, 0.24**; Phe, 0.88 (1); Pro, 2.95 (3); Ile, 1.09 (1); Asp, 1.07 (1).
II	Gly, 0.08; **Phe, 0.12**; Pro, 2.89 (3); Ile, 1.00 (1); Asp, 1.11 (1).
III	**Pro, 2.03** (2); Ile, 0.92 (1); Asp, 1.04 (1).
IV	Pro, 2.00 (2); **Ile, 0.00**; Asp, 1.15 (1).
V	**Pro, 0.91** (1); Asp, 1.09 (1).
VI	Pro, 1.00 (1); **Asp, 0.42.**
Carboxypeptidase A	Tyr, 1.00 (1) (2 hr, room temperature).
AAch1	Asp, 1.00 (1).
AAch2	Gly, 0.82 (1); Phe, 0.89 (1); Pro, 2.13 (2); Ile, 1.04 (1); Asp, 1.11 (1).
AAch3	Gly, 0.76 (1); Phe, 1.00 (1); Pro, 2.19 (2); Ile, 1.08 (1).
AAch4	Pro, 1.00 (1); Tyr, 0.35 (1).
T-61ch3	Val, 0.94 (1); Asp, 1.10 (1); Trp (1);[a] Pro, 0.96 (1); Ser, 1.09 (1); Phe, 0.92 (1); Arg, 0.97 (1) [12%; *el*, 0; *ch*, 21; B; Ehrlich].
Edman I	**Val, 0.07**; Asp, 1.03 (1); Pro, 1.10 (1); Ser, 0.97 (1); Phe, 1.10 (1); Arg, 0.90 (1); Trp, nd.
Carboxypeptidase B	Arg, 1.08. (1); Phe, 0.92 (1) (1 hr, room temperature).
Leucine aminopeptidase	Val, 1.11 (1); Ser, 0.98 (1); Phe, 0.94 (1); Arg, 1.00 (1) (preparation 1, 1 hr, 40°). Val, 1.00 (1); Ser, 0.25; Phe, 0.33; Arg, nd (preparation 2, 1 hr, 40°).
T-9ch2	Gly, 0.96 (1); Phe, 0.96 (1); Pro, 2.02 (2); Ile, 1.00 (1); Asp, 1.08 (1) [23%; *el*, −6; *ch*, 28; Gr to G].
Edman I	**Gly, 0.30**; Phe, 0.75 (1); Pro, 2.02 (2); Ile, 0.98 (1); Asp, 1.23 (1).
II	Gly, 0.18; **Phe, 0.26**; Pro, 1.92 (2); Ile, 1.09 (1); Asp, 1.50 (1).
T-16,17,18ch1	Pro, 1.12 (1); Tyr, 0.88 (1) [12%; *el*, 0; *ch*, 23; ninhydrin negative; Pauly].
T-39ch3	Val, 0.90 (1); Asp, 1.10 (1); Trp, 1.00 (1)[b] [35%; *el*, −7; *ch*, 26; B; Ehrlich].
Edman I	**Val, 0.05**; Asp, 1.00 (1); Trp, nd.
Carboxypeptidase A	Trp, +1 (30 min, room temperature). Trp, +2 (3 hr, room temperature).
T-57e1B2	Pro, 1.00 (1); Ser, 1.01 (1); Phe, 1.03 (1); Arg, 0.96 (1) [50%; *el*, +10; *ch*, 14; Sakaguchi].
Edman I	**Pro, 0.13**; Ser, 1.02 (1); Phe, 0.91 (1); Arg, 1.07 (1).
II	**Ser, 0.31**; Phe, 0.97 (1); Arg, 1.03 (1).
III	Ser, 0.23; **Phe, 0.40**; Arg, 1.00 (1).

[a] Assumed from positive Ehrlich reaction. [b] Analytical values obtained from a total leucine aminopeptidase digest. [c] See Nomenclature Section in text for abbreviations and their meanings.

fraction T-64, which contained the lone cysteine peptide.

An attempt was made, during the chromatography of the TT run, to assess the loss of peptides occurring at this stage. From the analyses of all the crude chromatographic fractions it was found that the recoveries of individual residues varied from 23 to 100%, with a numerical average of about 75% (see Table I). In view of these results, and considering also losses incurred during the subsequent manipulations, overall yields of peptides above 20% were considered to be adequate for placing a peptide with the others as part of the predominant subunit chain. That this view was justified is seen from the fact that the number

of residues thus placed corresponds well with the number determined by direct amino acid analysis of the protein (Groskopf *et al.*, 1966), with no apparent conflicting or extraneous sequences.

Amino Acid Sequences of Peptides

Peptides TT-5ch4, T-10ch2, T-16,17,18ch1, T-61ch3, T-39ch3, and T-57elB2. Composite sequence, **Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr-Val-Asp-Trp-Pro-Ser-Phe-Arg** (Table II). The unique compositions of these peptides indicate that they are all part of the same segment, a segment which moreover is clearly amino terminal in the subunits, as it includes the unique amino-terminal Gly-Phe-Pro sequence shown to occur in the intact protein (Groskopf *et al.*, 1966).

Peptide TT-5ch4 covers the entire sequence and contains all the four prolyl residues found in each subunit (Groskopf *et al.*, 1966). Its carboxyl-terminal sequence is Phe-Arg, as evidenced from the results of carboxypeptidase B digestion. It is remarkable that the leucine aminopeptidase preparation 1 released all the amino acids in the peptide other than the prolyl residues and the residues immediately preceding them. Similar results were obtained with peptide T-61ch3 and must be a consequence of contamination of the leucine aminopeptidase preparation with an enzyme cleaving at the carboxyl bond of proline. The leucine aminopeptidase preparation 2 released glycine from peptide TT-5ch4 in stoichiometric amounts, plus about 10% of an equivalent each of the other residues released by preparation 1, indicating that glycine was in fact amino terminal. The electrophoretic mobility of peptide TT-5ch4 showed that both aspartyl residues must be present in the nonamidated form.

Peptide T-10ch2, comprising the amino-terminal half of the over-all sequence, was shown to have the structure Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr by Edman degradation and carboxypeptidase A digestion, the third prolyl residue being placed by difference. This sequence was confirmed by the compositions of peptides recovered from a partial acid digest (peptides T-10ch-2AAch1, 2, 3, and 4). Peptide T-10ch2 was acidic, confirming the absence of an amidated residue in this part of the segment. It should be noted that at all steps of the Edman degradation no tyrosine was recovered, even after addition of hydrazine hydrochloride before acid hydrolysis (Sanger and Thompson, 1963). There is no obvious reason for this phenomenon, since normal recoveries of tyrosine were obtained in numerous other cases.

Peptides T-9ch2 and T-16,17,18ch1 comprise the amino-terminal and carboxyl-terminal segments, respectively, of peptide T-10ch2. Partial studies of their structure confirmed the sequence proposed for T-10ch2.

One step of the Edman degradation on peptide T-61ch3 gave valine as amino terminus, and digestion with carboxypeptidase B gave arginine and phenylalanine. Treatment with leucine aminopeptidase (preparation 1) gave results similar to those discussed above for peptide TT-5ch4, confirming, in particular, the presence of the aspartyl residue in its nonamidated form. From their unique compositions, peptides T-39ch3 and T-57elB2 are seen to be the amino-terminal and carboxyl-terminal segments, respectively, of peptide T-61ch3. The structure of T-39ch3 was determined as Val-Asp-Trp by Edman degradation and carboxypeptidase A digestion and the structure of T-57elB2 as Pro-Ser-Phe-Arg by three steps of Edman degradation, thus establishing the structure of this part of peptide TT-5ch4 as given.

Peptides TT-9elN, T-36ch5, and T-45ch1elA1. Composite Sequence, **Thr-Phe-Tyr-Ser-Ile-Ile-Asp-(Asp, Glu)-His-Lys** (Table III). Four steps of Edman degradation on peptide TT-9elN established the amino-terminal sequence as Thr-Phe-Tyr-Ser. The electrophoretic and chromatographic mobilities were nil, because of insolubility. Peptide T-36ch5 was shown to comprise the amino-terminal tripeptide sequence by Edman degradation and carboxypeptidase A digestion. The partial sequence of peptide T-45ch1elA1 was determined by Edman degradation and by carboxypeptidase B digestion. It may be noted that leucine aminopeptidase digestion established the presence of two isoleucines in this peptide, while the acid hydrolysates of residual peptides from Edman degradation consistently gave low values for isoleucine as expected from the presence of an Ile-Ile sequence (Whitfield, 1963). Digestion with carboxypeptidase B for 1.5 hr released only lysine, while treatment for 8 hr released also a small amount of histidine. Because of the small amount of purified peptide recovered from the parent fraction, the relative positions of the two residues shown in parentheses could not be established with certainty. Tentative evidence not listed in the table, obtained from a partial acid hydrolysis, appeared to indicate that the sequence of these residues is as noted.

Peptide T-27ch4. Sequence, **Thr-Leu-Phe.** This peptide (Thr, 0.99 (1); Leu, 1.03 (1); Phe, 0.98 (1) [22%; *el*, 0; *ch*, 36; G]) was subjected to two steps of Edman degradation, as follows. Edman I: **Thr, 0;** Leu, 1.00 (1); Phe, 1.00 (1). Edman II: **Leu, 0.07;** Phe, 1.00 (1).

Peptides T-21ch1 and T-5ch4. Composite Sequence, **Asn-Gly-Ile-Phe.** Peptide T-21ch1 (Asp, 1.05 (1); Gly, 1.11 (1); Ile, 0.97 (1); Phe, 0.88 (1) [13%; *el*, 0; *ch*, 23; Y to B]) was shown to have the amino acid sequence listed above by carboxypeptidase A digestion, which released phenylalanine (+3) and isoleucine (+1) after 45 min of digestion, and by Edman degradation as follows. Edman I: **Asp, 0.27;** Gly, 1.00 (1); Ile, 1.08 (1); Phe, 0.93 (1). Edman II: Asp, 0.22; **Gly, 0.43;** Ile, 1.08 (1); Phe, 0.91 (1). Digestion with leucine aminopeptidase followed by amino acid analysis showed that the peptide contained asparagine rather than aspartic acid, as was also indicated by the electrophoretic mobility.

Peptide T-5ch4 (Asp, 0.91 (1); Gly, 1.09 (1); Ile, 1.04 (1); Phe, 0.95 (1) [10%; *el*, −7; *ch*, 27; Or]) had the same amino acid composition as peptide T-21ch1 except that its electrophoretic mobility indicated the presence of aspartic acid rather than asparagine. This peptide thus appears to be a deamidated product of peptide T-21ch1. Digestion with carboxypeptidase

TABLE III

	Thr-Phe-Tyr-Ser-Ile-Ile-Asp-(Asp, Glu)-His-Lys ←——————TT-9elN——————→ ←T-36ch5→←————T-45ch1elA1————→
TT-9elN	Thr, 0.98 (1); Phe, 1.28 (1); Tyr, 0.74 (1); Ser, 1.07 (1); Ile, 0.97 (2); Asp, 2.11 (2); Glu, 0.95 (1); His, 1.00 (1); Lys, 0.98 (1) [43%; *el*, 0; *ch*, 0; B; Pauly].
Edman I	**Thr, 0.02;** Phe, 1.07 (1); Tyr, 0.93 (1); Ser, 1.00 (1); Ile, 1.21 (2); Asp, 1.96 (2); Glu, 1.00 (1); His, Lys, nd.
Edman II	Thr, 0.04; **Phe, 0.19;** Tyr, 0.69 (1); Ser, 0.96 (1); Ile, 1.25 (2); Asp, 2.07 (2); Glu, 0.96 (1); His, Lys, nd.
Edman III	Thr, 0.04; Phe, 0.10; **Tyr, 0.10;** Ser, 0.96 (1); Ile, 1.20 (2); Asp, 2.10 (2); Glu, 0.92 (1); His, Lys, nd.
Edman IV	Thr, 0.04; Phe, 0.19; Tyr, 0.10; **Ser, 0.28;** Ile, 1.12 (2); Asp, 2.10 (2); Glu, 1.00 (1); His, Lys, nd.
T-36ch5	Thr, 1.04 (1); Phe, 1.03 (1); Tyr, 0.93 (1) [48%; *el*, 0; *ch*, 31; G; Pauly].
Edman I	**Thr, 0.00;** Phe, 1.00 (1); Tyr, 0.20 (1).
Carboxypeptidase A	Tyr, +4 (30 min, room temperature).
T-45ch1elA1	Ser, 0.96 (1); Ile, 1.47 (2); Asp, 2.02 (2); Glu, 1.02 (1); His, 0.93 (1); Lys, 1.06 (1) [15%; *el*, −5; *ch*, 6; B; Pauly].
Edman I	**Ser, 0.03;** Ile, 0.31 (2); Asp, 1.97 (2); Glu, 1.09 (1); His, 1.03 (1); Lys, 0.72 (1).
II	Ser, 0.03; **Ile, 0.91** (1); Asp, 2.03 (2); Glu, 1.06 (1); His, 1.03 (1); Lys, 0.89 (1).
III	Ser, 0.10; **Ile, 0.16;** Asp, 2.10 (2); Glu, 1.06 (1); His, 0.90 (1); Lys, 0.71 (1).
IV	Ser, 0.03; Ile, 0.10; **Asp, 1.39** (1); Glu, 0.85 (1); His, 1.25 (1); Lys, 0.90 (1).
Leucine aminopeptidase	Ser, 0.98 (1); Ile, 2.04 (2) (preparation 2, 2 hr, 40°).
Carboxypeptidase B	Lys, +4 (1.5 hr, room temperature). Lys, +4; His, +1 (8 hr, room temperature).

A yielded results which were essentially identical with those obtained from peptide T-21ch1. However, T-5ch4 was anomalous in that it did not undergo Edman degradation, the residual peptide yielding the same analysis even after two cycles. Moreover, it gave an unusually strong orange color on paper with the ninhydrin–collidine reagent. Similarly, residual peptides observed during the paper chromatography of the carboxypeptidase A digests, which contained one and two residues less, respectively, at the carboxyl-terminal end, gave the same ninhydrin–collidine color. It is interesting to note that this peptide was clearly detected in the peptide maps of the original tryptic and chymotryptic digests of the whole protein. The possible reasons for this anomalous behavior are considered under Discussion.

Peptide T-47elA1. Sequence, **His-Leu-Ala-Ile-Asp-(Asp$_4$, Gly, Ala, Leu)-Glu-Leu-Arg.** This peptide (His, 0.91 (1); Leu, 2.86 (3); Ala, 2.04 (2); Ile, 0.97 (1); Asp, 5.11 (5); Gly, 1.00 (1); Glu, 1.05 (1); Arg, 1.04 (1) [34%; *el*, −9; *ch*, 13; Y; Pauly, Sakaguchi]) was subjected to five steps of Edman degradation as follows. Edman I: **His, 0.07;** Leu, 2.99 (3); Ala, 1.88 (2); Ile, 1.00 (1); Asp, 5.13 (5); Gly, 0.97 (1); Glu, 1.04 (1); Arg, 1.00 (1). Edman II: **Leu, 1.98** (2); Ala, 2.02 (2); Ile, 0.89 (1); Asp, 4.87 (5); Gly, 1.08 (1); Glu, 1.04 (1); Arg, His, nd.[1] Edman III: Leu, 2.22 (2); **Ala, 1.36** (1); Ile, 0.98 (1); Asp, 4.97 (5); Gly, 1.10 (1); Glu, 0.98 (1); Arg, His nd. Edman IV: Leu, 2.12 (2); Ala, 1.16 (1); **Ile, 0.27;** Asp, 5.01 (5); Gly, 1.01 (1); Glu, 0.99 (1); Arg, His, nd. Edman V: Leu, 2.14 (2); Ala, 1.18 (1); Ile, 0.23; **Asp, 3.98** (4); Gly, 1.02 (1); Glu, 0.98 (1); Arg, His, nd.

Analysis of the residual peptide after the sixth cycle of Edman degradation gave the same amino acid analysis as after the fifth cycle. Leucine aminopeptidase also failed to release any amino acids from this residual peptide. While the presence of an aspartyl residue alone would be sufficient to prevent action by the leucine aminopeptidase, the lack of action of both the agents could be explained by the presence of an aspartyl or, especially, an asparaginyl residue in the sixth place, while could have become converted to a β-aspartyl linkage by the acid treatment step of the Edman degradation. The modalities of such transformations are considered under Discussion. It may be noted that the electrophoretic mobility of the original peptide indicates that it contains at least one amide group, which cannot be on the lone glutamyl residue as this was liberated as glutamic acid by the action of carboxypeptidase B (see below), and which may well be on the residue in question. Digestion of the original peptide with leucine aminopeptidase yielded His, 1.10 (1); Leu, 1.04 (1); Ala, 0.96 (1); Ile, 0.91 (1), thus placing these four residues at the amino terminus, confirming to that extent the results of the Edman degradation.

Carboxypeptidase B (5 hr, room temperature)

TABLE IV

(His, Asp, Glu_2, Val, Leu, Phe)-Leu-Met-Gln-Ala-Ser-Gln-Tyr
←————T-0————→←——T-2ch2——→
←—T-14,15elA1—→
←————T-20ch2————→

T-0	His, 1.00 (1); Asp, 1.16 (1); Glu, 1.86 (2); Val, 1.03 (1); Leu, 2.02 (2); Phe, 1.10 (1); Met, 0.69 (1); Ser, 0.18; Gly, 0.14; Ala, 0.23; Ile, 0.13; Tyr, 0.14 [22%].
Carboxypeptidase A	Met, +4; Leu, +1 (2 min, room temperature). Met, +4; Leu, +3 (10 min, room temperature).
T-2ch2	Glu, 2.01 (2); Ala, 1.03 (1); Ser, 1.04 (1); Tyr, 0.94 (1) [37%; *el*, 0; *ch*, 18; ninhydrin negative; Pauly].
Carboxypeptidase A	Tyr, +4 (5 min, room temperature). Tyr, +4; Gln, +2; Ser, +1 (1hr, room temperature).
T-14,15elA1	Glu, 2.05 (2); Ala, 0.96 (1); Ser, 1.00 (1); Tyr, 1.00 (1) [4%; *el*, +6; *ch*, 14; B; Pauly].
Edman I	**Glu, 1.22** (1); Ala, 0.95 (1); Ser, 1.10 (1); Tyr, 0.95 (1).
II	**Ala, 0.26**; Ser, 0.96 (1); Glu, 1.29 (1); Tyr, 1.04 (1).
III	Ala, 0.21; **Ser, 0.27**; Glu, 1.02 (1); Tyr, 0.98 (1).
Carboxypeptidase A	Tyr, +4; Gln, +1 (15 min, room temperature). Tyr, +4; Gln, +3; Ser, +1 (2 hr, room temperature).
T-20ch2	Met, 0.74 (1); Glu, 1.99 (2); Ala, 1.01 (1); Ser, 1.05 (1); Tyr, 0.68 (1) [11%; *el*, +3; *ch*, 9; B; Pauly].
Edman I	**Met, 0.27**; Glu, 2.02 (2); Ala, 1.05 (1); Ser, 0.92 (1); Tyr, 0.31 (1).
Carboxypeptidase A	Tyr, +4 (15 min, room temperature). Tyr, +4; Gln, +3 (2 hr, room temperature).

released Arg, Leu, and Glu in that order from the peptide, in molar ratios relative to Arg of Arg, 1.00, Leu, 0.96, and Glu, 0.58, thus establishing the carboxyl-terminal sequence as given. Lack of further action of the carboxypeptidase B could be due to the presence of an aspartyl residue preceding the glutamyl residue. Extensive losses during the purification of the peptide and failure to separate well-characterized products following hydrolysis with chymotrypsin, pepsin, and dilute acetic acid precluded any further studies.

Peptide T-42ch1. **Arg.** This fraction was shown to consist of free arginine by amino acid analysis of an unhydrolyzed sample. It was obtained in 58% yield [*el*, +10; *ch*, 7; B; Sakaguchi].

Peptides T-0, T-20ch2, and T-14,15elA1. Composite Sequence, **(His, Asp, Glu_2, Val, Leu, Phe)-Leu-Met-Gln-Ala-Ser-Gln-Tyr** (Table IV). Peptide T-0 was obtained as the insoluble residue when the lyophilized initial tryptic digest was dissolved in the starting chromatographic buffer. This fraction was not further purified. Carboxypeptidase A digestion indicated the carboxyl-terminal sequence to be Leu-Met. Small amounts of the same peptide in yields of about 2% were recovered in fractions T-43 and TT-5.

The unique amino acid compositions of the other three peptides listed above clearly indicate that they belong to the same segment of the protein. Moreover, since the protein subunit contains a single methionyl residue (Groskopf *et al.*, 1966), T-20ch2 must follow and overlap with T-0. Carboxypeptidase A digestion of peptides T-2ch2, T-14,15elA1, and T-20ch2 showed that all three terminate in the sequence Glu-Tyr.

Successive Edman degradations of peptide T-20ch2 released only methionine, did not release any residues from peptide T-2ch2, and were blocked after three steps on peptide T-14,15elA1. In all cases this was presumably due to the cyclization of amino-terminal glutaminyl residues (see Hill, 1965). This conclusion is supported by the failure of leucine aminopeptidase to release any residues from peptide T-2ch2, while there was complete digestion of peptides T-20ch2 and T-14,15elA1.

Peptide T-14,15elA1 is apparently a deamidated product of T-2ch2, in which the amino-terminal residue is present as glutamic acid while the penultimate residue is glutamine, as indicated by its electrophoretic mobility and the fact that only three residues could be removed by Edman degradation. It may be noted that the electrophoretic mobility of T-2ch2 also clearly indicates that both acidic residues are amidated in this peptide.

The amino acid sequence of the carboxyl-terminal half of the over-all segment was established as detailed in Table IV while that of the amino-terminal half could not be pursued as peptide T-0 adsorbed irreversibly to paper and Sephadex G-25 during attempts at purification.

Peptide TT-12elB2. Sequence, **Cys-Thr-Gly-Lys.** Two steps of Edman degradation of this peptide (Cys, 0.85 (1); Thr, 1.04 (1); Gly, 1.02 (1); Lys, 1.13 (1) [25%;

TABLE V

Gln-Phe-Tyr-Asp-Glu-His-Lys-Lys
←————T-57elN————→
←T-5ch5→←— T-59elB1ch1 —→
←————T-41ch1————→
←T-47elNch1→

T-57elN	Glu, 1.87 (2); Phe, 1.01 (1); Tyr, 0.87 (1); Asp, 1.20 (1); His, 1.06 (1); Lys, 1.96 (2) [18%; *el*, 0; *ch*, 7; B; Pauly].
Edman I	No change in amino acid composition.
T-5ch5	Glu, 0.87 (1); Phe, 1.13 (1); Tyr, 1.00 (1) [10%; *el*, −7; *ch*, 38; ninhydrin negative; Pauly].
Edman I	No change in amino acid composition.
Carboxypeptidase A	Tyr, +4 (1 hr, room temperature).
T-59elB1ch1	Asp, 0.89 (1); Glu, 1.06 (1); His, 0.94 (1); Lys, 2.11 (2) [25%; *el*, +5; *ch*, 1; B; Pauly].
Edman I	**Asp, 0.20;** Glu, 1.01 (1); His, 0.99 (1); Lys, 1.48 (2).
T-41ch1	Glu, 2.02 (2); Phe, 0.92 (1); Tyr, 0.87 (1); Asp, 1.18 (1); His, 1.02 (1); Lys, 0.99 (1) [23%; *el*, −4; *ch*, 14; B; Pauly].
Edman I	No change in amino acid composition.
T-47elNch1	Asp, 1.00 (1); Glu, 0.98 (1); His, 0.90 (1); Lys, 1.13 (1) [24%; *el*, 0; *ch*, 1; B; Pauly].
Edman I	**Asp, 0.09;** Glu, 0.97 (1); His, 1.03 (1); Lys, 0.24 (1).
II	Asp, 0.05; **Glu, 0.35;** His, 1.15 (1); Lys, 0.85 (1).
III	Asp, 0.08; Glu, 0.24; **His, 0.40;** Lys, 1.00 (1).

el, +8; *ch*, 0; B]), the assumption of tryptic specificity to place lysine at the carboxyl terminus, and the location of the glycyl residue by difference suffice to establish the structure as given above. The Edman degradation results were as follows. Edman I: **Cys, 0;** Thr, 0.96 (1); Gly, 1.07 (1); Lys, 0.96 (1). Edman II: Cys, 0; **Thr, 0.45;** Gly, 1.00 (1); Lys, nd.

Peptides T-57elN, T-41Ch1, T-5ch5, T-59elB1ch1, and T-47elNch1. Composite Sequence, **Gln-Phe-Tyr-Asp-Leu-His-Lys-Lys** (Table V). Peptides T-57elN and T-41ch1 differed in amino acid composition by one lysine residue. Assuming tryptic specificity, one may postulate a Lys-Lys sequence as the carboxyl terminus of this segment of the protein. Both peptides were not degraded by the Edman procedure, indicating that glutamine, which had secondarily cyclized to pyrrolidonecarboxylic acid, probably occupied the amino terminus (Hill, 1965).

Peptide T-5ch5 behaved similarly to peptides T-57elN and T-41ch1 on Edman degradation. Its carboxyl-terminal residue was shown to be tyrosine by carboxypeptidase A digestion, thus establishing its amino acid sequence, assuming glutamine to have been the original amino-terminal residue. The structure of the remainder of the over-all sequence was established by Edman degradations of peptides T-59elB1ch1 and T-47elNch1. These two peptides clearly comprise the carboxyl-terminal segment since they contained the postulated Lys-Lys sequence and the lysine residue, respectively, while peptide T-5ch5 comprises the amino-terminal segment as evidenced by its composition and behavior on Edman degradation.

Chymotryptic digestions of peptides T-57elN and T-41ch1 yielded two fragments in each case. From both peptides, one of the chymotryptic fragments was identical by amino acid composition and electrophoretic and chromatographic mobility to peptide T-5ch5, while the second fragment obtained from peptide T-57elN was similarly identical with peptide T-59elB1ch1 and the second fragment recovered from peptide T-41ch1 was identical with peptide T-47elNch1. This confirms the over-all assignment of sequence listed above.

Peptide T-38elB1. **Lys.** This fraction (Lys, 1.00 (1) [45%; *el*, +13; *ch*, 5; B]) was shown to be free lysine by amino acid analysis of an unhydrolyzed sample.

Peptides TT-15elNch2, T-61ch1, T-44ch1elA1, and T-70ch2. Composite Sequence, **Lys-Glu-His-Glu-Thr-Phe-Ile-His-Ala-Leu-(Asp, Asn)-Trp-Lys** (Table VI). The unique compositions of these peptides indicate that peptides T-61ch1, T-44ch1elA1, and T-70ch2 represent segments of the over-all sequence contained in peptide TT-15elNch2. The single lysine in T-70ch2 must, from tryptic specificity, represent the carboxyl-terminal residue in the over-all sequence, placing this peptide at the carboxyl-terminal half of peptide TT-15elNch2 and, therefore, peptides T-61ch1 and T-44ch1elA1 at the amino-terminal half. This assignment was confirmed by carboxypeptidase A digestion of peptide TT-15elNch2, which released lysine and tryptophan, as it did in a similar experiment with peptide T-70ch2.

The sequence at the amino-terminal end of this segment of the protein was established by Edman degradation and carboxypeptidase A digestions of

TABLE VI

Lys-Glu-His-Glu-Thr-Phe-Ile-His-Ala-Leu-(Asp, Asn)-Trp-Lys
←——T-61ch1——→←——T-70ch2——→
←——TT-15elNch2——→
←— T-44ch1elA1 —→

T-61ch1	Lys, 1.22 (1); Glu, 2.02 (2); His, 0.92 (1); Thr, 0.89 (1); Phe, 0.93 (1) [7%; *el*, 0; *ch*, 5; B; Pauly].
Edman I	**Lys, 0.23;** Glu, 1.90 (2); His, 0.94 (1); Thr, 1.06 (1); Phe, 1.06 (1).
II	Lys, 0.20; **Glu, 1.26** (1); His, 0.93 (1); Thr, 1.09 (1); Phe, 1.00 (1).
III	Lys, 0.13; Glu, 1.16 (1); **His, 0.47**; Thr, 1.00 (1); Phe, 0.84 (1).
Carboxypeptidase A	Phe, +4 (3 hr, room temperature).
Dinitrophenylation	Di-DNP-lysine only.
T-70ch2	Ile, 0.92 (1); His, 0.92 (1); Ala, 1.13 (1); Leu, 1.05 (1); Asp, 1.91 (2); Trp (1);[a] Lys, 1.08 (1) [23%; *el*, +5; *ch*, 18; B; Ehrlich, Pauly].
Edman I	**Ile, 0.03;** His, 0.97 (1); Ala, 1.03 (1); Leu, 1.03 (1); Asp, 2.01 (2); Trp nd; Lys, 0.25 (1).
II	**His, 0.19;** Ala, 1.12 (1); Leu, 0.96 (1); Asp, 1.92 (2); Trp, nd; Lys, 0.25 (1).
III	His, 0.11; **Ala, 0.24;** Leu, 0.97 (1); Asp, 2.03 (2); Trp, nd; Lys, 0.22 (1).
IV	**Leu, 0.36;** Asp, 2.00 (2); Trp, nd; Lys, 0.00 (1).
Leucine aminopeptidase	Ile, 0.97 (1); Ala, 1.01 (1); Leu, 1.03 (1); His, Trp, Lys, nd (preparation 2, 4 hr, 40°).
Carboxypeptidase A	Lys, 1.00 (1); Trp, 1.00 (1) (4 hr, room temperature).
CG2	Ile, 0.89 (1); His, 0.85 (1); Ala, 1.15 (1); Leu, 1.10 (1) [*el*, +3; *ch*, 27; Pauly].
CG3a	Asp, 1.98 (2); Trp (1);[a] Lys, 1.02 (1) [*el*, 0; *ch*, 6; G; Ehrlich].
CG3b	Lys [*el*, +13; *ch*, 5].
CG4	Composition not determined [*el*, −6; *ch*, 15; Ehrlich].
TT-15elNch2	Glu, 1.86 (2); His, 1.89 (2); Thr, 0.89 (1); Phe, 1.00 (1); Ile, 0.89 (1); Ala, 1.05 (1); Leu, 1.00 (1); Asp, 2.16 (2); Trp (1);[a] Lys, 0.89 (1); Gly, 0.20 [16%; *el*, 0; *ch*, 3; B; Ehrlich, Pauly].
T-44ch1elA1	Glu, 2.06 (2); His, 1.25 (1); Thr, 0.99 (1); Phe, 0.97 (1) [26%; *el*, −8; *ch*, 14; B; Pauly].
Edman I	**Glu, 1.19** (1); His, 1.08 (1); Thr, 0.97 (1); Phe, 0.96 (1).
II	**His, 0.30;** Glu, 1.24 (1); Thr, 1.01 (1); Phe, 0.99 (1).
III	His, 0.30; **Glu, 0.65;** Thr, 1.05 (1); Phe, 0.95 (1).
Leucine aminopeptidase	Glu, 2.11 (2); His, 0.95 (1); Thr, 0.96 (1); Phe, 0.97 (1) (preparation 2, 4 hr, 40°).
Carboxypeptidase A	Phe, +4 (2 hr, room temperature). Phe, +4; Thr, +1 (4 hr, room temperature).

[a] Assumed from positive Ehrlich reaction.

peptides T-61ch1 and T-44ch1elA1. Amino acid analysis of a complete leucine aminopeptidase digest of peptide T-44ch1elA1 showed that both glutamyl residues were not amidated.

The structure of peptide T-70ch2 was determined by four steps of Edman degradation, by carboxypeptidase A digestion, and from the properties of fragments obtained by digestion with chymotrypsin (fragments CG2, CG3a, CG3b, and CG4). These were purified by Sephadex gel filtration and by paper electrophoresis. Fragment CG2 represented the amino-terminal tetrapeptide sequence, CG3a the carboxyl-terminal tetrapeptide, while a third fragment, CG3b, which appeared in the same Sephadex fraction as CG3a and was separated from it by electrophoresis, was shown to be free lysine by amino acid analysis of an unhydrolyzed sample. Fragment CG4 differed from fragment CG3a in the absence of the carboxyl-terminal lysine. CG4 was acid on electrophoresis while CG3a was neutral, indicating that one of the two aspartyl residues in these peptides and therefore in the over-all sequence was amidated. The relative position of the aspartyl and asparaginyl residues was not determined.

Peptide T-40ch1. Sequence, **Gly-Asp-Val-Lys.** The amino acid sequence of this peptide (Gly, 0.96 (1); Asp, 1.07 (1); Val, 1.01 (1); Lys, 0.98 (1) [50%; el, 0; *ch*, 7; G]) was determined by two steps of Edman degradation, assuming the lysyl residue to be carboxyl terminal from tryptic specificity and placing the valine by difference. The Edman degradation results were as follows. Edman I: **Gly, 0.12;** Asp, 0.99 (1); Val, 1.01 (1); Lys, 0.31 (1). Edman II: Gly, 0.03; **Asp, 0.27;** Val, 1.01 (1); Lys, 1.10 (1). Digestion with leucine aminopeptidase released only glycine, confirming the above results.

Peptides TT-13G3 and T-73,74ch1. Composite Sequence,

TABLE VII

Ser-Trp-Leu-Val-Asn-His-Ile-Lys
←————T-76ch1————→
←——T-65ch2G1——→
←——T-65ch1——→
←T-71ch2→

T-76ch1	Ser, 1.06 (1); Trp (1);[a] Leu, 1.05 (1); Val, 1.09 (1); Asp, 1.15 (1); His, 0.82 (1); Ile, 0.98 (1); Lys, 0.86 (1) [11%; *el*, +9; *ch*, 19; B; Ehrlich, Pauly].
Edman I	**Ser, 0.14;** Leu, 0.91 (1); Val, 1.00 (1); Asp, 1.09 (1); Ile, 0.91 (1); Trp, His, Lys, nd.
T-65-ch2G1	Leu, 1.07 (1); Val, 0.93 (1); Asp, 1.03 (1); His, 1.00 (1); Ile, 1.03 (1); Lys, 0.86 (1) [6%; *el*, +7; *ch*, 10; B; Pauly].
Edman I	**Leu, 0.07;** Val, 0.93 (1); Asp, 1.03 (1); His, 1.00 (1); Ile, 1.03 (1); Lys, 0.61 (1).
II	**Val, 0.00;** Asp, 0.94 (1); His, 1.06 (1); Ile, 1.00 (1); Lys, 0.61 (1).
III	**Asp, 0.35;** Ile, 1.00 (1); His, Lys, nd.
IV	**His, 0.22;** Ile, 1.00 (1); Lys, 0.18 (1).
T-65ch1	Val, 0.97 (1); Asp, 1.07 (1); His, 0.94 (1); Ile, 0.97 (1); Lys, 1.07 (1); Gly, 0.16 [6%; *el*, +7; *ch*, 4; B; Pauly].
Edman I	**Val, 0.23;** Asp, 1.02 (1); Ile, 0.98 (1); His, Lys, nd.
II	Val, 0.15; **Asp, 0.44;** His, 0.92 (1); Ile, 1.10 (1); Lys, 0.23 (1).
III	Val, 0.02; Asp, 0.30; **His, 0.50;** Ile, 1.00 (1); Lys, 0.20 (1).
T-71ch2	His, 0.92 (1); Ile, 0.99 (1); Lys, 1.08 (1) [12%; *el*, +10; *ch*, 6; B; Pauly].
Edman I	**His, 0.06;** Ile, 1.00 (1); Lys, 0.18 (1).
II	His, 0.38; **Ile, 0.00;** Lys, 1.00 (1).

[a] Assumed from positive Ehrlich reaction.

Trp-Ala-Lys. Peptide TT-13G3 (analysis of complete leucine aminopeptidase digest: Trp, 1.00 (1); Ala, 0.99 (1); Lys, 1.01 (1) [20%; *el*, +5; ch, 14; B; Ehrlich]) was subjected to dinitrophenylation. The acid hydrolysate of the dinitrophenyl (DNP) peptide contained ϵ-DNP-lysine, free alanine, and a typical fluorescent DNP-tyrptophan breakdown product. Digestion with carboxypeptidase A for 5 hr released only lysine, establishing the sequence as given.

Peptide T-73,74ch1 (Ala, 0.98 (1); Lys, 1.02 (1) [12%, *el*, +5; *ch*, 15; B]) was subjected to dinitrophenylation. The acid hydrolysate yielded DNP-alanine and ϵ-DNP-lysine, thus establishing the sequence Ala-Lys. Both of the above peptides are considered to arise from the same segment of the protein, peptide T-73,74ch1 deriving from a chymotryptic split at the tryptophanyl residue.

Peptides T-76ch1, T-65ch2G1, T-65ch1, T-71ch2. Composite Sequence, **Ser-Trp-Leu-Val-Asn-His-Ile-Lys** (Table VII). The amino acid sequence of this segment was determined by Edman degradations of the over-all peptide T-76ch1 and of the three other peptides, which contained different lengths of the carboxyl-terminal segment. The tryptophanyl residue was placed by difference. A complete leucine aminopeptidase digest of peptide T-76ch1 showed that it contained asparagine and not aspartic acid, in agreement with the cathodic mobility of all four peptides.

Peptides T45-ch2elN, T-59elB2, T-50ch2, T-45ch1elB1, and T-7elN. Composite Sequence, **Thr-Ile-Asp-Phe-Lys-Tyr-Lys-Gly-Lys-Ile** (Table VIII). These peptides were arranged in the above sequence because their unique amino acid compositions fit with the residues released by carboxypeptidase A from the denatured complete protein (Groskopf *et al.*, 1966). The assignment of sequence at the carboxyl-terminal end is consistent with the overlapping compositions of peptides T-50ch2 and T-45ch1elB1, the first of which contains the carboxyl-terminal isoleucine of the protein polypeptide chain and the other of which does not. Peptide T-59elB2 was placed between peptides T-45ch2-elN and T-50ch2 because, other than lysine and isoleucine, carboxypeptidase A liberated from hemerythrin glycine, tyrosine, and phenylalanine, in that order. The structure of peptide T-45ch2elN was determined by Edman degradation and by analysis and sequence determination of a dipeptide fragment, AAelB1, obtained from it by dilute acetic acid hydrolysis.

The structures of the dipeptides T-59elB2 and T-45ch1elB1 were established by Edman degradation. Peptide T-7elN was shown to be free isoleucine by analysis of an unhydrolyzed sample, and T-50ch2 was found to be Gly-Lys-Ile by Edman degradation and carboxypeptidase A digestion. Since after one cycle of Edman degradation lysine in the residual peptide from T-50ch2 was largely destroyed, the amino-terminal position of glycine was confirmed

TABLE VIII

Thr-Ile-Asp-Phe-Lys-Tyr-Lys-Gly-Lys-Ile

T-45

←—T-45ch2elN—→←T-59→←ch1→←T-7→

elB2 elB1 elN

←— T-50 —→

ch2

T-45ch2elN	Thr, 1.02 (1); Ile, 1.00 (1); Asp, 1.01 (1); Phe, 1.00 (1); Lys, 0.97 (1) [30%; *el*, 0; *ch*, 17; B].
Edman I	**Thr, 0.11;** Ile, 1.00 (1); Asp, 1.11 (1); Phe, 0.89 (1); Lys, 0.76 (1).
II	**Ile, 0.02;** Asp, 1.17 (1); Phe, 0.83 (1); Lys, 0.74 (1).
III	**Asp, 0.30;** Phe, 1.00 (1); Lys, 0.51 (1).
IV	**Phe, 0.02;** Lys, 1.00 (1).
AAelB1	Phe, 0.93 (1); Lys, 1.07 (1) [*el*, +7; *ch*, 14; B].
Edman I	**Phe, 0.35;** Lys, 1.00 (1).
T-59elB2	Tyr, 1.15 (1); Lys, 0.85 (1) [54%; *el*, +9; *ch*, 10; B; Pauly].
EdmanI	**Tyr, 0.08;** Lys, 1.00 (1).
T-45ch1elB1	Gly, 0.85 (1); Lys, 1.15 (1) [17%; *el*, +10; *ch*, 5; Y to G].
Edman I	**Gly, 0.30;** Lys, 1.00 (1).
Dinitrophenylation	DNP-Gly; ε-DNP-Lys
T-7elN	Ile, 1.00 (1) [21%; *el*, 0; *ch*, 27; B].
T-50ch2	Gly, 0.96 (1); Lys, 1.03 (1); Ile, 1.01 (1) [46%; *el*, +10; *ch*, 14; G].
Edman I	**Gly, 0.31;** Lys, 0.02 (1); Ile, 1.00 (1).
Carboxypeptidase A	Ile, +4 (6 hr, room temperature, pH 7.0).

by identifying PTH-glycine by paper chromatography.

Discussion

Enzymic Hydrolysis of Hemerythrin. Notwithstanding that the preparation of trypsin had been treated with 0.01 M HCl to destroy chymotryptic activity, the T digest contained a considerable number of peptides that had clearly resulted from the action of chymotrypsin. Thus, of 31 peptide bonds which were observed to be hydrolyzed, 15 involved tryptophan, phenylalanine, tyrosine, leucine, methionine, and asparagine, the usual range of chymotryptic specificity (Hill, 1965). That this was in fact due to contamination of the enzyme preparation with chymotrypsin and not to chymotryptic-like activity of the trypsin itself is clearly evidenced by the results of the TT digest. In this latter case the enzyme preparation was treated with TPCK, and every one of the 11 peptides that were isolated and characterized were ostensibly products solely of tryptic activity.

An unexplained peptide bond hydrolysis observed in the T digest was at a Trp-Pro bond, resulting in an amino-terminal segment (T-39ch3; Val-Asp-Trp) and a carboxyl-terminal segment (T-57elB2; Pro-Ser-Phe-Arg) of the parent peptide T-61ch3 in yields of 35 and 50%, respectively (see Table II). A 5% yield of hydrolysis at a Phe-Pro bond has been observed in ovine corticotropin following digestion with chymotrypsin (Leonis *et al.*, 1959). Whether such hydrolyses negate the commonly accepted rule that bonds involving the imino nitrogen of proline are not hydrolyzed by trypsin or chymotrypsin (Hill, 1965) or whether the digestion was due to a more specific enzymic contaminant of the preparation cannot readily be decided. Another factor which may have been operative in the hemerythrin peptide is the occurrence of four prolyl residues within a chain length of ten amino acids, resulting in a particular conformation which could conceivably have facilitated hydrolysis at the Trp-Pro bond.

Another anomalous point of hydrolysis occurred in the same segment of the protein at an Asp-Pro bond. The yields of the resultant peptides (T-9ch2 and T-16,17,18ch1) were 23 and 12%, respectively (see Table II). In this case not only is an imino nitrogen of proline involved but the other residue, aspartic acid, has never been observed to be susceptible to either trypsin or chymotrypsin action. The hydrolysis in this case may have been nonenzymatic, and again conformational factors may have played a role.

The amino-terminal peptide also served to detect an unusual enzymic contaminant of the leucine aminopeptidase preparations used. As shown in Table II, the only residues which were not liberated were the four prolines and the residues immediately preceding them in the sequence Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr-Val-Asp-Trp-Pro-Ser-Phe-Arg. Thus the action of leucine aminopeptidase, which is known not to split appreciably bonds involving the imino nitrogen of proline (Hill, 1965), must have been aided by hydrolyses at the Pro-Tyr and Pro-Ser bonds. Whether

TABLE IX: Sequences Established, and Their Over-All Yields, as Calculated from Yields of Constituent Peptides Recovered from Tryptic Digests of *G. gouldii* Hemerythrin.

	Sequence (represented by peptide or peptides)[a]	Over-All Yield (%)[b]
TT-5ch4	Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr-Val-Asp-Trp-Pro-Ser-Phe-Arg	66
TT-9elN	Thr-Phe-Tyr-Ser-Ile-Ile-Asp-(Asp, Glu)-His-Lys	48
T-27ch4	Thr-Leu-Phe	22
T-21ch1	Asn-Gly-Ile-Phe	23
T-47elA1	His-Leu-Ala-Ile-Asp-(Asp_4, Gly, Ala, Leu)-Glu-Leu-Arg	34
T-42ch1	Arg	58
T-0/T-2ch2	(His, Asp, Glu_2, Val, Leu, Phe)-Leu-Met-Gln-Ala-Ser-Gln-Tyr	52
TT-12elB2	Cys-Thr-Gly-Lys	25
T-57elN	Gln-Phe-Tyr-Asp-Glu-His-Lys-Lys	90
T-61ch1/T-70ch2	Lys-Glu-His-Glu-Thr-Phe-Ile-His-Ala-Leu-(Asp, Asn)-Trp-Lys	34
T-40ch1	Gly-Asp-Val-Lys	50
TT-13G3	Trp-Ala-Lys	20
T-76ch1	Ser-Trp-Leu-Val-Asn-His-Ile-Lys	35
T-45ch2elN/T-59elB2/T-50ch2	Thr-Ile-Asp-Phe-Lys-Tyr-Lys-Gly-Lys-Ile	67

[a] Sequences listed in order discussed under Results. [b] Yields are calculated by addition of the best combination of yield of nonconsecutive peptides from the same digest. For instance, the yield of the sequence represented by peptide TT-5ch4 is arrived at by adding the yields of peptides T-10ch2 (43%) and T-9ch2 (23%).

hydrolysis also occurred at the Pro-Ile and Pro-Asp bonds was not examined. The contaminating enzyme must therefore be specific for bonds involving the carboxyl group of proline and act either completely as an endopeptidase or at least be capable of attacking such bonds when the proline is the second residue from the amino-terminal end of the peptide. The former possibility is the more likely since the yield of free aspartic acid was only one-third that of the three terminal residues, implying that the Pro-Ser bond had been hydrolyzed even when present in a peptide the minimal structure of which must have been Asp-Trp-Pro-Ser-Phe-Arg. As far as we are aware, an enzyme having such a prolyl endopeptidase specificity has so far only been reported for a proteinase of *Bacillus thermoproteolyticus* (Ohta and Ogura, 1965). From these results it would appear that such an enzyme must also be present in animal tissues and when suitably purified may be of great value in protein structural studies.

Secondary Changes at Amino-Terminal Glutaminyl and Asparaginyl Residues. In every case in which glutamine occurred as an amino-terminal residue, the commonly observed cyclization to pyrrolidonecarboxylic acid (Sanger *et al.*, 1955; Hill, 1965) was observed to occur practically quantitatively, preventing both Edman degradation and leucine aminopeptidase digestion. Thus peptides T-57elN, T-41ch1, and T-5ch5 (see Table V) and peptide T-2ch2 (see Table IV) were isolated from the original digest with a cyclized residue at the amino terminus. Similarly, following Edman degradation when a glutaminyl residue became amino terminal in the residual peptide, cyclization prevented further Edman degradation cycles. This occurred in peptides T-20ch2 and T-14,15elA1 (see Table IV).

Peptide T-5ch4 (sequence, Asp-Gly-Ile-Phe; see Results) had an anomalous amino-terminal aspartyl residue. This residue was not liberated by Edman degradation, while the acidic electrophoretic mobility indicated the presence of two carboxyl groups in the peptide. It appears highly likely that this aspartyl residue was linked to the rest of the peptide through its β-carboxyl group. The transformation of α-aspartyl to α,β-aspartyl and to β-aspartyl linkages of amino-terminal aspartyl residues has been observed in peptides obtained by partial hydrolysis of a variety of proteins (Swallow and Abraham, 1958; Bernhard, 1958; Naughton *et al.*, 1960; Ambler, 1963). The bright orange color of this peptide with the ninhydrin–collidine color reagent used is also probably characteristic of the amino-terminal β-aspartyl residue. Stable yellow ninhydrin colors have previously been observed for such peptides with other ninhydrin reagents not containing collidine (John and Young, 1954; Naughton *et al.*, 1960; Ambler, 1963).

Peptide T-21ch1 was identical with peptide T-5ch4 except that its amino-terminal residue was a normally behaving asparagine, indicating that in the original protein the aspartyl residue in question was present as the amide. It is interesting to note that the presumed β-aspartyl peptide was consistently detected in good yield on the peptide maps of the initial digests before any purification procedures had been initiated. Acid catalysis can therefore not be invoked in its formation as has been the case for the β-aspartyl peptides detected by the authors cited above. It is possible that in this case the amidation of the β-carboxyl group had, by a mechanism similar to the cyclization of amino-terminal

glutaminyl residues, facilitated cyclization of the asparaginyl peptide to the α,β form, which subsequently converted to the pure β-aspartyl peptide. This may be explained by the postulate that water, acting as a nucleophile on an α,β-aspartyl structure, would preferentially attack the α-carbonyl position because of the electron-withdrawing properties of the adjacent protonated amino group, giving rise to the β- rather than to the α-aspartyl link. The absence of any electrophoretically neutral material with the same amino acid composition as peptides T-5ch4 or T-21ch1 among the purified peptides rules out the presence of any appreciable amount of α,β-cyclized aspartyl peptide. As mentioned under Results, the failure of further Edman degradation of peptide T-47elA1 after the fifth cycle may have been due to a similar conversion of a normal aspartyl or asparaginyl residue in sixth place to the resistant β-aspartyl linkage.

Yields of Peptides and the Composition of Hemerythrin. The over-all yields of peptide sequences accounting for different segments of the protein are listed in Table IX. They vary from 20 to 90% and account for 113 residues in the subunit chain, which compares with the figure of 112.3 residues obtained by direct analysis of the amino acid composition of the protein (Groskopf *et al.*, 1966). Only four sequences were recovered in yields below 30%, and because the over-all yields of amino acids in all fractions from the Dowex 50W-X2 column chromatography of the TT digest ranged from 23 to 100% with a numerical average of 75% (see Table I), it is highly probable that the sequences listed represent all the peptides of the predominant sequence of the *G. gouldii* hemerythrin subunits.

While it is true that the method of calculation of over-all yields by addition of the yields of overlapping constituent segments could conceivably give a false picture of the yields of certain peptides and tends to obscure the presence of variant residues, it is nonetheless felt that the conclusion above as to the essential predominance of the sequences listed is valid.

One of the peptide fractions listed under Results does not appear in Table IX. This is fraction T-38elB1, which was shown to be free lysine and which was recovered in 45% yield. It seems very likely that this residue represents the carboxyl-terminal lysine of peptide T-57elN (Table V), and is at the same time the amino-terminal residue of peptide T-61ch1 (Table VII). This residue would therefore be the point of overlap between the segments T-57elN and T-61ch1/T-70ch2. The lysine in this position is therefore counted only once in the count of the number of residues in the subunit chain.

As noted in the preceding paper (Groskopf *et al.*, 1966), there are reasons to believe that the subunits in *G. gouldii* hemerythrin are not all strictly identical and protein preparations obtained from mixed populations of worms of this species may well contain a proportion differing by one, two, or three residues from the common type (Manwell, 1963; Groskopf *et al.*, 1963). Several peptides obtained in low yields and not listed under Results apparently contain variant residues. For instance, one peptide recovered in about 7% yield from the TT digest (peptide TT-5ch5) contained equimolar amounts of alanine, tryptophan, valine, and leucine and was tentatively assigned the sequence Ala-Trp-Leu-Val by Edman degradation and carboxypeptidase A digestion. A final determination of the sequence was not possible because of the low yield of purified peptide. It is interesting to note that the tentative sequence fits the amino-terminal sequence of peptide T-76ch1 (Table VII), with the substitution of an alanyl residue in TT-5ch5 for the seryl residue in T-76ch1. Further studies will be required to determine unequivocally whether this sequence represents one of the possible hemerythrin subunit variations.

Acknowledgment

The authors are grateful to Dr. F. H. Carpenter of the University of California (Berkeley) for a generous gift of TPCK-treated trypsin and to Mr. Kenneth Thompson and Mrs. Marillyn B. Gragg of Northwestern University and Mr. Robert L. Gardisky of Abbott Laboratories for some of the amino acid analyses.

Added in Proof

Haley *et al.* (1966) have recently reported faster conversion of asparaginylglycine than of aspartylglycine to the β-aspartyl peptide, under mild conditions. These results are in accord with our discussion of the presumed origin of peptide T-5ch4 from peptide T-21ch1 by such a conversoin.

References

Ambler, R. P. (1963), *Biochem. J. 89*, 349.
Bernhard, S. A. (1958), *J. Cellular Comp. Physiol. 54, Suppl. 1*, 195.
Canfield, R. E. (1963), *J. Biol. Chem. 238*, 2698.
Chargaff, E., Levine, C., and Green, C. (1948), *J. Biol. Chem. 175*, 67.
Edman, P. (1950), *Acta Chem. Scand. 4*, 283.
Edman, P. (1953), *Acta Chem. Scand. 7*, 700.
Groskopf, W. R., Holleman, J. W., Klotz, I. M., Keresztes-Nagy, S., and Margoliash, E. (1963), *Science 141*, 166.
Groskopf, W. R., Holleman, J. W., Margoliash, E., and Klotz, I. M. (1966), *Biochemistry 5*, 3779 (this issue; preceding paper).
Haley, E. E., Corcoran, B. J., Dorer, F. E., and Buchanan, D. L. (1966), *Biochemistry 5*, 3229.
Hill, R. L. (1965), *Advan. Protein Chem. 20*, 37.
Hirs, C. H. W., Moore, S., and Stein, W. H. (1960), *J. Biol. Chem. 235*, 633.
Ingram, V. M. (1958), *Biochim. Biophys. Acta 28*, 539.
John, W. D., and Young, G. T. (1954), *J. Chem. Soc.*, 2870.
Konigsberg, W., and Hill, R. J. (1962), *J. Biol. Chem. 237*, 2547.
Leonis, J., Li, C. H., and Chung, D. (1959), *J. Am. Chem. Soc. 81*, 419.
Manwell, C. (1963), *Science 139*, 755.
Margoliash, E., and Smith, E. L. (1962), *J. Biol. Chem.*

237, 2151.
Michl, H. (1951), *Monatsh. Chem. 82*, 489.
Naughton, M. A., Sanger, F., Hartley, B. S., and Shaw, D. C. (1960), *Biochem. J. 77*, 149.
Northrop, J. H., and Kunitz, M. (1936), *in* Handbuch der Biologischen Arbeitsmethoden, Abderhalden, E., Ed., Berlin, Urban und Schwarzenberg, p 2213.
Ohta, Y., and Ogura, Y. (1965), *J. Biochem. (Tokyo) 58*, 607.
Olson, O. E., and Kuby, S. A. (1964), *J. Biol. Chem. 239*, 460.
Pan, S. C., and Dutcher, J. D. (1956), *Anal. Chem. 28*, 836.
Partridge, J. M., and Davis, H. F. (1950), *Nature 62*, 165.
Peterson, E. A., and Sober, H. A. (1959), *Anal. Chem. 31*, 857.
Sanger, F., and Thompson, E. O. P. (1963), *Biochim. Biophys. Acta 71*, 468.
Sanger, F., Thompson, E. O. P., and Kitai, R. (1955), *Biochem. J. 59*, 509.
Schoellmann, G., and Shaw, E. (1962), *Biochem. Biophys. Res. Commun. 7*, 36.
Schroeder, W. A., Shelton, J. R., Shelton, J. B., Cormick, J., and Jones, R. T. (1963), *Biochemistry 2*, 992.
Sjöquist, J. (1953), *Acta Chem. Scand. 7*, 447.
Smith, I. (1953a), *Nature 171*, 43.
Smith, I. (1953b), *Nature 172*, 1100.
Smith, I. (1960), Chromatographic and Electrophoretic Techniques, New York, N. Y., Interscience, p 96.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.
Stein, W., and Moore, S. (1954), *J. Biol. Chem. 211*, 907.
Swallow, D. L., and Abraham, E. P. (1958), *Biochem. J. 70*, 364.
Whitfield, R. E. (1963), *Science 142*, 577.

Optical Rotatory Dispersion of Phycocyanin*

L. J. Boucher,† H. L. Crespi, and J. J. Katz

ABSTRACT: The optical rotatory dispersion (ORD) (700–220 mμ) of phycocyanin isolated from the blue-green algae *Phormidium luridum*, *Plectonema calothricoides*, and *Synechococcus lividus* has been measured. ORD curves in the visible show multiple Cotton effects for the chromophore absorptions at ~620 mμ (+) and ~350 mμ (−). The ultraviolet ORD curves show a trough at 231 mμ with a crossover point at 222 mμ; the $[\alpha]_{231}$ of −3900 ± 200° indicates the protein to have a small amount of α-helix conformation. Urea denaturation irreversibly destroys the optical activity of the chromophore absorptions. The ORD behavior of phycocyanins from *Ph. luridum* and *P. calothricoides* resembles each other closely, while *S. lividus* phycocyanins are substantially different. The ORD curves of pairs of protio- and deuteriophycocyanins are essentially identical in both H_2O and D_2O. The rotatory dispersion of phycocyanin appears to be independent of the state of aggregation of the protein.

Phycocyanin, an important member of the class of biliproteins, is a blue, photosynthetic pigment found widely distributed in blue-green algae. Although a number of physical–chemical studies of the chromoprotein have been carried out, its detailed structure is largely unknown. The quaternary structure of phycocyanin involves a monomer, trimer, and hexamer equilibrium, which has been deduced from a study of the aggregation behavior of the protein (Hattori *et al.*, 1965b; Scott and Berns, 1965). Some additional information about the tertiary and secondary structure of phycocyanin can be deduced from thermal denaturation experiments (Hattori *et al.*, 1965a). This paper reports some new information about phycocyanin based on data obtained from the measurement of the ORD[1] properties of the protein in the native and in the urea-denatured state.

Experimental Section

Protein Preparation. The protio- and deuteriophycocyanins[2] from *Phormidium luridum*, *Plectonema calothricoides*, and *Synechococcus lividus* were isolated

* From the Chemistry Division, Argonne National Laboratory, Argonne, Illinois. *Received July 25, 1966.* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

† Resident Research Associate. Present address: Department of Chemistry, Carnegie Institute of Technology, Pittsburgh, Pa. 15213.

[1] Abbreviations used: ORD, optical rotary dispersion; CD, circular dichroism.

[2] The prefix "deuterio" refers to phycocyanin containing 99.6% deuterium at all nonexchangeable positions. The prefix "protio" refers to ordinary phycocyanin with hydrogen of mass 1 at all nonexchangeable positions.