Slater, E. C. (1967), *Methods Enzymol. 10*, 48.
Sulimovici, S., and Boyd, G. S. (1968), *Eur. J. Biochem. 3*, 332.
Užgiris, V. I., Whipple, C. A., and McIntosh, E. N. (1969), *Endocrine Soc. Program 51st Meeting*, 76 (Abstract).

Structural Studies on *Dendrostomum pyroides* Hemerythrin*

Robert E. Ferrell† and G. Barrie Kitto‡

ABSTRACT: The N-terminal sequence has been determined for the first 34 residues of the nonheme iron respiratory protein hemerythrin from the sipunculid worm *Dendrostomum pyroides* using an automated sequencer. Additionally, the composition and sequence of the tryptic peptides of *D. pyroides* hemerythrin have been compared to the previously determined sequence of *Golfingia gouldii* hemerythrin. The data indicate that there are four amino acid sequence differences between these two proteins and suggest that *Dendrostomum* and *Golfingia* separated from a common ancestor fairly recently in terms of the evolutionary history of the sipunculids.

In a previous study we reported the physicochemical properties of the hemerythrin from the sipunculid worm *Dendrostomum pyroides*. The pigment of this species was shown to have a molecular weight of approximately 100,000 and to consist of eight subunits. Amino acid analysis, peptide mapping, and immunological studies indicated that *D. pyroides* hemerythrin was very similar in structure to the hemerythrin from another sipunculid worm, *Golfingia gouldii*. The amino acid sequence of the *Golfingia* pigment has been reported by Klotz and his coworkers (Groskopf *et al.*, 1966a,b; Subramanian *et al.*, 1968; Klippenstein *et al.*, 1968).

This paper is concerned with the N-terminal sequence and with the composition and sequence of typtic peptides of *D. pyroides* hemerythrin and compares the structure of this protein to that from *G. gouldii*.

Materials and Methods

D. pyroides were obtained from Pacific Bio-Marine Supply Co., Venice, Calif. *G. gouldii* were from the Marine Biological Laboratory, Woods Hole, Mass. Hemerythrin was prepared as described previously (Ferrell and Kitto, 1970) and converted into the apoprotein by the procedure of Groskopf *et al.* (1966a). Chemicals for the automated sequence analysis were purchased as a kit from Beckman Instruments, Palo Alto, Calif.

Amino Acid Analysis. Amino acid analyses were performed according to the standard methods of Moore *et al.* (1958), using a Beckman–Spinco automatic amino acid analyzer. Iron-free hemerythrin, prepared according to Groskopf *et al.* (1966a), was hydrolyzed at 110° in 6 N HCl in sealed, evacuated tubes for 24, 48, and 72 hr. Serine and threonine were extrapolated to zero time of hydrolysis to account for their destruction. The highest values of valine and isoleucine were taken to represent complete liberation of these residues. Peptides were hydrolyzed under the same conditions for 24 hr. A mixture of β-2-thienylalanine and L-α-amino-β-guanidinopropionic acid hydrochloride at a concentration of 1 μmole/ml of each was used as an internal standard (Siegel and Roach, 1961; Walsh and Brown, 1962).

Enzymatic Digestion. Tryptic digests of iron-free hemerythrin were prepared by dissolving the apoprotein in deionized water at a concentration of 1 mg/ml and heated on a boiling-water bath for 10 min to denature the protein. This solution was allowed to cool to room temperature and made 0.5% in ammonium bicarbonate. TPCK[1]-treated trypsin (Worthington Biochemical Corp.) was added to give an enzyme to substrate ratio of about 5:100, and digestion allowed to proceed for 24 hr at 37°. The reaction was terminated by adjusting the pH of the reaction mixture to 2.0 by the dropwise addition of 2 N HCl. The resulting solution was reduced to a convenient volume on a rotary evaporator under reduced pressure and stored at −10° until used. Tryptic digestion was also carried out, in essentially the same manner, on hemerythrin in which the amino groups of lysine residues were trifluoroacetylated by the method of Goldberger and Anfinsen (1962).

High-Voltage Electrophoresis. All electrophoretic separations employed a modified Michl apparatus (Michl, 1951). Electrophoresis was carried out under three conditions of pH: pH 1.9 (glacial acetic acid–formic acid–water at 8:2:90, v/v), pH 3.5 (pyridine–glacial acetic acid–water at 1:10:89, v/v), and pH 6.0 (glacial acetic acid–pyridine–water at 1:10:90, v/v). The pH 1.9 and 3.5 electrophoresis tanks contained Savant mineral spirits (paraffin oil) as a coolant and the pH 6.0 tank contained xylene (Humble Oil and Refining Co).

Paper electrophoresis was first carried out at pH 6.0 following the procedures of Brown and Hartley (1966). Analytical separations utilized Whatman No. 1 paper, and preparative runs utilized Whatman No. 3MM paper. Samples were applied to the middle of the paper, in a narrow band, and the

* From the Clayton Foundation Biochemical Institute, Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. *Received January 28, 1971.* This work was supported in part by grants from the Clayton Foundation for Research in Biochemistry and from the National Institutes of Health (AI 07853 and HD 02889). A preliminary report of some of this work has already been published (Ferrell and Kitto, 1971).

† Present address: Department of Human Genetics, The University of Michigan Medical School, Ann Arbor, Mich. 48104.
‡ To whom to address correspondence.

[1] Abbreviations used are: TPCK, L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone; PTH, phenylthiohydantoin.

peptide band was narrowed as described by Brown and Hartley (1966). As an internal standard, 0.005 ml of a fluorescent marker containing a mixture of 1-(dimethylamino)naphthalene-5-sulfonylarginine and 1-(dimethylamino)naphthalene-5-sulfonic acid was added to the samples before application. A standard amino acid mixture containing lysine hydrochloride, histidine hydrochloride, glycine, alanine, serine, threonine, glutamic acid, aspartic acid, and cysteic acid monohydrate at a concentration of 10 μmoles of each amino acid per ml of H_2O, was spotted on either side of the sample band. Electrophoresis was carried out at 3 kV at a gradient of 60 V/cm for 20–45 min. The electrophoretogram was dried in a fume hood and examined under an ultraviolet lamp to detect fluorescent bands, which were outlined in pencil. A guide strip was cut from the edge of the dried sheet and developed with a cadmium–ninhydrin reagent described by Dryer and Bynum (1967).

For electrophoresis in the second dimension, a strip about 1.5 cm in width was cut out of the pH 6.0 electrophoretogram, and stitched about 12 cm from the bottom edge of a sheet of Whatman No. 3MM paper. The original paper behind the strip was cut out. Appropriate markers were applied, and electrophoresis carried out at pH 1.9 at 3 kV for 40 min or at pH 3.5 at 3 kV for 60 min. For peptide maps the full sheet was dipped in the cadmium–ninhydrin reagent, and exposed to the air at room temperature for 40 min. For preparative work, a guide strip was cut from the electrophoretogram, stained with cadmium–ninhydrin, and developed at room temperature. The sections of the electrophoretogram which contained peptides were then cut from the paper, and the peptides were eluted from the paper with glass-distilled water.

Preparative Scale Isolation of Tryptic Peptides. The tryptic peptides of the hemerythrin of *D. pyroides* were prepared from 50 μmoles of native hemerythrin in the following manner. The native protein was dissolved in 50 ml of deionized water and treated with a 10-fold molar excess of ethyl thioltrifluoroacetate as described by Goldberger and Anfinsen (1962). The protein was then precipitated with absolute ethanol (−20°), and the precipitate dialyzed extensively against deionized water. The protein was then subjected to tryptic digestion as described earlier. The tryptic digest was deblocked with piperidine (Goldberger and Anfinsen, 1962) and subjected to gel filtration on a 2.5 × 80 cm column of Sephadex G-25. High-voltage electrophoresis of a small sample of the resulting digest indicated that the trifluoroacetylation reaction had resulted in the blockage of few of the free lysine side chains. Also the piperidine, added as a deblocking reagent, did not separate from the peptide mixture on Sephadex G-25. The sample was placed on a rotary evaporator and reduced to a viscous oil under high vacuum. The pH of the resulting solution was adjusted to 2.0 and the mixture was subjected to ion-exchange chromatography on a 1 × 18 cm, jacketed, column containing Aminex A-4 resin (Bio-Rad). The column was prepared for sample application by regeneration with 2 N NaOH and equilibrated with 5% formic acid until the pH of the column effluent reached 2. The sample was pumped on the column at a flow rate of 18 ml/hr, and eluted at the same rate, as described below, with a Milton Roy Minipump (Milton Roy Co.). The effluent was collected in a fraction collector set at 20 min/tube. After the sample was applied to the column, it was followed by 25 ml of 5% formic acid (pH 2.0). The sample was then eluted with a gradient from a gradient maker consisting of four compartments linearly connected and stirred constantly. The gradient consisted of 100 ml of pH 2.5 buffer (8.06 ml of pyridine–600 ml of glacial acetic acid–1392 ml of water) in compartment one, 200 ml of pH 3.75 buffer (80.6 ml of pyridine–400 ml of glacial acetic acid–1520 ml of water) in compartments two and three, and 100 ml of pH 5.0 buffer (322.5 ml of pyridine–286.5 ml of glacial acetic acid–1391 ml of water). The column was then stripped using 500 ml of 1 N NH_4OH. The temperature was maintained at 50° throughout the run.

The column was monitored by spotting 0.02 ml from each tube on Whatman No. 1 paper at intervals of 1 cm and carrying out electrophoresis at pH 6.0 for 25 min. Amino acid and fluorescent markers were applied on each edge of the paper. The electrophoretogram was dried, examined under an ultraviolet lamp for fluorescence, and then stained with cadmium–ninhydrin reagent. The tubes were pooled according to the location of peptides, and the pooled fractions were then dried in a vacuum desiccator. The presence of piperidine, probably as piperidine acetate, interfered with the ion-exchange process and a large number of peptides were eluted from the column at the void volume. The material from the void volume was subjected to gel filtration on Sephadex G-50 in 0.1 N NH_4OH and purified by paper chromatography.

The peptides were further purified by high-voltage electrophoresis on paper, as described earlier, and the purified peptides stored at −10° until used.

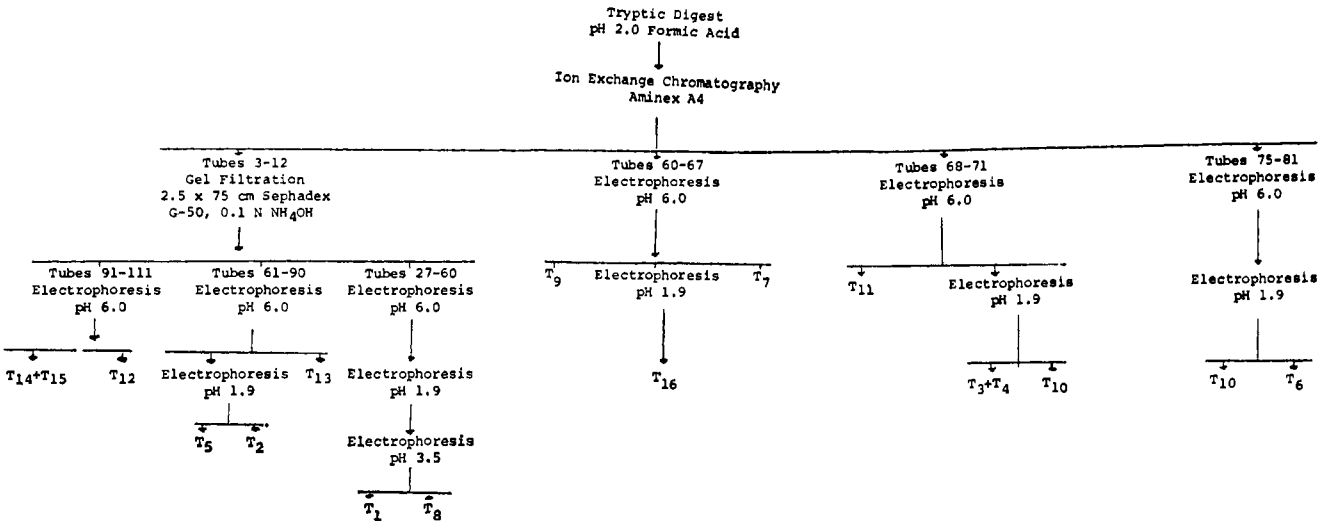
Automatic Edman Degradation. Samples of *D. pyroides* apohemerythrin were examined on a Beckman–Spinco protein–peptide sequencer, Model 890. The thiazolinone derivative of each amino acid was converted manually into the corresponding PTH derivative with 1 N HCl for 10 min at 80° (Edman and Begg, 1967). All PTH derivatives were analyzed by gas–liquid chromatography according to the procedure described by Pisano and Bronzert (1969). Those residues which could not be analyzed as the PTH derivative were silylated and analyzed by gas–liquid chromatography of the more volative trimethylsilyl derivatives (Pisano and Bronzert, 1969). Analysis was performed with a Beckman GC-45 chromatograph using a two-column system, with 10% DC-560 and OV-225.

All phenylthiohydantoin derivatives were examined by thin-layer chromatography in solvent system 1, and those which could not be unambiguously identified were then chromatographed in solvent system 2. PTH-glycine was confirmed by spraying the dry thin-layer plate with water and exposing it to concentrated ammonia vapors. A dark red spot is a positive test. Precoated silica gel plates (20 × 20 cm, Brinkmann No. F254) were used for thin-layer chromatography. The PTH derivatives of Thr, Asn, Asp, Glu, Gln, His, and Ser could be unambiguously identified with a solvent system containing chloroform, isopropyl alcohol, xylene, and propionic acid in the proportions 30:5:2:1. The PTH derivatives of Thr, Lys, Val, Pro, Gly, Ala, Trp, and Tyr were identified with a solvent system containing chloroform–formic acid (100:5, v/v) (Brenner *et al.*, 1961).

Subtractive Edman Degradation. Subtractive degradation was carried out by a modification of the Edman procedure as described by Gray (1967), with the omission of the dansylation step. Sufficient peptide was used so that 0.02–0.05 μmole of peptide could be hydrolyzed at each step. The sample was placed in a 10 × 75 mm test tube and dried in a vacuum desiccator over NaOH pellets and concentrated sulfuric acid. The dried peptide was dissolved in 200 μl of 50% aqueous pyridine (v/v), and 100 μl of 5% phenyl isothiocyanate was added. Sequencer grade triethylamine (20 μl) was added to each tube to increase the buffering capacity of the reaction mixture.

The tube was flushed with N_2 for three minutes, tightly

CHART I: PURIFICATION OF TRYPTIC PEPTIDES.



closed, and incubated at 45° for 1 hr. The reaction mixture was lyophilized to dryness, and 200 μl of anhydrous trifluoroacetic acid was added. The tube was again flushed with N_2, tightly closed, and incubated at 45° for 30 min. Excess trifluoroacetic acid was removed on a vacuum desiccator. Deionized water (150 μl) was added and the peptide was extracted three times with 1.5 ml of *N*-butyl acetate. The two phases were mixed thoroughly on a Vortex mixer and then separated by centrifugation. The *N*-butyl acetate layer (top) was removed with a Pasteur pipet and discarded. The aqueous phase, containing the peptide, was dried under vacuum, and the peptide was redissolved in 200 μl of 50% aqueous pyridine. An aliquot was removed for amino acid analysis, and the volume was restored to 200 μl with aqueous pyridine for the next cycle. In cases where the peptide contained four residues or less, the extraction step was sometimes omitted in order to reduce the amount of loss by extraction. If the peptide contained histidine, this was not possible since the nonextracted material gave a peak indistinguishable from histidine on the amino acid analyzer. When histidine was present in smaller peptides they were sequenced in the same manner as the larger peptides.

Assignment of Amides. The mobility of the small peptides on high-voltage electrophoresis at pH 6.0 was used to detect the presence of amide groups. Those amides positioned in the first 34 residues from the amino terminus were confirmed directly by analysis of the PTH derivative from the sequencer.

Results

N-Terminal Sequence. The sequence of the N-terminal region of the hemerythrin of *D. pyroides* was determined by Edman degradation using the Beckman–Spinco protein–peptide sequencer. The sequence was determined using 225 nmoles of iron-free hemerythrin. The results obtained are shown in Table I.

Tryptic Peptides. The purification of the tryptic peptides is shown schematically in Chart I. The composition of the tryptic peptides is given in Table II. For convenience in comparing the tryptic peptides of *D. pyroides* to those of *G. gouldii*, the amino acid composition of the tryptic peptides was compared to the known composition of the tryptic peptides from the hemerythrin of *G. gouldii* (Klippenstein *et al.*, 1968) and were assigned the number of that peptide to which they appeared most similar. A peptide with an amino acid composition

TABLE I: Automated Edman Degradation of *D. pyroides* Hemerythrin.

Cycle No.	Identification Glc[a]	Identification Tlc[b]	% Yield	Residue Assigned
1	Gly	Gly	31	Gly
2	Phe	Phe	96	Phe
3	Pro	Pro	53	Pro
4	Ile	Ile	58	Ile
5	Pro	Pro	50	Pro
6	N.d.[c]	Asp	N.d.	Asp
7	Pro	Pro	30	Pro
8	Tyr	Tyr	64	Tyr
9	Gly	Gly	28	Gly
10	Trp	Trp	25	Trp
11	Asp	N.d.	N.d.	Asp
12	Pro	Pro	60	Pro
13	Ser	Ser	N.d.	Ser
14	Phe	Phe	72	Phe
15	N.d.	Arg	N.d.	Arg
16	Thr	Thr	N.d.	Thr
17	Phe	Phe	33	Phe
18	Tyr	Tyr	30	Tyr
19	Ser	Ser	N.d.	Ser
20	Ile	Ile	21	Ile
21	Ile	Ile	41	Ile
22	N.d.	Asp	N.d.	Asp
23	N.d.	Asp	N.d.	Asp
24	N.d.	Glu	N.d.	Glu
25	N.d.	His	N.d.	His
26	N.d.	Lys	N.d.	Lys
27	Thr	Thr	N.d.	Thr
28	Leu	Leu	49	Leu
29	Phe	Phe	63	Phe
30	N.d.	Asn	N.d.	Asn
31	Gly	Gly	23	Gly
32	Ile	Ile	19	Ile
33	Phe	Phe	11	Phe
34	N.d.	His	N.d.	His

[a] Gas–liquid chromatography (glc). [b] Thin-layer chromatography (tlc). [c] Not determined (N.d.).

TABLE II: Amino Acid Composition of Tryptic Peptides from *D. pyroides* Hemerythrin.

Amino Acid	T_1	T_2	$T_3 + T_4$	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{10}'	T_{11}	T_{12}	T_{13}	$T_{14} + T_{15}$	T_{16}
Lysine		1 (1)				1 (1)		2 (2)		1 (1)	1 (1)	1 (1)	1 (1)	2 (2)	1 (1)
Histidine		0.96 (1)		0.72 (1)			0.91 (1)	1 (1)	1.45 (1)				0.84 (1)		
Arginine	1.0 (1)			1 (1)	1 (1)										
Aspartic acid	2.3 (2)	2.07 (2)	1.30 (1)	4.80 (5)			1.12 (1)	1.1 (1)	1 (1)	2.27 (2)	1.21 (1)		1.20 (1)	1 (1)	
Threonine		1.02 (1)	0.97 (1)			0.81 (1)								1.08 (1)	
Serine	0.86 (1)	0.92 (1)					0.87 (1)		2.29 (2)				0.93 (1)		
Glutamic acid		1.32 (1)		0.96 (1)			3.79 (4)	2.1 (2)	2.29 (2)						
Proline	3.6 (4)														
Glycine	1.2 (1)		1.4 (1)	1.12 (1)		0.91 (1)					1.32 (1)				1.05 (1)
Alanine				2.44 (2)			1.0 (1)		1 (1)	0.97 (1)		1.04 (1)			
Cysteine						1.10 (1)									
Valine	0.78 (1)						0.86 (1)				1.0 (1)		0.86 (1)		
Methionine							1 (1)								
Isoleucine	1.12 (1)	2.22 (2)	1.07 (1)	1.16 (1)					0.88 (1)				1.23 (1)	1.24 (1)	1 (1)
Leucine			1.13 (1)	3 (3)			2.3 (2)			1 (1)			1.07 (1)		
Tyrosine	0.96 (1)	1.02 (1)					1.2 (1)	1.0 (1)						1.08 (1)	
Phenylalanine	1.89 (2)	1.23 (1)	2 (2)				1.31 (1)	0.87 (1)	1.08 (1)					1.40 (1)	
Tryptophan	(1)									(1)		(1)	(1)		
Number of residues	15	11	7	15	1	4	14	8	7	6	4	3	8	7	3
Position in sequence	1–15	16–26	27–33	34–48	49	50–53	54–67	68–75	76–82	83–88	89–92	93–95	96–103	104–110	110–113

corresponding to that of T_3 plus T_4 and one of peptides T_{14} plus T_{15} were isolated and are referred to as $T_3 + T_4$ and $T_{14} + T_{15}$, respectively, for ease of comparison to the hemerythrin of *Golfingia* (Klippenstein *et al.*, 1968).

Peptides T_1–T_4. The amino acid sequence of the tryptic peptides was determined using the automatic sequencer and only their composition was determined from the isolated peptides. These peptides occupy positions 1–33 from the amino-terminal end of the protein. Peptides T_2, T_3, and T_4 of *Dendrostomum* hemerythrin were identical in sequence with the corresponding peptide from *Golfingia*. The differences found in peptide T_1 are discussed below.

Peptide T_5. This peptide was shown to have the amino acid composition (Arg,His,Aspl,Glu,Gly,Ala^2,Leu^3,Ile). Five cycles of Edman degradation were carried out on this peptide, and the sequence of the remaining residues is assumed from its similarity to the peptide T_5 from *G. gouldii*. This sequence is (His-Leu-Ala-Ile-Asp-Asp-Asn-Ala-Asp-Asn-Leu-Gly-Glu-Leu-Arg).

Peptide T_6. Peptide T_6 was found to contain only free arginine and was placed in the sequence according to its similarity to peptide T_6 of *Golfingia*.

Peptide T_7. Peptide T_7 was found to have the amino acid composition (Lys,Thr,Gly,half-Cys). Three cycles of Edman degradation gave the sequence (Cys-Thr-Gly-Lys), which is identical with the peptide T_7 of *Golfingia*.

Peptide T_8. The amino acid composition of peptide T_8 was found to be (His,Asp,Glu^4,Ser,Ala,Val,Leu^2,Tyr,Phe,Met). Five steps of Edman degradation were carried out on this peptide, and the proposed sequence for this peptide is (His-Phe-Leu-Asn-Gln-Glu-Val-Leu-Met-Glu-Ala-Ser-Gln-Tyr).

Peptide T_9. Peptide T_9 was found to have the amino acid composition (Lys^2,His,Asp,Glu^2,Tyr,Phe). Four steps of Edman degradation were carried out on this peptide. The proposed sequence is (Gln-Phe-Tyr-Asp-Glu-His-Lys-Lys). After four steps the yields were too low to determine the remainder of the sequence by Edman degradation.

Peptide T_{10}. Amino acid analysis of peptide T_{10} gave the following composition (Arg,His,Glu^2,Ala,Ile,Phe). Five cycles of Edman degradation established the following sequence for peptide T_{10} (Ala-His-Glu-Glu-Phe-Ile-Arg). Although this peptide did not appear obviously homologous to peptide T_{10} of *G. gouldii*, it is placed in this position for reasons to be discussed later.

Peptide $T_{10'}$. Peptide $T_{10'}$ was isolated from peptide maps of *D. pyroides* hemerythrin after staining with trinitrobenzenesulfonic acid. The quantities of peptide isolated were not sufficient to permit sequence analysis by the method of Edman. The amino acid composition of this peptide was found to be (Lys,Asp^2,Ala,Leu,Trp). A value of one tryptophan residue was assigned to this peptide because of its fluorescence under ultraviolet light. The amino acid sequence of this peptide is proposed to be (Ala-Leu-Asp-Asn-Trp-Lys) because of its similarity to such a sequence in peptide T_{10} of *G. gouldii*.

Peptide T_{11}. Peptide T_{11} was found to have the amino acid composition (Lys,Asp,Gly,Val). Edman degradation of this peptide established the sequence as (Gly-Asp-Val-Lys).

Peptide T_{12}. Peptide T_{12} was shown to be fluorescent when exposed to ultraviolet light and to contain the amino acids alanine and lysine. One cycle of Edman degradation resulted in the loss of fluorescence and no change in amino acid composition. Alanine was removed on the second Edman cycle and the sequence (Trp-Ala-Lys) is proposed for this peptide.

Peptide T_{13}. Peptide T_{13} was found to have the amino acid composition (Asp,His,Ile,Leu,Lys,Ser,Trp,Val) and was car-

TABLE III: Comparison of Residues 76–88 of *G. gouldii* and *D. pyroides* Hemerythrins.

	76	77	78	79	80	81	82	83	84	85	86	87	88
G. gouldii	Glu	His	Glu	Gly	Phe	Ile	His	Ala	Leu	Asp	Asn	Trp	Lys
	Thr												
	←———						T_{10}						———→
D. pyroides	Ala	His	Glu	Glu	Phe	Ile	Arg	Ala,Leu,Asp,Asn,Trp,Lys					
	←———				T_{10}		———→	←———			$T_{10'}$		———→

ried through four steps of Edman degradation. The sequence of the first four residues corresponds to the sequence of these residues in peptide T_{13} of *G. gouldii* and the sequence (Ser-Trp-Leu-Val-Asn-His-Ile-Lys) was assigned to this peptide.

Peptide T_{14} + T_{15}. The peptide T_{14} + T_{15} was found to have the amino acid composition of peptides T_{14} plus T_{15} of the *Golfingia* hemerythrin. This composition was (Lys^2,Asp,Thr,-Ile,Tyr,Phe). Four cycles of Edman degradation were carried out and the sequence (Thr-Ile-Asp-Phe-Lys-Tyr-Lys) was assumed for this peptide. Although it is obvious that efforts to limit cleavage by trifluoroacetylation of the ε-amino groups of lysine were not very successful, it is probable that the failure of trypsin to cleave at lysine in position 108 is due to blockage of this lysine residue.

Peptide T_{16}. This peptide was found to have the composition (Lys,Gly,Ile). Edman degradation was carried out on this peptide and the sequence (Gly-Lys-Ile) was established.

With this exceptions of peptides T_1, T_{10}, and $T_{10'}$, all of the tryptic peptides isolated from *D. pyroides* hemerythrin were found to have the identical amino acid compositions as individual tryptic peptides reported for the hemerythrin of *G. gouldii* (Klippenstein *et al.*, 1968). A comparison of the sequence of these peptides of identical amino acid composition shows that all sequences determined for the tryptic peptides of *Dendrostomum* are identical with those sequences in the *Golfingia* peptides.

Discussion

Analysis of the first 34 residues from the N terminus of *D. pyroides* hemerythrin using the automated sequencer gave a sequence which was identical with the sequence of that region in *G. gouldii* hemerythrin (Klippenstein *et al.*, 1968), except for residues 9–11. We have discussed elsewhere (Ferrell and Kitto, 1971) the evidence which suggests that residues 10 and 11 are tryptophan and aspartic acid, respectively, in both *D. pyroides* and *G. golfingia* hemerythrin, rather than the reverse, as reported by Klippenstein *et al.* (1968) for the *Golfingia* pigment. Our studies indicate that the only difference in sequence between the first 34 N-terminal residues of *Dendrostomum* and *Golfingia* hemerythrins occurs at residue 9 which is glycine in *Dendrostomum* hemerythrin and valine in the *Golfingia* protein.

Analysis of the tryptic peptides of *D. pyroides* hemerythrin again emphasized the similarity between the two Sipunculid hemerythrins which was noted in our earlier studies (Ferrell and Kitto, 1970). An examination of the *D. pyroides* hemerythrin peptide map had shown two ninhydrin-positive spots not present in a *Golfingia* hemerythrin digest while one prominent *Golfingia* peptide was absent from the *D. pyroides* hemerythrin map. All other peptides occupied identical positions on the peptide maps.

Other than those peptides responsible for the differences in the peptide maps and the single difference in peptide T_1 noted above, all the *Dendrostomum* hemerythrin tryptic peptides had amino acid compositions identical with those found with *Golfingia* hemerythrin. In addition, the sequences determined for these *Dendrostomum* peptides of identical composition were the same as those reported for the *Golfingia* peptides (Klippenstein *et al.*, 1968).

The peptides of *Dendrostomum* hemerythrin which differed from the corresponding *Golfingia* peptides were obtained as follows. An arginine-containing peptide (T_{10}) with an amino acid composition similar to part of peptide T_{10} of *G. gouldii* was isolated from the original tryptic digest (Chart I) and subjected to Edman degradation (Table III). No tryptic peptide was isolated in the original tryptic digest which had an amino acid composition similar to the remainder of peptide T_{10} of *Golfingia* hemerythrin.

In an attempt to clarify this point, tryptic digests of iron-free hemerythrin of *G. gouldii* and *D. pyroides* were prepared and peptide maps were prepared as described previously (Ferrell and Kitto, 1970). These maps were examined under ultraviolet light for fluorescence and then stained with 1% trinitrobenzenesulfonic acid in acetone. The chromatograms were then developed in a pyridine–water atmosphere and examined for differences. Differences corresponding exactly to those described earlier were found (Ferrell and Kitto, 1970). Those spots representing the differences were cut from the peptide maps and eluted with distilled water, hydrolyzed, and their amino acid compositions determined. The single spot in *G. gouldii* hemerythrin but absent in *D. pyroides* hemerythrin was found to have an amino acid composition identical with that of peptide T_{10} from the reported sequence of *Golfingia* hemerythrin. The nonfluorescent spot present in *Dendrostomum* hemerythrin and absent in *Golfingia* hemerythrin had an amino acid composition different from the peptide T_{10} of *Golfingia* but identical with the peptide from the original tryptic digest of *Dendrostomum*, labeled T_{10}, and assumed from amino acid sequence to be related to residues 76–82 of *Golfingia* hemerythrin. The fluorescent spot isolated from the *Dendrostomum* map and absent in the *Golfingia* map was found to have the amino acid composition (Lys,Asp^2,Ala,Leu), and the fluorescence of this peptide under ultraviolet light indicates the presence of a tryptophan residue. This peptide ($T_{10'}$) of *Dendrostomum* hemerythrin has an amino acid composition identical with residues 83–88 of *Golfingia* hemerythrin. From this information it appeared that a difference existed between peptide T_{10} of *Golfingia* and the similar region of the *Dendrostomum* sequence which provided an additional cleavage point for trypsin. This difference is the presence of arginine in position 82 of the sequence of *Dendrostomum* while position 82 of the *Golfingia* sequence is occupied by histidine. Two other differences were found between peptide T_{10} of *Dendrostomum* and the corresponding portion of peptide T_{10} of *Golfingia*. Position 76 in the sequence of *Golfingia* hemerythrin

TABLE IV: A Comparison of the Tryptic Peptides of *D. pyroides* Hemerythrin to the Sequence of *G. gouldii* Hemerythrin.

	5 15
Dendrostomum[a]	Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr-Gly-Trp-Asp-Pro-Ser-Phe-Arg-Thr-
	T_1
Golfingia[b]	Gly-Phe-Pro-Ile-Pro-Asp-Pro-Tyr-Val-Trp-Asp-Pro-Ser-Phe-Arg-Thr-
	20 25 30
Dendrostomum	Phe-Tyr-Ser-Ile-Ile-Asp-Asp-Glu-His-Lys-Thr-Leu-Phe-Asn-Gly-Ile-
	T_2 T_4
Golfingia	Phe-Tyr-Ser-Ile-Ile-Asp-Asp-Glu-His-Lys-Thr-Leu-Phe-Asn-Gly-Ile-
	35 40 45
Dendrostomum	Phe-His-Leu-Ala-Ile-Asp[Asp,Asn,Ala,Asp,Asn,Leu,Gly,Glu,Leu,Arg]
	T_5
Golfingia	Phe-His-Leu-Ala-Ile-Asp-Asp-Asn-Ala-Asp-Asn-Leu-Gly-Glu-Leu-Arg-
	50 60
Dendrostomum	Arg Cys-Thr-Gly-Lys His-Phe-Leu-Asn-Gln[Glu,Val,Leu,Met,Gln,Ala,
	T_6 T_7 T_8
Golfingia	Arg-Cys-Thr-Gly-Lys-His-Phe-Leu-Asn-Gln-Glu-Val-Leu-Met-Gln-Ala-
	65 70 75 80
Dendrostomum	Ser,Gln,Tyr] Gln-Phe-Tyr[Asp,Glu,His,Lys,Lys]Ala-His-Glu-Glu-Phe-
	T_9 T_{10}
Golfingia	Ser-Gln-Tyr-Gln-Phe-Tyr-Asp-Glu-His-Lys-Lys-Glu-His-Glu-Gly-Phe-
	85 90 95
Dendrostomum	Ile-Arg[Ala,Leu,Asp,Asn,Trp,Lys]Gly-Asp-Val-Lys Trp-Ala-Lys Ser-
	$T_{10'}$ T_{11} T_{12}
Golfingia	Ile-His-Ala-Leu-Asp-Asn-Trp-Lys-Gly-Asp-Val-Lys-Trp-Ala-Lys-Ser-
	T_{10}
	100 105 110
Dendrostomum	Trp-Leu-Val[Asn,His,Ile,Lys]Thr-Ile-Asp-Asp-Phe[Lys,Tyr,Lys]
	T_{13} T_{14} T_{15}
Golfingia	Trp-Leu-Val-Asn-His-Ile-Lys-Thr-Ile-Asp-Phe-Lys-Tyr-Lys-
Dendrostomum	Gly-Lys-Ile
	T_{16}
Golfingia	Gly-Lys-Ile

[a] → represents the results of direct sequence analysis. Brackets enclose those residues which were assigned by homology Refer to the text for additional nomenclature. [b] From Klippenstein *et al.* (1968).

is occupied by glutamic acid while in *Dendrostomum* this residue was found to be alanine. In the *Golfingia* sequence either glycine or threonine was found in position 79. Unique peptides with either glycine or threonine in this position were reported by Klippenstein *et al.* (1968). In the *Dendrostomum* hemerythrin, glutamic acid was found to occupy position 79 of the proposed sequence. As noted previously (Ferrell and Kitto, 1970) no variants of *Dendrostomum* hemerythrin have been detected. These findings are summarized in Table III and a comparison of the total sequence of *Golfingia* hemerythrin to the present information obtained about *Dendrostomum* hemerythrin is given in Table IV.

An examination of the accepted mRNA codons for the amino acids given by Dayhoff (1969) reveals that the differences observed between the sequence of the hemerythrin of *Dendrostomum* and that of *Golfingia* can each occur by a single mutation if glycine is assumed to occupy position 79 of the *Golfingia* sequence. The change involved if threonine occupies this position would require two mutations. The change from glycine to threonine, which has apparently occurred in heme-

rythrins isolated from populations of the same species in *Golfingia* requires two base changes.

In presenting these data it is recognized that the amino acid sequence of the hemerythrin of *D. pyroides* has not been unequivocally determined. This would require the complete sequencing of all tryptic peptides as well as the determination of overlaps using an enzyme of different specificity. With the larger peptides, the assignment of amides is open to question because the charged state of larger peptides is not always obviously reflected by their electrophoretic mobility. In spite of these reservations some discussion of the structure is possible in a comparative sense.

Fan and York (1969), using the hemerythrin of *G. gouldii*, have reported the reaction of hemerythrin with 5-diazo-1*H*-tetrazole in which they observed the formation of 7 moles of a bis(diazo-1*H*-tetrazole) derivative of histidine with the iron-free hemerythrin and the formation of only 3 moles of the typical bis(diazo-1*H*-tetrazole) derivative with native hemerythrin. This led them to postulate that histidine residues might be important in the binding of iron to the native hemerythrin. The present work indicates that the histidine in position 82 of the *Golfingia* sequence is replaced by an arginine in the *Dendrostomum* hemerythrin. This indicates that histidine-82 is not one of those believed to be involved in iron binding.

Any discussion of the function or evolution of hemerythrin in terms of comparative sequence depends, of course, on the availability of sequence data on hemerythrins from other species. Dr. Bolling Sullivan (Duke University, personal communication) has indicated that the sequence of the hemerythrin from the brachipod *Lingula* is quite different from that of *Golfingia*. The sequence of the hemerythrin from *Lingula*, a phylogenetically distant organism, should provide useful information regarding the rate of evolutionary change in this pigment.

If one accepts the amino acid sequence proposed for *Dendrostomum* hemerythrin then there are four amino acid sequence differences between this protein and the hemerythrin from *G. gouldii*. Comparison of sequence data on cytochrome *c* and hemoglobin from various organisms has shown that differences in amino acid sequence of a given protein from various organisms are related to their taxonomic positions, as established by classical means (Dayhoff, 1969; Zuckerkandl and Pauling, 1962). It has been calculated for cytochrome *c*, a very slowly evolving protein, that the average rate of change is one amino acid substitution per 22.6 million years (Margoliash and Smith, 1965). For the hemoglobins, which seem to change at a faster rate, one substitution per 7 million years has been estimated (Buettner-Janusch and Hill, 1965). If we assume that that hemerythrins evolve at comparable rates then *Dendrostomum* and *Golfingia* probably separated from a common ancestor between approximately 30 and 50 million years ago. This is a fairly recent separation in terms of evolutionary time, since the fossil records of Sipunculids extend back some 500 million years.

Acknowledgments

The authors thank Dr. James R. Brown, Dr. Austen Riggs, Thomas O. Baldwin, James Lemberg, Mrs. Florence Waddill, and Vivian Sellers for their advice and assistance.

References

Brown, J. R., and Hartley, B. S. (1966), *Biochem. J. 101*, 214.

Buettner-Janusch, J., and Hill, R. L. (1965), *in* Evolving Genes and Proteins, Bryson, V., and Vogel, H., Ed., New York, N. Y., Academic Press, p 167.

Dayhoff, M. O. (1969), *Atlas Protein Sequence Structure 4*, 90.

Dreyer, W. J., and Bynum, E. (1967), *Methods Enzymol. 11*, 32.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.

Fan, C. C., and York, J. L. (1969), *Biochem. Biophys. Res. Commun. 36*, 365.

Ferrell, R. E., and Kitto, G. B. (1970), *Biochemistry 9*, 3053.

Ferrell R. E. and Kitto G. B. (1971), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 12*, 322.

Goldberger, R. F., and Anfinsen, C. B. (1962), *Biochemistry 1*, 401.

Gray, W. R. (1967), *Methods Enzymol. 11*, 469.

Groskopf, W. R., Holleman, J. W., Margoliash, E., and Klotz, I. M. (1966a), *Biochemistry 5*, 3779.

Groskopf, W. R., Holleman, J. W., Margoliash, E., and Klotz, I. M. (1966b), *Biochemistry 5*, 3873.

Klippenstein, G. L., Holleman, J. W., Klotz, I. M. (1968), *Biochemistry 7*, 3868.

Margoliash, E., and Smith, E. L. (1965), *in* Evolving Genes and Proteins, Bryson, V., and Vogel, H., Ed., New York, N. Y., Academic Press, p 221.

Michl, H. (1951), *Monatsh. Chem. 82*, 489.

Moore, S., Spackman, D. H., and Stein, W. H. (1958), *Anal. Chem. 30*, 1185.

Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244*, 5597.

Siegel, F. L., and Roach, M. K. (1961), *Anal. Chem. 33*, 1628.

Subramanian, A. R., Holleman, J. W., and Klotz, I. M. (1968), *Biochemistry 7*, 3859.

Walsh, K. A., and Brown, J. R. (1962), *Biochim. Biophys. Acta 58*, 596.

Zuckerkandl, E., and Pauling, L. (1962), *in* Horizons in Biochemistry, Kasha, M., and Pullman, B., Ed., New York, N. Y., Academic Press, p 189.