The Primary Structure of *Golfingia gouldii* Hemerythrin. Order of Peptides in Fragments Produced by Tryptic Digestion of Succinylated Hemerythrin. Complete Amino Acid Sequence*

G. L. Klippenstein,† J. W. Holleman,‡ and I. M. Klotz

ABSTRACT: Hemerythrin from pooled samples of blood of the sipunculid worm *Golfingia gouldii* was succinylated under denaturing conditions to block free amino groups and thus limit tryptic attack to the three arginyl residues of the monomer chain. Digestion with L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone treated trypsin produced four main fragments. These were separated by gel filtration. The fragments together accounted for the amino acid content of the protein. Some minor fragments were also present, arising from the action of residual chymotrypsin in the trypsin preparation or from chymotryptic-like action of the trypsin itself. Peptides obtained in previous work from tryptic digests of nonsuccinylated hemerythrin, the sequences of which had been partially determined (Groskopf, W., Holleman, J. W., Margoliash, E., and Klotz, I. M. (1966), *Biochemistry 5*, 3783), were assigned to the large fragments on the basis of composition. Overlaps between these peptides, to arrange them in unique linear order within the large fragments, were established by characterization of peptides obtained from the succinyl-derived fragments by further cleavage by enzymatic and chemical means. Overlaps between the sections themselves were established by isolation and characterization of the arginine-containing bridge peptides produced by digestion of succinylated hemerythrin with chymotrypsin. The results show the four fragments to be: (1) an N-terminal section comprising residues 1–15 of the hemerythrin monomer; (2) a middle section comprising residues 16–48; (3) free arginine, which is residue 49; and (4) a C-terminal section comprising residues 50–113 of the polypeptide chain.

With the results reported in the preceding paper (Subramanian, A. R., Holleman, J. W., and Klotz, I. M. (1968), *Biochemistry 7*, 3859 (this issue; preceding paper)), the primary structure of the hemerythrin subunit is established.

As one approach to obtaining a new family of fragments into which the tryptic peptides previously obtained (Groskopf *et al.*, 1966b) could be fitted, it was decided to use the specificity of trypsin again but in such a way as to produce only a few large fragments of the protein by restricting the action of the enzyme to the arginyl bonds. This can be done by blocking the ε-amino groups of the lysyl residues by, for instance, acylation. There are three arginyl residues per subunit (Groskopf *et al.*, 1966a), and it was known that these are in the interior of the chain. Four fragments (of unknown size) were, therefore, to be expected from the above type of digestion. It was thought that one of these fragments might be arginine, since free arginine had been obtained from the previous tryptic digests, implying an arginyl-arginine or a lysyl-arginine sequence. However, some of the peptides from the previous work were clearly the result of chymotryptic or chymotryptic-like cuts, and so this point was still in doubt.

It was expected that once the fragments were separated, end-group analyses would identify the ones corresponding to the amino and carboxyl termini of the hemerythrin subunit, which were known (Groskopf *et al.*, 1966a). Alignment of the remaining two sections would be effected by isolation and characterization of the arginine-containing bridge peptides from another type of enzymatic digest of the acylated protein. Meanwhile, comparison of the analyses of the tryptic peptides, of the fragments, and of peptides formed from them by further enzymatic and chemical treatments, along with a limited amount of sequence work, would allow one to place all the residues of the protein in unique linear order.

The main considerations in selection of a reagent for modification of the lysyl residues were solubility of the modified protein to allow good enzyme attack, and ease of removal of the modifying group following the enzyme digestion. Trifluoroacetylation would yield masked ε-amino groups which could be easily regenerated, but the solubility of this derivative of hemerythrin was not known. Succinylation would be expected to produce a soluble modified protein and soluble frag-

* From the Biochemistry Division, Department of Chemistry, Northwestern University, Evanston, Illinois. *Received March 13, 1968.* 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-00626) from the National Institute of General Medical Sciences.

† Present address: Department of Biochemistry, Spaulding Life Science Building, University of New Hampshire, Durham, N. H. 03824.

‡ Present address: Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tenn. 37830.

ments, but removal of the succinyl group would not be practicable. Both methods would give derivatives stable to normal manipulations of separation and analysis. Because of the possibility of removal of the modifying group, modification by trifluoroacetylation was first tried, but was discarded because of difficulties with the digestion of the relatively insoluble trifluoroacetyl derivative. Modification by succinylation was then chosen, and turned out to be the effective approach to obtaining the large fragments needed to place the tryptic peptides in linear sequence.

Experimental Section

Preparation of Chemically Modified Hemerythrins. Native and iron-free hemerythrin were prepared as described previously (Groskopf *et al.*, 1966a), the protein being isolated from the sipunculid worms by the method described in the preceding article (Subramanian *et al.*, 1968). For modification by trifluoroacetylation, 0.250 g (17.5 μmoles of subunit) of iron-free hemerythrin was treated with ethyl thiotrifluoroacetate (prepared as described by Hauptschein *et al.*, 1952) by the method of Goldberger and Anfinsen (1962), and kept as a lyophilized powder.

In first experiments with succinylation, the protein was treated directly with succinic anhydride in the presence of 4.0 M guanidinium chloride. However, due to loss of cysteine (as cystine) on acid hydrolysis this method was abandoned. In later experiments the cysteine was protected either by treatment with free cystine (disulfide exchange) in the presence of sodium azide, essentially as in the preceding article (Subramanian *et al.*, 1968), or by alkylation with iodoacetamide, in a manner patterned after the method of Crestfield *et al.* (1963) for treatment of proteins with iodoacetate. In a typical preparation, 168 mg (12.4 μmoles of subunit) of hemerythrin was dissolved in 12 ml of a 0.36 M Tris-HCl (pH 8.6) buffer which was 8.0 M in urea and which contained 1.25 mg of disodium EDTA/ml. An excess of mercaptoethanol was added, and the solution was allowed to stand for 3 hr at room temperature. At the end of this time, 275 mg of iodoacetamide (Pierce Chemical Co.) was added, reaction was allowed to proceed for 20 min, and the mixture was then immediately treated with succinic anhydride.

Succinylation was carried out in solutions which were 8.0 M in urea in the case of carboxyamidomethylated hemerythrin as above, or 4.0 M in guanidine hydrochloride in the case of the native and cystine-treated proteins. A large excess (50 moles/mole of lysine) of succinic anhydride was added in small portions to the rapidly stirred hemerythrin solutions. The pH was maintained at 7.5 by addition of 5.0 M NaOH. Excess reagents and by-products were then removed either by passage of the mixture over a 2.5 × 75 cm column of Sephadex G-25 (Pharmacia) equilibrated with 0.05 M NH_4HCO_3 or by dialysis against several changes of 0.05 M NH_4HCO_3. The protein was then lyophilized.

The now completely iron-free, succinylated protein was then taken up in buffer and succinylated again as above with a 50-fold molar excess of succinic anhydride but without the presence of guanidine hydrochloride or urea. In early experiments, some formation of succinyl ester bonds was noted (besides the formation of the desired succinylamide bonds), and so to break these the product of subsequent preparations was treated with hydroxylamine. The mixture was made 1.0 M in hydroxylamine by the addition of solid hydroxylamine hydrochloride, the pH was adjusted to 10.0 with 5.0 M NaOH, and the solution was allowed to stand at room temperature for 3 hr. Finally the twice-succinylated hemerythrin was again desalted by dialysis or Sephadex filtration and lyophilized.

Tryptic Digestion of Modified Hemerythrins and Separation of Peptides. The modified proteins were digested with trypsin (Worthington) which had been treated with TPCK[1] according to the procedure of Kostka and Carpenter (1964) to inactivate any chymotrypsin present (Schoellman and Shaw, 1963). The trifluoroacetylhemerythrin was difficult to digest because of its low solubility. In a typical digestion of this derivative, 36 mg of it was suspended in 8 ml of 0.1 M NH_4HCO_3 (pH 8.6) buffer and 2% enzyme by weight was added. Another 2% enzyme was added at 10 hr, and digestion was continued for a total of 21 hr, at room temperature. After lyophilization of the mixture, the product was taken up in 0.5 ml of 0.1 M NH_4HCO_3 and the soluble portion was applied to a 1.0 × 55 cm column of Sephadex G-50 (Pharmacia) equilibrated with the same solvent. The effluent was monitored by measurement of the absorption of the fractions at 280 nm (mμ) and by the alkaline hydrolysis–ninhydrin method of Hirs *et al.* (1956).

In one experiment the peptide mixture produced from such a digest was treated with 1.0 M piperidine according to the procedure of Goldberger and Anfinsen (1962) to remove the trifluoroacetyl groups from the lysines. The supernatant from this treatment (mixed with an equal volume of glacial acetic acid) was subjected to gel filtration on a 2.5 × 150 cm column of Sephadex G-50 with 50% acetic acid as the eluting solvent. The effluent was monitored by the alkaline hydrolysis–ninhydrin method as above.

In the case of succinylated hemerythrin, the substrate was completely soluble at the pH of the digestion and essentially complete digestion was effected in a relatively short time, especially with the iodoacetamide-treated derivative. A typical digestion was as follows. To 178 mg (12.1 μmoles of subunit) of CAM-succinylated hemerythrin in 30 ml of 0.2 M NH_4HCO_3 (pH 8.3) buffer was added 0.2 ml of enzyme solution containing 1.6 mg of enzyme. After 3 hr at 31° another 0.1 ml (0.8 mg) of this solution was added. When 5 hr had elapsed, the digest was lyophilized.

Given amounts of this and similar digests were dissolved in 4.0 ml of 0.2 M NH_4HCO_3 and applied to a 1.9 × 200 cm column of Sephadex G-50 equilibrated with the same solvent. The column was eluted at a

[1] Abbreviations used that are not listed in *Biochemistry* 5, 1445 (1966), are: TPCK, L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone; CAM, carboxyamidomethyl. See nomenclature section for peptide designations.

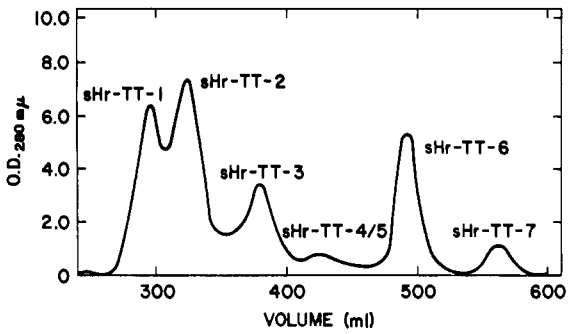

FIGURE 1: Elution pattern of peptides from tryptic digest of cystine-treated, succinylated hemerythrin (sHr) (see text for details).

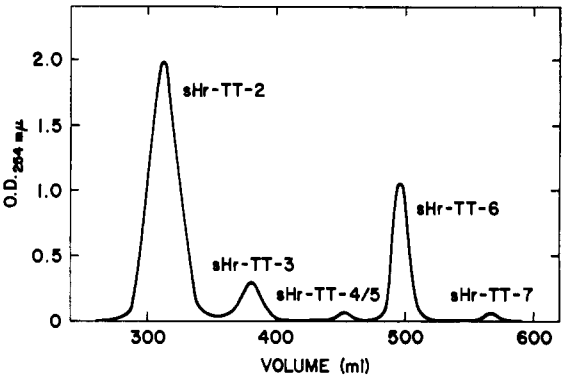

FIGURE 2: Elution pattern of peptides from tryptic digest of iodoacetamide-treated, succinylated hemerythrin (sHr) (see text for details).

flow rate of 15 ml/hr, and fractions of 4.0 ml were collected. The column effluent was monitored either by measurement of absorption of the fractions at 280 nm in a spectrophotometer or by continuous analysis of the flow stream at 254 nm using an Isco Model UA recording ultraviolet analyzer (Instrument Specialties Co.). Fractions were combined on the basis of the absorption pattern of the effluent, and lyophilized.

Chymotryptic Digestion of Succinylated Hemerythrin and Isolation of Arginine-Containing Peptides. In order to obtain sequences which would overlap the fragments expected from tryptic digestion of succinylated hemerythrin, chymotryptic digestion of the modified protein was carried out. CAM-succinylated hemerythrin (70 mg) was dissolved in 40 ml of 0.2 M NH_4HCO_3 (pH 8.3) and 0.7 mg of chymotrypsin (Worthington) was added. After 4 hr, an additional 0.7 mg of enzyme was added. After 7-hr digestion, the mixture was lyophilized. The arginine-containing peptides were separated by paper electrophoresis and paper chromatography.

Further Fragmentation of Peptides. PEPTIDE-SUCCINYLATED HEMERYTHRIN TT-2. The largest of the peptides arising from the action of TPCK-treated trypsin on succinylated hemerythrin was designated succinylated hemerythrin TT-2 (see Figures 1 and 2; section on Nomenclature). After analysis of its composition, further fragmentation of this peptide was effected by chymotryptic hydrolysis and by hydrolysis in dilute acid. The chymotryptic hydrolysis was carried out as follows. To 15 μmoles of the peptide in 40 ml of 0.2 M NH_4HCO_3 (pH 8.3) buffer was added 1.2 mg (0.9% by weight) of chymotrypsin (Worthington). The reaction was allowed to proceed for 2.5 hr at room temperature and was then terminated by lyophilization.

This digest was fractionated on a column of DEAE-Sephadex A-25 fine (Pharmacia) (0.9 × 92 cm) equilibrated with 0.1 M ammonium acetate (pH 7.9). The sample was applied in 3.0 ml of this buffer. The flow rate was 5.7 ml/hr. Fractions of 2.6 ml were collected. The column was eluted, first with 60 ml of starting buffer, then with a linear gradient of 500 ml to a 1.0 M ammonium acetate (pH 5.9) buffer, followed by 100 ml of this latter buffer and finally a linear gradient of 200 ml to a 2.0 M ammonium acetate (pH 5.9) buffer.

Dilute acid hydrolysis (Partridge and Davis, 1950) of succinylated hemerythrin TT-2 was carried out either in 0.03 M HCl at 110° (Grannis, 1960; Schultz *et al.*, 1962) or in 0.25 M acetic acid under reflux (Schroeder *et al.*, 1963). For the hydrolysis with HCl, 5.4 μmoles of the peptide was suspended in 62 ml of 0.03 M HCl in a 200-ml round-bottomed flask. The flask was evacuated on a vacuum pump and was then sealed off with a flame. Hydrolysis was allowed to proceed for 11 hr at 110°. The sample was then taken to dryness in a rotary evaporator, redissolved in a small amount of water, and lyophilized. For hydrolysis with acetic acid, 20 μmoles of succinylated hemerythrin TT-2 was suspended in 180 ml of 0.25 M acetic acid and the mixture was refluxed for 8.5 hr and then taken to dryness and lyophilized as above.

One part of the peptide mixture from hydrolysis with 0.03 M HCl was subjected to gel filtration on a 1.2 × 200 cm column of Sephadex G-25 with 0.2 M NH_4HCO_3 as eluent. Another part was chromatographed on CM-Sephadex. The CM-Sephadex C-25 fine (Pharmacia), equilibrated with sodium phosphate (pH 5.8) buffer of 0.1 ionic strength, was packed into a 0.9 × 96 cm column. The peptide mixture was applied to the column in 2.0 ml of the same buffer. The flow rate was 5.7 ml/hr, and fractions of 2.8 ml were collected. The column was developed with 100 ml of the starting buffer, followed by a linear gradient of 800 ml to a sodium phosphate buffer of the same ionic strength but of pH 6.8. Elution was continued with 175 ml of this buffer followed by 225 ml of a solution 0.1 M in NaOH and 0.5 M in NaCl. The fractions obtained were reduced in volume by rotary evaporation, desalted on a small column of Sephadex G-10, and lyophilized.

Peptides from the dilute acetic acid hydrolysis of succinylated hemerythrin TT-2 were subjected to gel filtration on a 1.9 × 206 cm column of G-25 Sephadex with 0.25 M acetic acid as the eluting solvent. A 4.0-ml sample was applied to the column and a flow rate of 15 ml/hr was maintained. Fractions of 5.3 ml were collected.

The main fraction from the above gel filtration, which contained several peptides, was further fractionated on a column of CM-Sephadex equilibrated with 0.1 M ammonium acetate (pH 7.9) buffer. The peptide mixture, dissolved in 2.0 ml of this buffer, was applied to a 0.8 × 45 cm column of the above ion-exchange

material, and elution was begun with a flow rate of 4.3 ml/hr. Fractions of 1.5 ml were collected. The column was developed with 45 ml of the starting buffer followed by a linear gradient to 0.5 M ammonium acetate (pH 9.4) over a volume of 120 ml; elution was continued with 30 ml of this latter buffer, followed by 30 ml of 1.0 M ammonium hydroxide.

The above separations were monitored either by measurement of absorption at 254 nm or by the alkaline hydrolysis–ninhydrin method.

Fragmentation of Succinylated Hemerythrin TT-3. This peptide was digested with chymotrypsin as follows. To 3.9 μmoles of the separated peptide in 10 ml of 0.2 M NH_4HCO_3 (pH 8.3) buffer was added 0.34 mg (2.3% by weight) of chymotrypsin and the digestion was allowed to proceed at room temperature for 3 hr. The digest was lyophilized and then dissolved in 3.0 ml of 0.2 M NH_4HCO_3 and applied to a 1.9 × 206 cm column of G-50 Sephadex with 0.2 M NH_4HCO_3 as eluting solvent. Fractions of 4.0 ml were collected at a flow rate of 18 ml/hr and the effluent was monitored by absorption of the fractions at 220 nm in a spectrophotometer.

Methods of Analysis and Characterization. Amino acid compositions were determined according to Spackman *et al.* (1958) using a Beckman Model 120B or 120C amino acid analyzer. Both instruments were equipped with high-sensitivity cuvet systems and with Beckman custom spherical resins for accelerated analysis.

Aliquots of peptides corresponding to 0.08–0.2 μmole usually were hydrolyzed with 3.0 ml of triply glass-distilled 5.7 M HCl plus 0.1 ml of a 0.05 M hydrazine solution to prevent destruction of tyrosine (Sanger and Thompson, 1963), in sealed, evacuated tubes for 24 hr at 107°. Hydrolysis of some of the larger peptides was carried out for three different lengths of time and the quantities of threonine and serine, which are progressively destroyed (Hirs *et al.*, 1954; Mahowald *et al.*, 1962), were estimated by extrapolation to zero time. Methionine is also destroyed and was estimated by a similar extrapolation. In some peptides in which it was present isoleucine was released slowly, and the value at the longest time was taken. For the remaining amino acids, the values at all three times were averaged. Tryptophan was determined in some cases spectrophotometrically by the method of Goodwin and Morton (1946); in others its presence was detected by the Ehrlich reaction.

Analytical and preparative paper electrophoresis and chromatography were done as described before (Groskopf *et al.*, 1966b). Gel electrophoresis (Smithies, 1955) was done using starch gel (Connaught) in the Shandon apparatus. The discontinuous buffer system of Ferguson and Wallace (1961) was used.

Digestions with carboxypeptidase A or B (Worthington) to determine carboxyl-terminal sequences were done as described previously (Groskopf *et al.*, 1966b). Aliquots were removed at various times, mixed with pH 2.2 sodium citrate buffer, and applied to the amino acid analyzer. Edman degradation to determine amino-terminal sequences was also done as described previously (Groskopf *et al.*, 1966b), by the subtractive method (Hirs *et al.*, 1960).

Nomenclature. The tryptic and chymotryptic peptides obtained in previous work are designated T1, T2, etc., and C1, C2, etc., as in the preceding article (Subramanian *et al.*, 1968), according to their position in the protein as determined by the present work. The fragments obtained from succinylated hemerythrin or derivatives of it (cystine-treated and carboxyamidomethylated hemerythrin) by digestion with TPCK-treated trypsin are designated succinylated hemerythrin TT-1, succinylated hemerythrin TT-2, etc., according to the order of elution of the fragments from tryptic digestion of cystine-treated succinylated hemerythrin from a column of Sephadex G-50. Peptides obtained from the tryptic fragments are designated C where chymotrypsin was used for further fragmentation, and A where dilute acid was used, followed by the numbers (in parentheses) of the amino-terminal and carboxyl-terminal residues of the peptide. Peptides obtained from chymotryptic digests of whole succinylated protein are similarly designated C, followed by the inclusive residue numbers.

Results

Enzymic Digestion of Modified Hemerythrins. Separation of Main Fractions. The trifluoroacetyl derivative of hemerythrin was relatively insoluble at the pH used for digestion. Some solubilization did occur on long incubation, but examination of the digests by paper electrophoresis–chromatography and by gel filtration revealed a complex mixture of peptides, indicating that cleavage at bonds other than the three arginyl linkages had occurred. This was confirmed by amino acid analysis of the peptides, which showed that some of them were products, ostensibly, of cleavage at chymotrypsin-sensitive bonds.

When the trifluoroacetyl group was removed and the pH was lowered, a large part of the peptide mixture precipitated. Most of the mixture was insoluble below pH 10 and did not become appreciably soluble even in 50% acetic acid.

Because of the above difficulties, no further work was done with the trifluoroacetyl derivative.

In contrast to the above, all the succinylated derivatives of hemerythrin and the peptides derived from them were found to be soluble at pH values over 5.5. Consequently, digestion with trypsin at pH 8 and subsequent separation of the peptides by gel filtration at the same pH could be carried out on completely soluble mixtures. Some complications in digestion and in separation and analysis, however, arose from the state of the cysteine residue of the monomer. When no attempt was made to protect the cysteine residue prior to succinylation low recoveries of cysteine (as cystine) on amino acid analysis resulted. The hemerythrin was therefore treated with cystine to form the cysteine adduct by disulfide exchange as described. Examination of the treated protein in the ultracentrifuge revealed that it had been completely dissociated into subunits, and it was therefore concluded that the cysteine residue

TABLE I: Amino Acid Composition of Main Fragments Obtained on Tryptic Digestion of Succinylated Hemerythrin.

Amino Acid	sHrTT6[e]	sHrTT3[e]	sHrTT7[e]	sHrTT2[e]
Lysine		1.12 (1)		10.10 (10)
Histidine		2.01 (2)		4.70 (5)
Arginine	1.02 (1)	1.21 (1)	1[b] (1)	
Aspartic acid	2.04 (2)	8.25 (8)		7.24 (7)
Threonine		1.91 (2)		2.40[c] (2, 3)
Serine	1.03 (1)	0.98 (1)		1.63[c] (1, 2)
Glutamic acid		2.07 (2)		8.00 (8)
Proline	4.00 (4)			
Glycine	0.99 (1)	2.01 (2)		3.61[c] (3, 4)
Alanine		1.94 (2)		3.40[c] (3, 4)
Cysteine				1.00[d] (1)
Valine	0.98 (1)			3.05 (3)
Methionine				0.96 (1)
Isoleucine	0.99 (1)	3.79 (4)		3.97 (4)
Leucine		3.97 (4)		4.02 (4)
Tyrosine	0.95 (1)	0.90 (1)		2.99 (3)
Phenylalanine	2.00 (2)	2.84 (3)		3.98 (4)
Tryptophan	1.0[a] (1)			3.1[a] (3)
Total: 113	15	33	1	64

[a] Determined spectrophotometrically (Goodwin and Morton, 1946). [b] Free arginine. [c] For explanation of non-integral result, see discussion of variant chains. [d] Determined as *S*-carboxymethylcysteine. [e] Moles per mole of fragment; residues per molecule accounted for by previously obtained tryptic peptides. sHr = succinylated hemerythrin.

had been quantitatively converted into the mixed disulfide. However, after succinylation, two peaks appeared on Sephadex G-75 chromatography, and two bands were found on starch gel electrophoresis. Material from each of the two G-75 peaks was reduced and carboxymethylated, and after this treatment the two species were found to be indistinguishable on starch gel electrophoresis. It was therefore apparent that both the mixed disulfide (cysteine adduct) and a disulfide dimer of succinylated hemerythrin had been present. As a result, upon tryptic digestion of cystine-treated succinylated hemerythrin, both a dimer (succinylated hemerythrin TT-1; see Figure 1) and a monomer (succinylated hemerythrin TT-2; see Figures 1 and 2) of the cysteine-containing peptide were found. The compositions of these fragments were identical, except for the presence in succinylated hemerythrin TT-2 of an extra half-cystine, thus corroborating the above conclusions.

The complications encountered with cystine-treated succinylated hemerythrin were obviated by the use of iodoacetamide as the blocking agent for cysteine. The CAM-cysteine derivative of hemerythrin was stable to the manipulations involved in succinylation and the cysteine was recovered quantitatively (as CM-cysteine) from acid hydrolyses of both CAM-succinylated hemerythrin and of the CAM-cysteine peptide.

A further advantage in the use of iodoacetamide as the blocking agent appeared when it was learned that the cysteine was next to one of the arginine residues of the protein. Canfield and Anfinsen (1963) have shown that the presence of an acid residue next to a trypsin-sensitive bond slows hydrolysis at that point. This was the case with the succinylated cystine-treated protein. Blocking with iodoacetamide eliminated this inhibition and made it possible to achieve essentially complete digestion with less enzyme and in a shorter time. Comparison of Figures 1 (fragments from cystine-treated protein) and 2 (fragments from carboxyamidomethylated protein) shows this clearly. Along with the absence of the dimer peak, the peaks succinylated hemerythrin TT-4/5 and succinylated hemerythrin TT-7, which are the main peaks produced by the action of residual chymotrypsin in the preparation or by the chymotryptic-like action of the trypsin itself, are present in considerably smaller amounts in Figure 2 than in Figure 1.

It may be remarked that while some chymotryptic or chymotryptic-like cuts occurred in all the digestions, the absence of any detectable cuts at lysyl residues testifies to the completeness of succinylation.

Removal of Succinyl Groups from Ester Linkage. Succinylation blocked some hydroxyl as well as the amino groups of hemerythrin, as was shown, for instance, by the fact that the fraction succinylated hemerythrin TT-6, derived from tryptic digestion of the succinylated cystine-treated hemerythrin (Figure 1), could be separated into two peptides on paper electrophoresis, each of which gave the same amino acid analysis. Treatment with hydroxylamine resulted in the conversion of the two electrophoretic bands into one. Since under the conditions used only ester bonds are hydrolyzed (Gallop *et al.*, 1959), it was concluded that

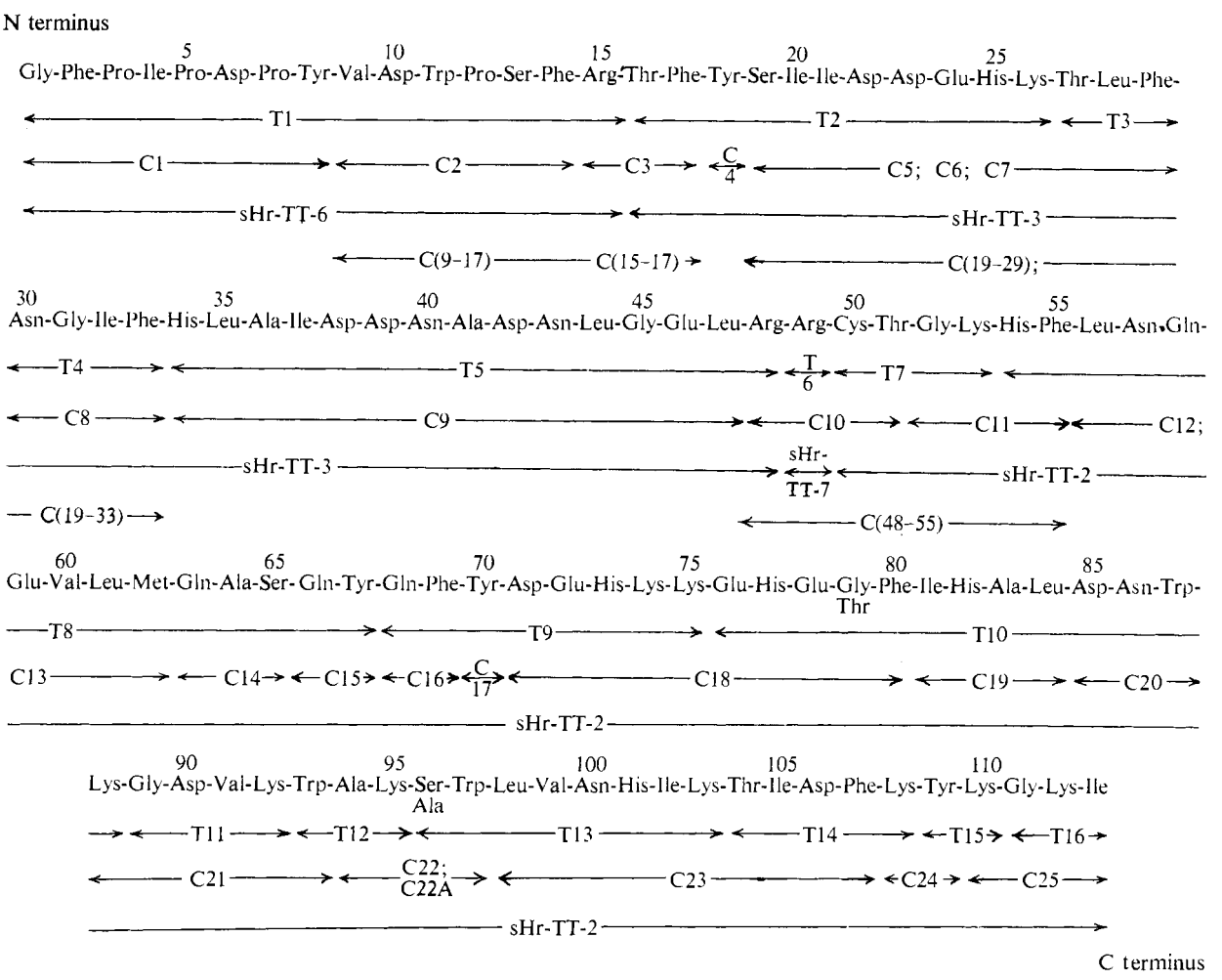

FIGURE 3: The primary structure of hemerythrin. Tryptic and chymotryptic peptides from previous work are included and may now be identified by residue numbers. From the present work, only the main fragments arising from tryptic digestion of succinylated hemerythrin (designated by sHr), the arginine-containing "bridge" fragments connecting these fragments, and the peptides containing the overlaps which remained to be determined, are shown. The other peptides mentioned in the text may be located by reference to their inclusive residue numbers.

the heterogeneity had indeed been caused by formation of a succinate ester bond with a hydroxyamino acid, of which there were two in the peptide, namely, tyrosine and serine. The weight of evidence favors succinylation of the serine hydroxyl since another peptide derived from the same digest and consisting of part of the sequence of succinylated hemerythrin TT-6, contained serine but not tyrosine and also showed this electrophoretic heterogeneity. When hydroxylamine treatment was used routinely after succinylation to break succinate ester bonds, no such heterogeneities showed up.

Compositions, Yields, and Placement of Peptides. The compositions of the main tryptic fragments are given in Table I. The yields of the purified main fragments recovered from the tryptic digest of cystine-treated succinylated hemerythrin were: succinylated hemerythrin TT-1, 25%; succinylated hemerythrin TT-2, 40%, thus the two together, 65%; succinylated hemerythrin TT-3, 62%; succinylated hemerythrin TT-6, 75%; and succinylated hemerythrin TT-7, 52%. The yields from the tryptic digest of carboxyamidomethylated succinylated hemerythrin were not determined, but were undoubtedly higher since in this case cleavage at chymotrypsin-sensitive bonds was markedly reduced.

The analyses of the compositions of the peptides obtained from the main fragments by further fragmentation with chymotrypsin or dilute acid are not generally given, as the results were in all cases unequivocal. The level of impurities varied from 2% in the case of peptide C (81–87) to 21% in the case of peptide A (90–105). Yields were not generally obtained, as in many cases the peptides were brought to final purification in only small amounts. The results of the detailed characterization of fractions, from whatever digest, are given in the following sections. Reference to the placement of the peptides may be had from Figure 3, which shows the completed sequence of the protein.

Fraction Succinylated Hemerythrin TT-6. RESIDUES 1–15. Fraction succinylated hemerythrin TT-6 as obtained from the column separations consisted of a single peptide in essentially pure form, as determined by electrophoresis–chromatography and by amino acid analysis. This peptide corresponds to the peptide

T1 which was shown in previous work (Groskopf *et al.*, 1966b) to be amino terminal in the protein. The electrophoretic and chromatographic mobility of succinylated hemerythrin TT-6 was different from that of T1 to the extent expected from the presence of the succinyl group on the amino-terminal glycine.

A small amount of a peptide, the composition of which was compatible with the sequence Val-Asp-Trp-Pro-Ser-Phe-Arg, residues 9–15, was separated from fraction succinylated hemerythrin TT-7 by paper chromatography. This peptide was clearly produced as a result of a chymotryptic or chymotryptic-like cleavage at the tyrosyl–valine bond of succinylated hemerythrin TT-6.

Fraction Succinylated Hemerythrin TT-3. Residues 16–48. This fraction was found to be made up principally of a 33-residue fragment which had a composition (Table I) consistent with the inclusion of the previously obtained tryptic peptides T2–T5. Two steps of the Edman degradation established the amino-terminal sequence as Thr-Phe- and thereby identified peptide T2 as the amino-terminal segment. From tryptic specificity, arginine was postulated to be the carboxyl-terminal residue. This was confirmed by digestion with carboxypeptidase B, which released predominantly arginine followed by some leucine, thereby establishing T5 as the carboxyl-terminal segment of the fragment.

The relative positions of peptides T3 and T4 within succinylated hemerythrin TT-3 were demonstrated by isolation and analysis of two peptides from the chymotryptic digest of succinylated hemerythrin TT-3. One of these, C (19–29), had the composition Ser, 0.95 (1); Ile, 1.96 (2); Asp, 2.17 (2); Glu, 1.02 (1); His, 0.94 (1); Lys, 0.98 (1); Thr, 0.96 (1); Leu, 1.03 (1); Phe, 1.01 (1). It was evident that this fragment was an overlap between T2 and T3 in that order, including all of T3. This overlap also placed T4 next in sequence by difference. This assignment was confirmed by the composition of C (19–33) which was as follows: Ser, 0.89 (1); Ile, 2.48 (3); Asp, 3.07 (3); Glu, 1.06 (1); His, 0.98 (1); Lys, 0.98 (1); Thr, 1.00 (1); Leu, 1.00 (1); Phe, 1.94 (2); Gly, 0.98 (1). In other words, this peptide had the same composition as C (19–29), plus aspartic acid (asparagine), glycine, isoleucine, and phenylalanine (which is the composition of T4), and was clearly a result of incomplete hydrolysis at the Phe–Asn bond between T3 and T4. T5 makes up the rest of the peptide, and the order T2–T5 is therefore established, and the two "inner fragments" of the preceding article are placed in proper order.

Fraction Succinylated Hemerythrin TT-2. Residues 50–113. The amino acid analysis of the peptide which was the main constituent of the fraction succinylated hemerythrin TT-2 is shown in Table I. It is seen that this peptide makes up more than half of the hemerythrin subunit molecule and contains the methionine, the cysteine, and all but one of the lysines. It contained no arginine, however, and was therefore postulated to be the carboxyl-terminal segment of the protein. This was confirmed by the results of digestion with carboxypeptidase A. After 1.5 min, 0.93 mole of isoleucine/mole of peptide was liberated; no other amino acids were released, even after 1.5 hr.

One step of the Edman degradation carried out on succinylated hemerythrin TT-2 reduced the amount of CM-cysteine recovered on acid hydrolysis from 1.0 to 0.49 residue, while all the other amino acids were found in essentially the same amounts as in the original peptide. Cysteine therefore appeared to be amino terminal.

Based on their amino acid composition, the peptides T7-T16 obtained from previous work could be fitted exactly within the succinylated peptide. This represents the remaining composition of the protein, following succinylated hemerythrin TT-6, succinylated hemerythrin TT-3, and succinylated hemerythrin TT-7.

Results of Enzymatic and Chemical Fragmentation of Succinylated Hemerythrin TT-2. The peptides formed by chymotryptic digestion of succinylated hemerythrin TT-2 were for the most part acidic as expected from the presence of ten succinylated lysines in this fragment of the protein. The mixture was therefore fractionated on a column of DEAE-Sephadex.

In contrast to the above, electrophoresis–chromatography of hydrolysates from treatment of succinylated hemerythrin TT-2 with dilute acid revealed a preponderance of either basic or neutral peptides. It was evident from this that the succinyl groups had been completely removed from the lysines by the acid treatment. This result is reasonable since the succinyl group has a structure similar to that of aspartic acid and could catalyze its own removal from amide linkage to ϵ-amino groups just as aspartic acid catalyzes its own liberation from peptide linkage. Electrophoresis chromatography of mixtures produced by hydrolysis of succinylated hemerythrin TT-2 with dilute acid for short periods of time indicated that removal of the succinyl groups was quite rapid compared with release of aspartic acid.

Peptides from the dilute acid hydrolyses were separated by passage of the hydrolysates through a column of Sephadex G-25, followed, in the case of the principal peak obtained, by passage through a column of CM-Sephadex, and finally by paper electrophoresis or paper chromatography.

Peptides from both types of hydrolyses which contributed to the establishment of overlaps between the tryptic peptides of succinylated hemerythrin TT-2 are described below, in order, from the amino terminus of the peptide to the carboxyl terminus of the protein.

Peptide C (50–55). This peptide contains cysteine and therefore is the amino-terminal peptide of succinylated hemerythrin TT-2. In addition to the amino acids of T7 it contains one residue each of histidine and phenylalanine. The only tryptic peptide in hemerythrin which has the possibility of these two amino acids in sequence at an amino terminus is T8. Chymotryptic specificity would predict the sequence at the carboxyl terminus of C (50–55) to be His-Phe, and this sequence was indeed established by carboxypeptidase A digestion which released predominantly phenylalanine and some histidine.

Peptide Fraction C (56–62). This fraction did not

contribute to an overlap since it was derived entirely from tryptic peptide T8; nonetheless it contributed to elucidation of the sequence of T8, which was only partly known at the time. The fraction appeared to be homogeneous in electrophoresis–chromatography. However, carboxypeptidase A digestion revealed that some methionine was present near the carboxyl terminus even though none of this amino acid had been detected after acid hydrolysis. On long digestion the amount of methionine released approached a mole ratio level of 0.29. When the values for methionine are normalized to those of leucine and valine liberated from the mixture, its rate of release is faster than valine but somewhat slower than that of leucine. These results are consistent with the presence in this fraction of two peptides one of which has the carboxyl-terminal sequence -Val-Leu-Met and the other -Val-Leu.

After one step of the Edman degradation, the residual peptide gave an analysis which indicated the disappearance of a leucine. A second step failed to remove any other residues. The electrophoretic mobility of the parent peptide indicated that two of the three acidic residues of the peptide were amidated. One of these could have been next to the leucine and could have become resistant to further Edman degradation by cyclization to pyrrolidonecarboxylic acid in the case of a glutamine, or by α,β rearrangement to a β-aspartyl linkage in the case of asparagine, under the conditions of treatment used. In any event, the results gave as partial sequence of C (56–62), Leu-(Asp, Glu_2)-Val-Leu-Met, and extended the sequence of T8 to His-Phe-Leu-(Asp, Glu_2)-Val-Leu-Met-Gln-Ala-Ser-Gln-Tyr. Meanwhile, the work done in the preceding article (Subramanian *et al.*, 1968) gave the sequence as shown in Figure 3.

Peptide(s) C (63–67). Two peptides representing this section of the chain were isolated. The electrophoretic mobilities indicated that in one of them one of the glutaminyl residues had been deamidated.

Peptides C (68–80), C (71–80), and A (72–84). A peptide overlap between T8 and T9 was not actually obtained, either in this or in previous work, except in the sense that succinylated hemerythrin TT-2 overlaps all peptides contained within itself. Nevertheless, connection between T8 and T9 became compelling as peptides were fitted into the remaining portion of succinylated hemerythrin TT-2. Overlap between T9 and T10 was obtained from the three peptides of this section. These peptides showed the nonintegral results for analysis of glycine and threonine, reflecting the presence in this part of the protein of this pair of variant residues.

Peptides A (86–89) and C (88–93). Both these peptides link T10 to T11, and C (88–93) also includes the amino-terminal tryptophan of T12. Peptide A (86–89) is clearly a result of cleavage at the aspartic acid residue 85 in T10. The electrophoretic mobility was consistent with the presence of asparagine rather than aspartic acid, which is also consistent with the fact that asparagine is released more slowly than aspartic acid on dilute acid hydrolysis (Tsung and Fraenkel-Conrat, 1965). The sequence of aspartic acid and asparagine at residues 85 and 86, which was in doubt, is thus seen to be -Asp-Asn. Carboxypeptidase A digestion of A (86–98) released predominantly glycine and some lysine and therefore confirmed the sequence for this peptide of Asn-Trp-Lys-Gly.

The composition of the other peptide, C (88–93), indicates that it contains all of T11, the amino-terminal tryptophan of T12 (which is the only tryptic peptide of the protein with amino-terminal tryptophan), and a lysine which could only have come from the carboxyl terminus of T10 since this is the only tryptic peptide of hemerythrin in which the carboxyl-terminal lysine is preceded by a chymotrypsin-susceptible bond. The sequence of C (88–93) is therefore Lys-Gly-Asp-Val-Lys-Trp.

Peptides C (94–97, Ser) and C (94–97, Ala). These two peptides had identical electrophoretic mobilities and both gave the Ehrlich test for tryptophan. Their chromatographic mobilities were slightly different, however, and therefore separation of the two was feasible. Both were tetrapeptides but their compositions differed by one amino acid, the difference being serine in one and alanine in the other. It is evident that the existence of these peptides localizes the second pair of variants in pooled hemerythrin.

Peptide(s) A (90–105). Two peptides including this sequence were isolated which had identical amino acid compositions but different electrophoretic mobilities due to the deamidation of an asparagine in one of them. They overlap peptides T11, T12, and T13.

Peptides C (98–107) and C (98–113). The composition of C (98–107) indicated an overlap between peptides T13 and T14, in that order, from cleavage at the tryptophanyl bond in T13 between residues 97 and 98 and the phenylalanyl bond in T14 between residues 107 and 108. Peptide C (98–113), extending to the end of the protein, was evidently produced by failure of cleavage of this phenylalanyl bond.

Peptide A (107–113). On the basis of compositions, part of peptide T14 and all of peptides T15 and T16 could be fitted into this peptide. These three peptides had been postulated (Groskopf *et al.*, 1966a) to be the carboxyl-terminal sequence of the protein. If this is true, A (107–113) should have the sequence Phe-Lys-Tyr-Lys-Gly-Lys-Ile, produced by cleavage at the aspartyl residue 106 of T14. This sequence was indeed confirmed by isolation of two small peptides from a chymotryptic digest of the peptide, one with the composition Phe, 0.96 (1); Lys, 1.04 (1); Tyr, 0.08 (1) (tyrosine was largely destroyed but its presence was indicated by a strong Pauly reaction before hydrolysis). The phenylalanyl–lysine bond was not cleaved, which is consistent with the position of the phenylalanyl residue at the N terminus in this peptide. The other peptide had the composition Lys, 1.84 (2); Gly, 1.20 (1); Ile, 0.99 (1). In view of chymotryptic specificity, the above compositions are consistent with the sequence of A (107–113) stated above, and the postulated carboxyl-terminal sequence of the protein is therefore confirmed.

Peptide A (91–113). Incomplete hydrolysis of the aspartyl bond between residues 106 and 107 gave rise

to a large peptide which extends from residue 91 to the end of the protein.

Chymotryptic Peptides of Succinylated Hemerythrin. Overlaps between the Large Fragments. The above work arranged all of the previously obtained tryptic peptides of hemerythrin in linear order within the tryptic fragments of succinylated hemerythrin. Furthermore, it identified succinylated hemerythrin TT-6 as the amino-terminal fragment and succinylated hemerythrin TT-2 as the carboxyl-terminal fragment, leaving succinylated hemerythrin TT-3 and free arginine as inner fragments. In order to obtain peptides which would overlap these fragments and thereby determine the relative positions of succinylated hemerythrin TT-3 and arginine, chymotryptic digestion of the succinylated protein was carried out. Because of the presence of succinyl groups on the lysines, it was expected many of the chymotryptic peptides produced would be more acidic than the few arginine-containing peptides, and accordingly separation of these latter on the basis of charge was indicated. The arginine-containing peptides indeed were simply isolated by electrophoresis on paper and then purified further by paper chromatography.

Peptide C (9–17). Composition: Val, 0.97 (1); Asp, 1.08 (1); Trp (1) (positive by Ehrlich reaction); Pro, 1.00 (1); Ser, 0.90 (1); Phe, 1.97 (2); Arg, 0.98 (1); Thr, 0.96 (1). While only slightly basic on electrophoresis, this peptide gave a Sakaguchi test for arginine. The amino acid analysis indicated it to be an overlap between succinylated hemerythrin TT-6 and succinylated hemerythrin TT-3, and the sequence is indicated to be Val-Asp-Trp-Pro-Ser-Phe-Arg-Thr-Phe. A tripeptide C (15–17) produced by cleavage at the phenylalanyl–arginine bond between residues 15 and 16 was also recovered and was assigned the sequence Arg-Thr-Phe. These two peptides not only fix the relative positions of succinylated hemerythrin TT-6 and succinylated hemerythrin TT-3 but also place the free arginine between peptides succinylated hemerythrin TT-3 and succinylated hemerythrin TT-2 by difference. The actual overlap between these fragments was established by characterization of the peptide described below.

Peptide C (48–55). Composition: Arg, 2.10 (2); Cys, 0.70 (1); Thr, 1.00 (1); Gly, 1.04 (1); Lys, 0.96 (1); His, 0.98 (1); Phe, 0.91 (1). Cleavage at the leucine (residue 47) adjacent to the carboxyl-terminal arginine in succinylated hemerythrin TT-3 and at the first phenylalanine (residue 55) of succinylated hemerythrin TT-2 produced this peptide. It overlaps fragments succinylated hemerythrin TT-3, succinylated hemerythrin TT-7, and succinylated hemerythrin TT-2. The presence of two arginines confirms the placement of the free arginine that was deduced above. The amino-terminal cysteine of succinylated hemerythrin TT-2 was recovered mostly as CM-cysteine but partly as cystine, due most likely to incomplete carboxyamidomethylation in this particular preparation of CAM-succinylated hemerythrin. The sequence of C (48–55) is Arg-Arg-Cys-Thr-Gly-Lys-His-Phe.

Based on the work described above and the results in the preceding paper (Subramanian *et al.*, 1968), the complete sequence of hemerythrin is established as shown in Figure 3.

Discussion

The technique of limited tryptic hydrolysis proved to be an effective procedure for obtaining a small number of fragments of the protein in which to place the previously obtained tryptic peptides in linear order. After masking of the ε-amino groups of the lysines, cleavage by trypsin was limited, with the exceptions discussed below, to the arginyl linkages of the hemerythrin subunit.

Trifluoracetylation was not found to be useful for blocking the lysines in this protein since the trifluoroacetyl protein was insoluble at alkaline pH, and as a result tryptic digestion was slow. Furthermore, removal of the blocking groups from the peptides after tryptic digestion did not result in significant solubilization of these peptides, and therefore their separation and purification were not feasible. Succinylation, on the other hand, produced a modified protein that was soluble and readily digested by trypsin.

As a general method for restricting tryptic action to the arginyl linkages in a protein, succinylation has the advantage of producing a derivative which is stable and also soluble. The fact that the succinyl group is difficult to remove is, however, a drawback. It would be desirable to break a protein into a few large fragments by limited tryptic hydrolysis, to separate these, and then, after removal of the blocking groups, to digest again with trypsin in order to obtain peptides produced by cleavage at the lysines, as one of the options. In this way, the complex problem of separation of peptides would be divided into two less complex steps. Furthermore, peptides could be assigned immediately to a zone or fragment of the protein. This method would be of particular interest in the determination of the sequences of larger proteins. If, however, a nonremovable group such as the succinyl group is used for blocking, trypsin cannot be used for further fragmentation of the blocked peptides and one would have to rely on other enzymes or on chemical methods of cleavage.[2]

In the case of hemerythrin, not only the succinyl derivative itself but also the peptides derived from it were soluble. The fragments produced by tryptic digestion of the succinylated hemerythrin were of quite different sizes (succinylated hemerythrin TT-6, 15 residues; succinylated hemerythrin TT-3, 33 residues; succinylated hemerythrin TT-7, 1 residue; and succinylated hemerythrin TT-2, 64 residues). This was a circumstance which could not be foreseen. Because of this size distribution, the peptides were conveniently separated by gel filtration. The charge distribution of the frag-

[2] Hartley (1967) has reported on the use of maleic anhydride as a blocking reagent for the ε-amino groups of lysine in peptide sequence. This reagent is quite specific for the amino group, which obviates the necessity for treatment with hydroxylamine to break any ester bonds. In addition, the maleyl group can be removed easily by mild treatment to regenerate the original ε-aminolysine groups.

ments was such that separation by ion exchange might also have been effected, but difficulty was encountered in eluting the larger peptides from the ion exchanger. However, the chymotryptic peptides of succinylated hemerythrin TT-2, which were smaller, were conveniently separated on columns of ion-exchange Sephadex. As for the peptides produced by treatment with dilute acid, the ε-amino groups were regenerated during the hydrolysis, and the peptides were separated by a variety of means. Regeneration of the ε-amino group by acid hydrolysis allowed lysine to be determined as such on amino acid analysis.

Of the main fragments, peptide succinylated hemerythrin TT-2 (64 residues) is big enough that in unmodified form it might not have been soluble. This fragment contains 10 of the 11 lysines of the protein, and succinylation of these was no doubt a factor in keeping this large fragment soluble.

The masking of the ε-amino groups of the lysines in a protein represents only a part of the potential use of restricted hydrolysis by trypsin. The other aspect involves the blocking of arginyl residues in order to limit tryptic cleavage to lysyl bonds. Techniques for the blocking of arginine have been reported, using benzil (Itano and Gottlieb, 1963), 1,2-cyclohexanedione (Toi *et al.*, 1965), and malonaldehyde (King, 1966). These techniques have not been extensively used as yet in protein sequence studies.

Even with the use of trypsin treated with TPCK to inactivate any chymotrypsin present, some chymotrypsin-like cuts occurred. Whether these were due to chymotryptic-like action of the trypsin itself or to the action of chymotrypsin which had escaped inactivation, or to the action of chymotrypsin which had become activated from chymotrypsinogen, is not known. In the case of a modified protein, the length of time needed to achieve good cleavage of the arginyl bonds when they are the only sensitive bonds allows secondary cleavages to become more manifest than they normally would be. Two bonds in hemerythrin, noted in this and in previous work, seem to be particularly sensitive to chymotryptic or intrinsic proteolytic activity. These are the tyrosyl–valine bond between residues 8 and 9 and the tyrosyl–glutamine bond between residues 67 and 68.

Some cases of incomplete cleavage may be noted. One such is the peptide C (19–33), which is the result of incomplete cleavage at the phenylalanyl–asparagine bond between residues 29 and 30. This incomplete cleavage can be explained if it is assumed that some of the asparagine was deamidated. Because of the negative charge of an aspartyl residue, binding of this portion of the peptide to the active site of chymotrypsin would be weak (Wallace *et al.*, 1963), and hydrolysis would therefore be slow. The failure of chymotrypsin to cleave completely the tyrosyl–aspartic acid bond (residues 70 and 71) of peptide T9 is probably similarly to be ascribed to the acidic nature of the aspartyl side group, and the partial cleavage of the phenylalanyl bond between residues 107 and 108, giving rise to peptide C (98–113), is probably due to a similar effect of the succinyl group of the adjacent modified lysine residue. The phenylalanyl–arginine bond between residues 15 and 16 was only partially cleaved, giving rise to peptide C (9–17). The failure of chymotrypsin to cleave completely the phenylalanyl bond between residues 69 and 70 is unexplained.

The leucyl–methionine bond between residues 61 and 62 was only partially cleaved. Slowness of action of the enzyme at this point in the chain was also observed in previous work (Subramanian *et al.*, 1968). The leucyl–aspartic acid bond between residues 84 and 85 was apparently hardly cleaved, which was also the case for the leucyl–asparagine bond between residues 56 and 57.

With the unmodified protein, digestion with carboxypeptidase A removes essentially all the residues back to residue 106 (which is an aspartyl residue) (Groskopf *et al.*, 1966a). When the succinylated fragment succinylated hemerythrin TT-2 was subjected to digestion by carboxypeptidase A, only the terminal isoleucine was removed. This failure of carboxypeptidase A to release other residues is presumably due to the fact that the penultimate residue, known to be lysine, is in this case succinylated. Blocked lysyl residues are known to be resistant to the action of carboxypeptidase (Hofmann and Bergmann, 1940), and the acidic acid side chain (the succinyl function) might also be expected to be inhibiting since glutamyl residues are known to be resistant to carboxypeptidase action (Green and Stahmann, 1952).

Two sites of amino acid interchange were identified in the hemerythrin polypeptide chain. Substitution of a threonine for a glycine at residue 79 and of an alanine for a serine at residue 96 were found. These substitutions had been suggested by the amino acid composition of the protein (Groskopf *et al.*, 1966a), in which the residue ratios for these four amino acids are closer to half than to integral residues. Similar results on amino acid analysis of fragment succinylated hemerythrin TT-2 localized the points of variation within this fragment. Finally, the identity of these interchanges was established by isolation of the two variant peptides in the case of the alanine-serine substitution and by observation of complementary non-integral residue ratios in several peptides containing the threonine-glycine substitution. Similar results are described in the preceding article (Subramanian *et al.*, 1968). The question of the distribution of these interchanges in the hemerythrin of individual worms is under study.

Some grouping of residues in the sequence may be remarked on. The subunit chain contains four prolines, and these are all in the first tryptic peptide, at positions 3, 5, 7, and 12. Since peptide chains are synthesized from the N-terminal end, this grouping may condition the folding of the chain in a special way.

The lysines of the chain are grouped toward the carboxyl terminus. The first lysine occurs at position 26. The remaining ones are at positions 53, 74, 75, 88, 92, 95, 103, 108, 110, and 112. Some concentrations of dicarboxylic amino acid residues can also be noted. One such is at residues 22–24, another at residues 38–46, and a third at residues 71–78.

The subunit molecule contains only a single cysteine. This residue is present in the sulfhydryl form in the native molecule. On denaturation, the SH becomes exposed, and dimer formation through a disulfide bond may occur.

The hemerythrin subunit binds two atoms of iron. Elucidation of the primary structure makes possible further studies on the nature and identity of the residues providing ligands to the iron. If the positions of the residues involved in the iron binding can be determined, it may be possible to bring parts of the sequence together and in this way obtain some idea of the conformation of the binding sites and some approximations to the three-dimensional structure of the molecule. It may also be possible to obtain some insight into the nature of the forces holding the subunits together into an octamer.

Acknowledgment

We are grateful to Dr. E. Margoliash for many helpful comments particularly during the early stages of this investigation.

References

Canfield, R. E., and Anfinsen, C. B. (1963), *Proteins 1*, 311.

Crestfield, A. M., Moore, S., and Stein, W. H. (1963), *J. Biol. Chem. 238*, 622.

Ferguson, K. A., and Wallace, A. L. C. (1961), *Nature 190*, 629.

Gallop, P. M., Seifter, S., and Meilman, E. (1959), *Nature 183*, 1659.

Goldberger, R. F., and Anfinsen, C. B. (1962), *Biochemistry 1*, 40.

Goodwin, T. W., and Morton, R. A. (1946), *Biochem. J. 40*, 628.

Grannis, G. F. (1960), *Arch. Biochem. Biophys. 91*, 255.

Green, M., and Stahmann, M. A. (1952), *J. Biol. Chem. 197*, 771.

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*, 3783.

Hartley, B. S. (1967), *Federation Proc. 26*, 246.

Hauptschein, M., Stokes, C. S., and Nodiff, E. A. (1952), *J. Am. Chem. Soc. 74*, 4005.

Hirs, C. H. W., Moore, S., and Stein, W. H. (1960), *J. Biol. Chem. 235*, 633.

Hirs, C. H. W., Stein, W. H., and Moore, S. (1954), *J. Biol. Chem. 211*, 941.

Hirs, C. H. W., Stein, W. H., and Moore, S. (1956), *J. Biol. Chem. 219*, 623.

Hofmann, K., and Bergmann, M. (1940), *J. Biol. Chem. 134*, 225.

Itano, H. A., and Gottlieb, A. J. (1963), *Biochem. Biophys. Res. Commun. 12*, 405.

King, T. P. (1966), *Biochemistry 5*, 3454.

Kostka, V., and Carpenter, F. H. (1964), *J. Biol. Chem. 239*, 1799.

Mahowald, T. A., Noltmann, E. A., and Kuby, S. A. (1962), *J. Biol. Chem. 237*, 1138.

Partridge, S. M., and Davis, H. F. (1950), *Nature 165*, 62.

Sanger, F., and Thompson, E. O. P. (1963), *Biochim. Biophys. Acta 71*, 468.

Schoellmann, G., and Shaw, E. (1963), *Biochemistry 2*, 252.

Schroeder, W. A., Shelton, J. B., Cormick, J., and Jones, R. T. (1963), *Biochemistry 2*, 992.

Schultz, J., Allison, H., and Grice, M. (1962), *Biochemistry 1*, 694.

Smithies, O. E. (1955), *Nature 175*, 307.

Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.

Subramanian, A. R., Holleman, J. W., and Klotz, I. M. (1968), *Biochemistry 7*, 3859 (this issue; preceding paper).

Toi, K., Bynum, E., Norris, E., and Itano, H. A. (1965), *J. Biol. Chem. 240*, PC 3455.

Tsung, C. M., and Fraenkel-Conrat, H. (1965), *Biochemistry 4*, 793.

Wallace, R. A., Kurtz, A. N., and Niemann, C. (1963), *Biochemistry 2*, 824.