1978

New evidence supporting the assignment of glutamic acid as an iron ligand in hemerythrin

Patricia M. Gormley
Portland State University

AN ABSTRACT OF THE THESIS OF Patricia Gormley for the Master of Science in Chemistry presented August 30, 1978.

Title: New Evidence Supporting the Assignment of Glutamic Acid as an Iron Ligand in Hemerythrin.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

[redacted]
Joann S. Loehr, Chairperson

[redacted]
Gordon Kilgour

[redacted]
Raymond Lutz

[redacted]
Mary Taylor

The amino acid sequence determination of Phascolopsis gouldii hemerythrin in the region of the proposed iron ligand at position 58 was the main objective of this research endeavor. Generation of a large peptide was pursued by trypsin digestion of citraconylated hemerythrin producing peptide 50-113 for sequenator analysis. Detection of the phenylthiohydantoin amino acid derivatives by gas-liquid and high-performance-liquid chromatography yielded unambiguous sequence elucidation through the region of interest identifying residue 58 as glutamic acid.

Further specific digestion of peptide 50-113 by cyanogen bromide yielded peptides 50-62 and 63-113 which were chromatographically resolved

using Sephadex G-50 superfine. The identity of peptide 50-62 was verified by amino acid analysis. Sequenator analysis of this peptide substantiated the assignment of residue 58 as glutamic acid.

Attempts to sequence peptide 54-74, derived from trypsin digestion of peptide 50-113, were frustrated in the ion exchange purification step. The peptide's high percentage of hydrophobic residues as well as its isoelectric form in the polar pyridine-acetate pH 5.0 buffer resulted in peptide insolubility.

Identification of residue 58 as glutamic acid in Phascolopsis gouldii hemerythrin is significant in the support it lends to the proposed model of the iron active site of Themiste dyscritum hemerythrin.

NEW EVIDENCE SUPPORTING THE ASSIGNMENT OF GLUTAMIC ACID AS AN IRON LIGAND IN HEMERYTHRIN

by

PATRICIA M. GORMLEY

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
CHEMISTRY

Portland State University

1979

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Patricia Mary Gormley presented August 30, 1978.

[redacted]
JoAnn S. Loehr, Chairperson

[redacted]
Gordon Kilgour

[redacted]
Raymond Lutz

[redacted]
Mary Taylor

APPROVED:

[redacted]
David W. McClure, Chairman
Department of Chemistry

[redacted]
Stanley E. Rauch, Dean of Graduate Studies and Research

ACKNOWLEDGMENTS

I would like to thank Dr. Joann S. Loehr for her valuable assistance provided throughout this investigation. I would also like to thank Dr. Bernie Brimhall of the University of Oregon Health Sciences Center for the amino acid analyses and Dr. Mark Hermodson of Purdue University for the sequenator analyses.

Finally, I thank my family and friends for their encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Hemerythrin and myohemerythrin are non-heme iron proteins whose function is reversible oxygen binding. Invertebrate marine worms of the sipunculid phylum contain hemerythrin in their coelomic fluid and myohemerythrin in their muscle cells. Hemerythrin of the Atlantic coastal Phascolopsis (syn. Golfingia) gouldii is the protein under investigation in this study. The protein is an octamer of molecular weight 108,000 containing two iron atoms per monomer (1).

IDENTIFICATION OF IRON LIGANDS

The x-ray crystal structure of P. gouldii hemerythrin at 5.5 Å resolution indicates that the two irons in each monomer are in close proximity and are coordinated by amino acid side chains from four, approximately parallel helices which comprise 70% of the protein structure (2). A 2.8 Å resolution crystal structure has been obtained for coelomic hemerythrin from a sipunculid indigenious to the Pacific coastal waters, Themiste (syn. Dendrostomum) dyscritum (3). The two iron atoms in the metaquo form of the T. dyscritum protein are in a trigonal antiprism with each iron octahedrally coordinated to five protein ligands (Fig. 1). Two of the ligands are common to both iron atoms and a third shared coordination site contains a water molecule which is presumably replaced by oxygen in the native form of the protein. The iron binding site in P. gouldii hemerythrin is very likely identical to that in T. dyscritum hemerythrin as the two proteins exhibit similar spectroscopic properties (4) as well as similar protein conformations.

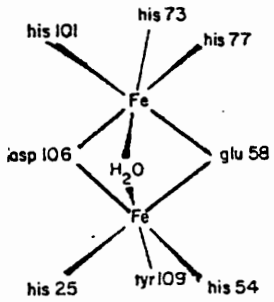


Figure 1. Proposed model for the active site of metaquo T. dyscritum hemerythrin at 2.8 Å resolution (3). The same iron ligands have been identified for myohemerythrin but with differing numbering (30). In the oxy form, the water is presumably replaced by O_2 which is bound as the peroxide, $O_2^{=}$ (31).

The 2.8 Å resolution electron density map of T. dyscritum hemerythrin in conjunction with the protein's amino acid sequence (5) identified the iron ligands as His-25, His-54, Glu-58, His-74, His-77, His-101, Asp-106, and Tyr-109. These data were corroborated by the x-ray diffraction studies (6) and amino acid sequence (7) of myohemerythrin from T. zostericolum, another Pacific Coast sipunculid (formerly misidentified as T. pyroides). The amino acid sequence of P. gouldii hemerythrin (8) showed that seven of the amino acid ligand identities coincided. However, in the eighth ligand assignment, the sequences of T. dyscritum hemerythrin and T. zostericolum myohemerythrin indicated a glutamic acid at position 58 while residue 58 of P. gouldii hemerythrin was identified as glutamine. The apparent lack of conservation of an iron ligand seemed unlikely and prompted the present investigation.

This study was undertaken to more accurately determine the identity of the amino acid at position 58 in P. gouldii hemerythrin. If glutamic acid proved to be the correct amino acid at residue 58, this would lend support to the trigonal antiprism model in which glutamic acid acts as one of the bridging ligands. However, should glutamine be verified as the amino acid at position 58, it would be the first example of glutamine as a metal ligand. It is known that amides have a low basicity and are, therefore, poor metal ligands (32). An active iron center containing an inferior ligand most probably would be disgarded by natural selection. The γ-carboxyl group of glutamic acid, however, is deprotonated above pH 4.2. Thus, there is a sound chemical basis for predicting glutamic acid as the correct iron ligand.

AMINO ACID SEQUENCE DETERMINATION

Facile elucidation of an amino acid sequence is attributable to the innovative automated protein sequenator of Edman and Begg (9). Utilization of the sequenator allows derivatization, cleavage and isolation of amino acids sequentially from a peptide's NH_2-terminus. The method emphasizes generation of a few large peptides from the parent protein instead of the many small peptides previously required for the manual Edman degradation procedure. The requirement of few large peptides necessitates availability of highly specific cleavage agents or suitable blocking groups to hone an agent's specificity.

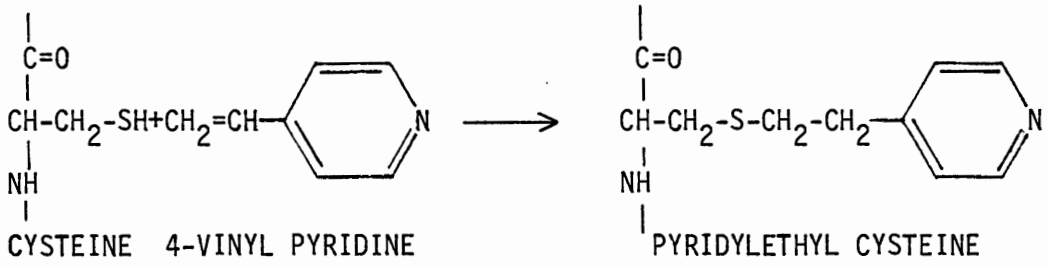
Trypsin is an endopeptidase specific for cleavage at the COOH-terminus of lysine and arginine. The enzyme's mode of action is illustrated in the following reaction:

$$H_3^+N\text{-}X\cdots ARG\text{-}Y\cdots LYS\text{-}R\cdots Z\text{-}COO^- \xrightarrow{\text{TRYPSIN}} H_3^+N\text{-}X\cdots ARG\text{-}COO^- + H_3^+N\text{-}Y\cdots LYS\text{-}COO^- + H_3^+N\text{-}R\cdots Z\text{-}COO^-$$

Trypsin is often contaminated with chymotrypsin, an enzyme that cleaves at the COOH-terminus of aromatic residues. Thus, trypsin preparations generally contain an inhibitor of chymotrypsin such as L-(tosyl-2-phenyl) chloroethyl methyl ketone (TPCK) to ensure an enzyme preparation of reliable specificity.

For trypsin digestion it is beneficial to block cysteine residues since the presence of an acidic residue adjacent to arginine or lysine reduces the effective cleavage (10). Furthermore, blocking cysteine residues prevents oxidation of sulfhydryl groups to intra- or inter-

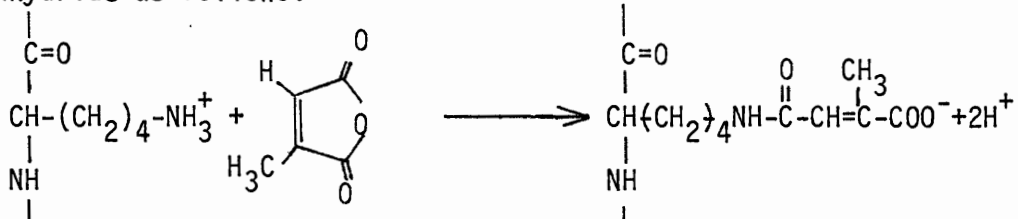
molecular disulfide moities which yield anomalous sequences. A blocking method of choice is represented by the reaction of 4-vinyl pyridine with a cysteinyl residue (11):



CYSTEINE 4-VINYL PYRIDINE PYRIDYLETHYL CYSTEINE

The protein subunit is reduced with dithiothreitol prior to reaction with 4-vinyl pyridine since derivatization occurs only with the sulfhydryl form of cysteine. In the subsequent amino acid composition and sequence determination, cysteine is detected as pyridylethyl cysteine.

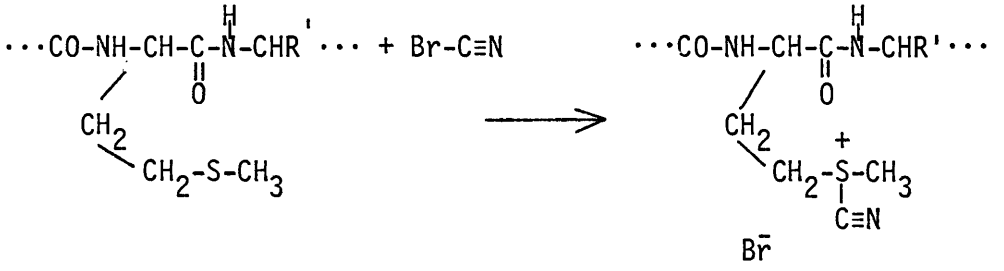
Derivatization of the ε-amino group of lysine confers enhanced specificity on trypsin, such that it only cleaves after arginine. Citraconic anhydride is a blocking agent of choice due to the high solubility of its protein derivative at neutral pH and the ease of removal of the blocking group in acid. Lysine reacts with citraconic anhydride as follows:



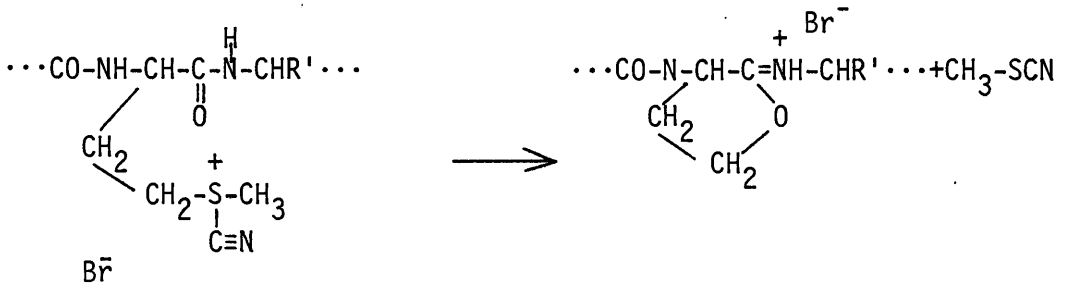
LYSINE CITRACONIC ANHYDRIDE CITRACONYL PEPTIDE

Other lysine blocking groups such as succinic anhydride and trifluoroacetic acid are less useful due to the difficulty of removal of the blocking group and solubility of the protein derivative, respectively (12).

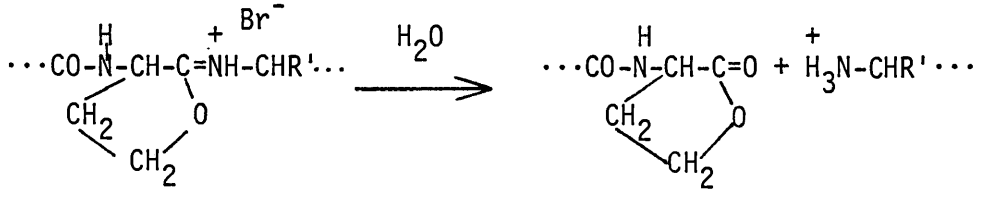
Cyanogen bromide is known to specifically cleave the COOH-terminal bond of methionine (13):



METHIONINE CYANOGEN BROMIDE CYANOSULFONIUM INTERMEDIATE (I)



CYANOSULFONIUM INTERMEDIATE (I) IMINOLACTONE (II) METHYL THIOCYANATE



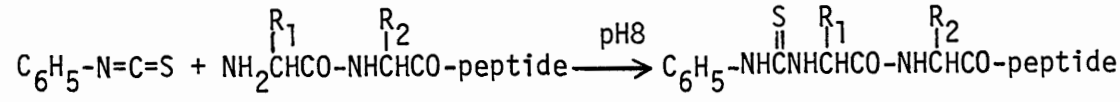
IMINOLACTONE (II) PEPTIDYL HOMOSERINE LACTONE (III)

The starting material is first reduced to convert any methionine sulfone or sulfoxide to methionine (14). Upon addition of CNBr, an unstable cyanosulfonium intermediate forms and then converts to the iminolactone (II) with elimination of the methyl thiocyanate as by-product. Hydrolysis of compound II produces the aminoacyl peptide and the peptidyl homoserine lactone.

Once specific peptides are isolated and purified, degradation

from the NH_2-terminal is performed automatically in the sequenator designed by Edman and Begg (9). The method involves the following reaction scheme:

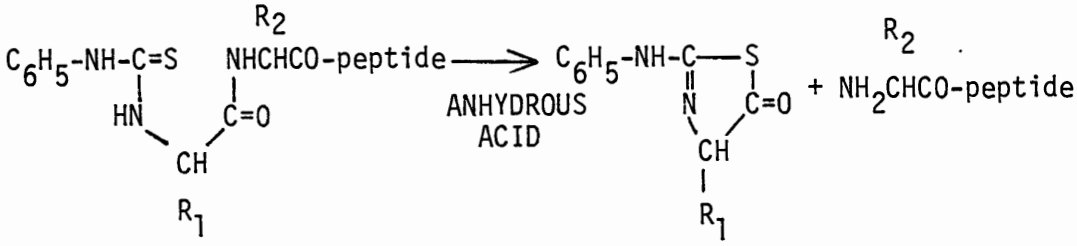
COUPLING STEP



PHENYLISOTHIOCYANATE

PHENYLTHIOCARBAMYL (PTC) PEPTIED (I)

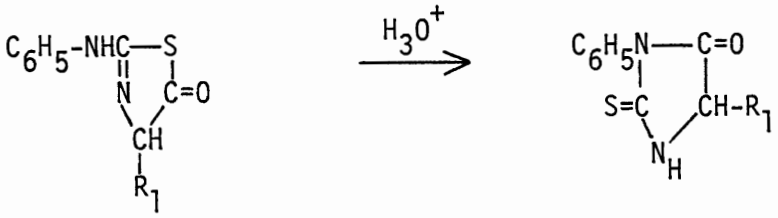
CYCLIZATION STEP



PTC-PEPTIDE

ANILINO-THIAZOLINONE DERIVATIVE (II)

REARRANGEMENT STEP



ANILINO-THIAZOLINONE DERIVATIVE (II)

PHENYLTHIOHYDANTOIN (PTH) DERIVATIVE (III)

The initial step involves formation of the phenyl thiocarbamyl derivative (I) which subsequently forms the thiazolinone (II) and peptide with a new

NH_2-terminal. Once (II) is removed by extraction, the peptide is ready for another round of coupling and cyclization. The extracted compound (II) is slowly rearranged to the more stable PTH-amino acid (III) in an acid-catalyzed reaction. The resulting PTH amino acids are detected by gas-liquid and high-performance-liquid chromatography. These detection methods offer greater sensitivity and accuracy than was previously possible with paper electrophoresis, thin-layer chromatography or amino acid analyzers.

The methods outlined above, in conjunction with gel permeation and ion exchange chromatography, were used to produce peptides of purity suitable for unambiguous determination of the sequence in the region of residue 58 in P. gouldii hemerythrin.

EXPERIMENTAL PROCEDURES

PREPARATION OF HEMERYTHRIN

The sipunculid worm, Phascolopsis gouldii, was obtained from the Marine Biological Station in Woods Hole, Massachusetts. Erythrocytes, which contain hemerythrin, were separated from coelomic hemolymph, purified and lysed (15). Hemerythrin crystallization was effected by dialysis at 4°C against a 100-fold excess of 20% ethanol, 0.01 M sodium azide and 0.01 M Tris-Cl (tris hydroxymethyl-aminomethane) pH 7.5. Conversion of hemerythrin to the iron-free apo form was accomplished by the method of Loehr, et al. (5).

CHEMICAL MODIFICATION OF HEMERYTHRIN

Apohemerythrin was reacted with 4-vinyl pyridine to covalently block cysteine and thus, stabilize the residue to oxidation (11). The hemerythrin modification procedure involved reduction of protein subunits which contain 3.7 μmoles cysteine with 0.12 mmole dithiothreitol, followed by reaction with 0.36 mmole 4-vinyl pyridine. Pyridylethylated hemerythrin and peptides were used in all subsequent reactions. The reaction was monitored by the presence of an absorbance maximum at 260 nm due to pyridylethly cysteine (16).

Pyridylethylated hemerythrin was further modified to prevent trypsin digestion at lysyl residues by reaction with citraconic anhydride. Approximately 3.7 μmole of hemerythrin subunits containing

41 μmole lysine (17) were covalently blocked with 1.9 mmole citraconic anhydride was vacuum distilled at 75°C and 2.5 mm Hg. Alkaline conditions were maintained throughout the reaction and dialysis, as the citraconyl-lysyl bond is acid labile.

SPECIFIC CLEAVAGE OF HEMERYTHRIN

Citraconylated pyridylethylated hemerythrin was digested specifically at the COOH-terminus of arginyl residues by TPCK-trypsin (5). The reaction mixture contained 3 mg trypsin per 100 mg hemerythrin. The reaction temperature was maintained at 37°C by a Lauda circulator. Following the digestion, the solution was acidified to pH 1.5 with 88% formic acid to hydrolyze the citraconyl-lysyl bonds. The protein solution was shelled in dry ice-acetone and lyophilized using a Virtis, Model 10-020. The lyophilized protein was dissolved by wetting with glacial acetic acid, followed by dilution with distilled water to approximately 10% acetic acid. The tryptic peptides were separated on Sephadex G-50 superfine in 9% formic acid by gel permeation chromatography.

One of the tryptic peptides (peptide 50-113) was treated with cyanogen bromide to obtain specific cleavage at the COOH-terminus of the methionyl residue (13). Approximately 1.0 meq cyanogen bromide was reacted with 1.8 μeq methionyl residue in the peptide in 70% formic acid for 24 hours. Following the reaction, the material was lyophilized and peptides were separated by gel permeation on Sephadex G-50 superfine in 9% formic acid. The sample on which amino acid composition was determined was subjected to reduction with 30% mercaptoethanol prior to

CNBr cleavage (14). The reduced protein solution was dialyzed against increasing concentrations of formic acid (10%, 20%, 70%) to remove excess mercaptoethanol and 5 M guanidine hydrochloride. The sample on which sequenator analysis was performed was not subjected to reduction prior to CNBr cleavage.

Peptide 50-113 was also subjected to a second TPCK-trypsin digest (3 mg trypsin per 50 mg peptide) without prior citraconylation to obtain proteolysis at the COOH-terminus of lysyl residues. In the preliminary runs, the tryptic hydrolysate was immediately shelled and lyophilized. Poor dissolution of the lyophilized protein in the chromatographic solvent prompted dialysis against 9% formic acid instead of lyophilization in later runs. The major tryptic peptide from this digestion (peptide 54-74) was purified by successive passage through columns of Sephadex G-50 and G-25 superfine and Bio-Rad AG 50W-X2.

SPECTROPHOTOMETRIC ANALYSES

All spectrophotometric analyses were made using a Cary 14 spectrophotometer. Matched 1 ml quartz curvettes with a 1 cm path length were used. The reference cell contained 9% formic acid. Absorbances at 280 nm were used to determine protein concentrations and resolve chromatographic fractions at certain key stages of the sequence determination. The following extinction coefficients were used: apohemerythrin and pyridylethyl hemerythrin, E_{280} = 2.0 ml/mg; peptide 50-113, E_{280} = 2.5 ml/mg; peptide 54-74, E_{280} = 0.8 ml/mg; peptide 50-62, E_{280} = 0.3 ml/mg, E_{260} = 3.3 ml/mg. These values were calculated from the known composition of tyrosine, tryptophan and pyridylethyl cysteine

per 13,500 molecular weight subunit (18). Ultraviolet spectra were used to identify products from the characteristic absorbance profiles of tyrosine, tryptophan and pyridylethylcysteine.

CHROMATOGRAPHY

Gel permeation chromatography was performed in glass columns, Pharmacia (1.5 x 86 cm) and Kontes Chromaflex (2.5 x 86 cm). The matrices used were Sephadex G-25 and G-50, superfine grades, purchased from Pharmacia and Sigma. The equilibrium and eluant solvent was 9% formic acid. Products were generally detected by their absorbances at 280 nm, although peptide 50-62 was detected by its absorbance at 260 nm.

Ion exchange chromatography was used to purify tryptic peptide 54-74 prior to amino acid analysis. The column diameter was 0.6 cm. The height of the Bio-Rad AG 50W-X2 resin bed was 56 cm. The column temperature was maintained at 47°C. A continuous pH gradient (3.1 to 5.0) effected column elution. This procedure was done using the method and equipment developed by Jones (19). Ion exchange fractions were resolved by reaction with ninhydrin (20) and absorbance at 570 nm of the Ruhemann's purple complexes.

AMINO ACID ANALYSIS

Amino acid analyses (21) were performed on the tryptic peptide 54-74 and cyanogen bromide peptide 50-62. Each peptide was hydrolyzed under vacuum in 6 N HCl at 110°C for 24 hours. The amino acid hydrolysates were evaporated to dryness and dissolved in 3.0 ml buffer. Separation and detection of the amino acids were done using a Beckman amino acid

analyzer, Model 120C. Methionine of the CNBr peptide was detected as homoserine lactone since the equilibrium between homoserine and the lactone is shifted to the lactone on heating in acid. Cysteine was detected as the pyridylethylated derivative.

AMINO ACID SEQUENCE DETERMINATION

Amino acid sequences were determined for tryptic peptide 50-113 and CNBr peptide 50-62 using an automated Edman degradation procedure and a Beckman Sequencer, Model 890 as described previously (9), but modified to improve yields from peptides (22). Sequenator products were identified by gas chromatography, thin layer chromatography, and spot tests, as described previously (22) and by high-performance liquid chromatography (Bio-Analytical Systems, Inc.) as described by Zimmerman, et al. (23).

RESULTS

The amino acid at position 58 in P. gouldii hemerythrin was identified as glutamic acid using specific peptide cleavage techniques in conjunction with amino acid analysis and automated Edman sequence determination. Tryptic hydrolysis of citraconylated pyridylethylated hemerythrin yielded three peptides (Fig. 2) as observed previously (12). These peptides were identified as 50-113, 16-48 and 1-15 in order of decreasing molecular weight (12). Since the amino acid composition and manual sequence information have already been reported for peptide 50-113 (12), this peptide was subjected only to sequenator analysis. The data for peptide 50-113 are summarized in Table I. Further digestions of peptide 50-113 and the resulting peptides are shown in Figure 3. Cyanogen bromide cleavage yielded peptide 50-62 which was identified by amino acid analysis (Table II) and utilized for sequence determination (Table I). Trypsin digestion of pyridylethylated 50-113 produced peptide 54-74 which was characterized by amino acid analysis (Table II), but was not available in sufficient quantity or purity for sequenator analysis.

PEPTIDE 50-113

Peptide 50-113 was isolated by Sephadex chromatography of the tryptic hydrolysate of citraconylated pyridylethylated hemerythrin. The elution pattern (Fig. 2) appears identical to that reported previously (12) and fractions from the earliest elution peak 0.45 V_T (peptide 50-113) were pooled, lyophilized and used for sequenator analysis. The peptide

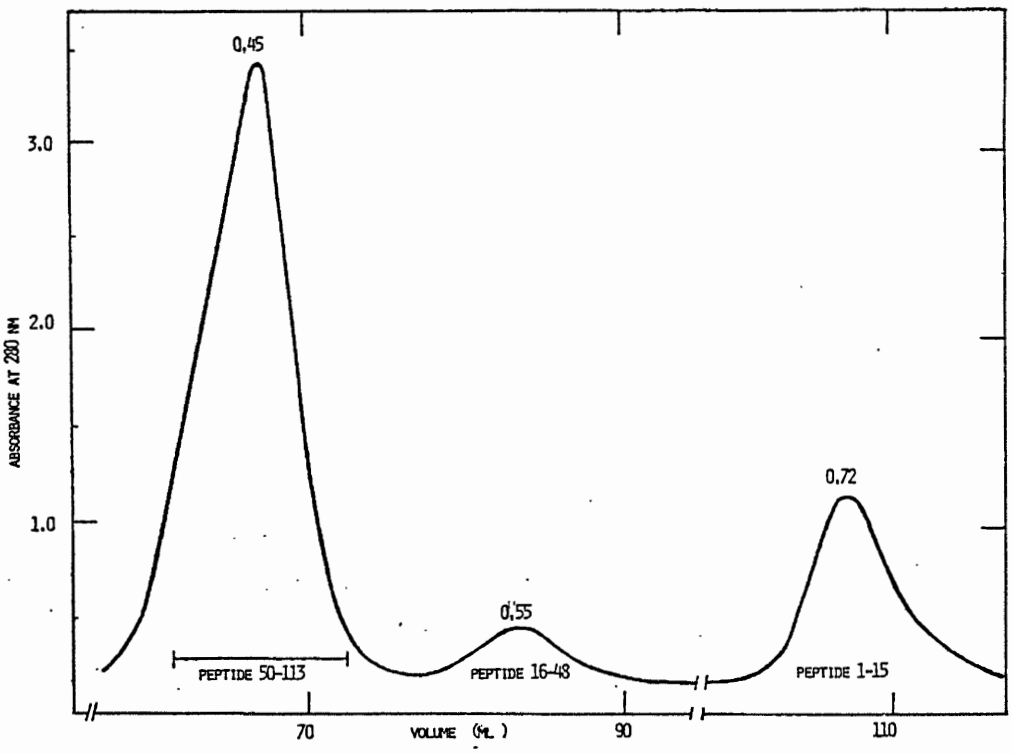


Figure 2. Gel permeation chromatography of trypsin digestion products of citraconylated hemerythrin on Sephadex G-50 superfine. A 50 mg sample was applied to the column (1.5 x 86 cm) and eluted with 9% formic acid at a flow rate of 8 ml/hr. Fractions were pooled as indicated by horizontal line. Numbers above each peak refer to the fraction of the total column volume (V_t) at which the material is eluting. Peptide designations are based on previous identification (12).

TABLE I

SEQUENATOR ANALYSES OF HEMERYTHRIN FRAGMENTS[a]

POSITION	PEPTIDE 50-113		PEPTIDE 50-62	
	PTH DERIVATIVE	DETECTION[b]	PTH DERIVATIVE	DETECTION[b]
1	CYS	T	CYS,ARG	H
2	THR	G	THR	G,H
3	GLY	G	GLY	G,H
4	LYS	G	LYS	G,H
5	HIS	S	HIS	H
6	PHE	G	PHE	G,H
7	LEU	G	LEU	G,H
8	ASN	G	ASN	G,H
9	GLU	G	GLU	G,H
10	GLN	G	GLN	G,H
11	VAL	G	VAL	H
12	LEU	G	LEU	H
13	MET	G		

[a]At any given position, all amino acids which were unambiguously identified are reported, with the exception of amino acids carried over from previous steps. Repetitive step-wise yields were 94-97% based on quantitation of gas chromatography data.

[b]G = gas chromatography (22), H = high-performance-liquid chromatography (23), S = spot test (22), T = thin layer chromatography (22).

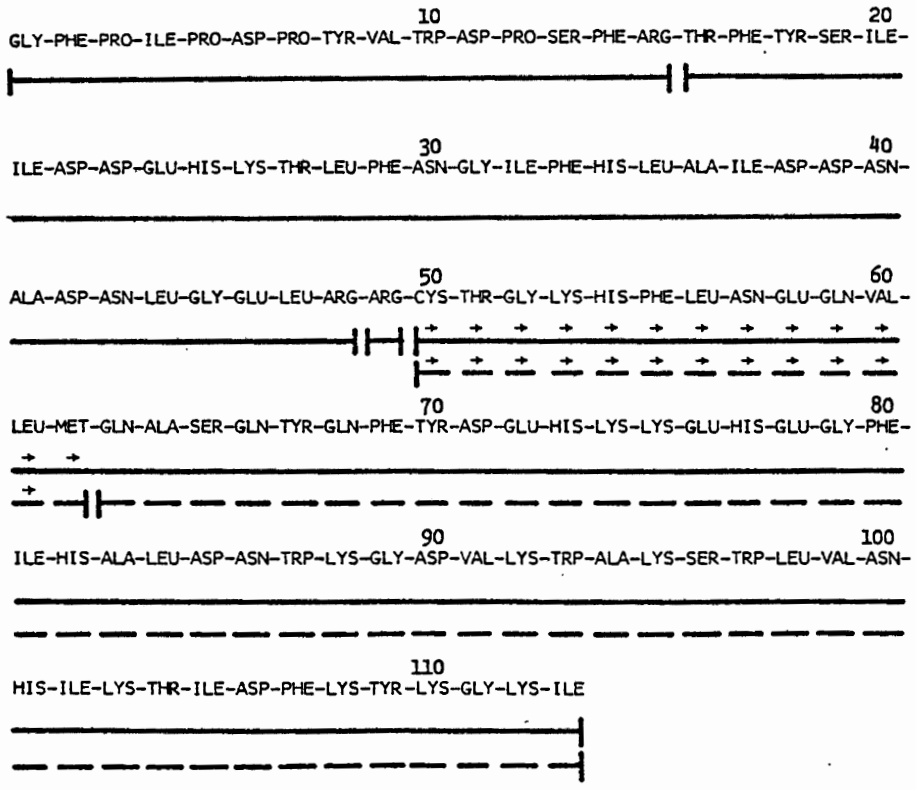


Figure 3. The amino acid sequence of P. gouldii hemerythrin as determined by Subramanian, et al. (8) with a change in the identification of residues 58 and 59. Peptides were produced by specific cleavage of citraconylated hemerythrin with trypsin (——), and peptide 50-113 with CNBr (---). Arrows indicate region sequenced by automated Edman degradations.

TABLE II

AMINO ACID COMPOSITION OF FRAGMENTS GENERATED BY TRYPSIN AND CYANOGEN BROMIDE DIGESTION OF HEMERYTHRIN[a]

AMINO ACID	PEPTIDE 54-74		PEPTIDE 50-62
	GEL PERMEATION[b]	ION EXCHANGE[c]	
LYS	1.6 (1)	2.0 (1)	0.8 (1)
HIS	2.0 (2)	2.2 (2)	0.8 (1)
CYS[d](b)	e	e	1.0 (1)
ARG	e	e	0.1 (0)
ASP+ASN	2.0 (2)	2.2 (2)	0.9 (1)
THR	0.5 (0)	0.4 (0)	0.8 (1)
SER	1.8 (1)	1.1 (1)	0.4 (0)
GLU+GLN	4.4 (6)	5.2 (6)	1.6 (2)
GLY	1.4 (0)	0.9 (0)	1.0 (1)
ALA	1.3 (1)	1.0 (1)	0.5 (0)
VAL	1.3 (1)	0.9 (1)	1.1 (1)
MET[f](c)	0.4 (1)	0.9 (1)	0.8 (1)
ILE	1.0 (0)	0.3 (0)	0.4 (0)
LEU	2.0 (2)	1.8 (2)	1.2 (2)
TYR	1.3 (2)	1.8 (2)	0.5 (0)
PHE	1.5 (2)	2.3 (2)	1.2 (1)
TOTAL	22.5 (21)	23.0 (21)	12.3 (13)

[a]Trypsin and CNBr digests were hydrolyzed in 6 N HCl at 110 C in vacuo for 24 hr and resolved by ion exchange chromatography on Bio-Rad AG 50W-X2 resin according to the method of Spackman, et al. (21).

[b]Gel permeation chromatography was done using Sephadex G-50 and G-25 superfine with 9% formic acid as solvent.

[c]Ion exchange chromatography on resin cited in a, with eluent pyridine-acetate buffer (19).

[d]Cysteine was detected as pyridylethyl cysteine.

[e]Not determined.

[f]Methionine was detected as homoserine lactone in analysis of peptide 50-62.

region beginning with cysteine at position 50 was sequenced cleanly through methionine 62 (Table I). These data unequivocally positioned glutamic acid at residue 58, followed by glutamine at 59. Although the sequence of 50-113 answered the original question of position 58 identity, subsequent digestions were initiated to yield a shorter peptide for both amino acid composition and sequence determination.

PEPTIDE 50-62

Peptide 50-113 was reacted with CNBr and the cleavage products were resolved by gel permeation chromatography into a series of peaks shown in Figure 4. The peak eluting at 0.62 V_T was tentatively assigned to contain CNBr peptide 50-62 on the basis of high 260 nm and low 280 nm absorbances, expected from the presence of PE-cysteine (16) and the absence of tyrosine and tryptophan. Fractions from this peak were pooled, lyophilized and subjected to amino acid sequence determination. Sequenator analysis (Table I) showed that peptide 50-62 was the major component and verified the expected sequence of 12 amino acids beginning with cysteine and ending with the penultimate amino acid, leucine. Glutamic acid was again clearly identified as residue 58 followed by glutamine at 59 (Fig. 5 and 6). The only contaminating peptide apparent in the sequencing was peptide 49-62, arising from incomplete trypsin cleavage between arginine 49 and cysteine 50. This caused a greater amount of carry-over in each sequence cycle than is normally observed.

A sample of CNBr peptide 50-62 was subjected to amino acid analysis (Table II). This material was prepared by reduction of peptide 50-113 prior to CNBr treatment. The amino acid analysis corroborated the sequence

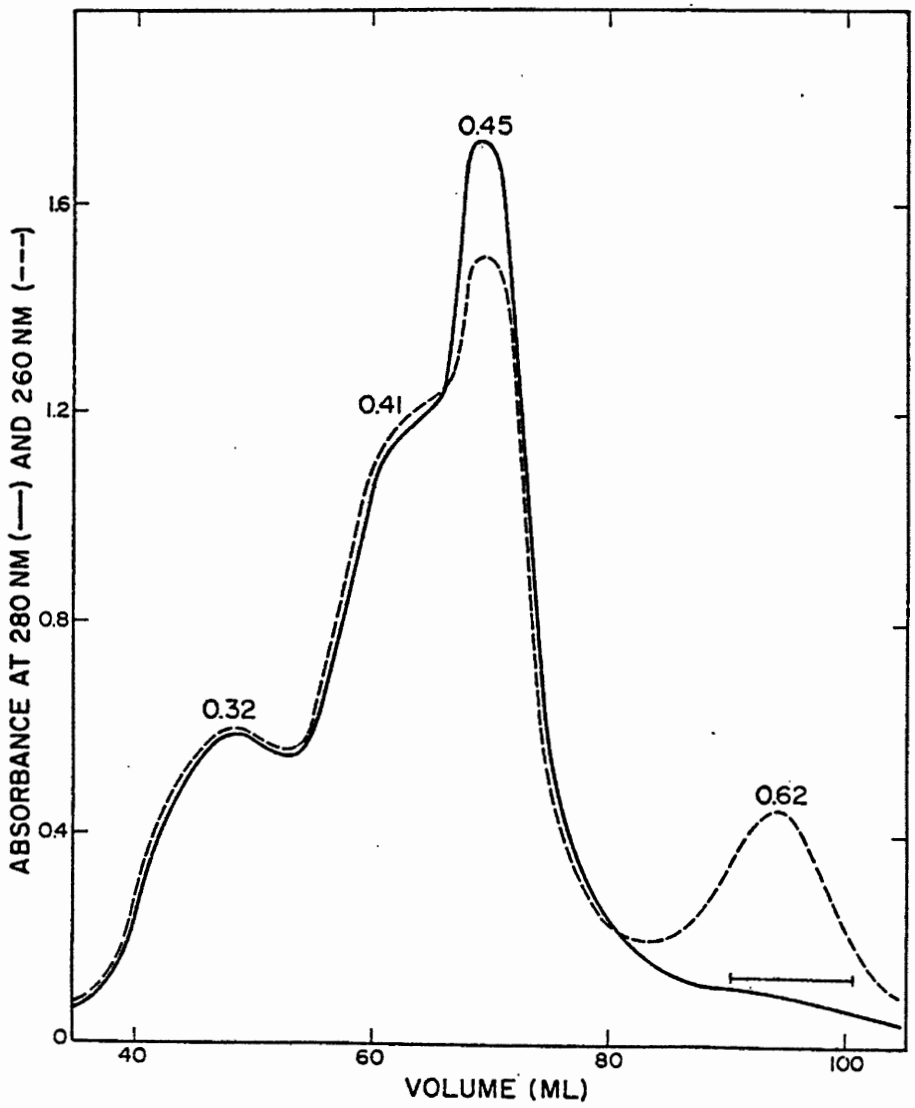


Figure 4. Gel permeation chromatography of CNBr digestion products (16 mg) of peptide 50-113 on Sephadex G-50 superfine. Conditions as described in Figure 2.

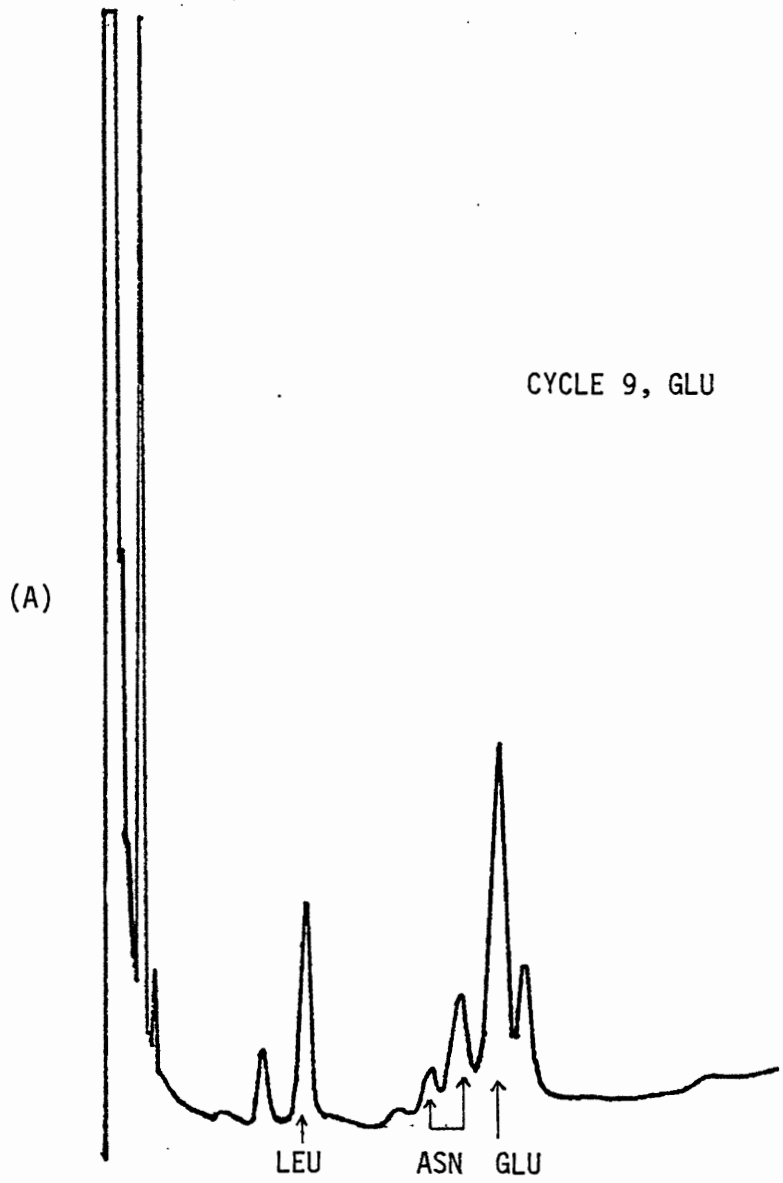


Figure 5a. Gas-liquid chromatography of PTH-derivative from sequenator analysis of peptide 50-62. (A) Glutamic acid is the cycle 9 (residue 58) amino acid. Small carry-over from cycles 7 (LEU) and 8 (ASN) are present.

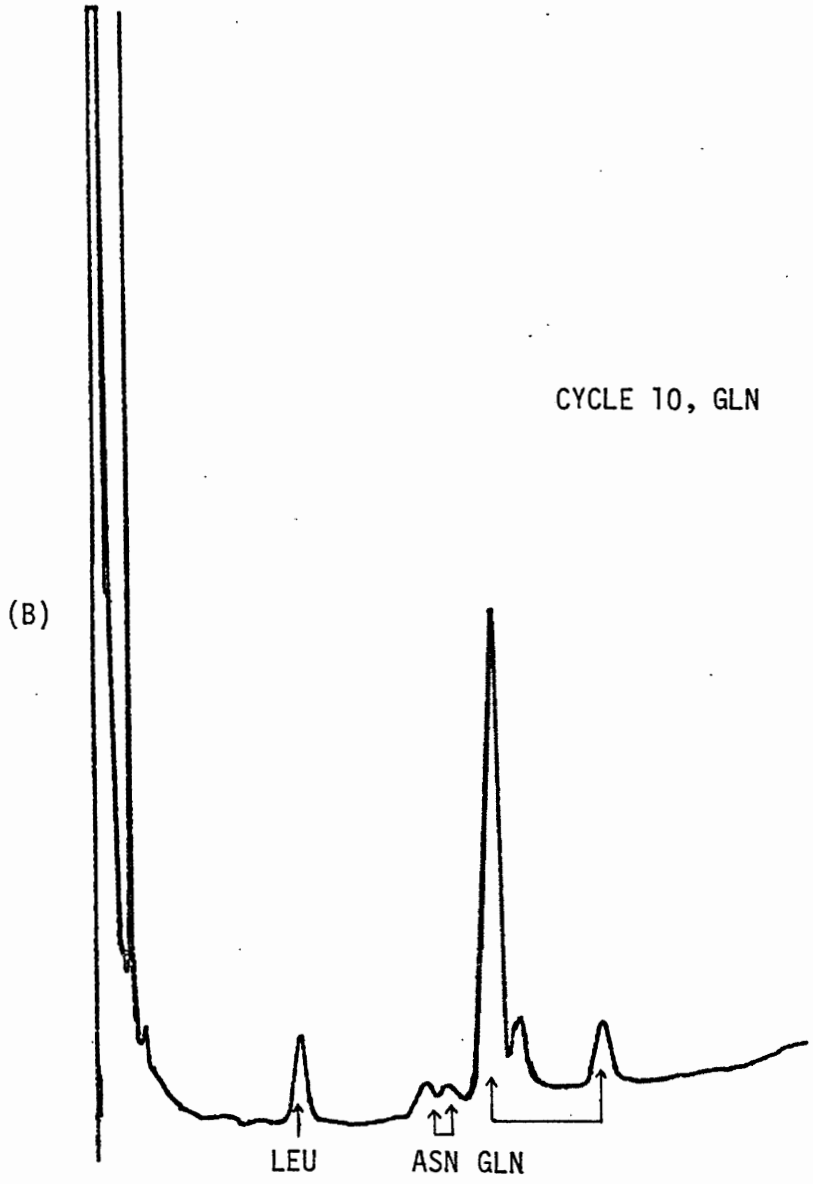


Figure 5b. Gas-liquid chromatography of PTH-derivative from sequenator analysis of peptide 50-62. (B) Glutamine is the cycle 10 (residue 59) amino acid. Intensity of the glutamic acid is characteristic of glutamine. Diminished LEU and ASN carry-over are noted.

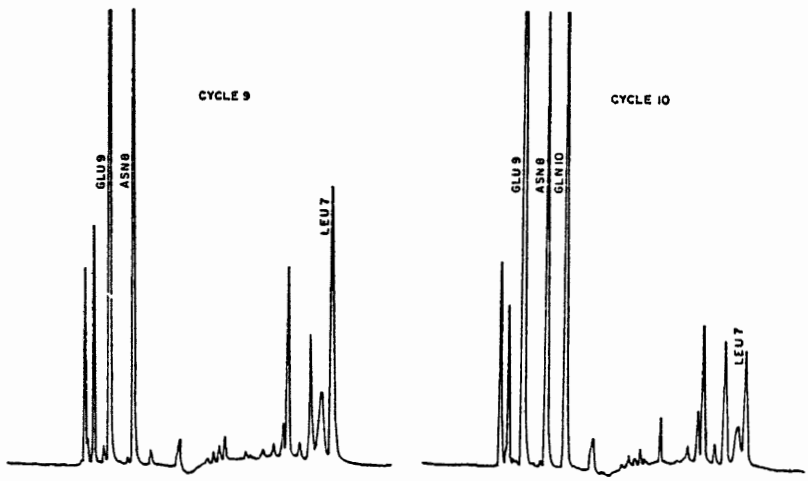


Figure 6. High-performance-liquid chromatograms of PTH-derivatives from sequenator analysis of peptide 50-62. Cycle 9 (residue 58) shows glutamic acid as the new peak. Overlap of Leu and Asn from cycles 7 and 8, respectively, are present. Cycle 10 (residue 59) shows a new peak which corresponds to glutamine. Leu and Asn carry-over is noticed although the Leu intensity is diminished.

obtained for CNBr peptide 50-62. However, serine, alanine, isoleucine and tyrosine occurred to the extent of 0.4-0.5 residue when none was anticipated. Thus, some other peptide material must be present, although at too low a level to appear in the sequenator analysis. The COOH-terminus of peptide 50-62, methionine, was modified in the CNBr reaction to homoserine lactone which is in equilibrium with homoserine. Acid hydrolysis converted all homoserine to the lactone (13) allowing indirect methionine quantitation by amino acid analysis.

The eultion profile of the CNBr hydrolysate (Fig. 4) contained a minimum of three earlier eluting peaks other than the one assigned to peptide 50-62. The majority of this material is probably the other cleavage product, peptide 63-113, in various states of aggregation. However, unreacted starting material may also be present.

PEPTIDE 54-74

Trypsin digestion of peptide 50-113 cleaved at lysyl residues produces a number of smaller fragments which could be partially resolved on a Sephadex G-50 superfine column (Fig. 7). The shoulder at 0.69 V_T was tentatively identified as containing peptide 54-74 on the basis of its tyrosine-like ultraviolet spectrum. This spectrum is compared to the spectrum of the tryptophan rich starting material and model compounds (Fig. 8,9). Fractions exhibiting the tyrosine-rich spectrum were pooled, lyophilized and rechromatographed on a Sephadex G-25 superfine column (Fig. 10). The material eluting at 0.44 V_T was subjected to amino acid analysis (Table II). Comparison of the experimental versus expected amino acid composition indicated considerable contamination. The remaining

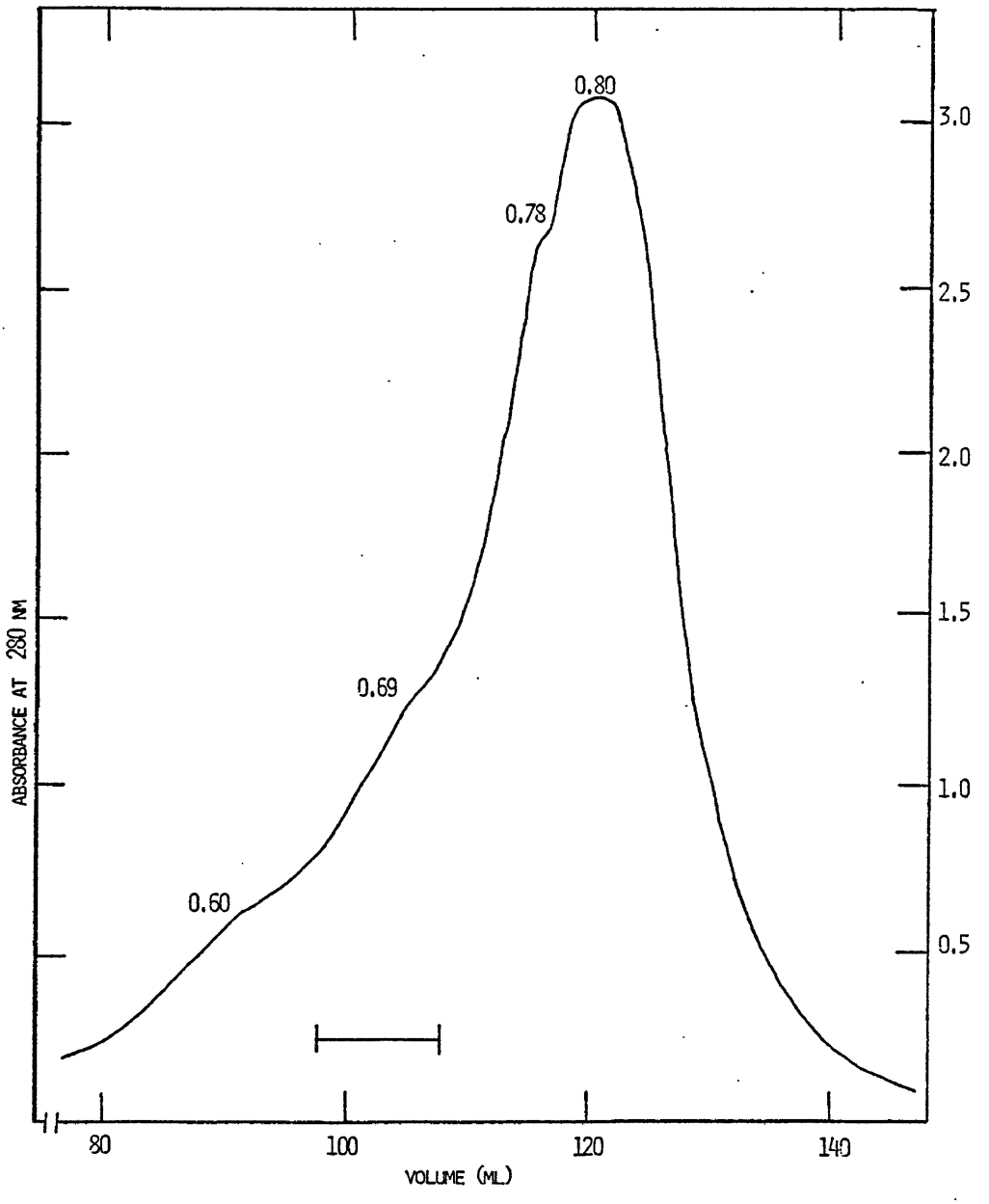


Figure 7. Gel permeation chromatography of trypsin digestion products of peptide 50-113 (48 mg) on Sephadex superfine. Conditions as described in Figure 2.

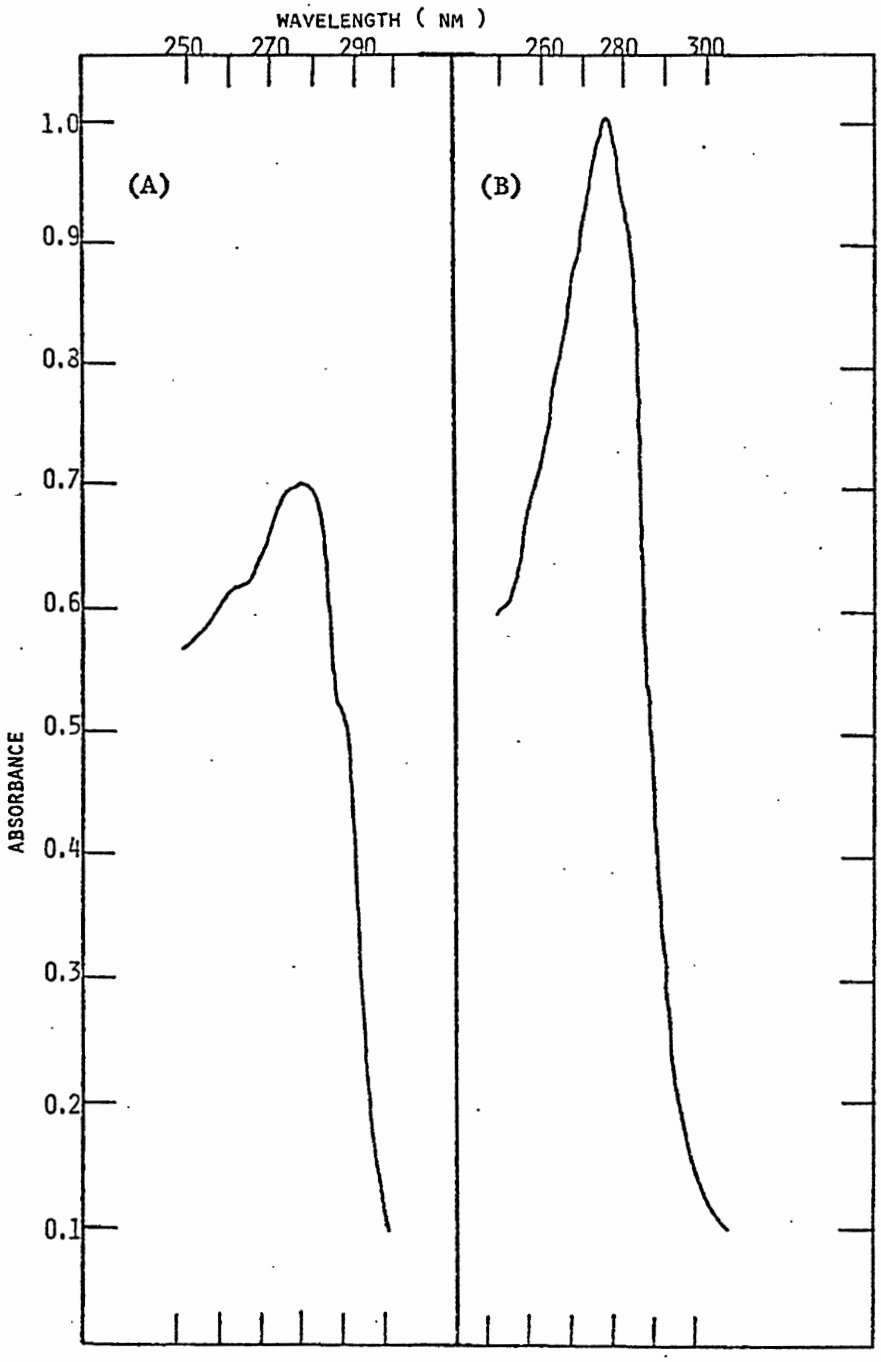


Figure 8. Ultraviolet spectra of P. gouldii hemerythrin peptide 50-113 (A) and peptide 54-74 (B).

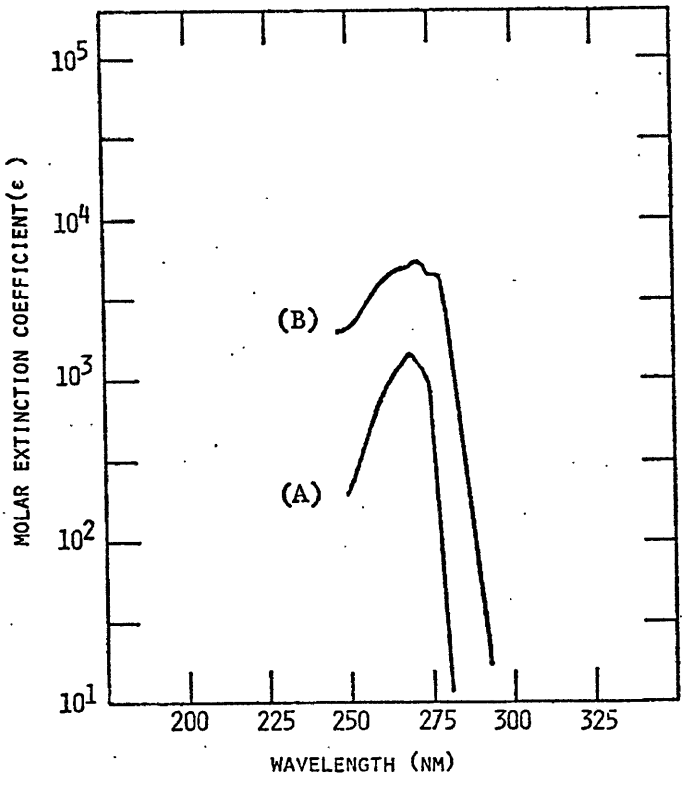


Figure 9. Ultraviolet spectra of N-acetyl-L-tyrosine ethyl ester (A) and N-acetyl-DL-tryptophan methyl ester (B) taken from (28).

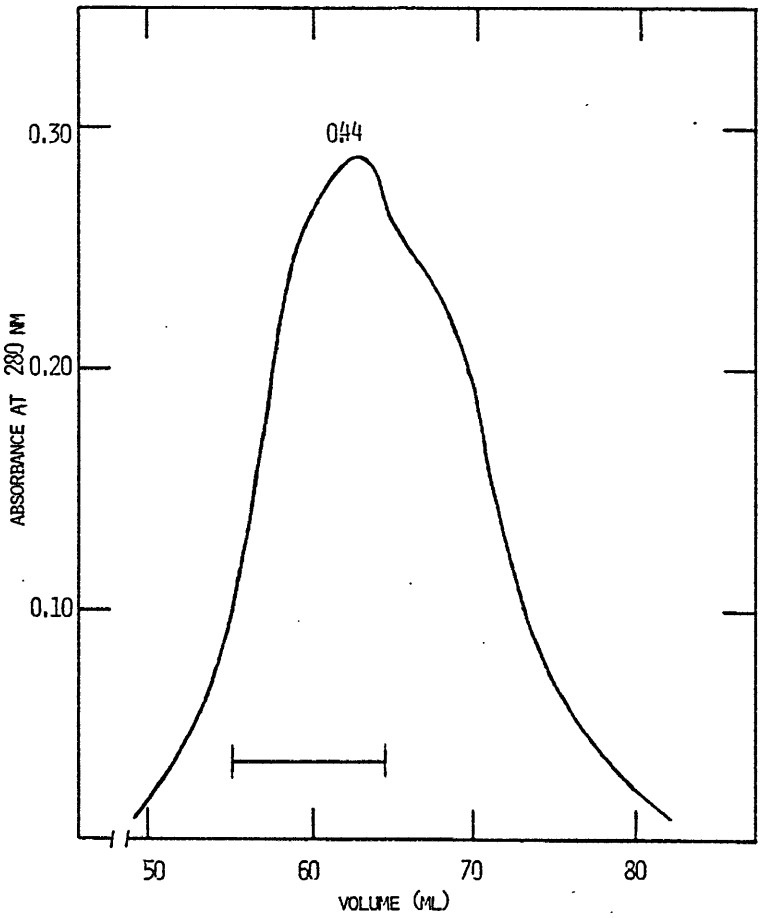


Figure 10. Gel permeation chromatography of peptide 54-74 (5 mg) on Sephadex G-25 superfine. Conditions as described in Figure 2.

fractions were pooled and chromatographed on an AG 50W-X2 ion exchange column. This significantly improved the peptide purity as evidenced by an amino acid analysis more consistent with the theoretical (Table II). However, fractional residues of threonine, lysine, and glycine persisted, suggesting incomplete trypsin cleavage at lysine 53, such that a significant amount of peptide 50-74 remained in the sample. Subsequently, trypsin cleavage was enhanced by inclusion of 2 M urea in the digestion reaction. However, the resulting product, peptide 54-74, had reduced solubility in pyridine-acetate buffer and precipitated during the ion exchange purification. Thus, the insolubility of peptide 54-74 precluded its use in amino acid sequence determination.

DISCUSSION

Trypsin digestion of P. gouldii citraconylated hemerythrin followed by gel permeation chromatography yielded peptide 50-113 in sufficient quantity and purity for sequenator analysis. The peptide's NH_2-terminus was cysteine and the amino acid identities were sequentially determined as indicated in Table I through methionine at position 62. This sequence clearly identified position 58 as glutamic acid. Of importance was verification of methionine-62 since its presence allowed an alternative specific cleavage procedure of peptide 50-113 with CNBr. Since an amino acid sequence determination corroborated by an amino acid composition is accepted as being reliable, formation and isolation of peptide 50-62 was pursued using CNBr cleavage and gel permeation chromatography.

Digestion of tryptic peptide 50-113 with CNBr yielded cleavage products, peptides 50-62 and 63-113. No detectable advantage was obtained from prior mercaptoethanol reduction of peptide 50-113, indicating a low occurrence of methionine sulfone and sulfoxide residues. Sequenator analysis on the fractions pooled and lyophilized from the 0.62 V_T peak of Figure 4 unequivocally assigned glutamic acid and glutamine to positions 58 and 59, respectively, terminating with the penultimate residue at 61, leucine. This sequence is corroborated by amino acid analysis which is in good agreement with the expected composition for peptide 50-62.

Purity of peptides 50-113 and 50-62 is evidenced by lack of extraneous amino acid peaks in the detection chromatograms, indicating

effective cleavage by trypsin and CNBr, respectively, as well as reliable chromatographic resolution of the hydrolysates with Sephadex G-50 superfine.

Attempts at purification and sequence determination of peptide 54-74 from trypsin digestion of peptide 50-113 were abandoned when the peptide precipitated in the pH 3.1 pyridine-acetate ion exchange chromatography buffer and failed to elute from the column. Apparently, the peptide's high percentage of hydrophobic residues (57%) confers a peptide conformation unfavorable to dissolution in the polar solvent. The peptide would be even less soluble at pH 5, the pH required to elute the peptide from the negatively charged column, since it is essentially isoelectric at this pH.

The amino acid sequence determinations of peptides 50-113 and 50-62 unambiguously identified glutamic acid at position 58 followed by glutamine at 59 in P. gouldii hemerythrin. Use of sequenator analysis coupled with direct asparagine and glutamine identification strongly suggests that the proposed sequence is correct. The high reliability of our sequencing method resides in precise regulation of temperature, reagents and pressure at each step of a cycle, enhanced extraction of the thiazolinone intermediate and sensitive detection of the PTH-amino acid derivative by gas-liquid (22) and high-performance-liquid chromatography (23). The greater than 95% recovery which is generally achieved in each derivatization, cleavage and detection cycle minimizes overlap from adjacent residues, making it possible to sequence 30-50 residues of a peptide at a time (22). The early position of glutamic acid 58 and glutamic 59 within the first 10 residues of the peptides which we sequenced lends further credibility to our identification.

The earlier P. gouldii hemerythrin sequence which reported the reverse order of glutamic acid and glutamine at positions 58 and 59 utilized a manual Edman degradation in conjunction with older detection procedures (12). Peptide 56-61, whose partial sequence was previously reported (18), was shown to contain 2 of the 3 acidic residues (asp, glu_2) in the amidated form on leucine aminopeptidase digestion. The manual Edman degradation of peptide 56-61 yielded the sequence, Leu-Asp-Glu-Glu, according to the amino acid analysis from each cycle following alkaline regeneration. Since the amide group is hydrolyzed under alkaline conditions, methods that preserve the amides' chemical integrity were adopted to determine the presence of glutamine in this region. Peptide 56-61, generated by chymotrypsin digestion of peptide 50-113, was subjected to carboxypeptidase A cleavage. This exopeptidase lacks high specificity as indicated in Table III and cleaves from the COOH-terminal until stopped by proline or arginine. Glutamic acid is very slowly released by this exopeptidase whereas, glutamine is rapidly cleaved. Assignment of position 59 as glutamic acid (and by inference position 58 as glutamine) resulted from the slow appearance of glutamic acid and no glutamine on carboxypeptidase A cleavage. More exact kinetic information concerning the rate of glutamic acid release was not included in the sequence report.

The failure of Subramanian and coworkers to observe the release of glutamine after treatment could have been due to experimental difficulties and should not be considered as definitive evidence for the presence of glutamic acid at position 59. The more direct and positive identification of glutamic acid at position 58 and glutamine at position 59 by gas-liquid and high-performance-liquid chromatography in the present investigation

TABLE III

RATES OF HYDROLYSIS OF PEPTIDE SUBSTRATES BY CARBOXYPEPTIDASE A[a]

RAPID	SLOW	VERY SLOW[c]	NOT RELEASED[c]
Tyr	Asn	Gly	Pro
Phe	Ser	Asp	Arg
Trp	Lys[b]	Glu	
Leu		$CySO_3H$	
Ile		S-carboxymethyl Cys	
Met			
Thr			
Gln			
His			
Ala			
Val			

[a]Adapted from R. P. Ambler, Methods Enzymol., 25, 262 (1972). Reported by Light (29).

[b]The rate of release of lysine may be influenced by the pH of the hydrolysis mixture.

[c]The presence of a "very slow" or "not released" amino acid as penultimate residue will generally decrease the rate of release of the COOH-terminal amino acid.

indicates that the previously reported order must have been in error.

The presence of glutamic acid at position 58 of P. gouldii hemerythrin supports the Stenkamp, Sieker and Jensen model (3) for the active site of hemerythrin because it shows complete conservation of the amino acids proposed as iron ligands in the three proteins which have been completely sequenced. The identification of glutamic acid as an iron ligand is chemically sensible as the deprotonated carboxyl group at neutral pH is capable of electron density donation to the electrophilic iron. Knowledge of the iron environment in hemerythrin is important as a potential model for oxygen-activating enzymes as well as for industrial development or modification to a useful synthetic catalyst.

CONCLUSIONS

Knowledge derived from the Stenkamp, Stieker and Jensen active site model of T. dyscritum hemerythrin has potentially diverse applications. Identification of the iron ligand at position 58 as glutamic acid in P. gouldii hemerythrin is, therefore, significant in the support it lends to the Stenkamp model.

Atkins, et al. (24) reported spectral and magnetic similarities between hemerythrin's active center and the iron-containing catalytic center of ribonucleotide reductase. Use of hemerythrin as a model system may be helpful in elucidating the catalytic mechanism of the reductase whose biochemical function is reduction of ribonucleotides to their deoxy form. Ribonucleotide reductase is a strategic regulatory point in DNA synthesis and is under fine feedback control of substrates and allosteric effectors. Agents capable of reductase inhibition are actively sought for their chemotherapeutic value. The most effective inhibitory drugs tested in vitro have been strong iron chelators, i.e., thiosemicarbazones of 1-formylisoquinoline and 2-formylpyrdine, (25) which would likely interfere with general iron metabolism in vivo. Structural information obtained from hemerythrin's iron active site in conjunction with the reductase's known substrate affinity has the potential for yielding a highly specific anticancer drug.

Use of hemerythrin synthetic analogs may improve the performance and longevity of artificial respirators. Currently, a synthetic analog of hemoglobin, Collman's tetraphenyl porphyrin picket fence (26), is utilized

in respirators but a high affinity for CO leads to poisoning of the analog. Since hemerythrin has little affinity for CO, an analog of structural design similar to the protein's active site would be medically beneficial.

Chemical models of oxygen activating proteins are industrially desirable as catalysts capable of high specificity under mild conditions (27). Elucidation of the complex principles underlying oxygen activation is a prerequisite for development of a suitable catalyst. This achievement will be possible only through the concerted efforts among the diverse branches of chemistry. An alternative to the development of synthetic catalysts capable of oxygen activation is modification of the protein allowing immobilization on a solid support. Ideally, such a system would be amenable to stereoselectivity and large-scale operation.

REFERENCES

1. Okamura, M. Y., and Klotz, I. M. (1973) Inorganic Biochem, 1, 320.

2. Ward, K. B., Hendrickson, W. A., and Klippenstein, G. L. (1975) Nature (London), 257, 818.

3. Stenkamp, R. E., Sieker, L. C., and Jensen, L. H. (1976) Proc. Natl. Acad. Sci. USA, 73, 349

4. Loehr, T. M., Dunn, J. B. R., and Loehr, J. S. (1976), Handbook of Biochemistry and Molecular Biology, 3rd Ed., 3, 588.

5. Loehr, J. S., Lammers, P. J., Brimhall, B., and Hermodson, M. A., (1978) J. Biol. Chem., in press.

6. Hendrickson, W. A. (1978) Naval Research News, 31, 1

7. Klippenstein, G. L., Cote, J. L., and Ludlam, S. E. (1976) Biochemistry, 15, 1128.

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

9. Edman, P. and Begg, G. (1967) European J. Biochem., 1, 80.

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

11. Hermodson, M. A., Ericsson, L. H., Neurath, H., and Walsh, K. A. (1073) Biochemistry, 12, 3146.

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

13. Gross, E. (1967) Methods Enzymol., 11, 238.

14. Weihing, R. R., and Korn, E. D. (1972) Biochemistry, 11, 1539.

15. Klotz, I. M., Klotz, T. A., and Feiss, H. A. (1957) Arch. Biochem. Biophys., 68, 284.

16. Friedman, M., and Krull, L. H. (1969) Biochem. Biophys. Res. Commun., 37, 630.

17. Atassi, M. Z., and Habeeb, A. F. S. A. (1972) Methods Enzymol., 25, 546.

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

19. Jones, R. T. (1970) Methods Biochem. Anal., 18, 205.

20. West, R. (1965) J. Chem. Educ., 42, 386.

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

22. Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972) Biochemistry, 11, 4493.

23. Zimmerman, C. L., Apella, E., and Pisano, J. J. (1977) Analyt. Biochem., 77, 569.

24. Atkins, C. L., Thelander, L., Reichard, P., and Lang, G. (1973) J. Biol. Chem., 248, 7464.

25. Hogenkamp, H. P. C., and Sando, G. N. (1974) Struct. and Bond., 20, 23.

26. Collman, J. P., and Reed, C. A. J. (1973) Amer. Chem. Soc., 95, 2048.

27. Sheldon, R. A., and Kochi, J. K. (1976) Adv. in Catal., 25, 272.

28. Weinheim, and Butterworths (1966) DMS UV Atlas of Organic Compounds, 2, J1/1 and J1/3.

29. Light, A. (1974) Proteins: Structure and Function, Prentice-Hall, Inc., N. J., 119.

30. Hendrickson, W. A., Klippenstein, G. L., and Ward, K. B. (1975) Proc. Natl. Acad. Sci. USA, 72, 2160.

31. Dunn, J. B. R., Shriver, D. F., and Klotz, I. M. (1975) Biochemistry, 14, 2689.

32. Breslow, E. (1973) in Inorganic Biochemistry, Eichhorn, G. L., Ed., New York, N.Y., Elsevier, 227.