THE AMINO ACID SEQUENCE OF THE β CHAIN OF ASIAN ELEPHANT, *ELEPHAS MAXIMUS*, HEMOGLOBIN

by

Thomas S. Vedvick, B. S., M. S.

A THESIS

Presented to the Department of Biochemistry
and the Graduate Division of the University of Oregon Medical School
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
June 1972

APPROVED [redacted]

(Professor in Charge of Thesis)

[redacted]

(Chairman, Graduate Council)

Dedication

To Claudia

Acknowledgements

A graduate education has been considered by many to be an apprenticeship. It has been a privilege apprenticing under Dr. Richard T. Jones and for this opportunity I am extremely grateful. The instruction, guidance and encouragement which he gave made this thesis possible.

It is a pleasure to acknowledge the helpful suggestions of Drs. Bernadine Brimhall, Michael Coates, and Peter Stenzel throughout this study. A special thank you is extended to Miss Marie Duerst and Mr. Jim Joyce for their excellent technical assistance.

The interest shown in this project by Drs. James Metcalfe and Dharim Dhindsa is appreciated. The generous samples of elephant blood from both the Portland and the San Diego Zoos were also appreciated.

I also wish to thank Mrs. Mary Buck for her help in typing the thesis.

Finally, I thank my wife, Claudia, for her help, encouragement and interest in this project.

Financial support through a National Institute of Health predoctoral fellowship, 5 T01 GM1200, is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. Statement of the Problem

Nearly 20 years ago, in 1954, after a decade of pioneering research, Frederick Sanger (1) and his colleagues at Cambridge University were able to write the first structural formula of a protein; the hormone insulin. Techniques developed by Sanger's group were generally applicable to all proteins. Consequently, investigations in numerous laboratories on many proteins were either initiated or enhanced because of these developments in protein chemistry. Since that time many refinements, improvements, and additions have been made to the arsenal of tools available to the protein chemist. Today, the task of sequencing a protein is formidable but not insurmountable.

The purpose of this thesis is to elucidate the primary structure of the β chain of Asian elephant, *Elephas maximus*, hemoglobin. This hemoglobin was chosen for investigation because it provides the specific model necessary for testing the following: 1. The direct comparison of the amino acid sequence of the β chain of elephant hemoglobin with other known sequences may explain the high oxygen affinity of this hemoglobin (2). 2. Elephant hemoglobin sequence studies may also clarify the tenuous taxonomic relationship which exists at present between the two living elephant species (3). 3. The relationship of this species to other animals may be extended by comparisons of known sequences.

B. Structure and Function

1. Background

Hemoglobin is one of the most widely studied globular proteins because of its important physiological role, relative homogeneity, and availability. The molecular weight of human hemoglobin is approximately 64,500 (4). The dimensions of this molecule are 64 x 55 x 50 Å (5). With the exception of the Agnatha, (6-8) all known vertebrate hemoglobins are tetrameric. The molecule consists of two pairs of folded chains called alpha and beta. These pairs differ in amino acid composition. The alpha chains contain 141 amino acid residues and the beta chains 146 amino acid residues. The secondary and tertiary structure of the pairs of chains are similar, but not identical (9). The major configurational difference of these two pairs of polypeptides is that the alpha chains lack one of the alpha helices possessed by the beta chain. As in myoglobin, beta chains have eight regions (called A,B,C,D, etc.) in which the polypeptide chain has an alpha helical configuration (10). Between these helices are non helical sections in which the polypeptide chain changes its direction in space. The alpha chain does not have a D helix but has an extended non-helical region, called C-E.

All known vertebrate hemoglobins contain identical prosthetic heme groups, which are bound intimately with the protein moeity (11). These heme groups are situated in a pocket between the E and F helices (10) of each chain. For this reason the four iron

atoms of the heme groups are relatively far apart. The heme group is attached by a covalent bond to a histidine group of the polypeptide chain. This histidine group is in the eighty-seventh position of the alpha chain and the ninety-second position of the beta chain.

Kitchen (12) suggests that because all hemoglobins have identical prosthetic heme groups, it is the protein moeity that has the ability to alter the molecular environment surrounding the hemes. This may explain how different hemoglobins can function optimally under a wide variety of oxygen tensions, temperatures, and pH.

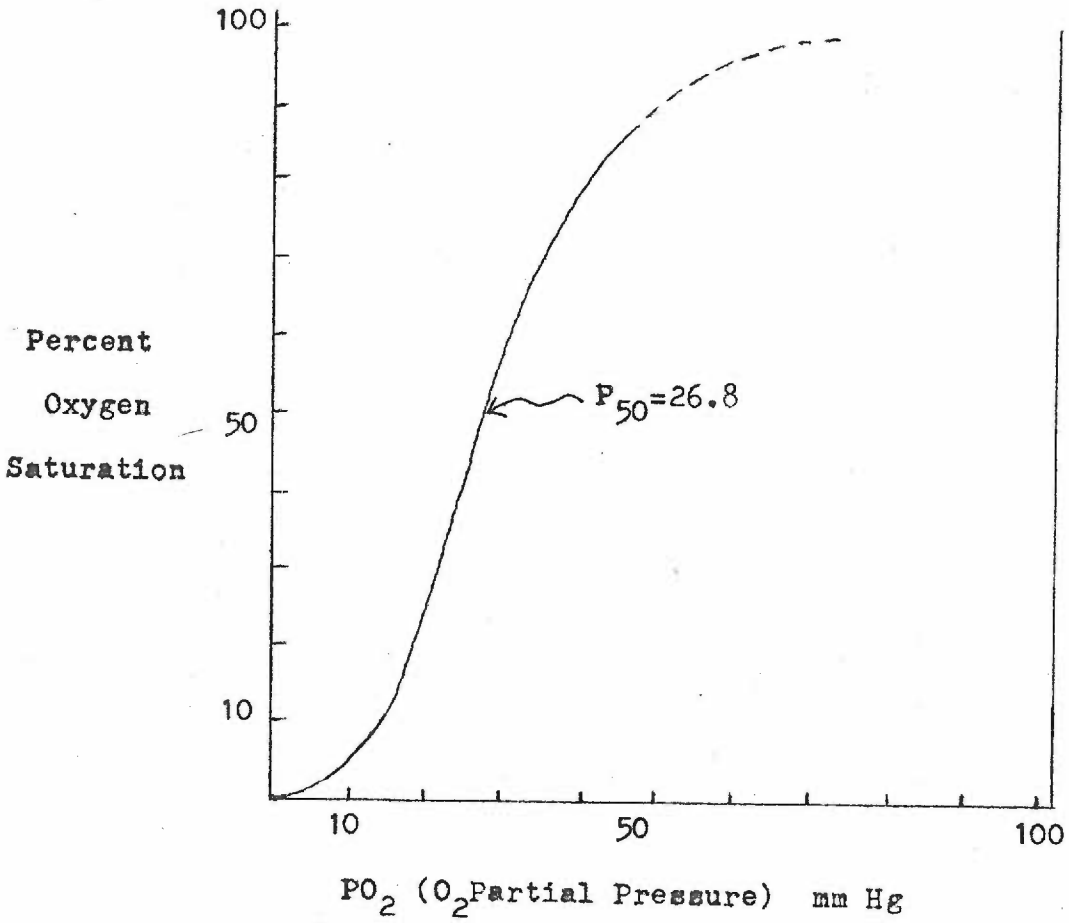
2. Function of Hemoglobin

The maintenance of cell functions and of life is dependent upon the continual supply of adequate amounts of oxygen. Hemoglobin is the respiratory protein of the red blood cells which carries this life-sustaining oxygen from the lungs to the tissues. In the cells where oxygen is utilized, carbon dioxide is produced. Hemoglobin also functions to transport carbon dioxide from the tissues to the lungs.

The oxygen affinity of hemoglobin increases with the degree of its saturation, the oxygen equilibrium curve has a sigmoidal shape, illustrated in Figure 1. This sigmoid shape is attributed to interaction between the four hemes in one molecule of hemoglobin. The position of the equilibrium curve is influenced by pH, known as the Bohr effect (13) and organic phosphates, especially 2,3 diphosphoglycerate (14-16). The affinities of hemoglobin for oxygen, carbon dioxide, hydrogen ion, and 2,3 diphosphoglycerate are interdependent,

Figure 1

Oxygen Affinity Curve of Asian Elephant Hemoglobin


100
50
10
Percent
Oxygen
Saturation
P_{50}=26.8
10
50
100
PO_2 (O_2Partial Pressure) mm Hg

so that the concentrations of any one of these ligands affects its combination with the others (17). The oxygenated and deoxygenated forms of hemoglobin have differential binding affinities for different effector ligands. The deoxygenated form of hemoglobin binds 2,3 diphosphoglycerate, carbon dioxide, and hydrogen ion firmly and when it does so it has a low affinity for oxygen. Conversely, the oxygenated form of hemoglobin binds oxygen well but has a low binding affinity for 2,3 diphosphoglycerate, carbon dioxide, and hydrogen ion.

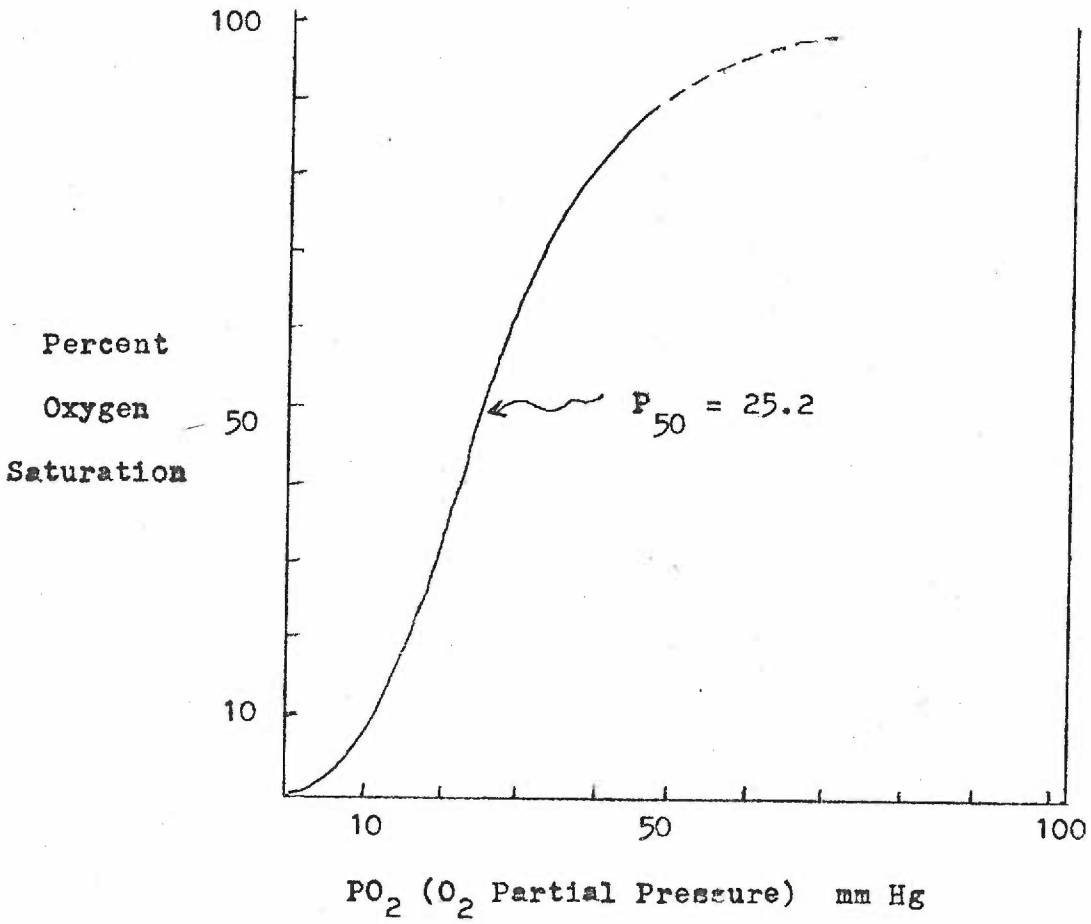
3. Elephant Hemoglobin

Elephants are the largest living land dwelling mammal. Prior to the work of Dhindsa et al. (2,18) little was known regarding the hemoglobin of the Asian elephant. Schmitt (19) had shown that Asian elephant hemoglobin, on starch gel electrophoresis, migrated faster than normal human A hemoglobin. Riggs (20) had determined the amino acid composition of Asian elephant globin. No attempt was made by Riggs to separate the α and β chains of Asian elephant globin.

The studies by Dhindsa et al. (2) revealed a mean P_{50} (the partial pressure of oxygen at which 50% of the whole blood's hemoglobin is saturated with oxygen) of 25.2 ± 0.5 mm Hg for Asian elephant hemoglobin. This investigation was completed on four adult female Asian elephants and the average oxygen affinity curve is shown in Figure 2. This P_{50} value, 25.2 mm Hg, is lower than the normal human P_{50} value of 26.8 mm Hg. Knowledge of these data prompted us to determine the sequence of the β chain of Asian

Figure 2

Oxygen Affinity Curve of Asian Elephant Hemoglobin


100
50
10
Percent
Oxygen
Saturation
P_{50} = 25.2
10
50
100
PO_2 (O_2 Partial Pressure) mm Hg

elephant hemoglobin. The purpose was to attempt an explanation for this high oxygen affinity by comparing the sequence of this hemoglobin with those hemoglobin sequences already reported (21).

4. Animal Hemoglobin Sequences

The structure, function, and development of mammalian hemoglobins can be related by a comparison of the amino acid sequences. To date, approximately 50 vertebrate hemoglobin chains have been partially or completely sequenced (21-36). Nearly all vertebrate hemoglobins studied are tetramers containing two alpha chains and two beta chains. The class Agnatha, which includes the lamprey eel as well as other primitive vertebrates, is the exception to this rule (31,32). The hemoglobin of the lamprey is characterized by a single polypeptide chain having a molecular weight of 17,500 with a single heme group (12). The oxygen equilibrium curve is hyperbolic and not sigmoid shaped as in the other vertebrate hemoglobins.

Most vertebrate hemoglobins consist of two pairs of chains, $\alpha_2\beta_2$. However, Tsuyuki and Ronald (37) have shown that several species of salmon contain tetrameric hemoglobin with 3 or 4 different chains. One species has at least eight different polypeptides which do combine to give mixed tetramers (i.e., $\alpha_2\beta\beta'$, $\alpha\alpha'\beta\beta'$, $\alpha\alpha'\beta_2$, etc.). Many vertebrate hemoglobin chain sequences have been or are being determined. These include: amphibians, Bullfrog (33), *Rana catesbeiana*, Frog (34), *Rana esculenta*, and Newt (38), *Taricha granulosa*; two marsupials, Kangaroo (25), *Macropus rufus* and Opossum (39), *Didelphis virginiana* and a host of other mammals including dog (35),

cow (27), sheep (28), llama (29), pig (30), horse (40), and others (21). Despite this apparent heterologous mixture of animal hemoglobin chain sequences, all whose three-dimensional structures have been determined are strikingly similar (41). In aligning these related globin sequences, six positions appear to be invariant. Four of these positions which do not appear to change are located near the site of the heme group. This suggests that the binding of the heme, the functional center of the molecule, is important for the maintenance of a useful, oxygen transporting molecule. A comparison of the elephant β chain sequence with those residues known to be involved in the heme contact region may be helpful in explaining the high affinity for oxygen that this hemoglobin possesses.

5. Abnormal Human Hemoglobins

The screening of blood samples all over the world has led to the discovery of over 100 mutant hemoglobins (42-44). Many of these abnormal hemoglobins have been shown to have altered functions. Some abnormal hemoglobins have a high affinity for oxygen. Examples include: 1. Hemoglobin Zürich (45,46), which has a substitution in the heme contact region; 2. Hemoglobins J Capetown (47), Yakima (48), and Kempsey (49), all of which have substitutions in the $\alpha_1\beta_2$ contact; and 3. Hemoglobin Little Rock (50), which has an amino acid substitution in the proposed binding site for 2,3 diphosphoglycerate (DPG). These abnormal hemoglobins have all been reported to have higher than normal oxygen affinities. These examples provide

alternate sites where amino acid substitutions may cause an increase in oxygen affinity. Namely, the α_1-β_2 contact region and DPG binding site must also be carefully examined in addition to the heme contacts.

C. Evolving Hemoglobin

A time scale for the origin of the earth and the evolution of many organisms has been proposed from fossil evidence and radioactive dating. Unfortunately, direct evidence of a species divergence is rarely found in fossil form. Biochemical evidence does permit the deduction of the change at the time of a species divergence (21). The information recorded in protein sequences is often used for the construction of phylogenetic trees (51-55) which show the relatedness of one species with another. The phylogenetic trees provide detailed information on patterns of protein-gene evolution among species.

Aguirre (3) has described in detail the known fossil evolutionary history of the elephant. There is presently much confusion among the fossil experts (56,57) in the systematics of the family Elephantidae (i.e., the common ancestor of the African elephant, *Loxodonta africana*, and the Asian elephant, *Elephas maximus*, is unclear). Perhaps, in long range plans, the comparison of the amino acid sequences of the hemoglobins of these two currently living species of elephants may clarify the present void in man's knowledge.

The direct comparison of the beta chain of Asian elephant hemoglobin with other previously determined hemoglobin chain sequences should elucidate the phylogenetic relationship of this species. This

should be a useful and independent test of the phylogenetic trees produced by classical taxonomic considerations.

D. Comments

More is known about the secondary, tertiary, and quaternary structure of hemoglobin than any other protein. Because this biologically active protein has a similar function regardless of species, that of oxygen transport, meaningful comparisons can be made of differences in primary sequences. Therefore, the direct comparison of the β chain sequence of Asian elephant hemoglobin with both sequences of other animal hemoglobins and the hemoglobinopathies may result in a better understanding of the high oxygen affinity of this molecule. Elephant hemoglobin has been shown to have an oxygen affinity much higher than normal human hemoglobin (2).

II. MATERIALS AND METHODS

A. Blood Samples

Samples were obtained from six adult female elephants by vein puncture. Two of the samples were generously supplied by the Portland Zoo. The other four samples were obtained from the San Diego Zoo in collaboration with Dhindsa et al. (2). These samples were obtained by immobilizing the elephants with M 99 (Oripavine (58) injected intramuscularly. The amount of M 99 used for each animal was approximately 1 mg/500 kg of body weight. Within 30 minutes after the administration of M 99 the animals laid down and were immobile but continued breathing regularly. Fifty to 100 ml of blood were withdrawn from an ear vein of each animal into syringes containing a solution of heparin and sodium fluoride to prevent coagulation. The blood was thoroughly mixed with the heparin solution and placed in ice water. Following these procedures each elephant was given an intravenous injection of M 285 (Cyprenorphine) to antagonize the effect of M 99. Within a few minutes of the M 285 administration each animal stood on its feet and walked away in a coordinated manner.

B. Preparation of Hemoglobin

The heparinized whole blood was centrifuged at 2,000 x g for twenty minutes and the plasma was removed by decanting. The red blood cells were washed three times with an isotonic saline solution of 7.65 gm sodium chloride (NaCl), 1.16 gm sodium phosphate dibasic (Na_2HPO_4), and 0.18 gm sodium phosphate monobasic (NaH_2PO_4) diluted

to one liter with distilled, deionized water. The red blood cells and saline solution were mixed thoroughly and centrifuged at 2,000 x g for ten minutes. The top or saline layer was removed and the process repeated. Hemolysates were prepared by mixing 1 volume of washed, packed erythrocytes, 1 volume of water, and 0.4 volume of toluene. This mixture was shaken vigorously for 2 minutes and centrifuged at 10,000 x g for thirty minutes. The hemoglobin solution was removed by gently aspirating into a syringe. Care was taken to keep the tip of the needle below the uppermost toluene layer but above the packed cell ghosts which were on the bottom of the centrifuge tube.

The concentration of the hemoglobin sample was determined by the cyanmethemoglobin method of Drabkin (59) at 540 mμ. The optical density was then multiplied by the color factor 1.44 and also by any dilution factor necessary to determine the solution optical density. This yielded the concentration of hemoglobin in terms of mg per ml. (see eq. 1)

$$\text{O.D.} \times 1.44 \times \text{dil. factor} = \text{mg Hb/ml} \qquad (1)$$

C. Separation of Hemoglobin

Hemolysate (500 mg) was applied to a 3.6 x 35 cm column of Amberlite IRC-50 (Bio-Rex 70, Bio Rad Laboratories) according to the procedure of Jones and Schroeder (60). The column had been equilibrated with 8 liters of developer #6. The column was developed at 60 ml/hour and was maintained at a temperature of 7°C except for the

last hemoglobin zone which was eluted from the column by raising the temperature to 22°C.

Starch gel electrophoresis was done at 6V/cm for four hours with pH 8.3, Tris EDTA-borate buffer (61). Samples with a hemoglobin concentration of 10 mg/ml were applied to the starch gels. The gels were stained with benzidine (62).

D. Preparation of Globin

The hemoglobin was converted to globin by precipitation with cold acid-acetone essentially as described by Anson and Mirsky (63). The hemoglobin was dialysed overnight against distilled, deionized water and then diluted with water to a concentration of 40 mg/ml. This hemoglobin solution was added as a fine spray from a 27 gauge hypodermic needle to the acid-acetone (1 ml of 1 N HCl per 100 ml acetone) cooled in a dry ice bath. The globin mixture was stirred for approximately 30 minutes and then centrifuged at 10,000 x g for 15 minutes at -10°C. The precipitated globin was washed once with cold acid-acetone and centrifuged again. The acid-acetone was decanted off and the globin precipitate was dissolved in water and lyophilized.

E. Chain Separation

Attempts to separate the α and β chains by the method of Clegg et al. (64) were unsuccessful. However, the α and β chains could be separated by a slight modification of the Dintzis (65) method. The modifications are that 16.1 ml of redistilled pyridine plus 98.7 ml of concentrated formic acid were mixed and diluted to one liter

with water. This solution, called Buffer B, was used as made without adjusting the pH. The pH of this buffer was between 2.44 and 2.51. Buffer A was prepared by diluting Buffer B 1:20 with water. The pH of Buffer A was between 2.67 and 2.70. A carboxymethyl cellulose column (cm-52) of 0.9 x 20 cm was equilibrated with Buffer A for several hours, until the effluent pH was identical to the initial buffer. Approximately 100 mg of elephant globin dialysed against Buffer A was applied to the column. The column was loaded by layering the sample onto the top of the column resin and forcing the sample into the resin by air pressure. Buffer A was layered carefully over the resin bed and a gradient consisting of 250 ml each of Buffers A and B was used. The flow rate was 15 ml per hour and fractions of 5 ml each were collected. The isolated α and β chains were chromatographed on a Sephadex G-25 column in order to remove urea and other contaminants and lyophilized.

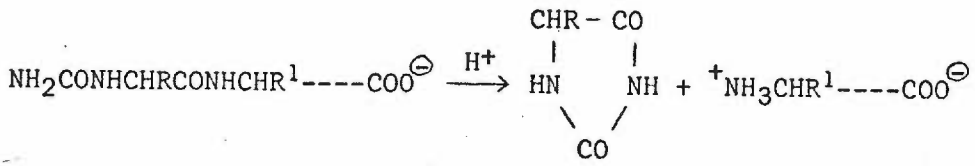
F. Chemical Treatment of the Asian Elephant β Chain

1. N-terminal β Chain Determination

The N-terminal amino acid of the β chain was determined by the method of Stark (66,67) without modification. Fifty mg of elephant β chain was treated with 250 mg of potassium cyanate in 8 M urea adjusted to pH 8 with n-ethylmorpholine and glacial acetic acid. The reaction is illustrated in equation 2.

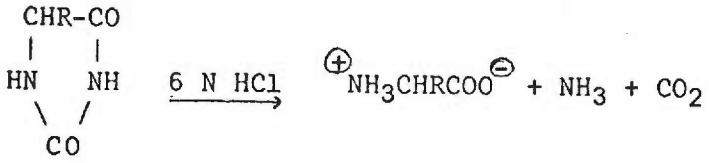
$$\overset{\oplus}{N}H_3CHRCONHCHR^1\text{-----}COO^{\ominus} + \overset{\ominus}{N}CO \xrightarrow{pH\ 8} NH_2CONHCHRCONHCHR^1\text{-----}CO\overset{\ominus}{O}_2 \qquad (2)$$

The carbamylated protein was separated from the unused reagents by gel filtration on a Sephadex G-25 column and converted to the hydantoin. The cyclization was performed by dissolving the carbamylated protein in 2 ml of 50% acetic acid. The sample was evacuated and sealed in a glass tube. The tube was immersed in a boiling water bath for one hour. This converted the carbamylated protein to the hydantoin as shown in reaction 3.



(3)

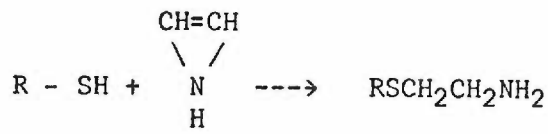
The hydantoin was isolated from a column of Dowex 50 W-X2 (0.9 x 10 cm, 200-400 mesh). The purified hydantoin was hydrolysed with 6 N HCl at 110° for 22 hours by the procedure of Spackman, Stein, and Moore (68) which results in converting the hydantoin to the free amino acid as shown in equation 4.



(4)

2. S-aminoethylation

Lyophilized β chains were usually aminoethylated. The chemical reaction involved in the conversion of cysteine residues to S-(β aminoethyl) cysteine residues is illustrated in equation 5 (69).



$$R - SH + \begin{matrix} CH{=}CH \\ \backslash\ / \\ N \\ H \end{matrix} \longrightarrow RSCH_2CH_2NH_2 \qquad (5)$$

The procedure used for S aminoethylation is essentially that of Cole (70) and Raftery and Cole (71) which was performed in the following manner. Approximately 100 mg of β chain was dissolved in 10 ml of 8 M urea and 1 gm of 2-amino-2 (hydroxymethyl)-1,3 propanediol (Tris). The pH was adjusted to 9.0 with hydrochloric acid. This solution was stirred continuously for one hour after the addition of 0.2 ml of mercaptoethanol, following which 1 ml of ethyleneimine was added and the solution was stirred for a second hour. At the conclusion of this time the pH was adjusted to 3.0 with concentrated hydrochloric acid and the aminoethylated elephant β chain was separated from the urea and unused reagents by passage through a 2.5 x 50 cm column of Sephadex G-25 equilibrated and eluted with 0.2 N acetic acid. The aminoethylated protein was then lyophilized.

3. Amino Acid Analysis of Aminoethylated β Chain

Six aliquots of 2.5 mg each of aminoethylated chain were dissolved in 2 ml of 6 N HCl containing 9 mg of phenol per 100 ml (72) and hydrolyzed in vacuo for 22, 48 or 72 hours. Quantitative amino acid analyses were made on each sample with a Spinco Model 120B Amino Acid Analyzer as modified by Jones and Weiss (73).

4. ε-amino Blocking Reagent

Several reversible blocking reagents for amino groups have

been reported (74-78). Singhal and Atassi (79) found that citraconic anhydride was the most satisfactory for their work on myoglobin. Therefore, elephant aminoethylated β chain was reacted with citraconic anhydride similarly to the methods of Singhal and Atassi (79). One gram of β chain was dissolved in 50 ml of water adjusted to pH 8.2 on the pH stat. Ten aliquots of 100 µl each of citraconic anhydride were added at 20 minute intervals to the stirred β-chain solution. The pH was maintained between 8.5 and 9.0 by the addition of 5 N NaOH. After adding the citraconic anhydride, the mixture was allowed to stand for 1 to 2 hours. This solution was dialysed against water, which had been adjusted to pH 8.5 with NaOH. After dialysis, the sample was ready for enzymatic digestion with trypsin.

G. Enzymatic Hydrolysis of β Chain

Initial enzymatic digestion of chemically treated elephant β chain, as previously described, was done with trypsin by the method of Baglioni (80). All samples were hydrolysed in the following manner: Chains were dissolved in distilled, deionized water in the proportion of 5 mg of chain per 1 ml of water. The pH was adjusted to a value between 8.0 and 9.0 by the dropwise addition of trimethylamine. Trypsin was added in a ratio of 1 mg trypsin per 100 mg of protein. After one-half hour the pH was checked and adjusted as needed to maintain it between 8.0 and 9.0. The elephant β chain was hydrolysed for a total of two hours. After hydrolysis the pH was adjusted to between 2.0 and 3.0 with 1 N HCl. The tryptic peptides

in this solution were separated by ion exchange chromatography.

H. Separation and Further Fragmentation of Tryptic Peptides

1. Separation by Ion Exchange

Tryptic peptides have been separated from each other by ion exchange chromatography (81-85). In our laboratory a sample equivalent to the load provided by the tryptic peptides of 50-100 mg of aminoethylated peptide chain was adjusted to a volume of about 1 ml by rotary evaporation. The pH was adjusted to one or two pH units below that of the initial chromatographic buffer. The sample was layered onto the top of a 0.9 x 20 cm column of Bio Rad Aminex A-5 resin. It was forced into the column by air pressure. The top of the column was then covered with initial buffer and development of the column was initiated. The gradient was produced in a two-chamber system with 250 ml of 0.2 M pyridine-acetic acid, pH 5.0, in the reservoir. The column flow rate was 30 ml per hour. Of the effluent, 3 ml per hour were diverted to a mixing block where it was mixed with ninhydrin and citrate buffer, pH 5.25, each flowing at 12 ml per hour. This reaction mixture was heated by passage through a coil of Teflon tubing 50 feet in length (size AWG 22) which was maintained at 100°C in a refluxing water bath. The resulting reaction products were detected photometrically at 570 mμ in a flow photometer and recorded directly with a strip chart recorder. The remainder of the column effluent was diverted to a fraction collector.

The peptides obtained from this initial separation on a

column of Aminex A-5 were further purified by rechromatography on a column (0.9 x 60 cm) of Aminex AG 50 W-X2 (270-325 mesh, Bio Rad Laboratories) with the use of the same developer, flow rates, and automatic detection system as previously described.

2. Separation by Gel Filtration

Separation of the peptides resulting from the tryptic digestion of citraconylated elephant aminoethylated β chain was carried out by gel filtration on Sephadex G-50 (74,86,87). The lyophilized sample was dissolved in 10 ml of 7% formic acid and applied to the top of the Sephadex G-50 column (5 x 190 cm). The sample was allowed to flow into the resin bed and the top of the column was rinsed three times with 5 to 10 ml aliquots of 7% formic acid. After the last rinse, the column top was covered with 30 ml of 7% formic acid. The column was pumped at 100 ml/hour and the total column effluent was monitored directly with a Beckman DB Spectrophotometer at 280 mμ and collected in a fraction collector.

3. Chemical Methods of Fragmenting Peptides

N-Bromosuccinimide (NBS) cleavage of peptide bonds has been demonstrated (88,89). Tryptophyl, tyrosyl, and histidyl bonds are all cleaved by NBS. It is not possible to cleave histidyl bonds without cleavage of tyrosine and tryptophan residues if they are present (90). However, if a peptide contains only one of these three amino acids, then the cleavage will be specific for that one tryptophyl, tyrosyl or histidyl residue in the peptide. The method described by Shaltiel and

Patchornik (91) for simple histidine containing peptides was used. The cleavage of a histidyl peptide procedure is as follows: One ml of peptide solution, peptide dissolved in a buffer containing pyridine-acetic acid-water (1:10:19), was treated with 3 ml NBS solution (0.905 gm NBS in 100 ml buffer previously mentioned) at room temperature for 30 minutes. Excess NBS was destroyed by the addition of a slight excess of imidazole. The mixture was then refluxed at 100°C for one hour to cleave the histidyl bond. This method was particularly useful for a peptide containing one histidine and devoid of tyrosine and tryptophan.

Cleavage of aspartic acid residues by dilute HCl (92) or acetic acid (93) have been documented. Although there are some problems with spurious cleavage of asparagine residues (92) this method is useful for sequence analysis.

In this laboratory 1 to 2 micromoles of the peptide were evaporated to dryness and 1 ml of 0.25 M acetic acid was added. This solution was transferred to a glass ampoule and the flask rinsed with another 1 ml aliquot. The sample was evacuated, sealed, and heated at 100°C for 16 hours. On completion of the 16 hours, the sample was removed from the ampoule and evaporated to dryness. The sample was then dissolved in pH 3.1 buffer. This sample was then loaded onto a Dowex AG 50W-X2 column (0.9 x 60 cm) and chromatographed as described previously.

3. Enzymatic Methods of Fragmenting Peptides

Many proteolytic enzymes are useful for further splitting of the tryptic peptides into shorter peptides. Routinely used were pepsin (94), chymotrypsin (95), thermolysin (96), and papain (95).

Peptides were further digested with pepsin (Worthington, twice crystallized at pH 2 for 18 hours at 25°C. This was accomplished by dissolving the peptide which had been dried by rotary evaporation with 0.5 ml of 0.01 N HCl and adding 0.1 ml of 10 mg % pepsin.

For chymotryptic digestion a modification of the procedure described by Konigsberg and Hill (95) was used. Peptides were dissolved in 1 ml of distilled, deionized water and adjusted to pH 8.0 to 8.5 with 2% trimethylamine. One mg of chymotrypsin was dissolved in 1 ml of 0.001 N HCl and 0.1 ml was added to the peptide solution. The digestion was carried out at 25°C for 16 to 24 hours and the enzyme action was halted by the dropwise addition of 1 N HCl to adjust the pH to between 2.0 and 3.0. The sample was evaporated to dryness and the chymotryptic peptides were separated on a Dowex AG 50W-X2 column (0.9 x 60 cm) as described previously.

Digestion of the peptides with papain was accomplished in a fashion similar to that of Konigsberg and Hill (95). One to 10 micromoles of peptide were dissolved in 1 ml of 0.16 M pyridine-acetate buffer, pH 5.5, and 0.1 ml of the papain suspension containing 20 mg of papain per ml was added to the peptide solution. The digestion mixture was maintained at 37°C for 16 hours. The sample was dried and the peptides separated as described earlier.

Some peptides were hydrolysed with thermolysin. The peptides were dissolved in 0.5 ml of distilled, deionized water and adjusted to pH 8.0 with 0.5 M Tris. One mg of thermolysin was dissolved in 5 ml of water containing 60 mg Tris and 15 mg $CaCl_2 \cdot 2H_2O$ and 0.2 ml of this solution was added to the peptide solution. The digestion was carried out at 35° for 3.5 hours. The hydrolysis was stopped by the dropwise addition of 1 N HCl to adjust the pH to between 2.0 and 3.0. The sample was evaporated to dryness and the thermolysin peptides were separated on a column of Dowex AG 50W-X2 as described previously.

I. Amino Acid Analysis

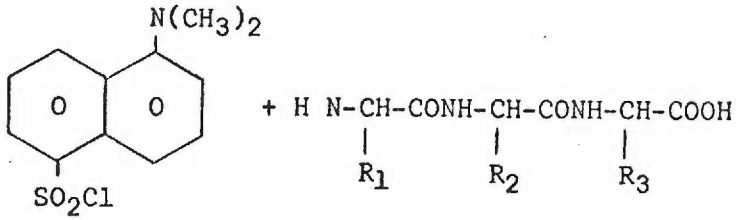
Purified peptides were prepared for quantitative amino acid analysis essentially as described by Moore, Stein and Spackman (68, 97,98). The sample was evaporated to dryness, then dissolved in 6 N HCl containing 9 mg phenol per 100 ml HCl and transferred to a hydrolysis ampoule. The ampoule was evacuated, sealed, and heated for 22 hours at 110°C. At the end of the hydrolysis time the samples were removed from the ampoules, evaporated to dryness, and 2.5 ml of pH 2.2 buffer was added. Quantitative amino acid composition of the samples was determined on a modified Beckman 120B amino acid analyser (73). The areas encompassed by the elution peak of each amino acid were determined by an Infotronics 12 AB integrator. The amino acid composition of the samples was calculated by multiplying the integrator value for each peak by a specific color factor determined for each

elution peak from standard mixtures of amino acids purchased from Beckman. A quantitative analysis of the amino acids is one of the first steps in the elucidation of the chemical structure of a protein molecule. Knowledge of the composition of a peptide is useful as a guide before sequence analysis is begun.

J. Sequencing of Peptides

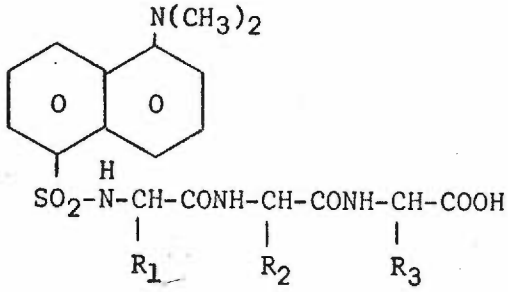
The tryptic peptides of the β chain of Asian elephant hemoglobin were determined by standard manual methods. These included treating the peptides with one or more of the following procedures: 1. Digestion with other proteolytic enzymes (94-96), 2. Chemical cleavage (90,93), or 3. Sequential dansyl-Edman degradation (99-104).

The procedure for dansylation of the N-terminal amino acid was a slight modification of the method reported by Gray and Smith (99). Dansylation, or the reaction of 1-dimethyl amino-naphthalene-5-sulfonyl chloride (dansyl chloride) with the amino group of the N-terminal amino acid, is shown in equations 6-8. (Reactant I is dansyl chloride; reactant II is a tripeptide; product III is the dansylated tripeptide, upon acid hydrolysis of III, the dansyl amino acid IV, and the free amino acids V are released.) This reaction of dansyl chloride with the peptides proceeds under relatively mild conditions and the derivatives formed are, with few exceptions, resistant to acid hydrolysis. These derivatives possess a very intense yellow fluorescence, allowing the detection of small amounts.



I + II

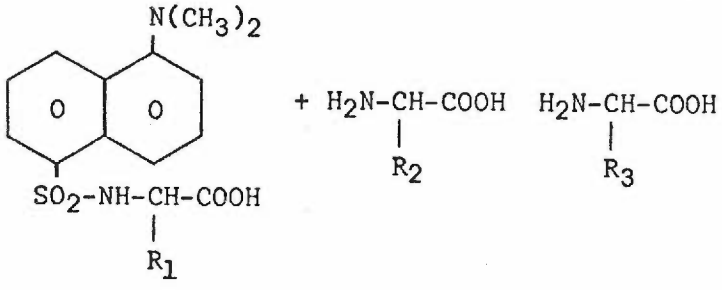
(6)



III

(7)

H^+/H_2O



IV + V

(8)

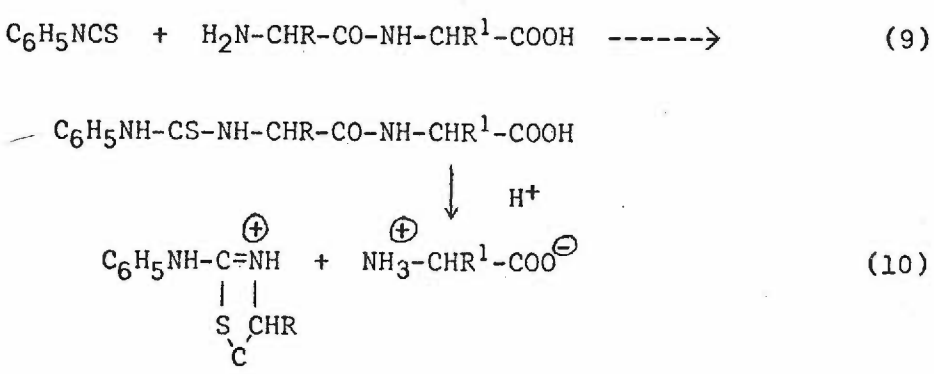
Five to ten nanomoles is sufficient. In addition, the dansyl derivatives are easily separated by thin layer chromatography (105-107).

Each sample taken for reaction with dansyl chloride was thoroughly dried in a small pyrex glass test tube (6mm O.D. x 50 mm). The sample was dissolved in 20 microliters of 0.2 M $NaHCO_3$ solution and dried; this drying facilitates the removal of ammonia which might be present. The samples were then redissolved in 20 microliters of distilled, deionized water and the pH was determined. The pH should be between 8.0 and 9.0. A pH of 8.5 was used because it was found to be optimal for this reaction. After determining and adjusting the pH as necessary, 20 microliters of a 10 mM dansyl chloride solution (3 mg/ml in acetone) were added to the sample. The sample was covered with parafilm and heated for 15 to 30 minutes at 45°C. One hundred microliters of 6 N HCl were added and the test tube was sealed in a fine flame. Hydrolysis proceeded overnight (16 hrs.) at 110°C.

The samples were then dried in vacuo and 1-2 drops of a acetone-water (9:1 v/v) mixture was added to each. These were spotted on silica gel G plates (Quantum Industries) with the appropriate standard dansylated amino acids. The plates were then developed by placing in the appropriate chromatographic tank. Two chromatographic systems were used in separate tanks. One system contained chloroform-methanol-acetic acid (95:10:1 v/v/v) and the other contained n-propanol-ammonium hydroxide (80:20 v/v). The former solvent system readily separates the dansyl derivatives of valine, alanine, glycine,

leucine, threonine, serine, phenylalanine, tyrosine, proline, and lysine. The latter solvent system was essential for identifying the dansyl derivatives of aspartic acid, glutamic acid, arginine, cysteine and histidine. An example of an amino acid identification by this method is depicted in Figure 3.

The sequential degradation of peptides with phenylisothiocyanate (PITC) (103) has become a standard procedure of protein chemistry. The mechanism of the Edman procedure is given in reactions 9 and 10.

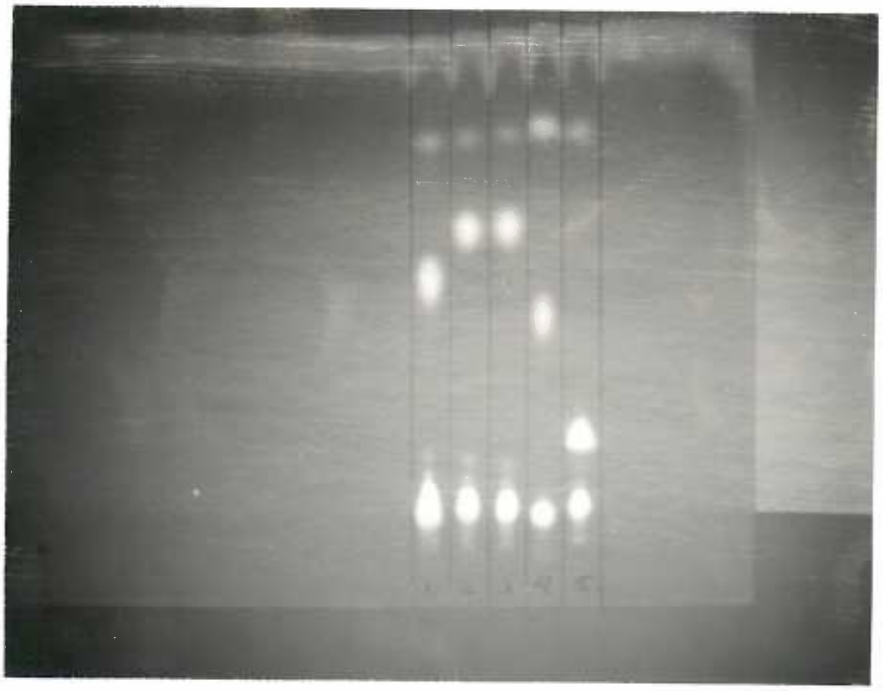


The first reaction (9), called the coupling reaction, is the formation of a phenylthiocarbamyl (PTC) derivative of the peptide. The next reaction (10) involves the cleavage of the PTC-peptide at the peptide bond nearest to the PTC substituent. This requires a strong acid medium and leads to the formation of a 2-anilino-5-thiazolinone derivative and a peptide with one less amino acid than the original sample. This is termed the cleavage reaction (104). If one uses the thiazolinone cleaved off for identification of the N-terminal amino

Figure 3

Thin Layer Chromatographic Identification

of Dansyl-Amino Acids



The columns are: 1. Standard dansyl-Leu. 2. Standard dansyl-Val. 3. The unknown. 4. Standard dansyl-Phe. 5. Standard dansyl-Gly. The unknown was identified as dansyl-Val.

acid, it must be converted to a more stable derivative (104).

Samples of approximately 10-20 nanomoles of peptide were placed in small test tubes (6 mm O.D. x 50 mm) for the Edman degradation. To each sample was added 100 µl of coupling mixture (30 µl Phenylisothiocyanate, Pierce sequential grade, 600 µl redistilled pyridine and 400 µl water). This mixture was made up fresh just prior to each degradative step. Samples were covered with parafilm and placed in a heating block at 50°C for 45 minutes. Following this step the samples were dried in vacuo at 65°-75°C. Thorough drying of the sample was an absolute necessity. The phenylthiocarbamyl derivative of the N-terminal amino acid was cleaved by the addition of 100 µl anhydrous trifluoroacetic acid (TFA), Pierce sequential grade. After addition of the TFA the samples were placed uncovered in a heated oven at 65-70°C for 10 minutes and then dried in vacuo. The PTC peptide had now been cleaved and a new N-terminal amino acid (residue #2) was identified as described under dansylation. In principle this reaction should allow the degradation of any sequence regardless of length. In practice 2 to 7 degradations have been possible.

K. Determination of Amides

The amide groups of glutamine and asparagine are released upon acid or alkaline hydrolysis. Therefore, enzymic procedures were used for the determination of the presence of asparagine and glutamine. A digest of the peptide with leucine aminopeptidase (108,109) or

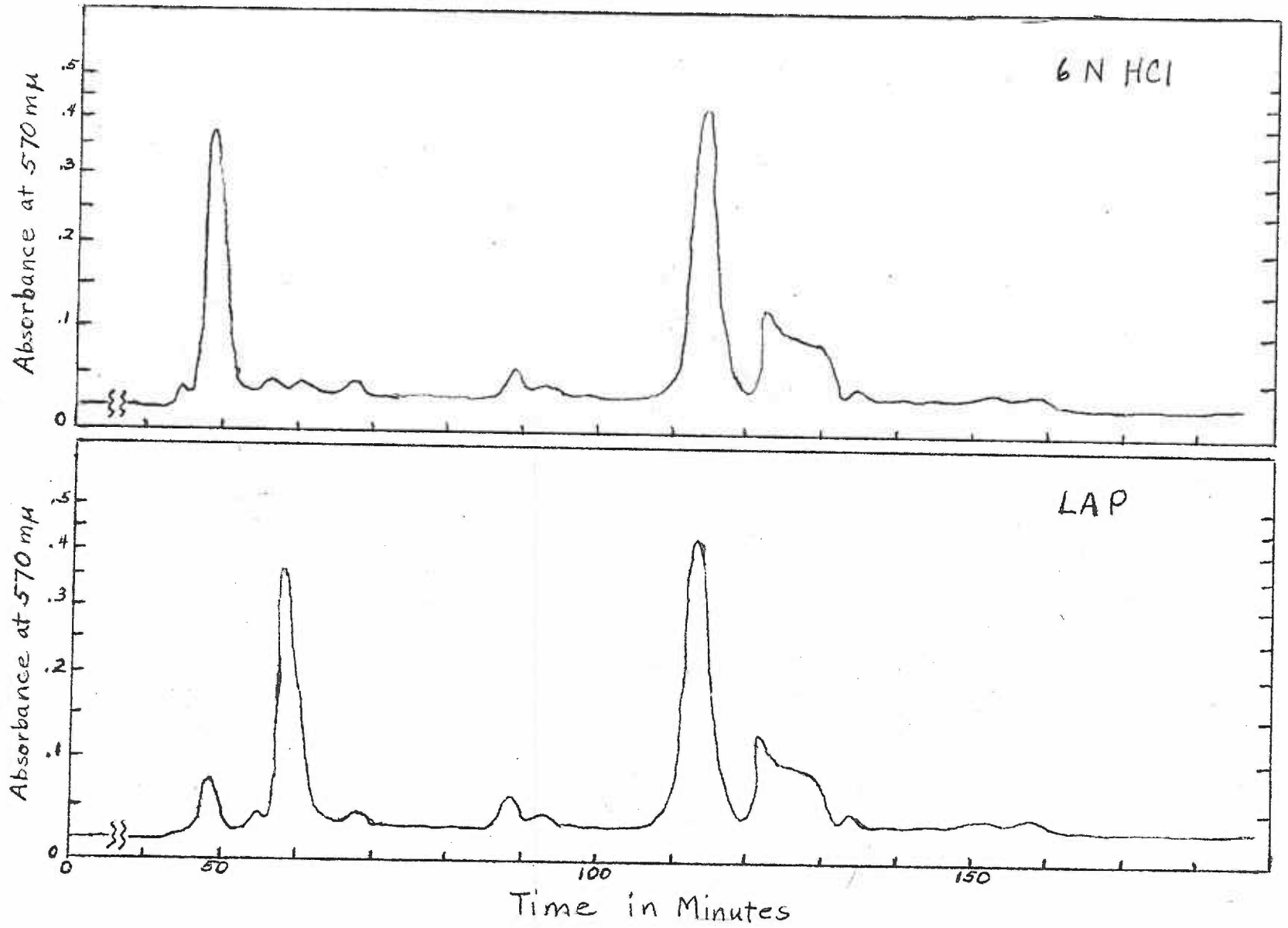
aminopeptidase M (110) can be analyzed by the conventional quantitative amino acid analysis procedure. Peptides which contain serine present a slight problem because serine is eluted together with asparagine and glutamine. The asparagine and glutamine content of a peptide can clearly be determined by comparing the amino acid analysis of both the acid and enzyme hydrolysed sample aliquots. An example of this is illustrated in Figure 4.

One-tenth to 0.5 micromoles of peptides were dissolved in water and the pH was adjusted to 8.6 with 0.5 M Tris solution. Ten λ of 0.25 M $MgCl_2$ and 10 λ of leucine aminopeptidase were added. The samples were hydrolysed at 37°C for 16 hours. The samples were then evaporated to dryness and 2.5 ml of pH 2.2 buffer was added. Analysis of the samples was done on the Beckman amino acid analyzer.

Aminopeptidase M was used in a similar fashion for the determination of amides as the leucine aminopeptidase except the Mg^{++} is not required. Leucine aminopeptidase is specific for the *l*-form of the amino acids; the complete hydrolysis of a peptide provides strong evidence that amino acids with this configuration are present.

Figure 4

Comparison of the Amino Acid Analyses of a Tryptic Peptide Hydrolysed with 6N HCl and Leucine Aminopeptidase (LAP).


6 N HCl
LAP
Absorbance at 570 mμ
Absorbance at 570 mμ
0
.1
.2
.3
.4
.5
0
50
100
150
Time in Minutes

III. RESULTS

A. Hemoglobin Separation

Asian elephant hemoglobin was chromatographed on an Amberlite IRC-50 column. Figure 5 presents a typical chromatogram illustrating the separation of Asian elephant hemoglobin. Zones I through V were concentrated by ultrafiltration and checked for homogeniety by gel electrophoresis. Figure 6 shows the electrophoretic relationship of these five zones. Zones I and II migrated by electrophoresis similar to one another and slightly faster than Zones III, IV, and V. Zones III, IV and V were electrophoretically identical and made up 85% of the total hemoglobin. Zone IV, the major zone, was about 60% of the total, as shown in Table I. Investigations of the β chain of all five zones were initiated. No detectable differences in these five beta-like chains could be demonstrated by amino acid analysis. Therefore, the bulk of the sequence studies were of β chains from whole, unchromatographed hemoglobin.

B. Chain Separation

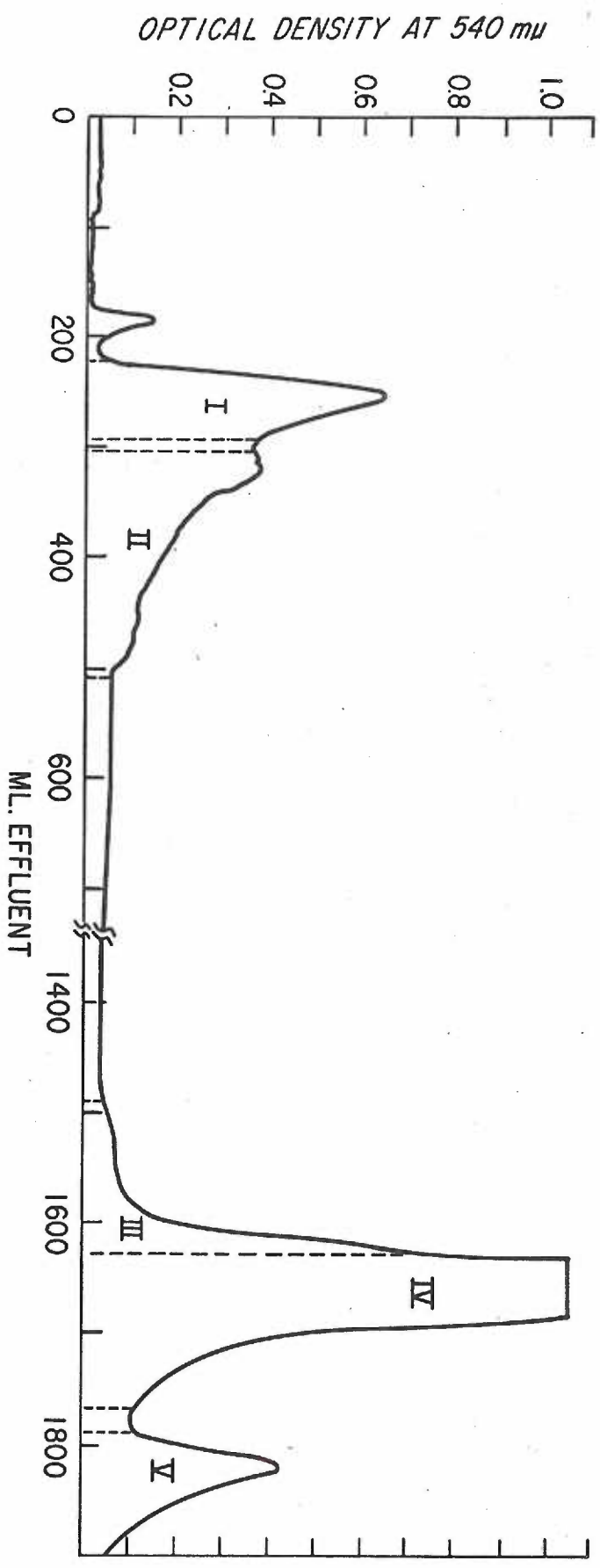
Alpha and beta chains were separated by a modified method of Dintzis (65) as already described. The separation of the two polypeptides is shown in Figure 7. The α chain elutes from the column first, followed by the β chain. The β chain was aminoethylated and stored for further studies.

C. Amino Acid Composition of Asian Elephant β Chain

The amino acid composition of Asian elephant β chain is

Figure 5

IRC-50 Column Separation of Asian Elephant Hemoglobin


OPTICAL DENSITY AT 540 mμ
0.2
0.4
0.6
0.8
1.0
0
200
400
600
1400
1600
1800
ML. EFFLUENT
I
II
III
IV
V

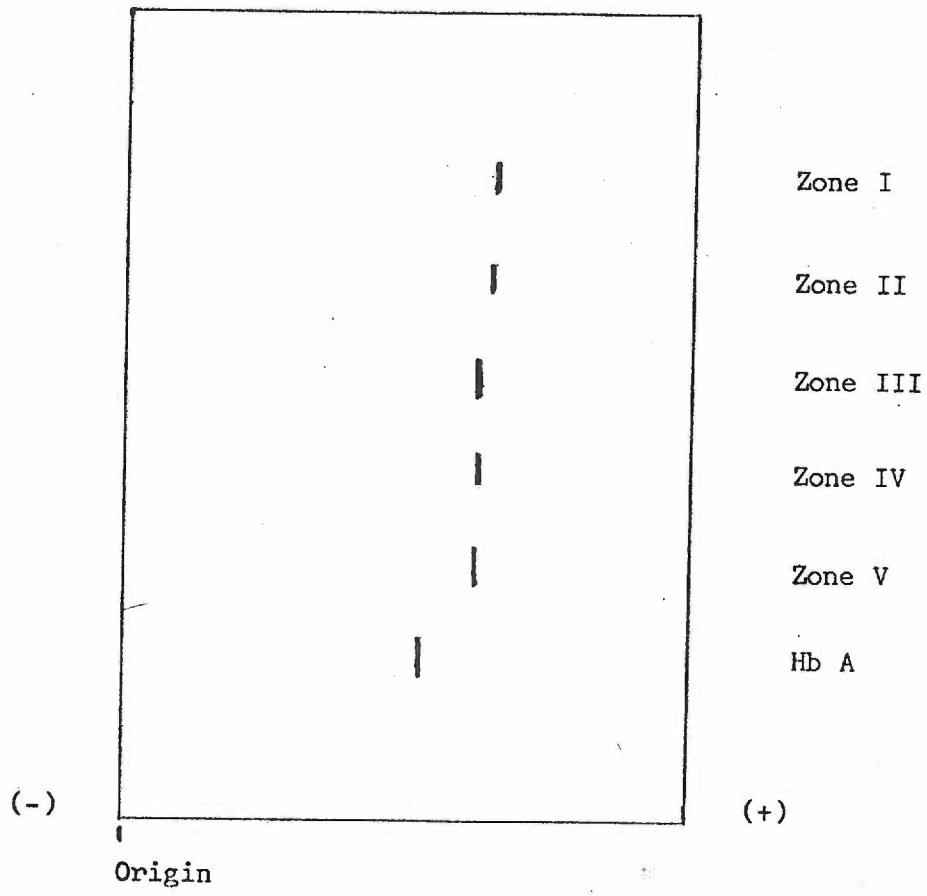


Figure 6

Gel Electrophoresis of the Five IRC-50 Zones

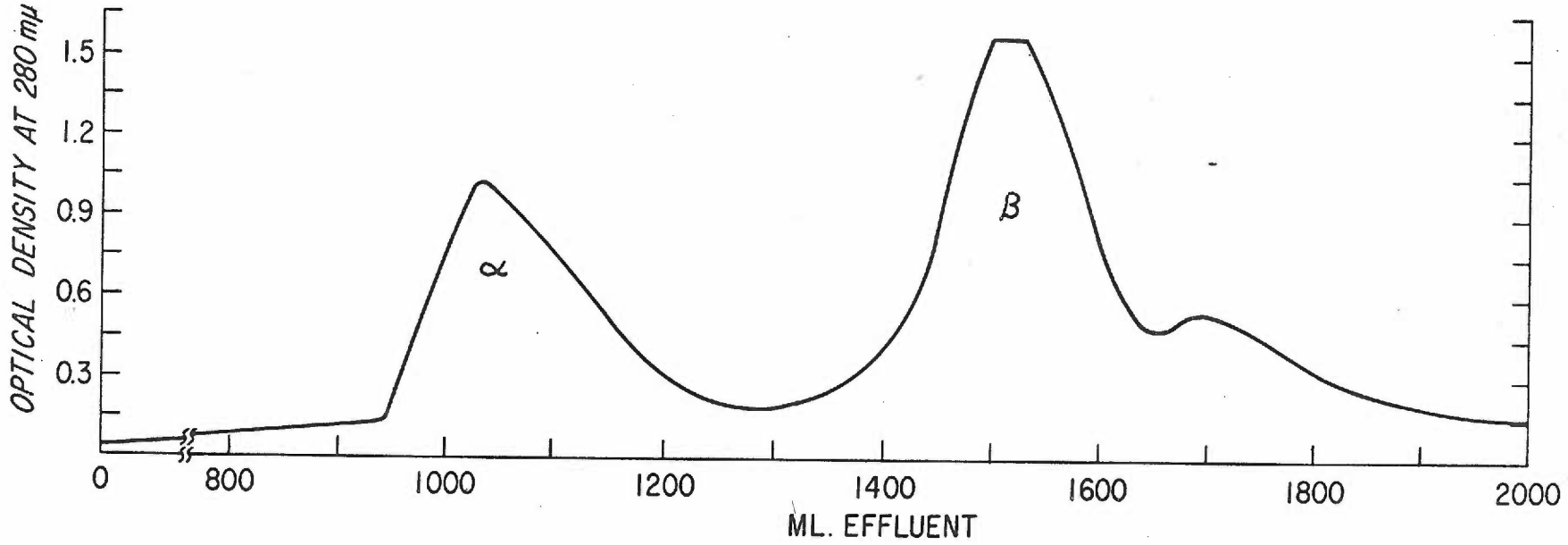
Table 1

PERCENTAGE COMPOSITION OF THE FIVE ZONES
OF ELEPHANT HEMOGLOBIN SEPARATED ON AN IRC-50 COLUMN

Zone	% of Total Hemoglobin
I	5.2
II	9.5
III	17.0
IV	59.5
V	8.8

Figure 7

Carboxymethyl Cellulose Separation of the α and β Chains of Asian Elephant Globin.


OPTICAL DENSITY AT 280 mμ
1.5
1.2
0.9
0.6
0.3
0
800
1000
1200
1400
1600
1800
2000
ML. EFFLUENT
α
β

presented in Table II. It compares satisfactorily with the composition as deduced from the sequence analysis.

D. N-terminal Determination

An aliquot of the β chain was used for N-terminal identification as described by Stark (67). A valine residue was determined as the N-terminus.

E. Separation and Identification of the Fragments Produced by Tryptic Hydrolysis of the Aminoethylated β Chain of Asian Elephant Hemoglobin

In a typical experiment 100 mg of aminoethylated β chain was digested with trypsin. Figure 8 shows the separation of the tryptic peptides on an Aminex A-5 column. Each of the zones were purified further by rechromatographing on a Dowex 50W-X2 column. The amino acid composition of the tryptic peptides which were isolated are presented in Table 3. In this and all subsequent figures and tables, the peptides are numbered consecutively from the N-terminus (T-1, T-2, etc.). To avoid confusion the numbering of the tryptic peptides has been modeled after the tryptic peptides of the aminoethylated β chain of human hemoglobin. The amino acid analyses presented in Table 3 are arranged in the sequencial order that the peptides occur in the β chain.

F. Sequencing of Peptides

The sequences of the tryptic peptides of the aminoethylated β chain were determined by the methods previously mentioned. These

Table 2

TOTAL AMINO ACID ANALYSIS OF AMINOETHYLATED

β-CHAIN OF ASIAN ELEPHANT HEMOGLOBIN

Amino Acid	22 hr Hydrolysis Calculated Values	48 hr Hydrolysis Calculated Values	72 hr Hydrolysis Calculated Values	Expected Values
Lysine	13.1	12.7	12.5	11
Histidine	10.7	9.9	9.7	10
A-E Cysteine	Trace	Trace	Trace	2
Arginine	5.0	5.0	5.1	4
Aspartic Acid	13.0	13.3	13.3	14
Threonine	7.4	7.5	6.7	7.5
Serine	4.7	4.5	4.2	4
Glutamic Acid	10.5	10.6	10.6	11
Proline	3.5	3.2	4.1	3
Glycine	10.4	10.5	10.4	10
Alanine	13.7	13.4	13.7	15.5
Valine	12.5	14.4	15.2	15
Leucine	19.0	17.9	18.2	18
Tyrosine	2.8	2.8	2.9	3
Phenylalanine	7.5	7.6	7.7	8

Figure 8

Separation of the Tryptic Peptides on a Aminex A-5 Column

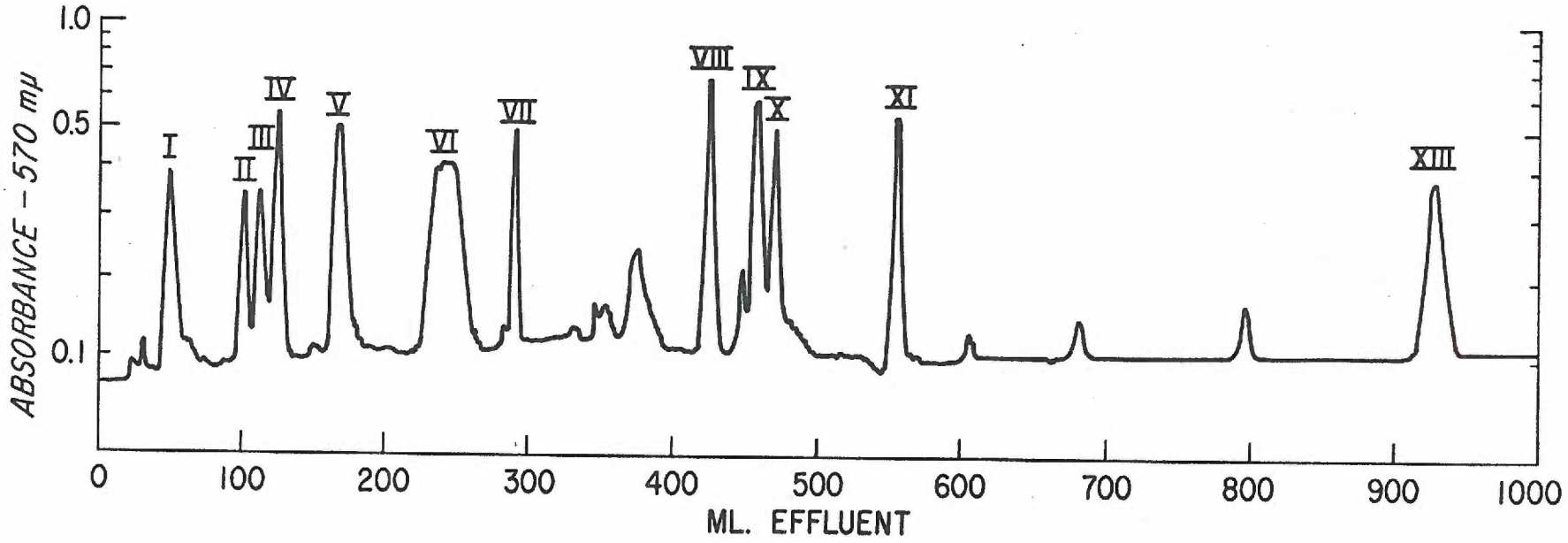

ABSORBANCE - 570 mμ
1.0
0.5
0.1
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XIII
0
100
200
300
400
500
600
700
800
900
1000
ML. EFFLUENT

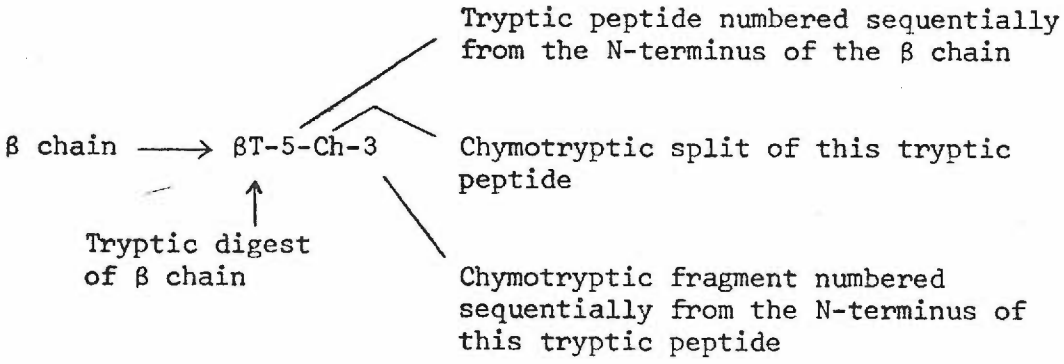
Table 3

AMINO ACID COMPOSITION OF THE TRYPTIC PEPTIDE

Amino Acid	βT-1	βT-2	βT-3a	βT-3b	βT-4a	βT-4b	βT-5	βT-6,7,8	βT-9a	βT-9b	βT-10a	βT-10b	βT-11	βT-12a	βT-12b	βT-13	βT-14	βT-15
Lysine	1.03	1.06	1.11				1.20	0.98	1.03	1.06		0.99			1.07	1.06	1.07	
Histidine							2.01	0.94		0.98	1.02		0.97		1.90		1.18	1.01
Arginine				1.06	1.02	1.00							0.82					
Tryptophan		Pres.			Pres.													
A-E Cysteine											0.22							
Aspartic Acid	1.04	1.07	1.04				3.26			2.00	1.24	1.01	2.04			1.12	1.09	
Threonine	0.91	1.42			0.99		1.08		1.05		1.00					1.12		
Serine				0.87			1.08		1.00		0.96							
Glutamic Acid	0.98	0.95		1.96			1.20	0.99	1.08		0.99		1.03			2.60		
Proline					0.99								1.16			1.14		
Glycine		1.03		2.06			1.20	1.10	2.06		1.12				1.01		1.20	
Alanine	1.94	0.72		1.09			3.14	0.97			0.95				0.95	1.95	3.78	
Valine	0.92	1.08	1.85		1.65		0.81	0.98	1.00				0.89		0.94	1.03	2.82	
Methionine																		
Isoleucine																		
Leucine	0.92	1.07		1.99	2.02		1.91	1.01	1.88	1.96	1.84		1.07		1.05		1.02	
Tyrosine					0.92											0.98		0.99
Phenylalanine							2.89		1.10		0.90		0.99		1.05	1.03		
Total	8	9	4	9	9	1	19	7	10	6	11	2	9		8	12	12	2

methods include sequentially identifying the N-terminal amino acid by either the dansyl-Edman or the subtractive Edman methods. Amino acid residues determined by the dansyl-Edman method are designated with an arrow ($\longrightarrow$) below the residue. An arrow ($\longrightarrow$) above the residue denotes subtractive Edman results. The aspartic acid, asparagine and glutamic acid, glutamine identities usually were determined by hydrolysing the unfragmented tryptic peptides with either leucine aminopeptidase or aminopeptidase M. Sometimes fragmentation of the tryptic peptide was necessary to unambiguously determine the position of these four amino acid residues in the sequence. Some tryptic peptides were fragmented by hydrolysis with other proteolytic enzymes or cleaved by chemical methods. The fragments produced were numbered sequentially from the N-terminus. The sequential order of the peptide fragments can be deduced easily in many cases. The peptide fragment which contains a lysine, arginine or aminoethylated cysteine residue must be the C-terminus of the parent tryptic peptide. This is because of the known specificity of trypsin which hydrolyses peptides involving carbonyl groups of lysine, arginine, or aminoethylated cysteine. If there are only two peptide fragments, then by deduction the portion which does not contain lysine, arginine or aminoethyl cysteine must be from the N-terminus of the original tryptic peptide. Additional composition or partial sequence information may be necessary to deduce the relative position of three or more fragments from a tryptic peptide.

The following standard nomenclature was adopted. The Greek letter "β" refers to the polypeptide chain under investigation. "T" refers to the tryptic hydrolysis of the β chain. A number 1,2,3, etc. designates the sequential order of the tryptic peptides in the β chain. The letters Ch,[1] Th,[1] HAc,[1] NBS,[1] and P[1] denote the further fragmentation of the tryptic peptides. These fragments of the tryptic peptides were numbered sequentially from the N-terminus. An example of this nomenclature is illustrated below.



The following format was normally used to present the sequence determination of the individual tryptic peptides.

1. Amino acid composition of the tryptic peptide.
2. Determination of amides.
3. Fragmentation of the tryptic peptide as needed.
4. Dansylation and Edman degradation of the complete

[1] Ch = chymotrypsin, Th = thermolysin, HAc = acetic acid, NBS = N-Bromosuccinimide, and P = papain

tryptic peptide.

5. Dansylation and Edman degradation of the peptide fragments.

6. Summary of data.

7. Final sequence of tryptic peptide.

Often a flow diagram was used to summarize these data in schematic form.

1. βT-1

Asian elephant βT-1 has the amino acid composition shown in Table 4. Digestion of this peptide with leucine aminopeptidase to determine the amides showed one residue each of asparagine and glutamic acid present in this peptide. βT-1 was fragmented further by hydrolysis with chymotrypsin. The amino acid analysis of the two peptide fragments obtained are presented in Table 5.

One aliquot of the whole peptide, βT-1, was dansylated. Another aliquot was subjected to one Edman degradation cycle followed by dansylation. The dansylated derivatives of these two aliquots were identified by thin layer chromatography. The first aliquot was a valine residue and the second aliquot an aspartic acid.[1]

The chymotryptic digestion of βT-1 produced two peptide fragments, βT-1-Ch-1 and βT-1-Ch-2. The pentapeptide, βT-1-Ch-2, contained lysine and therefore is the C-terminal fragment of βT-1. The

[1] Differentiation between the acid and amide form of aspartic and glutamic acid cannot be made by the dansylation procedure.

Table 4

AMINO ACID ANALYSIS OF βT-1

Amino Acid	Calculated Values	Integral Values
Lysine	1.28	1
Aspartic Acid	1.04	1
Threonine	0.91	1
Glutamic Acid	0.98	1
Alanine	1.94	2
Valine	0.92	1
Leucine	0.92	1

Table 5

AMINO ACID ANALYSIS OF CHYMOTRYPTIC FRAGMENTS 1 AND 2 OF βT-1

	Ch-1		Ch-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.12	1
Aspartic Acid	1.05	1		
Threonine			0.94	1
Glutamic Acid			1.06	1
Alanine			1.84	2
Valine	0.94	1		
Leucine	0.99	1		

tripeptide, βT-1-Ch-1, contained valine, leucine and aspartic acid. Because valine and asparagine were shown to be residues 1 and 2 of this peptide by the dansyl-Edman procedures, leucine must be residue 3. βT-1-Ch-2 was divided into three aliquots. One aliquot was dansylated, another subjected to one Edman degradation cycle and dansylated. The third aliquot underwent two Edman degradations and then dansylation. Upon identification of these three dansyl derivatives the N-terminus was concluded to be threonine; residue 2 was alanine and residue 3 was another alanine. This information coupled with the known specificity of trypsin allows the unambiguous conclusion that the sequence of this peptide, βT-1-Ch-2, is Thr-Ala-Ala-Glu-Lys. The arrow (⟶) denotes a residue position established by the dansyl-Edman sequential identification method.

Combining these results, the amino acid sequence of βT-1 is Val-Asn-Leu-Thr-Ala-Ala-Gly-Lys. Figure 9 summarizes in schematic form these results. The placement of this and all subsequent tryptic peptides in the elephant β chain was by homology with known β chain sequences of other species and other evidence provided by fragmenting elephant β chain at the arginine residues. These data will be presented later.

2. βT-2

Table 6 shows the amino acid composition of the second tryptic peptide, βT-2. Hydrolysis of this peptide with leucine aminopeptidase proved the presence of one asparagine and one glutamine

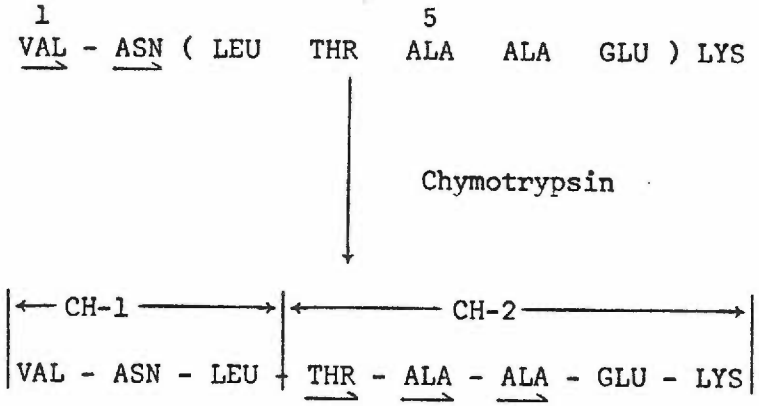


Figure 9

Flow Diagram for Degradation of βT-1

Table 6

AMINO ACID ANALYSIS OF βT-2

Amino Acid	Calculated Values	Integral Values
Lysine	1.03	1
Tryptophan	present	1
Aspartic Acid	1.02	1
Threonine	1.42	1.5
Glutamic Acid	0.84	1
Glycine	1.04	1
Alanine	0.72	0.5
Valine	1.04	1
Leucine	0.92	1

residue. Digesting this tryptic peptide with chymotrypsin gave three smaller peptides, the compositions of which are depicted in Table 7. Another preparation of βT-2 was hydrolysed with thermolysin which resulted in splitting the peptide into the two fragments shown in Table 8.

Dansylation and Edman degradations similar to those described for βT-1 were performed for four cycles on an aliquot of βT-2. The N-terminal sequence of this peptide was found to be Thr-Gln-Val-Asn. The fifth residue of βT-2 was identified to be an equal mixture of threonine and alanine and was determined from analyses of βT-2-Th-1. This was deduced by subtracting the four residues identified by the sequential dansyl-Edman procedure from the composition of βT-2-Th-1. The remaining one-half residue each of threonine and alanine must represent the C-terminal residues of two homologous βT-2-Th-1 peptides. This apparent heterogeneity at the fifth residue is similar to results obtained by other investigators (23,24,35) and will be discussed later.

Dansylation of thermolysin fragments, βT-2-Th-2, proved the N-terminus to be Leu. Recognizing the specificity of trypsin and utilizing the composition of peptides from the secondary splits by chymotrypsin plus these data just discussed, the sequence of βT-2 is Thr-Gln-Val-Asn-$\frac{\text{Thr}}{\text{Ala}}$-Leu-Try-Gly-Lys. This is schematically drawn in Figure 10.

Table 7

AMINO ACID ANALYSIS OF THE THREE CHYMOTRYPTIC FRAGMENTS OF βT-2

	Ch-1		Ch-2		Ch-3	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine					1.01	1
Tryptophan			present	1		
Aspartic Acid	0.94	1				
Threonine	1.52	1.5				
Glutamic Acid	0.89	1				
Glycine					0.96	1
Alanine	0.54	0.5	1.04	1		
Valine	0.89	1				
Leucine	1.14	1	0.95	1		

Table 8

AMINO ACID ANALYSIS OF THE THERMOLYSIN FRAGMENTS 1 AND 2 OF βT-2

	Th-1		Th-2	
Amino Acids	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.07	1
Tryptophan			present	1
Aspartic Acid	1.01	1		
Threonine	1.47	1.5		
Glutamic Acid	0.92	1		
Glycine			0.98	1
Alanine	0.52	0.5		
Valine	1.04	1		
Leucine			0.95	1

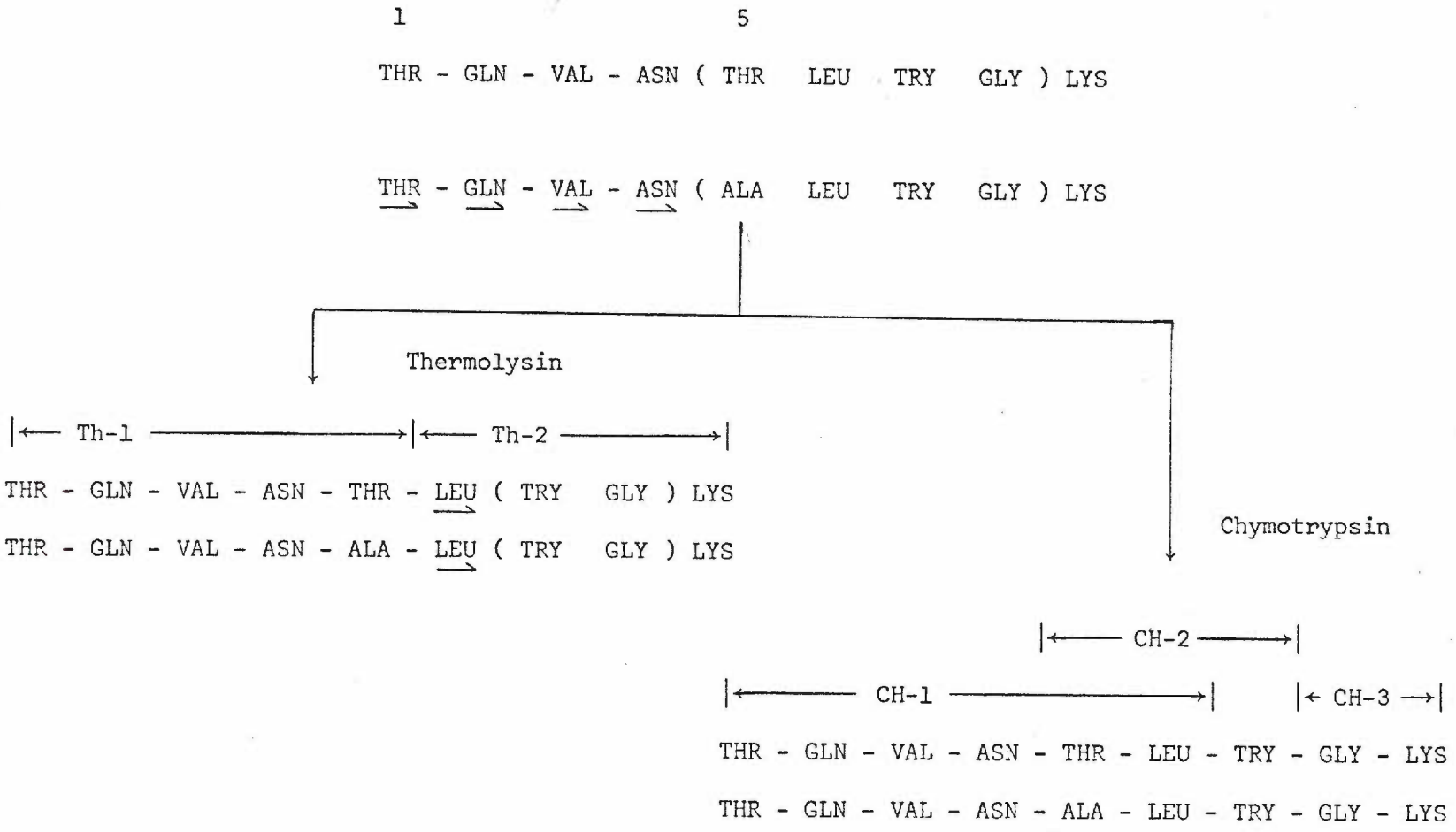


Figure 10

Flow Diagram for βT-2

3. βT-3a and βT-3b

The amino acid composition of βT-3a and βT-3b are given in Table 9. βT-3a, a tetrapeptide, was sequenced without further fragmentation. Digestion of this peptide with leucine aminopeptidase indicated the presence of an asparagine residue. The dansyl-Edman procedure gave the sequence Val-Asn-Val. Because this is a tetrapeptide, the fourth residue must be lysine. Therefore, the unambiguous sequence of βT-3a is Val-Asn-Val-Lys.

βT-3b was hydrolysed with leucine aminopeptidase and two glutamic acid residues were found to be present by analysis. The results of hydrolysis with thermolysin are given in Table 10.

The first three residues of the N-terminal region of βT-3b were determined by the dansyl-Edman method. These residues are Glu-Leu-Gly. The hydrolysis of βT-3b with papain produced several free amino acids plus one tripeptide, βT-3b-P-1, given in Table 11. Sequential Edman degradation and amino acid analysis proved the sequence of this tripeptide to be Gly-Glu-Ala. An arrow above an amino acid residue denotes position shown by sequential Edman degradation. Because these data were insufficient for proving the sequence of the first six residues of βT-3b, another portion of whole βT-3b peptide was used for sequence analysis by the subtractive Edman method through four residues. The results of this experiment are presented in Table 12. The N-terminal residues as shown by this method were Glu-Leu-Gly-Gly. Therefore, the sequence of the first six residues

Table 9

AMINO ACID ANALYSIS OF βT-3a AND βT-3b

	βT-3a		βT-3b	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine	1.11	1		
Arginine			1.06	1
Aspartic Acid	1.04	1		
Serine			0.87	1
Glutamic Acid			1.96	2
Glycine			2.06	2
Alanine			1.09	1
Valine	1.85	2		
Leucine			1.99	2

Table 10

AMINO ACID ANALYSIS OF THERMOLYSIN FRAGMENTS 1 AND 2 OF βT-3b

	Th-1		Th-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Arginine			0.96	1
Serine			0.92	1
Glutamic Acid	2.05	2		
Glycine	1.99	2		
Alanine	0.94	1		
Leucine	1.03	1	1.10	1

Table 11

AMINO ACID ANALYSIS FOR THE PEPTIDE OBTAINED FROM PAPAIN TREATMENT OF βT-3b

	βT-3b-P-1		1-Edman[a]		2-Edman[a]	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Glutamic Acid	0.99	1	1.02	1	0.18	0
Glycine	1.05	1	0.21	0	0.15	0
Alanine	0.97	1	0.99	1	1.05	1

[a] Amino acid analysis of this same peptide after 1 Edman and 2 Edman degradative cycles.

Table 12

AMINO ACID ANALYSIS FOR FOUR SUCCESSIVE EDMAN DEGRADATIONS OF βT-3b

	Edman 1		Edman 2		Edman 3		Edman 4	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Serine	0.87	1	0.82	1	0.73	1	0.84	1
Glutamic Acid	1.10	1	1.13	1	1.01	1	1.11	1
Glycine	2.02	2	1.92	2	1.22	1	0.61	0
Alanine	1.00	1	1.07	1	0.94	1	1.00	1
Leucine	2.01	2	1.06	1	1.02	1	1.03	1

of βT-3b is Glu-Leu-Gly-Gly-Glu-Ala.

Thermolysin fragment βT-3b-Th-2 was found to have a leucine N-terminus. Knowing the specificity of trypsin, arginine must be the C-terminus of this peptide and by deduction serine must be the middle residue. The summation of these data allow the conclusion that the complete sequence of βT-3b is Glu-Leu-Gly-Gly-Glu-Ala-Leu-Ser-Arg. The flow diagram for this peptide is shown in Figure 11.

4. βT-4a and βT-4b

The results of amino acid analysis of βT-4a and βT-4b are tabulated in Table 13. Often in hydrolysing the aminoethylated β chain of elephant with trypsin some chymotryptic-like splitting would occur. A tryptophan residue present in βT-4a appeared to be sensitive to this atypical hydrolysis. The composition of fragments produced by this cleavage are shown in Table 14.

The N-terminal sequence of βT-4a was found by dansyl-Edman degradation to be Leu-Leu-Val-Val-Tyr-Pro. These are six of the seven residues of βT-4-Ch-1 shown in Table 14. Subtracting these six residues from the composition of this peptide, the seventh residue in βT-4a must be tryptophan. βT-4a-Ch-2 provided sufficient information to complete the sequence determination of βT-4a. Because arginine must be the C-terminus of this dipeptide (βT-4a-Ch-2), threonine must be the N-terminus. The summation of these results allows the conclusion that the complete sequence of this peptide is Leu-Leu-Val-Val-Tyr-Pro-Try-Thr-Arg. These data are summarized in Figure 12.

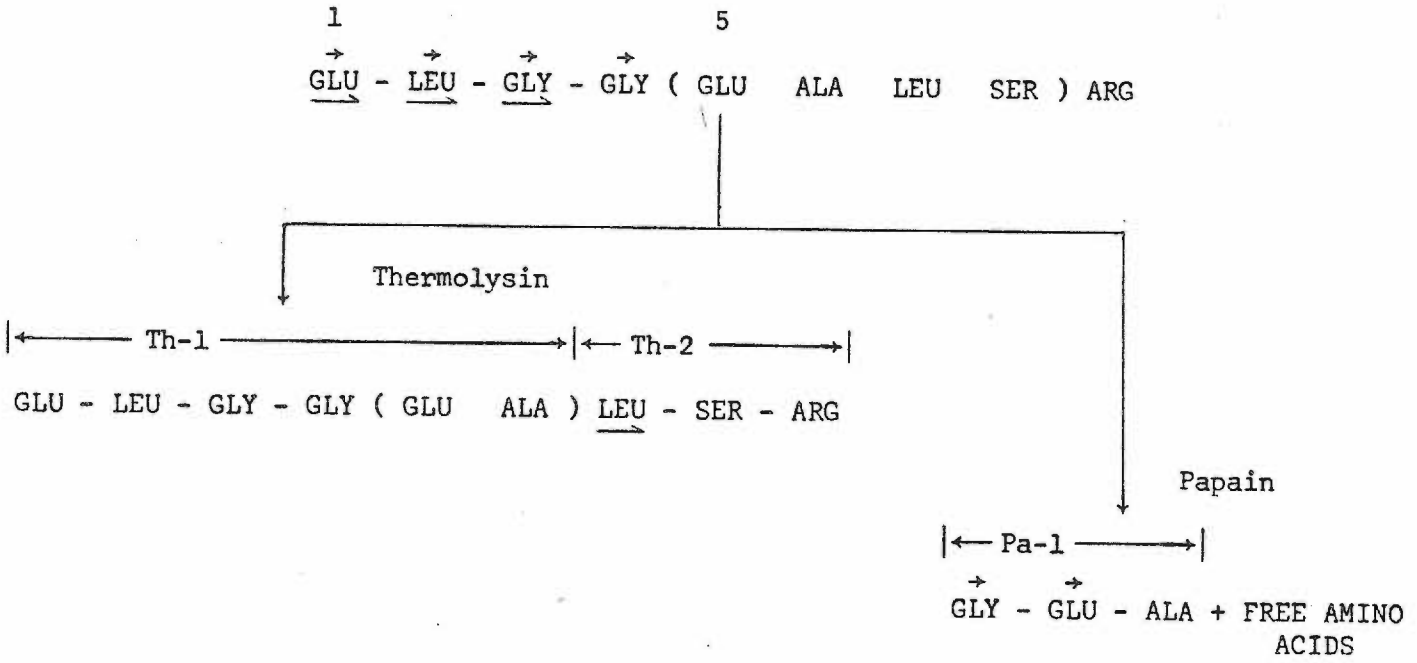


Figure 11

Flow Diagram for βT-3b

Table 13

AMINO ACID ANALYSIS OF βT-4a AND βT-4b

	βT-4a		βT-4b	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Arginine	1.02	1	1.00	1
Tryptophan	present	1		
Threonine	0.99	1		
Proline	0.99	1		
Valine	1.65	2		
Leucine	2.02	2		
Tyrosine	0.92	1		

Table 14

AMINO ACID ANALYSIS OF THE CHYMOTRYPTIC FRAGMENTS 1 AND 2 OF βT-4a

	βT-4a-Ch-1		βT-4a-Ch-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Arginine			1.02	1
Tryptophan	present	1		
Threonine			0.98	1
Proline	1.05	1		
Valine	1.64	2		
Leucine	2.21	2		
Tyrosine	1.07	1		

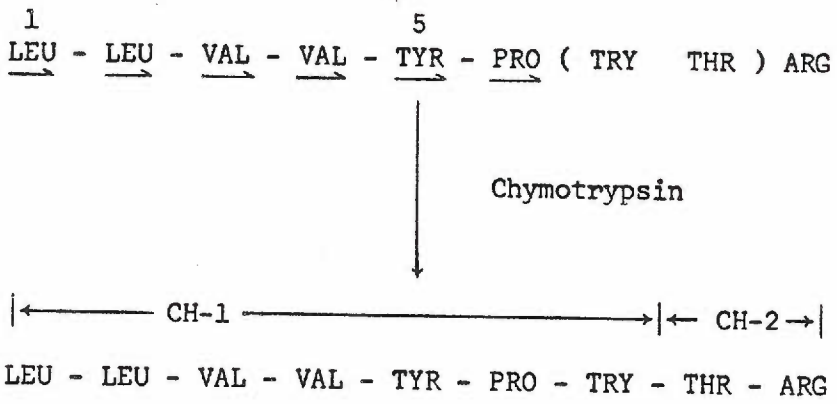


Figure 12

Flow Diagram for βT-4a

βT-4b, as shown in Table 13, chromatographed as arginine after acid hydrolysis. To test the special, highly unlikely, case that acid hydrolysis cleaved some substituent from arginine which was not detected by amino acid analysis, an aliquot of the βT-4b zone was analyzed without acid hydrolysis. The results showed that the unhydrolysed sample of βT-4b also chromatographed as arginine.

5. βT-5

The result of amino acid analysis of βT-5 is tabulated in Table 15. This peptide, which contains 19 amino acid residues, was fragmented further by thermolysin digestion and by dilute acetic acid hydrolysis. The results of these splittings are shown in Tables 16 and 17.

An aliquot of βT-5 was subjected to sequential Edman degradation and identification of the N-terminal amino acids produced by their dansyl derivatives. The N-terminal sequence of this peptide was found to be Phe-Phe-Glu. The N-terminus of βT-5-Th-2 was Phe. βT-5-Th-3 yielded Leu-His-Asp as its N-terminal sequence. βT-5-Th-5 had an alanine N-terminus. Analyses of βT-5-HAc-2 revealed an N-terminal sequence of Leu-Ser and finally βT-5-HAc-3 had an Ala-Val N-terminus.

The structural information obtained was sufficient to deduce the sequence of βT-5. The logic of this may be followed by utilizing the flow diagram in Figure 13. Residues 1 through 3 were proven by sequential dansyl-Edman degradation of the whole tryptic peptide.

Table 15

AMINO ACID ANALYSIS OF βT-5

Amino Acid	Calculated Values	Integral Values
Lysine	1.20	1
Histidine	2.01	2
Aspartic Acid	3.26	3
Threonine	1.08	1
Serine	1.08	1
Glutamic Acid	1.20	1
Glycine	1.20	1
Alanine	3.14	3
Valine	0.81	1
Leucine	1.91	2
Phenylalanine	2.89	3

Table 16

AMINO ACID ANALYSIS OF THE FIVE THERMOLYSIN FRAGMENTS OF βT-5

	Th-1		Th-2		Th-3		Th-4		Th-5	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine					1.26	1				
Histidine	1.17	1			0.95	1				
Aspartic Acid			1.87	2	0.93	1			1.16	1
Threonine			0.91	1			0.99	1		
Serine			0.95	1			0.92	1		
Glutamic Acid	0.97	1								
Glycine			1.15	1						
Alanine			1.95	2	1.00	1			0.85	1
Valine			1.15	1						
Leucine			1.02	1	0.82	1	1.06	1		
Phenylalanine	1.86	2	0.99	1						

Table 17

AMINO ACID ANALYSIS OF FOUR ACETIC ACID FRAGMENTS OF βT-5

	HAc-1		HAc-2		HAc-3		HAc-4	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine							1.04	1
Histidine	0.99	1			0.95	1		
Aspartic Acid								
Threonine			0.94	1				
Serine			0.89	1				
Glutamic Acid	1.04	1						
Glycine	1.14	1						
Alanine			1.10	1	1.01	1	0.95	1
Valine					1.04	1		
Leucine			1.05	1	1.02	1		
Phenylalanine	2.82	3						

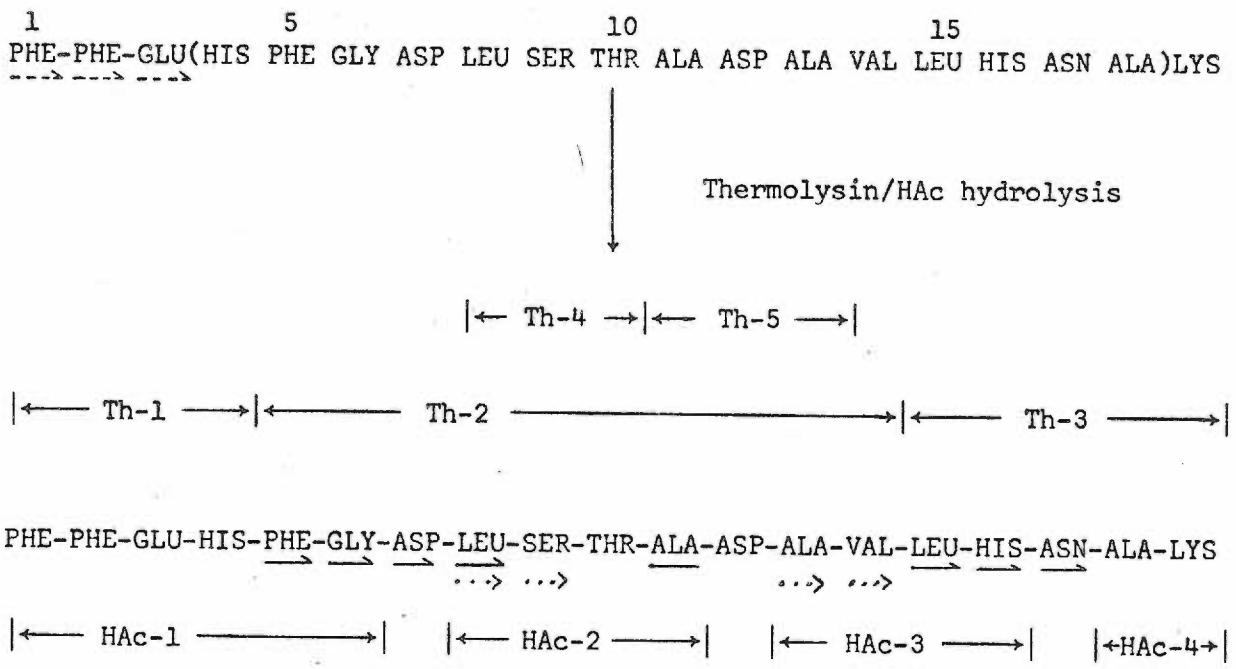


Figure 13

Diagrammatic representation of the thermolysin and acetic acid splitting of βT-5. The arrows denote: ---› dansylation of the whole tryptic peptide; ——› dansyl derivatives of the thermolysin fragments; · · ·› acetic acid fragment sequence data.

Residue 4 was placed by knowing the amino acid compositions of βT-5, the thermolysin fragments, and the sequence of the first three residues of βT-5. Sequential dansyl-Edman degradation of βT-5-Th-2 determined residues 5 through 8. βT-5-HAc-2 gave results for residues 8 and 9. Residues 10 and 11 were deduced from the compositions of βT-5-Th-4 and βT-5-HAc-2 and the sequence information already presented. βT-5-Th-5 proved residue 12. Residues 13 and 14 were determined from βT-5-HAc-3. βT-5-Th-3 was used to place residues 15 through 19. The sequence of residues 18 and 19 was confirmed by βT-5-HAc-4.

The amide determination was made by aminopeptidase M (APM) digestion of thermolysin fragments 1, 2 and 3. The results showed that residue 3 is a glutamic acid, residues 7 and 12 are aspartic acids, and residue 17 is an asparagine. Therefore, the sequence of this βT-5 is Phe-Phe-Glu-His-Phe-Gly-Asp-Leu-Ser-Thr-Ala-Asp-Ala-Val-Leu-His-Asn-Ala-Lys.

6. βT-6-7-8

The amino acid composition of βT-6-7-8 is shown in Table 18. Hydrolysis of the whole peptide with leucine aminopeptidase proved there to be one residue of glutamic acid in this peptide. All attempts at fragmenting this peptide by the enzymes thermolysin and pepsin were unsuccessful. The peptide finally was cleaved chemically by N-Bromosuccinimide as described under Materials and Methods. The results are tabulated in Table 19.

Table 18

THE AMINO ACID ANALYSIS OF THE TRYPTIC PEPTIDE DESIGNATED βT-6 - 7 - 8

Amino Acid	Calculated Values	Integral Values
Lysine	0.98	1
Histidine	0.94	1
Glutamic Acid	0.99	1
Glycine	1.10	1
Alanine	0.97	1
Valine	0.98	1
Leucine	1.01	1

Table 19

AMINO ACID ANALYSIS OF N-BROMOSUCCINIMIDE

FRAGMENTS 1 AND 2 OF βT-6 - 7-8

	NBS-1		NBS-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.04	1
Glutamic Acid			1.07	1
Glycine			0.91	1
Alanine	1.10	1		
Valine	0.87	1		
Leucine	1.03	1		

The N-terminal sequence of the whole peptide was determined to be Val-Leu. The tripeptide, βT-6-7-8-NBS-1, contains these two amino acids and one other which, because of the composition of the whole tryptic peptide, must be the C-terminus of this peptide. The N-terminal sequence of N-Bromosuccinimide fragment 2, βT-6-7-8-NBS-2, was determined to be Gly-Glu and, therefore, lysine must be the C-terminus of this peptide.

Combining these data, the sequence of βT-6-7-8 must be Val-Leu-Ala-His-Gly-Glu-Lys.

7. βT-9a

The amino acid analysis of this peptide is as shown in Table 20. A glutamic acid residue was found by digestion of this peptide with leucine aminopeptidase. Hydrolysis of βT-9a with thermolysin produced three fragments which were isolated by column chromatography as described previously. The amino acid compositions of these fragments are shown in Table 21. βT-9a was also digested with chymotrypsin. This resulted in splitting the tryptic peptide into two smaller peptides shown in Table 22. The N-terminal sequence of βT-9a was determined to be Val-Leu-Thr. βT-9a-Th-3 had a N-terminal Leu. The N-terminal sequence of βT-9a-Ch-2 was determined by the subtractive Edman method to be Gly-Glu.

The data for the sequence analysis of βT-9a may be summarized as follows: Residues 1, 2 and 3 were determined by the dansyl-Edman method applied to the whole tryptic peptide. Residue 4 was deduced

Table 20

THE AMINO ACID ANALYSIS OF βT-9a

Amino Acid	Calculated Values	Integral Values
Lysine	1.03	1
Threonine	1.05	1
Serine	1.00	1
Glutamic Acid	1.09	1
Glycine	2.06	2
Valine	0.81	1
Leucine	1.88	2
Phenylalanine	1.10	1

Table 21

AMINO ACID ANALYSIS OF THREE THERMOLYSIN FRAGMENTS OF βT-9a

	Th-1		Th-2		Th-3	
Amino Acids	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine					1.10	1
Threonine	0.96	1				
Serine	0.94	1				
Glutamic Acid			0.90	1		
Glycine			2.10	2		
Valine	1.02	1				
Leucine	1.08	1			0.91	1
Phenylalanine			1.04	1		

Table 22

AMINO ACID ANALYSIS OF THE CHYMOTRYPTIC FRAGMENTS 1 AND 2 OF βT-9a

	Ch-1		Ch-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.02	1
Threonine	1.07	1		
Serine	0.98	1		
Glutamic Acid			1.16	1
Glycine			1.93	2
Valine	1.02	1		
Leucine	1.10	1	1.05	1
Phenylalanine	0.91	1		

to be serine from the composition of βT-9a-Th-1 and the knowledge of residues 1, 2 and 3. Residue 5 of the tryptic peptide was placed in its position by similar reasoning as used for residue 4 and by knowing the composition of the pentapeptide, βT-9a-Ch-1 and the sequence of residues 1 through 4. Residues 6 and 7 were placed in position by the subtractive Edman data obtained for βT-9a-Ch-2. Residue 8 was identified by deduction from the sequence already presented and the total amino acid composition of βT-9a-Th-2. Residue 9 and 10 were determined from dansylation of the dipeptide, βT-9a-Th-3, and by the specificity of trypsin. These data are summarized schematically in Figure 14 where the sequence of βT-9a is shown as Val-Leu-Thr-Ser-Phe-Gly-Glu-Leu-Lys.

8. βT-9b

The amino acid composition of βT-9b is presented in Table 23. Hydrolysis of the whole peptide with leucine aminopeptidase showed one residue each of aspartic acid and asparagine. A portion of this peptide was subjected to dilute acetic acid hydrolysis. The results of this hydrolysis are shown in Table 24. Dansyl-Edman degradation of βT-9b revealed the N-terminal sequence to be His-Leu. Digestion of βT-9b-HAc-3 with leucine aminopeptidase showed the presence of one asparagine residue. Sequential dansyl-Edman degradation of this acetic acid fragment determined the N-terminal portion of this peptide to be Asn-Leu. The tryptic specificity places the lysine at the C-terminus of this tripeptide and this peptide as the C-terminal

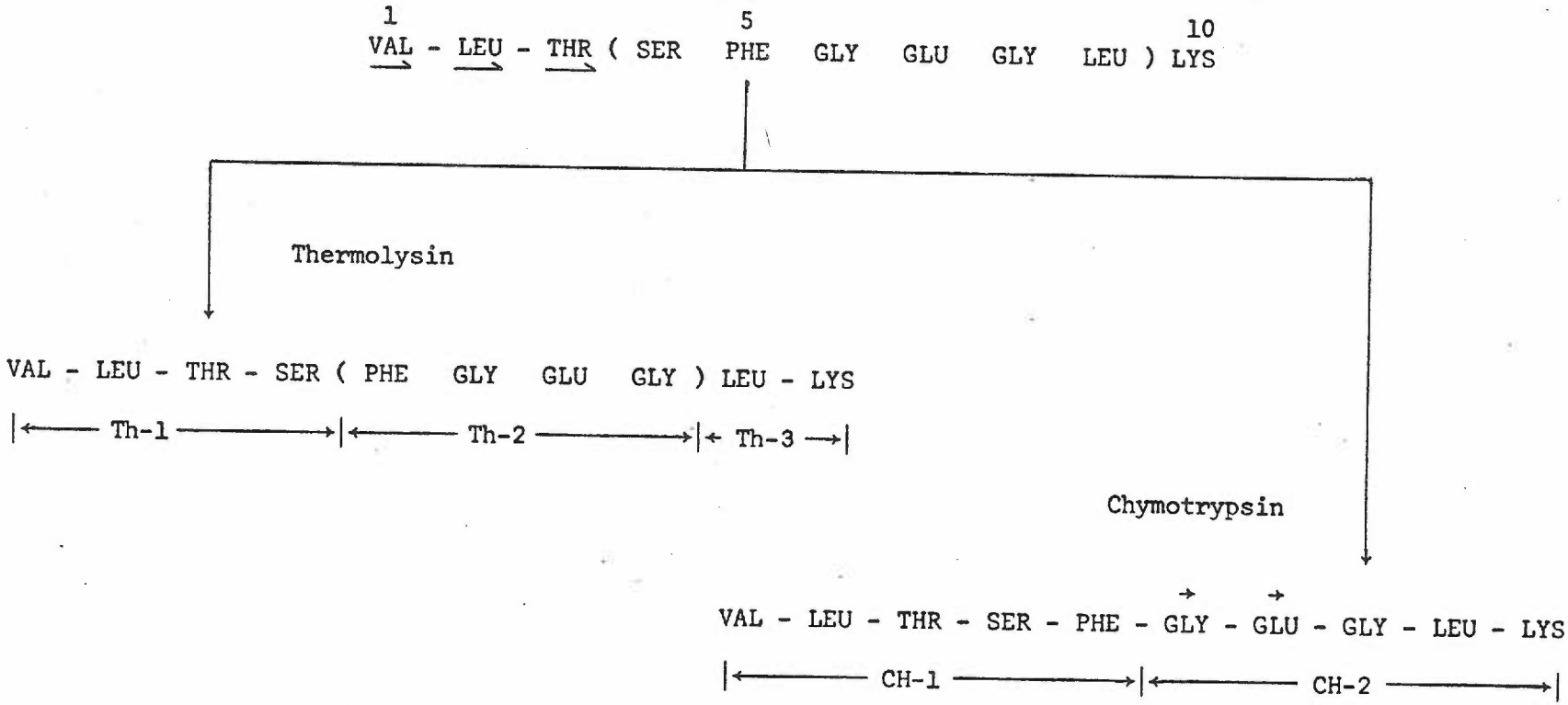


Figure 14

Schematic Representation of the Flow Diagram of βT-9a

Table 23

THE AMINO ACID ANALYSIS OF βT-9b

Amino Acid	Calculated Values	Integral Values
Lysine	1.06	1
Histidine	0.98	1
Aspartic Acid	2.00	2
Leucine	1.96	2

Table 24

AMINO ACID ANALYSIS OF ACETIC ACID FRAGMENTS 1 AND 2 OF βT-9b

	HAc-1		HAc-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			0.96	1
Histidine	0.96	1		
Aspartic Acid			0.92	1
Leucine	1.04	1	1.06	1

fragment of βT-9b. This sequence data allow the placement of residues 1,2, 4,5 and 6; therefore, Asp is residue 3. The sequence of βT-9b is His-Leu-Asp-Asn-Leu-Lys. A summary for the sequence of this peptide is presented in Figure 15.

9. βT-10

Tabulated in Table 25 are the amino acid compositions of βT-10 and two fragments, βT-10a and βT-10b, that result from incomplete tryptic hydrolysis. When βT-10a was hydrolysed with leucine aminopeptidase, asparagine and glutamic acid residues were found. Similarly, enzyme digestion of βT-10b yielded an aspartic acid residue. When βT-10 was incubated with thermolysin, four fragments were produced, as shown in Table 26. Both the whole peptide, βT-10, and the thermolysin fragment, βT-10-Th-4, were partially sequenced by the dansyl-Edman method. These sequences were: Gly-Thr-Phe for βT-10 and Leu-His for βT-10-Th-4. The sequences of βT-10-Th-2 and βT-10-Th-3 were determined using the Edman difference method. The thermolysin fragment, βT-10-Th-2, has a N-terminal phenylalanine and an alanine for the second residue. Thermolysin fragment, βT-10-Th-3, yielded the sequence Leu-Ser. The accumulated data is sufficient to establish the sequence of βT-10 as: Gly-Thr-Phe-Ala-Asn-Leu-Ser-Glu-Leu-His-Cys-Asp-Lys. These data are summarized in Figure 16.

10. βT-11

The amino acid composition of βT-11 is shown in Table 27. A secondary split of this peptide was obtained by hydrolysing the

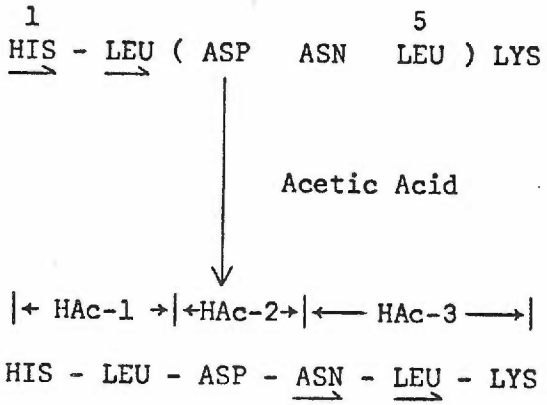


Figure 15

Flow Diagram for βT-9b

Table 25

AMINO ACID ANALYSES OF βT-10, βT-10a AND βT-10b

	βT-10		βT-10a		βT-10b	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine	1.10	1			0.99	1
Histidine	1.30	1	1.10	1		
A-E Cysteine	0.50	1	present	1		
Aspartic Acid	2.04	2	1.24	1	1.01	1
Threonine	1.00	1	1.00	1		
Serine	0.88	1	0.96	1		
Glutamic Acid	1.08	1	0.99	1		
Glycine	1.07	1	1.12	1		
Alanine	0.98	1	0.95	1		
Leucine	1.93	2	1.84	2		
Phenylalanine	1.02	1	0.90	1		

Table 26

AMINO ACID ANALYSIS OF THE FOUR THERMOLYSIN FRAGMENTS OF βT-10

	Th-3		Th-2		Th-4		Th-1	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine					1.11	1		
Histidine					0.92	1		
A-E Cysteine					present	1		
Aspartic Acid			1.15	1	1.11	1		
Threonine							0.91	1
Serine	0.95	1						
Glutamic Acid	1.01	1						
Glycine							1.02	1
Alanine			1.05	1				
Leucine	0.93	1			0.90	1		
Phenylalanine			0.80	1				

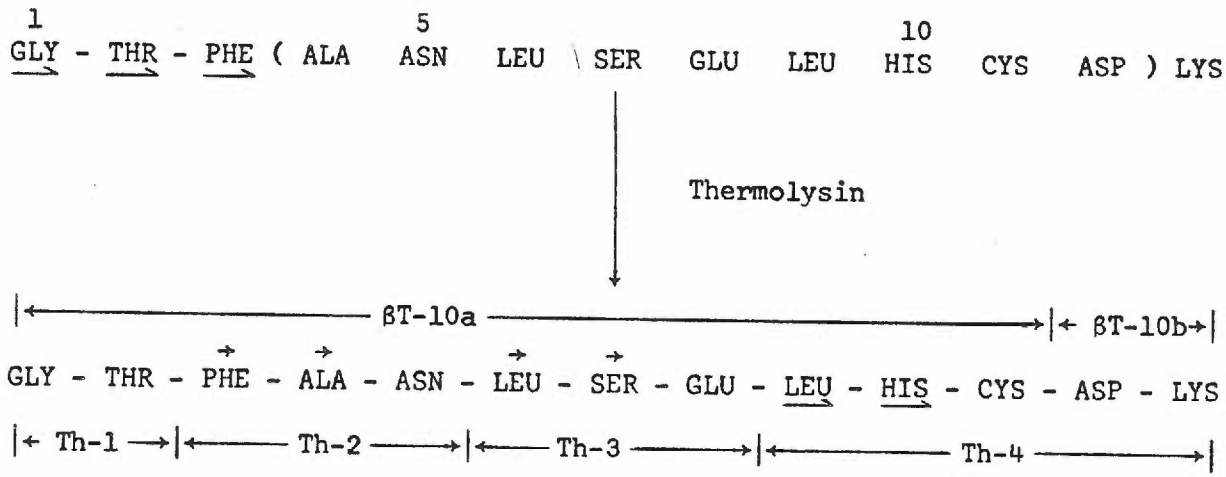


Figure 16

Flow Diagram for βT-10

Table 27

AMINO ACID ANALYSIS OF βT-11

Amino Acid	Calculated Values	Integral Values
Histidine	0.97	1
Arginine	0.82	1
Aspartic Acid	2.04	2
Glutamic Acid	1.03	1
Proline	1.16	1
Valine	0.89	1
Leucine	1.07	1
Phenylalanine	0.99	1

sample in dilute acetic acid. The results of that hydrolysis are tabulated in Table 28. Aminopeptidase M digestion of acetic acid fragment 1, βT-11-HAc-1, yielded one aspartic acid residue. Digestion of βT-11-HAc-2 with aminopeptidase M proved there to be one residue each of glutamic acid and asparagine. Sequential dansyl-Edman degradation of the whole peptide and acetic acid fragment 2 produced the following results. The N-terminal sequence of βT-11 was Leu-His-Val and acetic acid fragment 2 was Pro-Glu-Asn. These data are sufficient to establish the sequence of βT-11 as Leu-His-Val-Asp-Pro-Glu-Asn-Phe-Arg. These results are summarized in Figure 17.

11. βT-12a

At this writing βT-12a has not been found by the author. Several methods of attacking this problem peptide have been tried without success. These methods include: exhaustively searching through the tryptic peptide pattern zones of the aminoethylated β chain; investigating the tryptic peptides obtained from non-aminoethylated β chain; and finally attempting to block the epsilon amino groups of lysine with citraconic anhydride and isolating the large polypeptide which should have contained βT-12a. This problem has received a lot of attention in the past and experiments are planned for the future which should prove the presence or absence of this peptide in elephant hemoglobin. These experiments include carboxymethylation of the cysteine residues of elephant β chains with ^{14}C, iodoacetic acid. This will provide an additional method of tracing

Table 28

AMINO ACID ANALYSIS OF ACETIC ACID FRAGMENTS 1 AND 2 OF βT-11

	HAc-1		HAc-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Arginine			0.99	1
Histidine	1.01	1		
Aspartic Acid	0.94	1	0.87	1
Glutamic Acid			1.07	1
Proline			1.02	1
Valine	1.05	1		
Leucine	0.99	1		
Phenylalanine			1.05	1

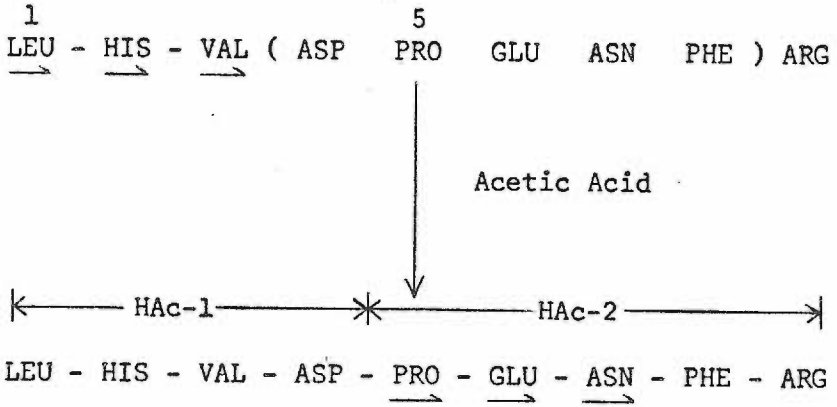


Figure 17

Flow Diagram for βT-11

this peptide because the radioactivity of the labelled peptide can be monitored.

12. βT-12b

The amino acid analysis of βT-12b is shown in Table 29. βT-12b was incubated with thermolysin. Two fragments were isolated by ion exchange chromatography and their amino acid compositions are shown in Table 30. Sequential dansyl-Edman degradation of βT-12b revealed the sequence Val-Leu-Ala-His. Similar sequence analysis of the thermolysin fragment, βT-12b-Th-2, proved the N-terminal region to be Phe-Gly. These data, as summarized in Figure 18, show the complete sequence of βT-12b to be Val-Leu-Ala-His-His-Phe-Gly-Lys.

13. βT-13

Quantitative amino acid analysis of βT-13 is tabulated in Table 31. This peptide was susceptible to hydrolysis by both chymotrypsin and thermolysin. The results of these enzymatic digestions are shown in Tables 32 and 33.

Digesting thermolysin fragments βT-13-Th-1, βT-13-Th-2, and βT-13-Th-3 with aminopeptidase M revealed one glutamine residue in thermolysin fragment 3, one glutamic acid residue in thermolysin fragment 2, and one residue each of glutamic acid, glutamine and aspartic acid in thermolysin fragment 1.

Sequential dansyl-Edman degradation on βT-13, chymotryptic fragments βT-13-Ch-1, βT-13-Ch-2, and βT-13-Ch-3 produced the following sequence results: βT-13 N-terminal sequence was Glu-Phe;

Table 29

AMINO ACID ANALYSIS OF βT-12b

Amino Acid	Calculated Values	Integral Values
Lysine	1.07	1
Histidine	1.90	2
Glycine	1.01	1
Alanine	0.95	1
Valine	0.94	1
Leucine	1.05	1
Phenylalanine	1.05	1

Table 30

AMINO ACID ANALYSIS OF THE THERMOLYSIN FRAGMENTS 1 AND 2 OF βT-12b

	Th-1		Th-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.12	1
Histidine	1.84	2		
Glycine			0.91	1
Alanine	0.97	1		
Valine	1.06	1		
Leucine	1.12	1		
Phenylalanine			0.96	1

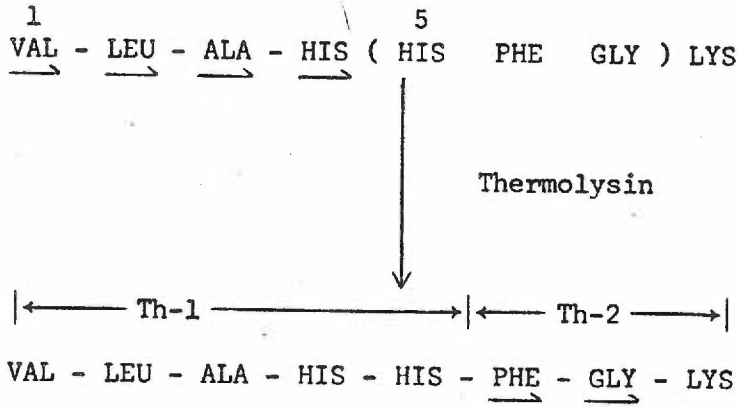


Figure 18

Flow Diagram for βT-12b

Table 31

AMINO ACID ANALYSIS OF βT-13

Amino Acid	Calculated Values	Integral Values
Lysine	1.06	1
Aspartic Acid	1.12	1
Threonine	1.04	1
Glutamic Acid	2.60	3
Proline	1.20	1
Alanine	1.93	2
Valine	1.00	1
Tyrosine	1.02	1
Phenylalanine	0.93	1

Table 32

AMINO ACID ANALYSIS OF CHYMOTRYPTIC FRAGMENTS 1, 2 AND 3 OF βT-13

	Ch-1		Ch-2		Ch-3	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.07	1		
Aspartic Acid	1.04	1				
Threonine	0.98	1				
Glutamic Acid	1.91	2	0.94	1		
Proline	1.06	1				
Alanine	2.04	2			2.10	2
Valine	1.01	1				
Tyrosine	0.93	1			0.90	1
Phenylalanine	1.04	1				

Table 33

AMINO ACID ANALYSIS OF THE THERMOLYSIN FRAGMENTS 1, 2 AND 3 OF βT-13

	Th-1		Th-2		Th-3	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.12	1		
Aspartic Acid	0.99	1				
Threonine	0.85	1				
Glutamic Acid	1.98	2	1.03	1	1.06	1
Proline	1.10	1				
Alanine	1.10	1	1.00	1	1.94	2
Valine	0.94	1			1.01	1
Tyrosine			0.91	1		
Phenylalanine	1.04	1				

chymotryptic fragment 2 had a glutamic acid residue as its N-terminus; chymotryptic fragment 1 revealed the N-terminal sequence of Glu-Phe-Thr-Pro-Asp-Val; and chymotryptic fragment 3 had an alanine residue as its N-terminus and another alanine for residue 2. This composite of data, summarized in Figure 19, establishes the sequence of βT-13 as: Glu-Phe-Thr-Pro-Asp-Val-Gln-Ala-Ala-Tyr-Glu-Lys.

14. βT-14

The amino acid analysis of βT-14 is shown in Table 34. Exhaustive digestion of aminoethylated β chains with trypsin occasionally produced two fragments of βT-14 which are shown in Table 35. This split is probably the result of some contamination of the trypsin used by chymotrypsin. Hydrolysis of βT-14 with leucine aminopeptidase revealed the presence of an asparagine residue.

Dansyl-Edman degradation of the whole tryptic peptide revealed the N-terminal sequence to be Val-Val-Ala-Gly-Val-Ala. Similarly, the N-terminal sequence of fragment 2 (Table 35) was found to be Ala-Leu-Ala.

The combination of these data established the sequence of this peptide to be Val-Val-Ala-Gly-Val-Ala-Asn-Ala-Leu-Ala-His-Lys. These results are summarized in Figure 20.

15. βT-15

The amino acid sequence of βT-15 is recorded in Table 36. Dansylation of this peptide revealed a tyrosine residue. Therefore, the sequence of this peptide is Tyr-His.

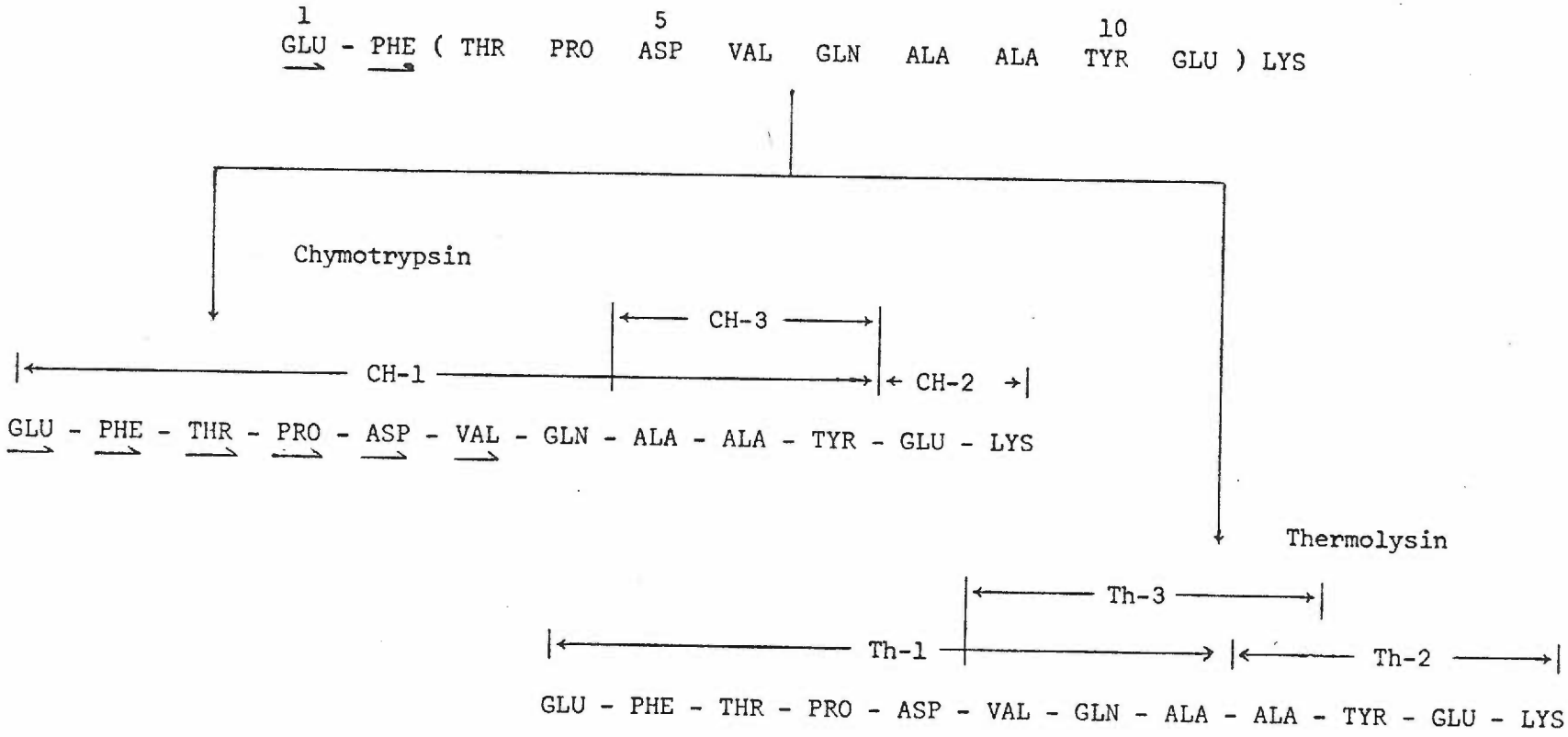


Figure 19

Flow Diagram for βT-13

Table 34

AMINO ACID ANALYSIS OF βT-14

Amino Acid	Calculated Values	Integral Values
Lysine	1.07	1
Histidine	1.18	1
Aspartic Acid	1.09	1
Glycine	1.20	1
Alanine	3.78	4
Valine	2.82	3
Leucine	1.02	1

Table 35

AMINO ACID ANALYSIS OF THE CHYMOTRYPTIC FRAGMENTS 1 AND 2 OF βT-14

	Ch-1		Ch-2	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values
Lysine			1.09	1
Histidine			0.98	1
Aspartic Acid	1.20	1		
Glycine	1.18	1		
Alanine	2.34	2	1.94	2
Valine	2.25[a]	3		
Leucine			0.98	1

[a] This peptide contains a Val-Val bond which is resistant to acid hydrolysis. Therefore this value is lower than expected.

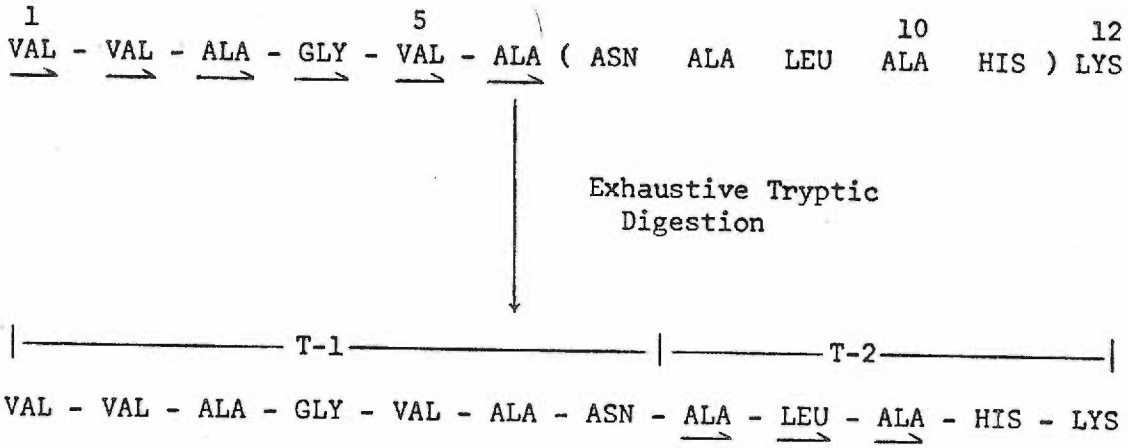


Figure 20

Flow Diagram for βT-14

Table 36

AMINO ACID ANALYSIS OF βT-15

Amino Acid	Calculated Values	Integral Values
Histidine	1.02	1
Tyrosine	0.98	1

This peptide was placed at the C-terminus of the elephant β chain for three reasons: 1. This is the only tryptic peptide that does not contain lysine, arginine or aminoethyl cysteine. 2. All known mammalian β chains have Tyr-His as the C-terminal residues. 3. Data from experiments on treating the elephant β chains with citraconic anhydride and the hydrolyzing with trypsin prove that βT-12b through 15 are connected.

G. Citraconylated β Chain

Part of the aminoethylated β chain was reacted with citraconic anhydride. Citraconic anhydride forms an acid labile covalent bond to the epsilon nitrogens of the lysine amino acid residues in peptides. Subsequent digestion with trypsin will cleave only at the arginyl residues. The citraconylated aminoethylated β chain digested with trypsin was chromatographed on a Sephadex G-50 column. Figure 21 shows the chromatogram of this separation.

Zone I is at the column void volume and contained a complex mixture which was not resolved further. Quantitative amino acid analysis of zone II, CA-4, revealed 64 amino acid residues and is presented in Table 37. This analysis corresponded exactly to residues tentatively placed in positions 41 through 104 by homology of tryptic pieces with other known β chain sequences. This zone was hydrolysed with trypsin after removing the blocking groups as described previously. The hydrolysis produced βT-5, βT-6-7-8, βT-9a, βT-9b, βT-10 and βT-11 which were separated by ion exchange chromatography.

Figure 21

Sephadex G-50 Separation of Citraconylated, Tryptic Digested β-Chain

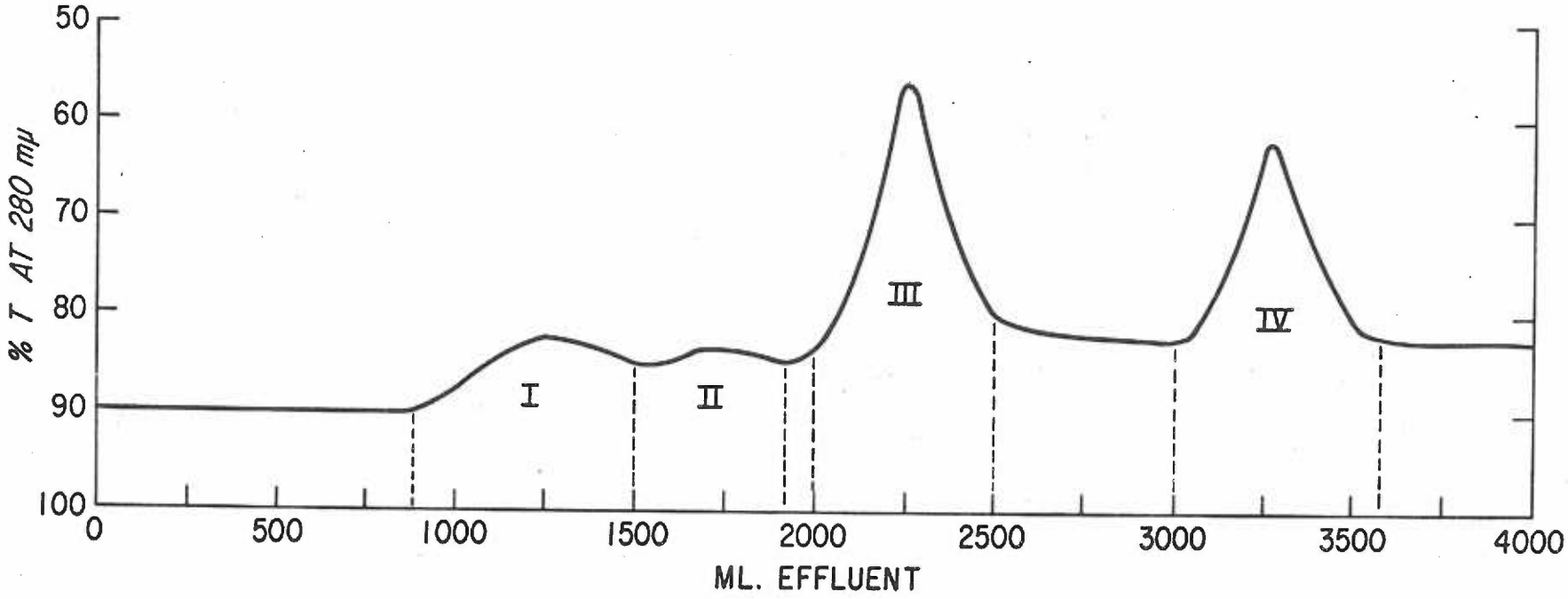

50
60
70
80
90
100
% T AT 280 mμ
I
II
III
IV
0
500
1000
1500
2000
2500
3000
3500
4000
ML. EFFLUENT

Table 37

AMINO ACID ANALYSIS OF PEAK II FROM FIGURE 21. RESIDUES 41-104 OF ASIAN ELEPHANT β CHAIN

Amino Acid	Calculated Values	Integral Values	Expected Values
Lysine	4.9	5	5
Histidine	5.4	5	6
A-E Cysteine	0.22	1	1
Arginine	1.1	1	1
Aspartic Acid	8.8	9	9
Threonine	2.9	3	3
Serine	2.8	3	3
Glutamic Acid	5.5	5	5
Proline	1.3	1	1
Glycine	5.3	5	5
Alanine	5.3	5	5
Valine	4.1	4	4
Leucine	10.2	10	10
Phenylalanine	6.0	6	6
Total	63.7	63	64

None of the other tryptic peptides of the elephant β chain were detected in this zone.

Peak III was a mixture of two peptides, one containing 30 amino acids, CA-1, and another thought to be 41 amino acids in length, CA-5. A partial separation of these two peaks was achieved by recycling this mixture on a Sephadex G-50 column. As shown in Table 38, the 30 residue peptide has a composition which corresponds to residues 1-30 of the elephant β chain. It is quite pure. However, the second peptide which is suspected to be 41 amino acid residues in length could not be obtained pure. Tryptic digestion of the peptide containing 30 amino acids, CA-1, after removing the blocking groups resulted in 4 tryptic peptides. These peptides were separated on an Aminex A-5 column and were shown to be βT-1, βT-2, βT-3a and βT-3b. When the other impure peptide, containing approximately 41 residues, CA-5, was hydrolysed with trypsin and chromatographed, tryptic peptides corresponding to βT-1,2,3a,3b, and βT-12b, 13, 14, and 15 were isolated. The first four tryptic peptides were concluded to be from contamination by the 30 residue, CA-1, peptide.

Peak IV contained a mixture of βT-4a, βT-4b, and βT-4ab. βT-4ab, CA-2-3, was shown to be a composite peptide of βT-4a, CA-2, and βT-4b, CA-3. These three peptides were separated on a Dowex 50W-X2 column. Table 39 shows the amino acid analysis of these three peptides. The position of βT-4a in the β chain of elephant hemoglobin was ambiguous. βT-4a could correspond to either residues 31-39 or

Table 38

AMINO ACID ANALYSIS OF RESIDUES 1-30 OF THE ASIAN ELEPHANT HEMOGLOBIN β CHAIN

Amino Acid	Calculated Values	Integral Values	Expected Values
Lysine	3.29	3	3
Arginine	0.85	1	1
Tryptophan	present	1	1
Aspartic Acid	3.06	3	3
Threonine	2.20	2	2.5
Serine	0.80	1	1
Glutamic Acid	3.73	4	4
Glycine	3.01	3	3
Alanine	3.86	4	3.5
Valine	4.12	4	4
Leucine	4.30	4	4
Total	30.22	30	30

Table 39

AMINO ACID ANALYSIS OF βT-4a, βT-4b AND βT-4ab

	βT-4a		βT-4b		βT-4ab	
Amino Acid	Calculated Values	Integral Values	Calculated Values	Integral Values	Calculated Values	Integral Values
Arginine	1.02	1	1.00	1	2.15	2
Threonine	0.99	1			0.96	1
Proline	0.99	1			0.99	1
Valine	1.65	2			1.46	2
Leucine	2.02	2			2.08	2
Tyrosine	0.92	1			0.94	1
Tryptophan	present	1			present	1

32-40 of the β chain. Table 40 shows the amino acid analysis of βT-4ab before and after one Edman degradation. In Table 40 the mole ratio of leucine decreases from 2 to 1 clearly revealing that leucine must be the N-terminus of this peptide. Therefore, residues 31-39 of the β chain correspond to βT-4a and residue 40 is βT-4b, arginine.

H. Sequence of β Chain

Figure 22 shows the amino acid sequence of the β chain of elephant hemoglobin. The numbers above and below the sequence identify the tryptic and citraconic anhydride peptides, respectively.

Table 40

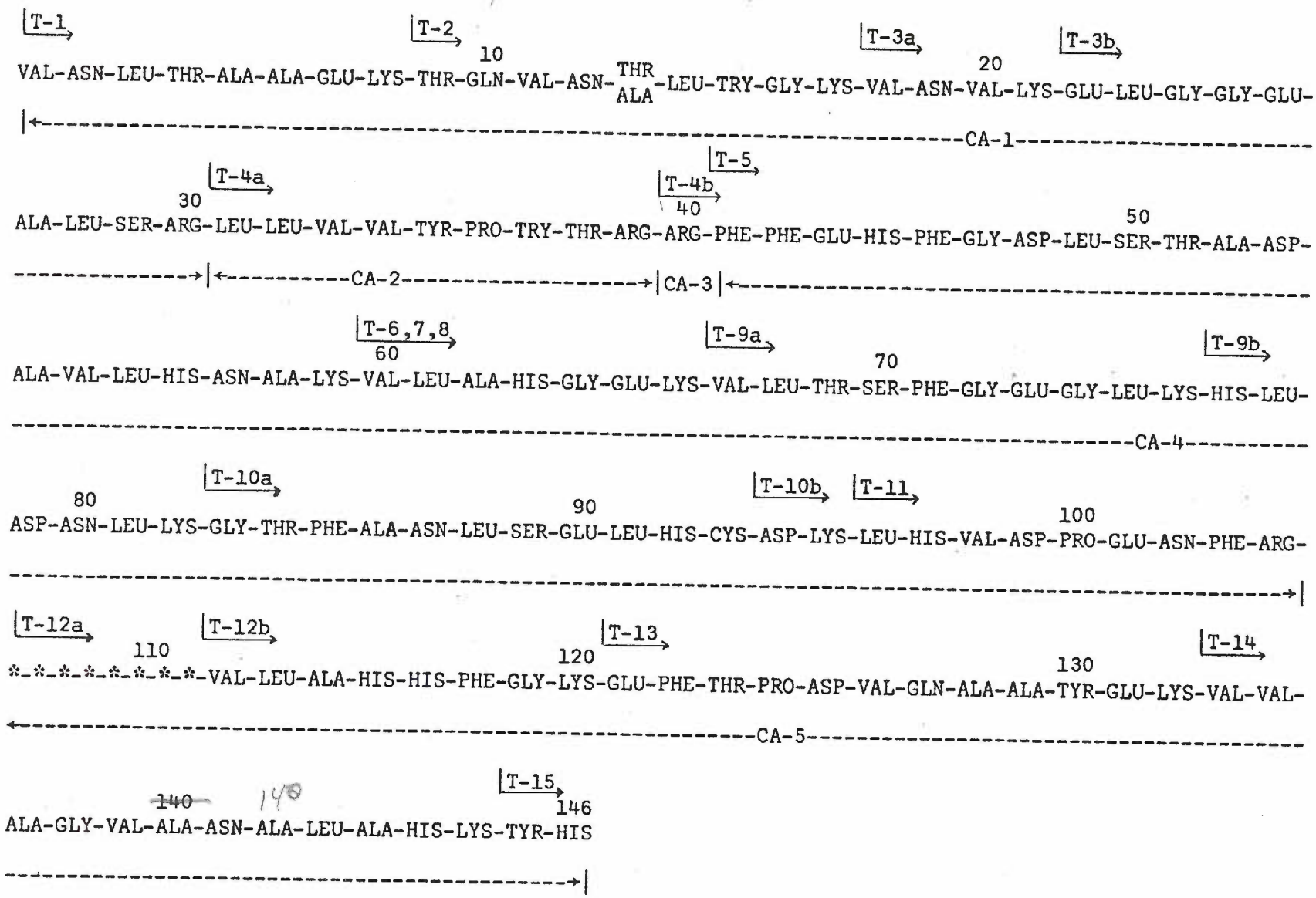
AMINO ACID ANALYSIS OF βT-4ab BEFORE AND AFTER EDMAN DEGRADATION

	βT-4ab		I-Edman	
Amino Acid	Calculated[a] Values	Integral Values	Calculated[a] Values	Integral Values
Arginine	2.15	2	2.00	2
Threonine	1.09	1	1.15	1
Proline	1.21	1	1.03	1
Valine	1.51	2	1.95	2
Leucine	2.17	2	1.20	1
Tyrosine	1.03	1	0.76	1

[a] Both samples were hydrolysed for 48 hours.

Figure 22

Amino Acid Sequence of Asian Elephant β-Chain


T-1
T-2
T-3a
T-3b
10
20
VAL-ASN-LEU-THR-ALA-ALA-GLU-LYS-THR-GLN-VAL-ASN-THR/ALA-LEU-TRY-GLY-LYS-VAL-ASN-VAL-LYS-GLU-LEU-GLY-GLY-GLU-
CA-1
T-5
T-4a
T-4b
30
40
50
ALA-LEU-SER-ARG-LEU-LEU-VAL-VAL-TYR-PRO-TRY-THR-ARG-ARG-PHE-PHE-GLU-HIS-PHE-GLY-ASP-LEU-SER-THR-ALA-ASP-
CA-2
CA-3
T-6,7,8
T-9a
T-9b
60
70
ALA-VAL-LEU-HIS-ASN-ALA-LYS-VAL-LEU-ALA-HIS-GLY-GLU-LYS-VAL-LEU-THR-SER-PHE-GLY-GLU-GLY-LEU-LYS-HIS-LEU-
CA-4
T-10a
T-10b
T-11
80
90
100
ASP-ASN-LEU-LYS-GLY-THR-PHE-ALA-ASN-LEU-SER-GLU-LEU-HIS-CYS-ASP-LYS-LEU-HIS-VAL-ASP-PRO-GLU-ASN-PHE-ARG-
T-12a
T-12b
T-13
T-14
110
120
130
--*-*-*-*-*-*-VAL-LEU-ALA-HIS-HIS-PHE-GLY-LYS-GLU-PHE-THR-PRO-ASP-VAL-GLN-ALA-ALA-TYR-GLU-LYS-VAL-VAL-
CA-5
T-15
140
140
146
ALA-GLY-VAL-ALA-ASN-ALA-LEU-ALA-HIS-LYS-TYR-HIS

IV. DISCUSSION

A. Methodology in Determining the Amino Acid Sequence

The methods used were either identical to or slight modifications of reported procedures. The peptides that resulted from the hydrolysis of the Asian elephant hemoglobin β chain with trypsin were separated by ion exchange chromatography and further purified by rechromatographing on other ion exchangers. All peptides studied had less than 0.2 of a residue contamination of spurious amino acids.

Figure 18 shows the 138 residues definitely established for the β chain of hemoglobin from the Asian elephant. The missing portion of the sequence, residues 105-112, is the location for the as yet elusive peptide, βT-12a. Additional experiments are planned to clarify the presence or absence of this peptide in the β chain. These include the technique of blocking the lysine residues with citraconic anhydride and subsequent cleavages at the arginine residues with trypsin. Hopefully, an improved purification method devised for isolating these fragments will aid in locating βT-12a.

B. Validity of the Sequence

Several animal hemoglobin chain sequences have been determined (21). The sequence of the Asian elephant β chain does show obvious homologous structure with these other completed sequences. As might be expected, it was relatively simple to assign the order of the tryptic peptides once their amino acid composition was obtained.

Other hemoglobin chains have been either completely or partially sequenced by this comparison method (21).

One of the reasons for attempting the citraconic anhydride blocking of the epsilon amino groups of lysine and cleavage at the arginine residues with trypsin was as a check on the alignment of the tryptic peptides. The isolated citraconic anhydride fragments, upon further digestion with trypsin, produced only the predicted tryptic peptides. For example, purified CA-1, a 30 residue polypeptide, contained only βT-1,2,3a and 3b. Therefore, the citraconic anhydride fragments are additional evidence that the sequence as shown in Figure 18 is correct.

C. Features of the Sequence

βT-2 was found to be heterogeneous. This was deduced from the presence of nearly equal molar amounts of threonine and alanine at residue 13 of the β chain. This suggests the presence of two different β chains. Although this heterogeneity in the Asian elephant β chain was an unexpected result, these data are consistent with and similar to those previously reported (21,23,24,35). For example, dog hemoglobin has been reported to have two α chains (35). This was proposed from the determination of a threonine-alanine heterogeneity in position 130. It has been suggested that rabbit hemoglobin has several α chains. These results were proposed from the determination of several amino acid multiplicities in the sequence.

Asian elephant β chain was completely devoid of isoleucine

and methionine. Many mammalian hemoglobin β chain sequences are missing at least one of these amino acids (21).

D. Structure and Function

Asian elephant hemoglobin has an oxygen affinity higher than normal human hemoglobin. A comparison of the sequences of Asian elephant hemoglobin β chain with other hemoglobin β chains may suggest the residue(s) which are responsible for this high oxygen affinity.

Several ligands are known to affect the oxygen affinity of hemoglobin. Among these ligands is 2,3 diphosphoglycerate (DPG). DPG has been shown to affect the oxygen affinity curve of hemoglobin in a physiologically advantageous manner. DPG shifts the oxygen affinity curve to the right, which means that it lowers the oxygen affinity of the hemoglobin.

Dhindsa et al. (2) could not demonstrate a significant difference in the DPG levels between elephant and human blood. Therefore, the high affinity for oxygen of Asian elephant hemoglobin may not be explained by reduced levels of 2,3 diphosphoglycerate in the blood.

The residues thought to participate in the binding site for DPG (16) were compared with the determined elephant β chain sequence. Both residues believed to bind DPG, the N-terminal valine and the histidine at position 143, are present in this β chain. Therefore, the high affinity for oxygen of Asian elephant hemoglobin may not be explained by changes in the binding site, which could alter the degree

of the DPG effect.

The oxygen equilibrium curve of hemoglobin has a sigmoidal shape. This has been attributed to interaction between the four hemes in one molecule of hemoglobin. Because the four hemes are rather widely separated, the interaction between the heme groups must be transmitted via the protein moeity surrounding the heme groups. Studies of mutant hemoglobins have shown several locations where protein modifications and amino acid substitutions cause changes which increase oxygen affinities. These sites are at the contacts of the protein residues with the heme group and the interatomic contacts between residues in the α_1-β_2 contact region.

Perutz (11) described the contacts between 21 residues of the β chain of horse hemoglobin and the heme group. In comparing these 21 residues with the amino acid sequence of the elephant β chain, all but one of these residues are identical. The single difference is a histidine in the elephant β chain for the serine residue present in the horse hemoglobin β chain at position 44. Other mammalian hemoglobin β chains have this same substitution (21) and none of these hemoglobins are reported to have unusually high oxygen affinities. So, although possible, it is highly unlikely that this substitution is the sole reason for the reported high oxygen affinity of Asian elephant hemoglobin.

There are two different types of contacts between unlike chains called α_1-β_1 and α_1-β_2. Thirty-four residues are involved in

the α_1-β_1 contact region. Because of this large number of residues, the subunits at the α_1-β_1 contact region have only a slight movement in the transition from the oxygenated to the deoxygenated form of hemoglobin. Twenty-one of these contact residues are common to all mammalian hemoglobin sequences known to date. There are five amino acid differences between the elephant β chain sequence and the α_1-β_1 contact residues described by Perutz (11) for horse hemoglobin β chain sequence. None of these five residues violate the invariable residues mentioned for the α_1-β_1 contact region. Because this contact has been shown to be only slightly affected by oxygenation and the five substitutions appear to be in non-critical locations it is unlikely, although possible, that this contact is the cause for the high oxygen affinity of this hemoglobin.

The α_1-β_2 contact comprises nineteen residues. Movement of the subunits at this contact region is quite large compared to that at the α_1-β_1 contact in the transition from the oxygenated to the deoxygenated form of hemoglobin. Also, the α_1-β_2 contact region is closely connected to the heme groups so that changes in this area would be expected to affect the heme environment. Of the 19 residues that comprise the α_1-β_2 contact, all but one of these have been common in the sequences of the hemoglobin chains. The one exception was reported for the β chain sequence of llama hemoglobin, which has an arginine residue in place of a glutamine at position 39. In comparing the elephant β chain sequence with the α_1-β_2 contact residues

this same substitution that was reported for llama is present in the elephant β chain. The other β chain contact residues were identical. Perutz (11) suggests that this replacement, arginine for a glutamine, at position 39 would have no effect on the stereochemistry of the contact. Hence, it is also unlikely that this substitution can explain the high oxygen affinity of this molecule.

From these data one can conclude that an attempt at deciding which single residue is responsible for altering the oxygen affinity of a molecule is very difficult. Perhaps an explanation of this high oxygen affinity may be the sum of many very slight changes at several residues.

E. Relationship to Other Species

All living organisms are related and for any two species alive today most of the common ancestors are extinct. Evidence for this evolutionary interpretation of relationship between species has been drawn from morphology, biochemistry, embryology, and paleontology. Zuckerkandl and Pauling (111) were first to suggest that amino acid sequences of homologous proteins might be the basis of quantitative phylogenetic schemes. Their idea stemmed from the observation that the hemoglobins of man and gorilla are very similar, but that horse hemoglobin, clearly homologous to the other two, differs at many sites. As more protein sequences are determined, additional support for this hypothesis is obtained.

Aguirre (3) has noted that the common ancestor of the Asian

and African elephants is unclear from fossil evidence. A comparison of the amino acid sequences of the hemoglobin chains of these two currently living species may contribute the necessary information to clarify this situation. Some differences in the hemoglobins of these two species have been noted. Investigations (19) have shown that the African elephant hemoglobin migrates faster than Asian elephant hemoglobin and both of them migrate faster than normal human hemoglobin by gel electrophoresis. Preliminary results have shown that the α and β chains of the African elephant hemoglobin can be separated (112). Other homologous amino acid sequences have been employed to deduce the phylogenetic relationship of species (113).

Comparisons of homologous proteins from diverse species have so many identical amino acids that they can be aligned with one another with full confidence. The probability that so many identities could have arisen independently seems negligible. The following compilation will consider only those animal hemoglobin β chain sequences which have been completely[1] determined. Table 41 shows the amino acid sequence differences between nine species. Six of these species, human, monkey, rabbit, horse, sheep and kangaroo, were compiled by Air et al. (25) but corrected to exclude the differences determined for residues 105-112 because these residues have not been

[1] Completely is defined as 90% certain of being correct by Dayhoff (21)

Table 41

NUMBER OF AMINO ACID SEQUENCE DIFFERENCES BETWEEN BETA GLOBINS

	Human	Monkey	Horse	Sheep	Rabbit	Dog	Elephant	Kangaroo	Frog
Human[a]	0	8	24	25	13	15	25	35	58
Monkey[a]	8	0	26	26	15	16	27	33	58
Horse[a]	24	26	0	32	23	26	33	42	54
Sheep[a]	25	26	32	0	28	25	34	40	56
Rabbit[a]	13	15	23	28	0	21	29	35	55
Dog	15	16	26	25	21	0	26	34	53
Elephant	25	27	33	34	29	26	0	38	52
Kangaroo[a]	35	33	42	40	35	34	38	0	63
Frog[b]	58	58	54	56	55	53	52	63	0

[a] Differences calculated by Air et al. (25) have been modified to exclude residues 105-112 for comparisons.

[b] The beta chain of frog hemoglobin has a deletion of the first six residues at the N-terminus. When the frog β chain was compared with the other species β chains these residues were not counted.

determined for the elephant β chain. Amino acid sequence differences were compiled by the author for the other three species which are elephant, dog and frog. The six residues deleted from the amino terminal end of the frog β chain were excluded from the difference calculation when the frog β chain sequence was compared with the other β chain sequences. Some investigators count deletions as differences, thus the comparison of the frog β chain sequence with the other β chain sequences, counting the deletion of the N-terminal residues as differences, would increase the differences shown in Table 41 by six.

Table 41 shows that when the elephant hemoglobin β chain sequence was compared with dog, rabbit, horse, sheep, monkey and human there were fewer differences than when compared to the kangaroo, a marsupial. This suggests, as might be expected, that the elephant is less related to marsupials than other mammals. These data are consistent with phylogenetic trees based on paleontological information.

F. Substitution in the α Chain

An alternative possibility for explaining the observed high oxygen affinity of this hemoglobin is a substitution(s) in the α chain. Such a substitution in the $\alpha_1\beta_2$ contact site could cause the noted high oxygen affinity of this hemoglobin.

V. SUMMARY AND CONCLUSION

The goal set forth in this thesis was to determine the primary structure of the β chain of Asian elephant, *Elephas maximus*, hemoglobin. The sequence of the Asian elephant β chain does show obvious homologous structure with other completed globin sequences. The sequential order of the tryptic peptides was assigned readily once the amino acid composition of the tryptic peptides was known. Additional evidence supporting the placement of the tryptic peptides was obtained from the isolated citraconic anhydride fragments. These fragments were produced by first blocking the lysine residues of the β chain with citraconic anhydride and then splitting the peptide at the arginine residues with trypsin. The citraconic anhydride fragments were purified and hydrolysed with trypsin again after removing the blocking groups. This hydrolysis contained only the tryptic peptides expected. For example: One citraconic anhydride fragment contained 64 amino acid residues and only the tryptic peptides assigned to βT-5 through βT-11 could be detected from the tryptic hydrolysis of this fragment.

The second tryptic peptide, βT-2, was found to be heterogeneous. This suggests that two separate β chains exist in Asian elephant hemoglobin.

Asian elephant hemoglobin has been shown to have an oxygen affinity much higher than human hemoglobin. It was proposed that the determination of the primary structure of this hemoglobin might

establish the residue(s) which cause the high oxygen affinity observed for this molecule. Several possible sites were shown to have the potential of altering oxygen affinities of hemoglobins. The elephant β chain sequence was compared to regions discussed. It was suggested that perhaps this high oxygen affinity was not caused by a single residue but rather a composite of many residues, each contributing a small fraction of the difference in the oxygen affinities observed for man and elephant.

The relatedness of the Asian elephant hemoglobin to other species was examined. The results indicated that the elephant is more related to man, monkey, dog, sheep and rabbit than to the kangaroo.

Bibliography

1. Sanger, R., Thompson, E. O. P., and Kitai, R. The amide groups of insulin. Biochem. J., 1955. 59, 509-518.

2. Dhindsa, D. S., Sedgwick, C. J., and Metcalfe, J. Comparative studies of the respiratory functions of mammalian blood. VIII. Asian elephant (Elephas maximus) and African elephant (Loxondonta africana africana). Respiration Physiology, 1972. 14, 332-342.

3. Aguirre, E. Evolutionary history of the elephant. Science, 1969. 164, 1366-1376.

4. Schroeder, W. A., and Jones, R. T. Some aspects of the chemistry and function of human and animal hemoglobins. Fortschritte der Chemie organischer Naturstoffe, 1965. 23, 115-194.

5. Perutz, M. F. The haemoglobin molecule. Proc. Roy. Soc. B., 1969. 173, 113-140.

6. Wald, G., and Riggs, A. J. The hemoglobin of the sea lamprey *Petromyzon marinus*. J. Gen. Physiol., 1951. 35, 45-53.

7. Hendrickson, W. A., Love, W. E., and Murray, G. C. Crystal forms of lamprey hemoglobin and crystalline transitions between ligand states. J. Mol. Biol., 1968. 33, 829-842.

8. Antonini, E., Wyman, J., Bellelli, L., Rumen, N., and Siniscalco, M. The oxygen equilibrium curve of some lamprey hemoglobins. Arch. Biochem. Biophys., 1964. 105, 404-408.

9. Perutz, M. F. The hemoglobin molecule. Scientific American, Nov. 1964. 211, 64-76.

10. Perutz, M. F. "The Croonian Lecture, 1968, The hemoglobin molecule." Proc. Roy. Soc. B., 1969. 173, 113.

11. Perutz, M. F. Stereochemistry of cooperative effects in haemoglobin. Haem-haem interaction and the problem of allostery. Nature, 1970. 228, 726-734.

12. Kitchen, H. Heterogeneity of animal hemoglobins. Adv. in Vet. Sci. and Comp. Med., 1969. 13, 247-330.

13. Perutz, M. F. The Bohr effect and combination with organic phosphates. Nature, 1970. 228, 734-739.

14. Brewer, G. J., and Eaton, J. W. Erythrocyte metabolism: Interaction with oxygen transport. Science, 1971. 171, 1205.

15. Berman, M., Benesch, R., and Benesch, R. E. The removal of organic phosphates from hemoglobin. Arch. Biochem. Biophys., 1971. 145, 236.

16. Bunn, H. F., Briehl, R. W., Larrabee, P., and Hobart, V. The interaction of 2,3-diphosphoglycerate with various human hemoglobins. J. Clin. Invest., 1970. 49, 1088.

17. Caldwell, P. R. B., Nagel, R. C., and Jaffe, E. R. The effect of oxygen, carbon dioxide, pH and cyanate on the binding of 2,3 diphosphoglycerate to human hemoglobin. Biochem. Biophys. Res. Commun., 1971. 44, 1504.

18. Dhindsa, D. S., Sedgwick, C. J., and Metcalfe, J. A comparison of the respiratory functions of blood of Asian and African elephants. Proceedings of the International Union of Physiological Sciences, 1971. Vol. IX, 142. (Abstract).

19. Schmitt, J. Hochspannungselektrophoretische Untersuchungen an Saugetierhamoglobinen. Biochem. Z., 1960. 333, 71-77.

20. Riggs, A. The amino acid composition of some mammalian hemoglobins: Mouse, guinea pig, and elephant. J. Biol. Chem., 1963. 238, 2983-2987.

21. Dayhoff, M. O. Atlas of Protein Sequence and Structure. Silver Spring, Md.: National Biomedical Research Foundation, 1969.

22. Chauvet, J. P., and Archer, R. The β chain of frog hemoglobin (*Rana esculenta*): the complete amino acid sequence. FEBS Lett., 1970. 10, 136-138.

23. Rifkin, D. B., Rifkin, M. R., and Konigsberg, W. The presence of two major hemoglobin components in an inbred strain of mice. Proc. Natl. Acad. Sci., 1966. 55, 586-592.

24. von Ehrenstein, G. Translational variations in the amino acid sequence of the α chain of rabbit hemoglobin. Cold Spring Harbor Symposia, 1966. 31, 705-714.

25. Air, G. M., Thompson, E. O. P., Richards, B. G., and Sharman, G. B. Amino acid sequences of kangaroo myoglobin and hemoglobin and the date of marsupial-eutherian divergence. Nature, 1971. 229, 391-394.

26. Air, G. M., and Thompson, E. O. P. Studies on marsupial proteins. II. Amino acid sequence of the β chain of hemoglobin from the grey kangaroo, *Macropus giganteus*. Aust. J. Biol. Sci., 1969. 22, 1437-1454.

27. Schroeder, W. A., Shelton, J. R., Shelton, J. B., Robberson, B., and Babin, D. R. Amino acid sequence of the α chain of bovine fetal hemoglobin. Arch. Biochem. Biophys., 1967. 120, 1-14.

28. Boyer, S. H., Hathaway, P., Pascasio, F., Bordley, J., Orton, C., and Naughton, M. A. Differences in the amino acid sequence of tryptic peptides from three sheep hemoglobin β chains. J. Biol. Chem., 1967. 242, 2211-2232.

29. Braunitzer, G., Hilse, K., Rudloff, V., and Hilschmann, N. The hemoglobins. Adv. Protein Chem., 1964. 19, 34-72.

30. Bunn, H. F. Subunit dissociation of certain abnormal human hemoglobins. J. Clin. Invest., 1969. 48, 126-138.

31. Hendrickson, W. A., and Love, W. E. Structure of lamprey hemoglobin. Nature New Biol., 1971. 232, 197-203.

32. Li, S. L., and Riggs, A. The amino acid sequence of hemoglobin V from the lamprey, *Petromyzon marinus*. J. Biol. Chem., 1970. 245, 6149-6169.

33. Baldwin, T. O., and Riggs, A. F. Amino acid sequence of bullfrog hemoglobin. Fed. Proc., 1971. 30, 1295. (Abstract).

34. Chauvet, J. P., and Archer, R. The β chain of frog hemoglobin *(Rana esculenta)*: the complete amino acid sequence. FEBS Lett., 1970. 10, 136-138.

35. Jones, R. T., Brimhall, B., and Duerst, M. Amino acid sequence of the α and β chains of dog hemoglobin. Fed. Proc., 1971. 30, 1259. (Abstract).

36. Beard, J. M., and Thompson, E. O. P. Studies on marsupial proteins. V. Amino acid sequence of the α chain of haemoglobin from the grey kangaroo, Macropus giganteus. Aust. J. Biol. Sci., 1971. 24, 765-786.

37. Tsuyuki, H., and Ronald, A. P. The subunit structure and the molecular basis of the multiple hemoglobins of two species of trout, *Salmo gairdneri* and *S. clarki clarki*. Comp. Biochem. Physiol., 1971. 39B, 195-202.

38. Brimhall, B., and Coates, M. Personal Communication. 1972.

39. Stenzel, P. Personal Communication. 1972.

40. Smith, D. B. Amino acid sequences of some tryptic peptides from the β chain of horse hemoglobin. Canad. J. Biochem., 1968. 46, 825-843.

41. Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. Three dimensional fourier synthesis of horse oxyhemoglobin at 2.8 Å resolution: The atomic model. Nature, 1968. 219, 131.

42. Huismann, T. H. J., and Schroeder, W. A. New aspects of the structure, function, and synthesis of hemoglobins. CRC Crit. Rev. Clin. Lab. Sci., 1970. 1, 471-526.

43. Lehmann, H., and Carrell, R. W. Variations in the structure of human haemoglobin. Brit. Med. Bull., 1969. 25, 14-27.

44. Morimoto, H., Lehmann, H., and Perutz, M. F. Molecular pathway of human haemoglobin: Stereochemical interpretation of abnormal oxygen affinities. Nature, 1971. 232, 408-412.

45. Muller, C. J., and Kingma, S. Haemoglobin Zürich: $\alpha_2^{A}\beta_2^{63Arg}$. Biochim. Biophys. Acta, 1961. 50, 595.

46. Winterhalter, K. H., Anderson, N. M., Amiconi, G., Antonini, E., and Brunori, M. Functional properties of Hemoglobin Zürich. Europ. J. Biochem., 1969. 11, 435-440.

47. Lines, J. G., and McIntosh, R. Oxygen binding by haemoglobin J-Capetown. Nature, 1967. 215, 297.

48. Jones, R. T., Osgood, E. E., Brimhall, B., and Koler, R. D. Hemoglobin Yakima. I. Clinical and biochemical studies. J. Clin. Invest., 1967. 46, 1840-1847.

49. Reed, C. S., Hampson, R., Gordon, S., Jones, R. T., Novy, M. J., Brimhall, B., Edwards, M. J., and Koler, R. D. Erythrocytosis secondary to increased oxygen affinity of a mutant hemoglobin, hemoglobin Kempsey. Blood, 1968. 31, 623-632.

50. Alben, J. Personal Communication. 1972.

51. Barnabas, J., Goodman, M., and Moore, G. W. Evolution of hemoglobins in primates and other therian mammals. Comp. Biochem. Physiol., 1971. 39B, 455.

52. Benson, A., Tomoda, K., Chang, J., Matsueda, G., Lode, E. T., Coon, M. J., and Yasunobu, K. T. Evolutionary and phylogenetic relationships of rubredoxin-containing microbes. Biochem. Biophys. Res. Commun., 1971. 42, 640-646.

53. Clarke, B. Darwinian evolution of proteins. Science, 1970. 168, 1009-1011.

54. Jukes, T. H. Comparisons of the polypeptide chains of globins. J. Mol. Evol., 1971. 1, 46-62.

55. Kimura, M. Evolutionary rate at the molecular level. Nature, 1968. 217, 624-626.

56. Oakley, K. P. Frameworks for dating fossil man. London: Weidenfeld and Nicolson, 1964. (page 35)

57. Osborn, H. F. Proboscidea, Vol. 2. New York: American Museum, 1942. (page 1602)

58. Wallach, J. D., and Anderson, J. L. Oripavine (M 99) combinations and solvents for immobilization of the African elephant. J. Am. Vet. Med. Assn., 1968. 153, 793-797.

59. Drabkin, D. Spectrophotometric studies. XIV. J. Biol. Chem., 1946. 163, 703-723.

60. Jones, R. T., and Schroeder, W. A. Chromatography of human hemoglobin. J. Chromatog., 1963. 10, 421-431.

61. Huisman, T. H. J. Normal and abnormal hemoglobins. Advan. Clin. Chem., 1963. 6, 231.

62. Baur, E. W. Thin layer starch-gel electrophoresis and plastification method. J. Lab. Clin. Med., 1963. 61, 166.

63. Anson, M. L., and Mirsky, A. E. Protein coagulation and its reversal. J. Gen. Physiol. 1930. 13, 469-476.

64. Clegg, J. B., Naughton, M. A., and Weatherall, D. J. An improved method for the characterization of human haemoglobin mutants. Nature, 1965. 207, 945-947.

65. Dintzis, H. M. Assembly of the peptide chains of hemoglobin. Proc. Natl. Acad. Sci., 1961. 47, 247.

66. Stark, G. R. Use of cyanate for determining NH_2-terminal residues in protein. In C. H. W. Hirs (Ed.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 125-138.

67. Stark, G. R., and Smyth, D. G. The use of cyanate for the determination of NH_2-terminal residues in proteins. J. Biol. Chem., 1963. 238, 214-226.

68. Spackman, D. H., Stein, W. H., and Moore, S. Chromatography of amino acids on sulfonated polystyrene resins. Anal. Chem., 1958. 30, 1190-1206.

69. Schroeder, W. A. The primary structure of proteins. New York: Harper and Row, 1968.

70. Cole, R. D. S-Aminoethylation. In C. H. W. Hirs (Eds.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 315-317.

71. Raftery, M., and Cole, R. Tryptic cleavage at cysteinyl peptide bonds. Biochem. Biophys. Res. Commun., 1962. 10, 467-472.

72. Brimhall, B. Personal Communication. 1968.

73. Jones, R. T., and Weiss, G. Long-path flow cells for automatic amino acid analysis. Anal. Biochem., 1964. 9, 377-391.

74. Elzinger, M. Amino acid sequence on rabbit skeletal muscle actin. Cyanogen bromide cleavage of the protein and determination of the sequences of seven of the resulting peptides. Biochemistry, 1970. 9, 1365-1374.

75. Marzotto, A., Pajetta, P., Galzinga, L., and Scoffone, E. Reversible acetoacetylation of amino groups in proteins. Biochim. Biophys. Acta, 1968. 154, 450-456.

76. Butler, P. J. G., Harris, J. I., Hartley, B. S., and Liberman, R. Reversible blocking of peptide amino acid groups by maleic anhydride. Biochem. J., 1967, 103, 78p. (Abstract).

77. Braunitzen, G., Beyreuther, K., Fujiki, H., and Schrank, B. Tetrafluorbernsteinsäurr-anhydrid, ein neues reagens zur spezifischen und reversiblen maskierung der aminogruppen in proteinen. Hoppe-Seyler's Z. Physiol. Chem., 1968. 349, 265.

78. Dixon, H. B. F., and Perham, R. N. Reversible blocking of amino groups with citraconic anhydride. Biochem. J., 1968. 109, 312-314.

79. Singhal, R. P., and Atassi, M. Z. Immunochemistry of sperm whale of myoglobin. IX. Specific interaction of peptides obtained by cleavage at arginine peptide bonds. Biochemistry, 1971. 10, 1756-1762.

80. Baglioni, T. Abnormal human haemoglobins. X. A study of hemoglobin Lepore (Boston). Biochem. Biophys. Acta., 1965. 97, 37.

81. Jones, R. T. Automatic peptide chromatography. In D. Glick (Ed.) Methods of Biochemical Analysis, Vol. 18. New York, N. Y.: Interscience, 1970.

82. Jones, R. T. Structure studies of aminoethylated hemoglobins by automatic peptide chromatography. Cold Spring Harbor Symp. Quant. Biol., 1964. 29, 297-308.

83. Jones, R. T., Brimhall, B., Huisman, T. H. J., Kleihauer, E., and Betke, K. Hemoglobin Freiburg: Abnormal hemoglobin due to deletion of a single amino acid residue. Science, 1966. 154, 1024-1027.

84. Bornstein, P., and Piez, K. A. The nature of the intramolecular cross-links in collagen. The separation and characterization of peptides from the cross-link region of rat skin collagen. Biochemistry, 1966. 5, 3461-3473.

85. Hamilton, P. B. Ion exchange chromatography of amino acids, a single column, high resolving, fully automatic procedure. Anal. Chem., 1963. 35, 2055.

86. Adelstein, R. S., and Kuehl, W. M. Structural studies on rabbit skeletal actin. I. Isolation and characterization of the peptides produced by cyanogen bromide cleavage. Biochemistry, 1970. 9, 1355-1364.

87. Steers, E., Jr., Craven, G. R., Anfinsen, C. B., and Bethune, J. L. Evidence for nonidentical chains in the β-galactosidase of Escherichia coli K12. J. Biol. Chem., 1965. 240, 2478-2484.

88. Witkop, B. Nonenzymatic methods for the preferential and selective cleavage and modification of proteins. Adv. Protein Chem., 1961. 16, 221-321.

89. Sarges, R., and Witkop, B. Gramicidin A. V. The structure of valine and isoleucine-gramacidin A. J. Am. Chem. Soc., 1965. 87, 2011-2027.

90. Ramachandran, L. K., and Witkop, B. N-Bromosuccinimide cleavage of peptides. In C. H. W. Hirs (Ed.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 283-299.

91. Shaltiel, S., and Patchornik, A. Cleavage of histidyl peptide bonds by N-bromosuccinimide. J. Am. Chem. Soc., 1963. 85, 2799-2806.

92. Schultz, J. Cleavage at aspartic acid. In C. H. W. Hirs (Ed.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 255-263.

93. Ingram, V. M. Sequence methods. In S. P. Colowick and N. O. Kaplan (Eds.) Methods in Enzymology, VII. New York, N. Y.: Academic Press, 1963. pp. 831-848.

94. Goldstein, J., Konigsberg, W., and Hill, R. J. The structure of human hemoglobin. VI. The sequence of amino acids in the tryptic peptides. J. Biol. Chem., 1963. 238, 2016-2027.

95. Konigsberg, W., and Hill, R. J. The sequence of amino acids in the tryptic peptides of the α chain. J. Biol. Chem., 1962. 237, 2547-2561.

96. Matsubara, H., Sasaki, H., Singer, A., and Jukes, T. H. Specific nature of hydrolysis of insulin and tobacco mosaic virus protein by thermolysin. Arch. Biochem. Biophys., 1966. 115, 324-331.

97. Moore, S., and Stein, W. H. Chromatographic determination of amino acids by the use of automatic recording equipment. In C. H. W. Hirs (Ed.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 819-829.

98. Spackman, D. H., Stein, W. H., and Moore, S. The disulfide bonds of ribonuclease. J. Biol. Chem., 1960. 235, 648-659.

99. Gray, W. R., and Smith, J. F. Rapid sequence analysis of small peptides. Anal. Biochem., 1970. 33, 36.

100. Seiler, N. Use of the dansyl reaction in biochemical analysis. In D. Glick (Ed.) Methods of Biochemical Analysis, Vol. 18. New York, N. Y.: Interscience, 1970. pp. 259-337.

101. Gray, W. R. Dansyl chloride procedure. In C. H. W. Hirs (Ed.) Methods in Enzymology, XI. New York, N. Y.: Academic Press, 1967. pp. 139-151.

102. Edman, P. On the mechanism of the phenyl isothiocyanate degradation of peptides. Acta Chem. Scand., 1956. 10, 761-768.

103. Edman, P., and Begg, G. A protein sequenator. European J. Biochem., 1967. 1, 80-91.

104. Edman, P. Sequence determination. In S. B. Needleman (Ed.) Protein sequence determination. New York, N. Y.: Springer-Verlag, 1970. pp. 211-255.

105. Cole, M., Fletcher, J. C., and Robson, A. The separation of 1-dimethylaminonapthalene-5 sulphonamido acids (DNS-amino acids) by thin layer chromatography. J. Chromatog., 1965. 20, 616-618.

106. Nedkov, P., and Genov, N. A method for identifying C-terminal amino acids as their DNS derivatives obtained after the hydrazinolysis of peptides. Biochim. Biophys. Acta, 1966. 127, 544-545.

107. Crowshaw, K., Jessup, S. J., and Ramwell, P. W. Thin layer chromatography of 1-dimethylaminonapthalene-S-sulfonyl derivatives of amino acids present in superfusates of cat cerebral cortex. Biochem. J., 1967. 103, 79-85.

108. Schroeder, W. A., Shelton, J. R., Shelton, J. B., Cormick, J., and Jones, R. T. The amino acid sequence of the γ chain of human fetal hemoglobin. Biochemistry, 1963. 2, 992.

109. Jones, R. T., Brimhall, B. Structural characterization of two δ chain variants. J. Biol. Chem., 1967. 242, 5141-5145.

110. Jones, R. T., Brimhall, B., and Lisker, R. Chemical characterization of hemoglobin Mexico and hemoglobin Chiapas. Biochim. Biophys. Acta, 1968. 154, 488-495.

111. Zuckerkandl, E., and Pauling, L. Evolution divergence and convergence in proteins. In V. Bryson and H. J. Vogel (Eds.) Evolving genes and proteins. New York, N. Y.: Academic Press, 1965. pp. 97-166.

112. Vedvick, T. S. Separation of α and β chains of African elephant. Unpublished results, Univ. of Oregon Medical School, 1972.

113. Hermodson, M. A., Tye, R. W., Reeck, G. R., Neurath, H., and Walsh, K. A. Comparison of the amino terminal sequences of bovine, dogfish, and lungfish trypsinogens. FEBS Lett., 1971. 14, 222-224.