INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

76-11,976

WEISER, Ellen Spatzer, 1949-
CHARACTERIZATION OF THE HETEROGENEITY OF RAT HEMOGLOBIN: AN INVESTIGATION OF PRIMARY GLOBIN STRUCTURE.

The City University of New York
Ph.D., 1976
Chemistry, biological

Xerox University Microfilms, Ann Arbor, Michigan 48106

© COPYRIGHT BY

ELLEN SPATZER WEISER

1975

CHARACTERIZATION OF THE HETEROGENEITY OF RAT HEMOGLOBIN: AN INVESTIGATION OF PRIMARY GLOBIN STRUCTURE

by

ELLEN SPATZER WEISER

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, City University of New York.

1975

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

11/14/75
Date

Abraham Mazur
Chairman of Examining Committee

11-14-75
Date

Aaron Lukton
Executive Officer

ABSTRACT

Characterization of the Heterogeneity
of Rat Hemoglobin: An Investigation
of Primary Globin Structure

by

Ellen Spatzer Weiser

Adviser: Professor Abraham Mazur

Comparisons of the primary sequences of homologous molecules have provided considerable information concerning the structure, function and evolution of proteins. The investigation of multiple hemoglobin chains in many mammalian species has helped to reveal the genetic mechanisms that cause chemical heterogeneity in proteins and also underscores the principles of molecular evolution. It has been demonstrated that many animals have multiple hemoglobins in circulating erythrocytes. Isoelectric focusing procedures have given conclusive evidence for the existence of six hemoglobins in the red blood cells of rats. Chemical studies of the six purified rat hemoglobin fractions prior to this investigation have demonstrated that the heterogeneity was due to differences in the primary structure of the globin chains and not to induced chemical modifications, artifacts of isolation or to differences in aggregation. Variations in the subunits easily account for the presence of multiple hemoglobin forms in the mammalian hemoglobin molecule - a tetramer composed of two pairs of nonidentical subunits usually referred to as α and β. Structural comparisons of the heterogenous hemoglobins in

some species have demonstrated multiple amino acid differences in the α and β chains. This investigation attempts to answer the question, "How many structurally different α and β chains does the rat possess?" through the examination of the amino acid content and the amino acid sequence of rat hemoglobins III, IV, V. To do this, rat hemoglobin was subjected to preparative isoelectric focusing followed by extraction and purification of hemoglobins III, IV, V from the polyacrylamide gels. The α and β chains corresponding to rat hemoglobin fractions III, IV and V were separated, purified and subjected to: (i) amino acid analysis, (ii) tryptic digestion and two-dimensional chromatography, (iii) amino acid analysis of selected tryptic peptides, and (iv) sequencing by homology. The results to be presented in this investigation will confirm the finding of at least three structurally different β chains and the presence of at least three structurally different α chains as well. These α chain differences and β chain differences are due to the existence of at least three different structural genes coding for α chains and three for β chains.

ACKNOWLEDGEMENTS

The author wishes to express her deep appreciation

To her husband Stephen for his support
and understanding

To her mentor Dr. Abraham Mazur for his
patience and guidance

The work in this investigation was supported by a grant from the National Institutes of Health.

TABLE OF CONTENTS

	PAGE
Copyright page	1
Approval page	2
Abstract	3
Acknowledgements	5
Table of Contents	6
List of Tables	7
List of Figures	8
List of Abbreviations	10
Introduction	11
Experimental	16
Results	25
Discussion	104
Appendix	121
References	144

LIST OF TABLES

		PAGE
TABLE I	Amino Acid Composition of the α Chains	44
TABLE IA	Differences in the Amino Acid Compositions of the α Chains	46
TABLE II	Amino Acid Composition of the β Chains	47
TABLE IIA	Differences in the Amino Acid Compositions of the β Chains	49
TABLE IIIa	Tryptic Peptide Comparisons by Map Positioning After Two-Dimensional Chromatography - α Chains	58
TABLE IIIb	Tryptic Peptide Comparisons by Map Positioning After Two-Dimensional Chromatography - β Chains	59
TABLE IV	Amino Acid Composition of the 3 α Chain Tryptic Peptides	61
TABLE V	Amino Acid Composition of the 4 α Chain Tryptic Peptides	62
TABLE VI	Amino Acid Composition of the 5α Chain Tryptic Peptides	63
TABLE VII	Amino Acid Composition of the 3β Chain Tryptic Peptides	64
TABLE VIII	Amino Acid Composition of the 4β Chain Tryptic Peptides	65
TABLE IX	Amino Acid Composition of the 5β Chain Tryptic Peptides	66
TABLE X	Amino Acid Composition of 4α Chain Tryptic Peptides Isolated From an Individual Rat Speciman ...	67
TABLE XI - TABLE XXII	Amino Acid Composition Comparisons of α Chain Peptides	122-133
TABLE XXIII - TABLE XXXII	Amino Acid Composition Comparisons of β Chain Peptides	134-143

LIST OF FIGURES

		PAGE
Figure 1	Isoelectric Focusing of Total Rat Hemoglobin	26
Figure 2	Separation of Total Rat Globin into α and β Chains. Urea-Phosphate System	27
Figure 3	Separation of Rat Globin #3 into α and β Chains. Urea-Phosphate System	29
Figure 4	Separation of Rat Globin #4 into α and β Chains. Urea-Phosphate System	30
Figure 5	Separation of Rat Globin #5 into α and β Chains. Urea-Phosphate System	31
Figure 6	Separation of Total Rat Globin into α and β Chains. Pyridine-Formate System	33
Figure 7	Separation of Rat Globin #3 into α and β Chains. Pyridine-Formate System	34
Figure 8	Rechromatography of Rat Globin 3α Chain. Pyridine-Formate System	35
Figure 9	Rechromatography of Rat Globin 3β Chain. Pyridine-Formate System	36
Figure 10	Separation of Rat Globin #4 into α and β Chains. Pyridine-Formate System	37
Figure 11	Rechromatography of Rat Globin 4α Chain. Pyridine-Formate System	38
Figure 12	Rechromatography of Rat Globin 4β Chain. Pyridine-Formate System	39

LIST OF FIGURES

		PAGE
Figure 13	Separation of Rat Globin #5 into α and β Chains. Pyridine-Formate System	40
Figure 14	Rechromatography of Rat Globin 5α Chain. Pyridine-Formate System	41
Figure 15	Rechromatography of Rat Globin 5β Chain. Pyridine-Formate System	42
Figure 16	Tryptic Peptide Maps of the Rat 3α, 4α and 5α Chains	53
Figure 17	Tryptic Peptide Maps of the Rat 3β, 4β and 5β Chains	54
Figure 18	Sequence Homology of the Rat 3α, 4α and 5α Chains	105
Figure 19	Sequence Homology of the Rat 3β, 4β and 5β Chains	108

LIST OF ABBREVIATIONS

IEF = Isoelectric Focusing

TPCK = L-(1-Tosylamido-2-phenyl)ethylchloromethyl ketone

The Amino Acids

Ala = Alanine
Arg = Arginine
Asp = Aspartic
Cys = Cysteine
Glu = Glutamic
Gly = Glycine
Hist = Histidine
Ile = Isoleucine
Leu = Leucine
Lys = Lysine
Met = Methionine
Phe = Phenylalanine
Pro = Proline
Ser = Serine
Thr = Threonine
Trp = Tryptophan
Tyr = Tyrosine
Val = Valine

INTRODUCTION

Comparisons of the primary sequences of homologous molecules have provided considerable information concerning the structure, function and evolution of proteins. The investigation of multiple hemoglobin chains in many mammalian species has helped to reveal the genetic mechanisms that cause chemical heterogeneity in proteins and also underscores the principles of molecular evolution.

It has been demonstrated that many animals, e.g., goat (2), horse(3), mouse (4), as well as human (5), have multiple hemoglobins in circulating erythrocytes. The occurrence of more than a single hemoglobin within a species is no longer considered unusual. Stein, et al (1), using isoelectric focusing procedures, gave conclusive evidence for the existence of six hemoglobins in the red blood cells of rats. The recent work of Ranney, et al (6), and Chua, et al (7,8), employing ion-exchange chromatography, have confirmed this initial observation. The relative distribution of rat hemoglobin components is constant from one animal to another, regardless of age, sex, or strain.

Studies on the primary structure of hemoglobin in other mammalian species have revealed genetic mechanisms that cause chemical heterogeneity in these proteins. Thus far it has been found that structural variability is caused by several mechanisms:

1) single amino acid substitutions resulting in changes in the amino acid sequence of one or more poly-

peptide chain types;

2) deletion mutations resulting in the loss of one or more of the amino acids in a given polypeptide chain type;

3) chain elongation or amino acid additions to one of the polypeptide chain types;

4) hybridization of the genes for different chains;

5) the existence of multiple and different templates for specific chains.

Also influencing hemoglobin heterogeneity is differential gene activity related to development - the sequential emergence and arrest of respective polypeptide chain types (derepression and repression, respectively).

The mammalian hemoglobin molecule is a tetramer composed of two pairs of nonidentical subunits - usually referred to as α and β . Each α and β chain is composed of a heme moiety coordinately bonded to a polypeptide chain consisting of 141 amino acid residues (as in the α chain) or 146 amino acid residues (as in the β chain). Variations in the subunits easily account for the presence of multiple hemoglobin forms. Structural comparisons of heterogeneous hemoglobins in some species have demonstrated multiple amino acid differences in the α and β chains.

In adult human hemoglobin, the non-allelic β and δ chains account for hemoglobin A and A_2 when they are combined with a structurally identical α chain (9). The occur-

rence of hemoglobin heterogeneity in the rabbit is found to be due to structural differences in the α chain (10). More recently, structural differences in the rabbit β chain have also been reported (11). Variability among the hemoglobins in different strains of mice has been studied. The differences in electrophoretic patterns of mouse hemoglobins are due to differences in the structure of the β chain (12,13) while the solubility properties of specific hemoglobin types are influenced by differences in the mouse α chains (4, 14). In the horse, there are two electrophoretically distinct hemoglobin components and evidence has been offered that the structural differences between the two reside in the α chain (3,15). In goats, the hemoglobin heterogeneity is the result of multiple structural differences in β chains that are products of allelic genes and in α chains that are products of non-allelic genes (2,9,16).

Chemical studies of the six purified rat hemoglobin fractions performed by Stein, et al (1) demonstrated that the heterogeneity was due to differences in the primary structure of the globin chains and not to induced chemical modifications, artifacts of isolation or to differences in aggregation.

The aim, then, of this investigation is to partially answer the question "How many structurally different α and β chains does the rat possess?" through the examination of the amino acid content and the amino acid sequence of rat hemoglobins III, IV and V. Then, depending upon the

nature of the amino acid sequence differences among the three α chains and three β chains, one would be able to speculate as to the mechanisms of synthetic control.

In order to accomplish this, rat hemoglobin was subjected to preparative isoelectric focusing by the method of Stein, et al (1) followed by extraction and purification of hemoglobins III, IV and V from the polyacrylamide gels. The α and β chains corresponding to rat hemoglobin fractions III, IV and V were then separated and, once purified, subjected to the following, in order to determine which of the α chains and which of the β chains are the same and which are different: a) comparison of the total amino acid compositions of each α and β chain: b) comparison of the two-dimensional peptide maps of the tryptic digests of 3α, 4α, 5α and 3β, 4β, 5β; c) comparison of the amino acid content of selected α and β tryptic peptides and sequence assigned by homology to those peptides.

The recent data of Ranney, et al (6) has confirmed the existence of at least two different α chains in rat hemoglobin and from preliminary findings, at least three different β chains are also indicated. The results to be presented in this investigation will confirm the finding of at least three structurally different β chains and the presence of at least three (and not two as proposed by Ranney, et al) structurally different α chains as well. These α chain differences and β chain differences are due to the existence of at least three different structural

EXPERIMENTAL

Materials

All but the following chemicals were purchased from Fisher Scientific Co. Unless otherwise stated, chemicals were reagent grade. Ampholines (pH 6-8) were from LKB Productor. TEMED, acrylamide, and bis-acrylamide for IEF were purchased from Eastman Chemicals. Of the analytical resins used, DEAE cellulose was supplied by Bio-Rad, Sephadex G-25 was from Pharmacia Fine Chemicals, and CM-52 cellulose was a product of the Whatman Co. Ultrapure urea was bought from Schwarz-Mann. The pyridine used for column chromatography was spectroscopic grade and purchased from Matheson-Coleman and Bell. β-Mercaptoethanol was a Sigma product. Ninhydrin was purchased from Pierce Chemicals as was the 4N Methane Sulfonic acid containing 0.2% 3-(2-Aminoethyl)indole used for the tryptophan determinations. The trypsine enzyme, treated with TPCK, was purchased from Worthington Biochemicals.

The isoelectric focusing apparatus and the high voltage electrophoresis equipment were from Shandon. The Wistar rats (CFN strain) were purchased from Carworth Farms.

Methods

Preparation of Hemoglobin

The preparation of crystalline rat hemoglobin from erythrocytes was after the fashion of Stein, et al (1) with the following modifications: after cell lysis, the rat

genes coding for α chains and three for β chains.

These studies are of great relevance to an understanding of the maturation of bone marrow erythroid cells, in light of the findings of Stein, et al (1) who demonstrated that the various rat hemoglobins are synthesized non-uniformly in such cells at various stages of their maturation. Stein, et al (1) followed the synthesis of rat hemoglobin in vivo using ^{59}Fe incorporation into red cell hemoglobin, and found that at different stages of erythroid cell maturation there was a continuous change in the extent of synthesis of different hemoglobins, with the youngest erythroid cell producing hemoglobin V and the oldest hemoglobin IV. These findings produced evidence that hemoglobin V was synthesized most actively in the "youngest" erythroid cells, whereas hemoglobin IV, the major hemoglobin of the circulating red cell, was synthesized most actively in the "oldest" erythroid cells.

hemoglobin was crystallized from 0.005M potassium phosphate buffer, pH 7.4, saturated with carbon monoxide in the cold. The crystalline hemoglobin was collected by centrifugation and dissolved at room temperature in 0.05M Tris-HCl, pH 8.6 buffer. The hemoglobin was then dialyzed against 0.005M Tris-HCl, pH 8.6 buffer, overnight in the cold. Final concentrations of hemoglobin ranged from 15-30 mg/ml. Hemoglobin concentrations were determined from optical densities at 540 mu on a Beckman DU Spectrophotometer. An OD_{540} of 0.9 was equivalent to 1.0 mg of hemoglobin/ml of solution.

Preparative Isoelectric Focusing of Crystalline Rat Hemoglobin

Crystalline rat hemoglobin was separated into its six components via isoelectric focusing (IEF) according to Stein, et al (1) with modifications.Preparative IEF was performed in polyacrylamide gels, using glass tubes 1x13 cm in size, along a pH gradient from 6-8. Gel solutions consisted of the following in a 1:2:2:0.3:1 ratio (V/V) respectively: 0.8% TEMED, 28% acrylamide-0.735% Bis-acrylamide, 40% sucrose, ampholine, 0.004% riboflavin. Hemoglobin samples ranging 15-18 mg/gel tube were incorporated into the gels prior to photopolymerization. IEF was carried out for 48 hours at 100 volts at 8°C.

After completion of electrophoresis, the gels were removed from the tubes, sliced as close to the bands as possible and stored at -20°C.

Extraction of Rat Hemoglobin Fractions from IEF Gel Slices

All procedures were performed at room temperature and in the presence of carbon monoxide, unless otherwise stated.

Gel slices corresponding to hemoglobin fractions III, IV, and V (see figure I) were homogenized in a Virtis Homogenizer with 0.05M potassium phosphate- 0.1M NaCl, pH 8.6 extraction buffer. After centrifugation at 3,000 rpm, precipitated acrylamide gel was discarded and the recovered rat hemoglobin fraction was exhaustively dialyzed against large volumes of 0.005M Tris-HCl, pH 8.6 buffer in the cold.

After dialysis, the hemoglobin solution was concentrated onto a DEAE-Sephadex column equilibrated with 0.005M Tris-HCL, pH 8.6. The concentrated hemoglobin was then eluted with 0.10M KP_i-0.10M NaCl, pH 7.4 buffer, desalted by passage through a Sephadex G-25 column equilibrated with 0.01M KP_i, pH 8.0 buffer and then dialyzed against 0.005M Tris-HCl, pH 8.6 buffer for three days in the cold.

Final rat hemoglobin concentrations for each fraction were 10-20 mg hemoglobin/ml of solution.

Preparation of Globin

Globins # 3, 4, and 5 were prepared from rat hemoglobin fractions III, IV, and V, respectively, via the acid-acetone procedure of Rossi-Fanelli, et al (17). Hemoglobin was added dropwise to a 2% HCl-acetone solution

containing 0.05M β-mercaptoethanol at -60°C. After centrifugation, the globin pellet was washed four times with cold acetone and once with ether.

Separation of Globin into α and β Chains

I. Urea-Phosphate Method (18)

Rat globins 3, 4, and 5 were separated into α and β chains using the method of Clegg, et al (18) as modified below. Rat globin (10-25 mg) was dissolved in Developer A buffer and allowed to dialyze against same at room temperature before application to a 1x10 cm Whatman CM-52 cellulose column already equilibrated with developer A. Elution of α and β chains was achieved using a linear Na^+ gradient consisting of equal volumes of Developer A and Developer B. Developer A: 0.005M Na_2HPO_4, 0.05M β-mercaptoethanol, 8M urea, adjusted to pH 6.5 or pH 6.7 with 1.0M H_3PO_4. Developer B: 0.03M Na_2HPO_4, 0.05M β-mercaptoethanol, 8M urea, adjusted to pH 6.5 or pH 6.7 with 1.0M H_3PO_4.

The 8M urea was filtered and deionized before use by passage through a 2x30 cm column of Bio-Rad AG501-X8(D) mixed bed resin (20-50 mesh). The conductivity of the 8M urea was monitored with a Lab-Line Lectro Mho^{-1} meter.

II. Pyridine-Formate Method (19)

Rat globins 3, 4, and 5 were separated into α and β chains via the method of Dintzis (19), modified as follows:

Before use, Whatman CM-52 cellulose was suspended

in 0.2M pyridine (spectroscopic grade) for two hours; filtered and resuspended in strong buffer for one hour; filtered and washed four times with weak buffer. Rat globin samples (10-25 mg) were dissolved in weak buffer containing 0.05M β -mercaptoethanol and dialyzed against same at room temperature before application to a 1x10 cm column of treated Whatman CM-52 cellulose, equilibrated with weak buffer. A gradient consisting of equal volumes of weak and strong buffer were used to elute the α and β chains. Weak buffer: 0.02M pyridine, 0.2N formic acid. Strong buffer: 0.2M pyridine, 2.0N formic acid.

An LKB Uvicord II was used to continuously monitor all effluents at 280 mu. 3.0 ml fractions were collected using an LKB fraction collector.

<u>Purification of the α and β Chains After Pyridine-Formate Separation</u>

After many chain separations, the recovered α chains (as well as β chains) corresponding to globin fractions 3, 4, or 5 were combined and lyophilized once from weak buffer and twice from 0.5% formic acid. Each α chain and each β chain was then re-chromatographed from fresh Whatman CM-52 cellulose in order to purify them further. For re-chromatography, the limits of the gradient were changed (see Figures 8,9,11,12,14,15 for details). The re-chromatographed α and β chains of each globin fraction was again lyophilized from weak buffer once and twice from 0.5% formic acid before storage at -20°C.

Amino Acid Analysis

Amino Acid analysis of the 3α, 3β, 4α, 4β, 5α and 5β chains was performed by the method of Moore and Stein (20) with a Beckman Model 118 Amino Acid Analyzer. Samples were hydrolyzed in evacuated sealed ampoules at 110°C in 6M HCl for 24 hours, 48 hours and 72 hours. The hydrolysis ampoules were washed before use with a mixture of HNO_3 and H_2SO_4 (1:3 V/V) and then rinsed with distilled water and dried.

Tryptophan was determined by the method of Lui and Chang (21,22). The α and β chains were hydrolyzed in 4N methane sulfonic acid containing 0.2% 3 - (2-Aminoethyl) indole for 24 hours at 115°C.

Cysteine was determined as cysteic acid and methionine as methionine sulphone after the performic acid oxidation method of Hirs (23).

Trypsin Digestion

The α and β chains of rat globin fractions 3, 4 and 5 were digested with trypsin, that had been first treated with TPCK in order to quench any chymotryptic activity, according to the procedure of Guidotti, et al (24).

The α and β chains were dissolved in water to give a 1% solution. The solution was then made 1% in ammonium bicarbonate and the pH adjusted to 9.0. An aliquot of a 1% solution of trypsin in 0.001N HCl was added to give a final concentration of trypsin of 0.01%. Digestion was allowed to proceed for eight hours at room temperature.

Digestion was terminated with the addition of acetic acid, until there was no more release of carbon dioxide - about pH 4.5. After centrifugation for ten minutes, the soluble supernatant was separated from the insoluble "core" and lyophilized twice from water.

Two-Dimensional Peptide Mapping

Two-dimensional peptide mapping was performed after the fashion of Bennet (25), Ingram (26,27) and Katz, et al (28).

Descending chromatography was performed on 18" x 18" Whatman 3 MM chromatography paper for 14½ hours at room temperature. The aqueous phase of the buffer n-butanol - acetic acid - water (4:1:5) was used to saturate the chromatography chamber prior to chromatography, while the organic phase was placed in the buffer troughs for descending chromatography. Solubilized trypsin digest, dissolved in an aliquot of water, was applied to the chromatography paper in an amount equivalent to 1.0 mg of original protein.

After the descending chromatogram was air dried, it was rotated 90° and lengthened by 4½" by sewing a strip of Whatman 3 MM chromatography paper to what was to become the anode side of the chromatogram. High voltage paper electrophoresis was then performed in a pyridine - acetic acid - water (10:0.4:90) pH 6.4 buffer for 2½ hours at 2000 volts. The pyridine reagent was distilled once from ninhydrin prior to use. Varsol was used as the coolant.

After high voltage electrophoresis, the papers

were air dried before staining.

Stains For Two-Dimensional Chromatograms

1. 0.5% ninhydrin in absolute ethanol-2N acetic acid (75/25 V/V). Stain developed at 60°C for 0.5 hour (25).

2. Pauly stain for imidazole (histidine) and phenolic compounds (tyrosine) (29). A 0.1% diazonium salt of sulfanilic acid in 10% Na_2CO_3. The stain is developed at room temperature.

3. Ehrlich stain for tryptophan (30). 1% p-dimethylaminobenzaldehyde in acetone-HCl (90:10). The stain is developed at room temperature.

Elution of Peptides From Two-Dimensional Peptide Maps (25)

Four to six two-dimensional chromatograms, each containing the equivalent of 1.0 mg of original protein, were prepared and sprayed lightly with a 0.05% ninhydrin solution. Spots were allowed to develop for five minutes at 60°C (to minimize destruction of the N-terminal amino acid) before being circled and cut out from the chromatography paper. Peptides were stored in glass vials at -20°C until they were eluted from the chromatography paper with 6N HCl. Eluted peptides were then hydrolyzed immediately for 24 hours at 110°C in sealed, evacuated ampoules. Released amino acids from each peptide were analyzed on a Beckman Model 118 Amino Acid Analyzer equipped with an expanded range scale.

Some peptides were later eluted with distilled water, lyophilized and then hydrolyzed for 24 hours at 110°C

to achieve better recovery of released amino acids.

Sequencing of Peptides by Homology (31)

From studies with other mammalian hemoglobins, it has been found, with few exceptions, that all α chains have 141 amino acids. and most β chains have 146 amino acids (31). Sequence alignments of known mammalian α chains and of known mammalian β chains have revealed that there are families of α globin chains as well as β globin chains that have similar overall amino acid sequence patterns. It was expected that the rat α and β chain sequences would also fit into these globin families as well. Therefore, once the amino acid content of the rat α and β peptides had been determined, comparisons with peptides in other mammalian α and β systems were made. The amino acids in the rat α and β peptides were then aligned by comparing them with the sequences found in other similar mammalian peptides. In this way, the rat α and β chains were sequenced by homology. The mammalian α and β chain sequences used for homology (comparison) studies were human (31), dog (31), mouse strain C57BL and mouse strain BALB/c (12,13,31) and rabbit (31).

RESULTS

Preparation of Globin

For the preparation of globin from hemoglobin, it was necessary to first dialyze the hemoglobin against a low salt buffer. This prevents fragmentation of the chains which occurs when a high salt concentration is present. Removal of the heme group from rat hemoglobin yielded a globin which was either used immediately for the separation of α and β chains or stored at -20°C. Total rat globin was prepared from total rat hemoglobin; i.e. hemoglobin that had not been separated into its six components by IEF. Rat globin #'s 3, 4 and 5 were prepared from rat hemoglobin fractions #'s 3, 4 and 5 respectively (Figure 1).

Separation of Rat Hemoglobin into α and β Chains

By use of the urea-phosphate system, total rat hemoglobin could be separated into a variety of α and β chains (Figure 2). In this system, β chains elute in front of α chains; but there is no absolute assurance at this point which peaks represent α and which β chains. This preliminary separation gives only a rough indication of the number of different α and β chains that can be expected. A minimum of five different α and β chains were revealed, with the possibility that a few other chains remained unresolved under the two major peaks - peaks I and II (Figure 2). To find the best conditions for the separation of the multiple peaks, the pH of the buffers and the limits of the Na^{+} gradient were varied. A Na^{+} gradient from 0.005M to

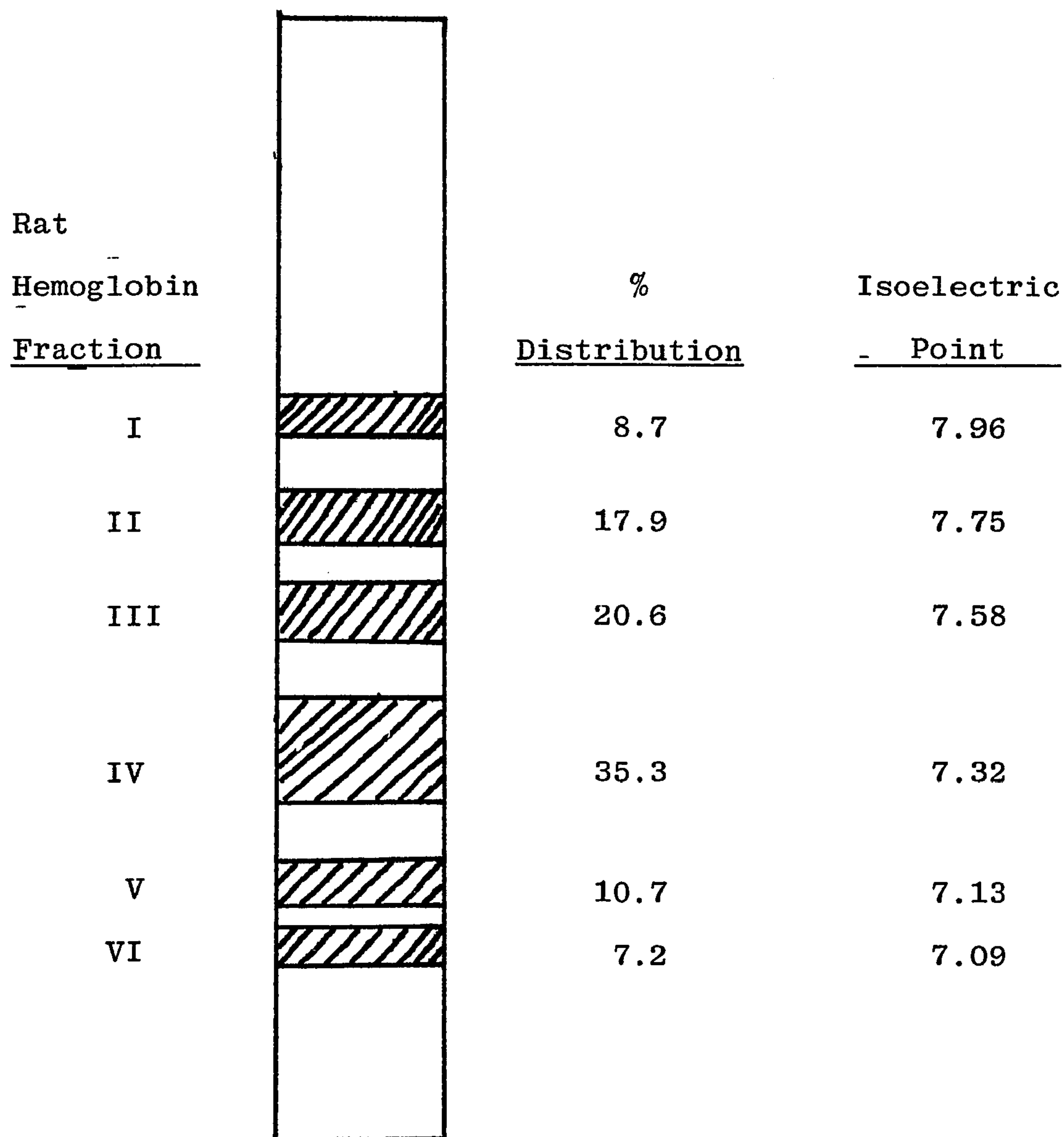

Figure 1 ISOELECTRIC FOCUSING OF TOTAL RAT HEMOGLOBIN

pH gradient: 6-8
Cathode at top
From Stein, et al (1)

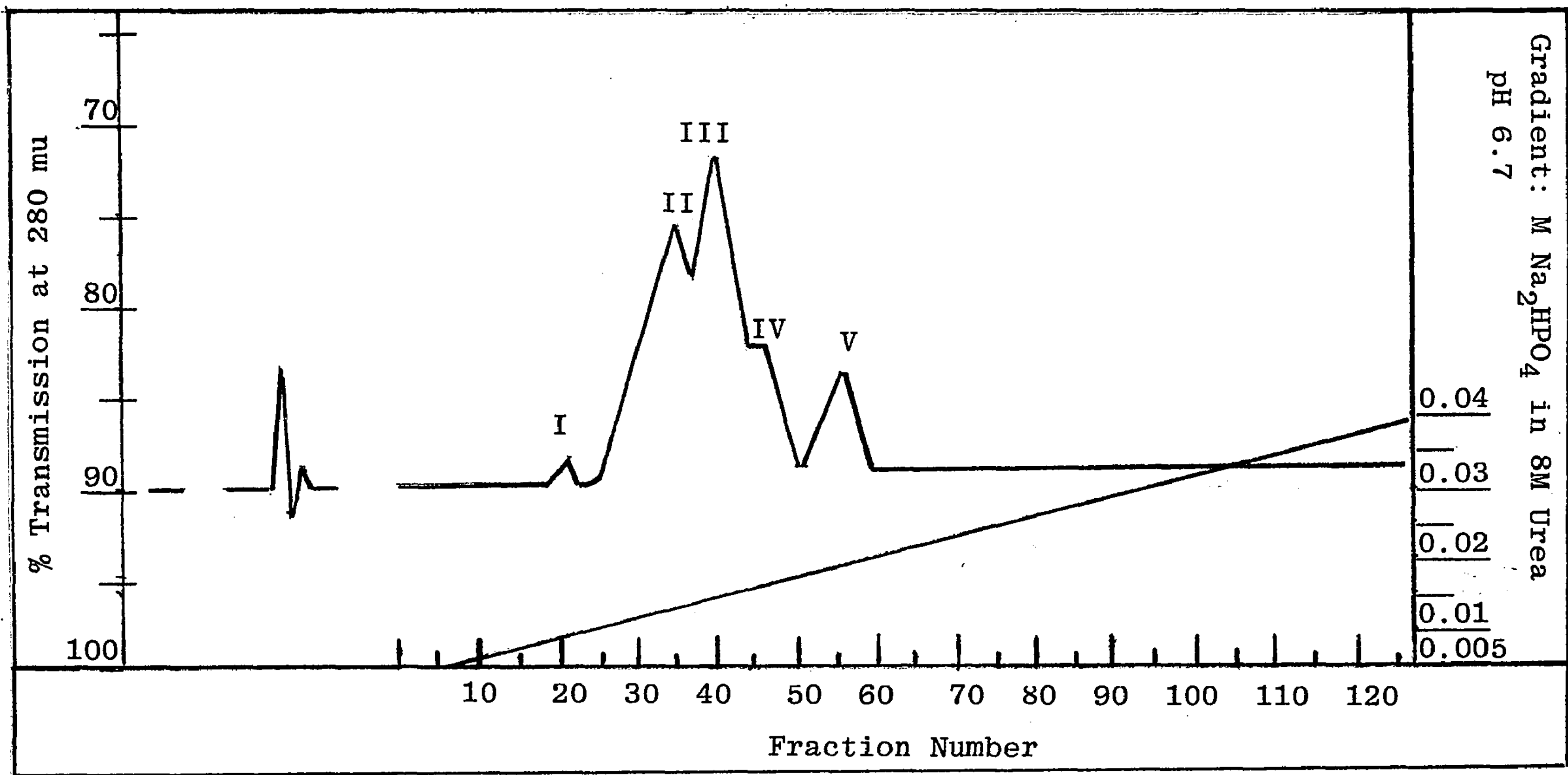

Figure 2

SEPARATION OF TOTAL RAT GLOBIN INTO α AND β CHAINS

UREA-PHOSPHATE SYSTEM

25 mg protein applied to a 1x10 cm carboxymethylcellulose column.

0.030M at a pH of 6.5 gave the best separation of the two major peaks. These conditions were then used to separate rat globin fractions 3, 4 and 5 into α and β chains (Figures 3, 4 and 5). Unlike human globin whose α and β chains elute at either extreme of the gradient, the α and β chains of each rat tetramer revealed close elution patterns. This indicated that the α and β chains from fractions 3, 4 and 5 share a similar overall net charge composition.

Separation of rat fraction #3 globin into α and β chains (Figure 3) was the best one achieved with the urea-phosphate system. Unfortunately there were drawbacks to this system. Once the separated α and β chains were collected, the task of removing the urea was a formidable one which resulted in a considerable loss of separated chains. For every 15-20 mg of starting globin material, only 1-2 mg of each lyophilized chain was recovered after urea removal. Since, in most instances, the separation of the α and β chains was not as effective as desired, it would have been necessary to re-chromatograph each chain on fresh CM cellulose leading to a further loss of protein during a second urea removal treatment.

For these reasons, another method was used which gave adequate separations and at the same time minimized loss of protein during recovery of individual chains. The pyridine-formate method satisfied these requirements. The separation of total rat globin (Figure 6) resulted in only two peaks as compared to the five revealed in the urea-

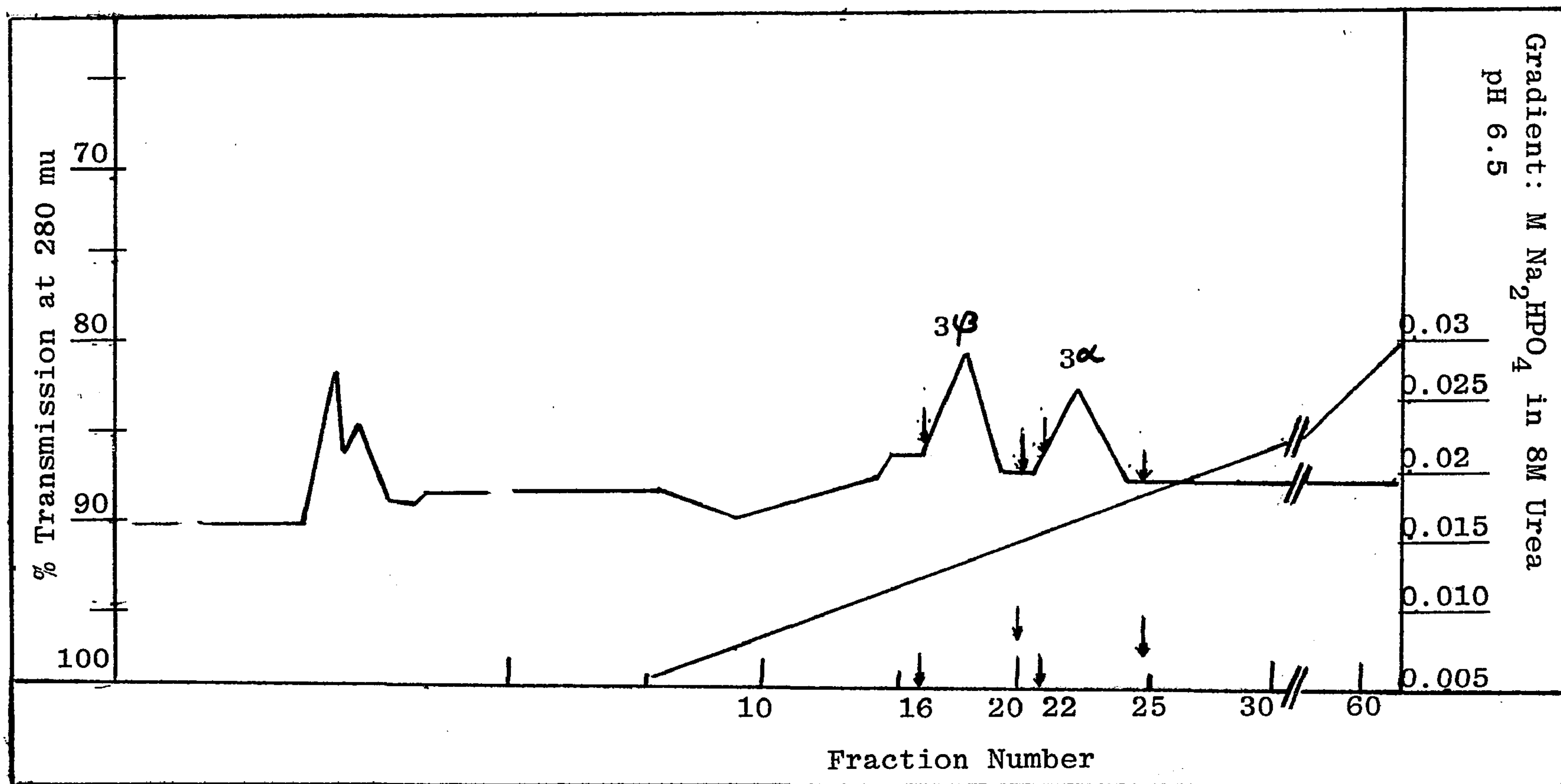

Figure 3

SEPARATION OF RAT GLOBIN #3 INTO α AND β CHAINS

UREA-PHOSPHATE SYSTEM

12 mg protein applied to a 1x10 cm carboxymethylcellulose column.

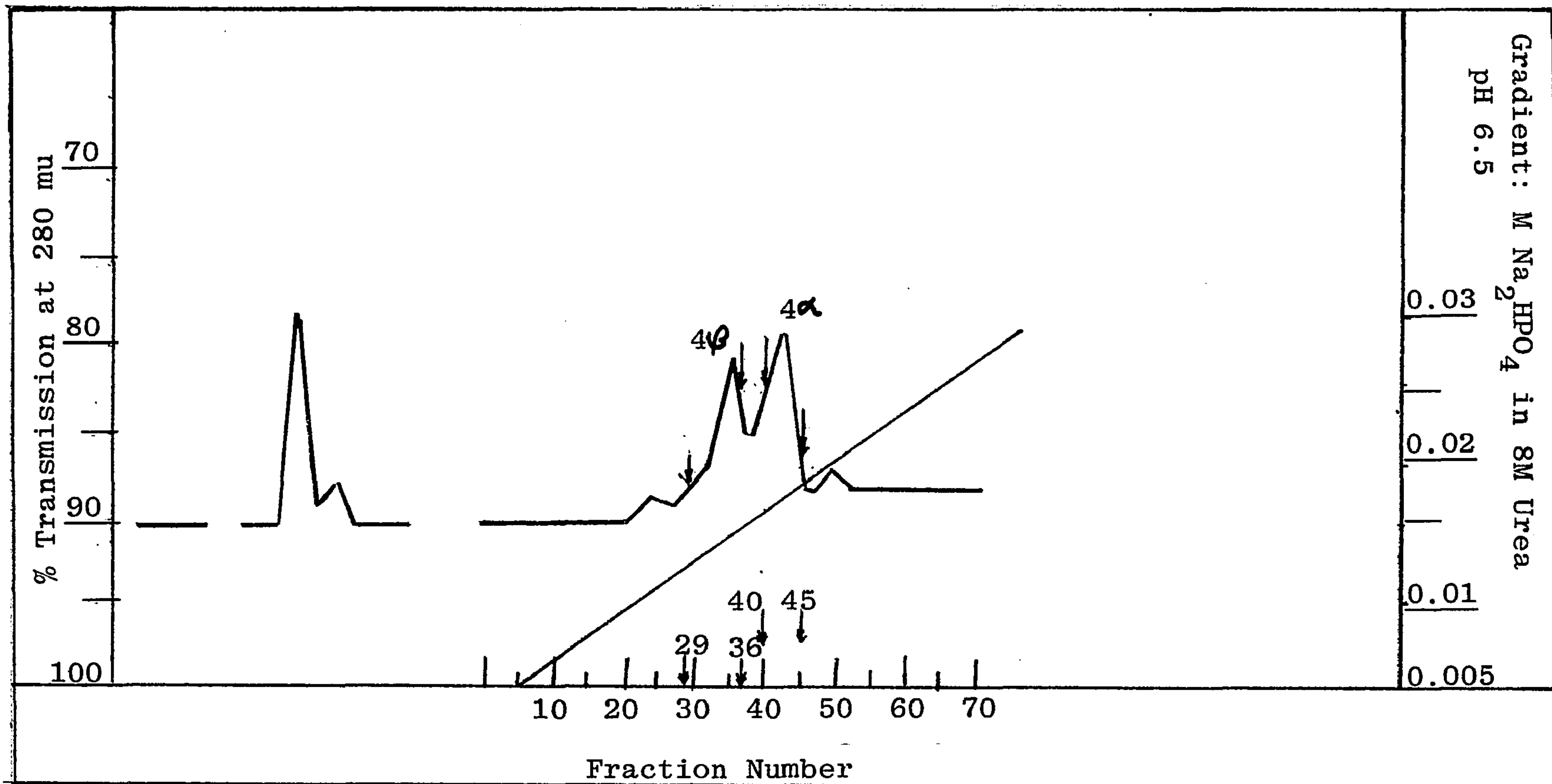

Figure 4

SEPARATION OF RAT GLOBIN #4 INTO α AND β CHAINS

UREA-PHOSPHATE SYSTEM

16 mg protein applied to a 1x10 cm carboxymethylcellulose column.

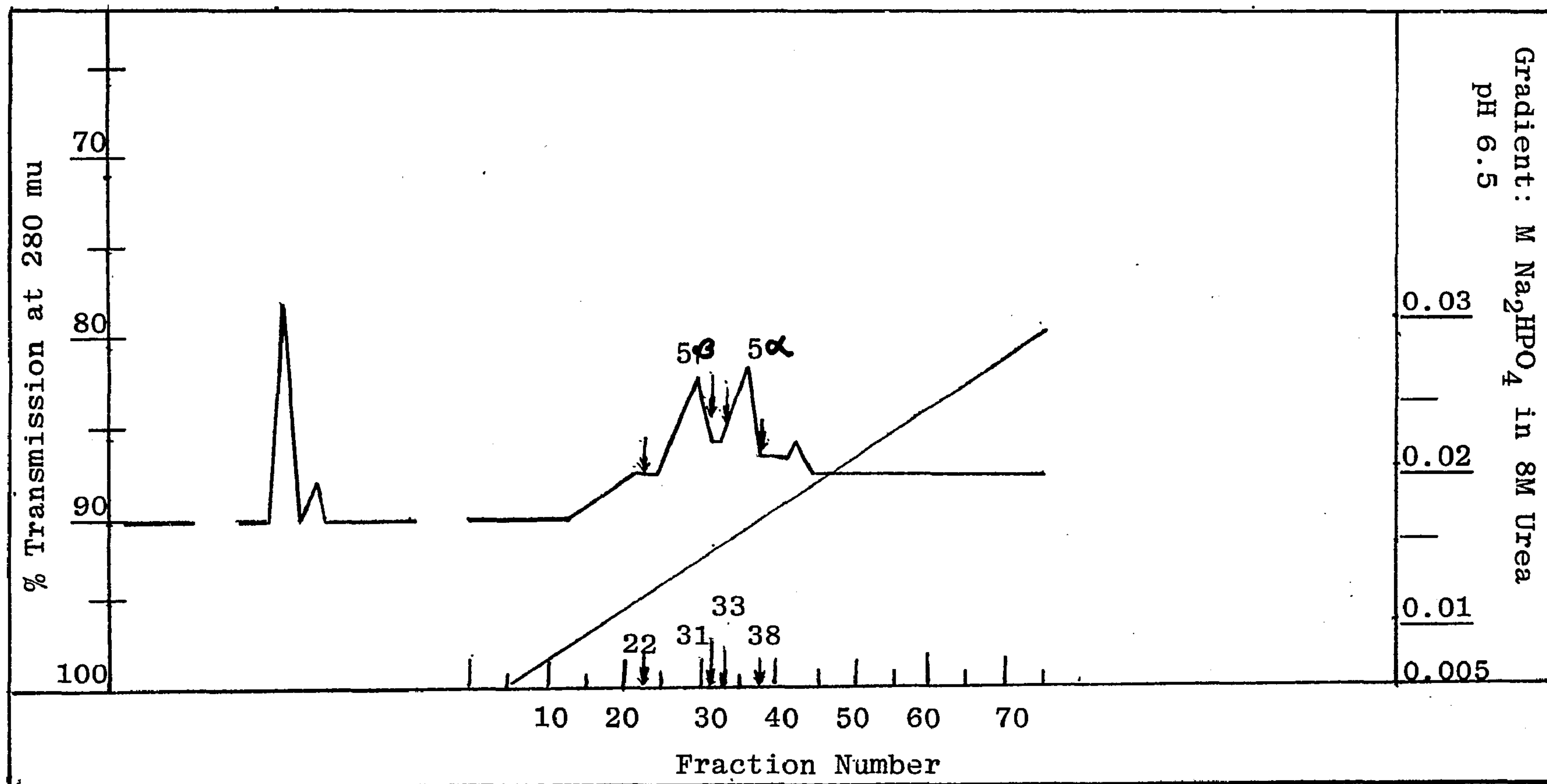

Figure 5

SEPARATION OF RAT GLOBIN #5 INTO α AND β CHAINS

UREA-PHOSPHATE SYSTEM

18 mg protein applied to a 1x10 cm carboxymethylcellulose column.

phosphate system. The first peak represents all the α chains present in total rat globin and the second represents all the β chains present. Separating rat globin 3, 4 and 5 by this method resulted in more of an overlap in the elution patterns of the α and β chains as compared with the separations achieved in the urea-phosphate system (Figures 7,10,13). The problem of protein loss during recovery after each chromatography was greatly reduced by the fact that the pyridine-formate buffer could be lyophilized away with minimal loss of material. After many chain separations from one particular globin fraction, the recovered lyophilized α and β chains were then pooled and re-chromatographed on fresh CM cellulose to yield α chains free of β chain contamination and vice versa (Figures 8,9,11,12,14,15). Re-chromatography of each α and β chain from each rat globin fraction compensated for the initial crude separations. After re-chromatography, the α and β chains were freed from 0.5% formic acid by several lyophilizations before storage at -20°C. For every 15 mg of globin used per separation, an estimated 5-6 mg of each chain was recovered.

Amino Acid Composition of the Chains

In order to determine the amino acid composition of purified α and β chains, acid hydrolysis was performed at 24, 48 and 72 hour intervals. The values in Table I and Table II each represent an average of six hydrolyzates. Arginine, isoleucine and valine with bulky side groups are released more slowly than the other amino acids and their

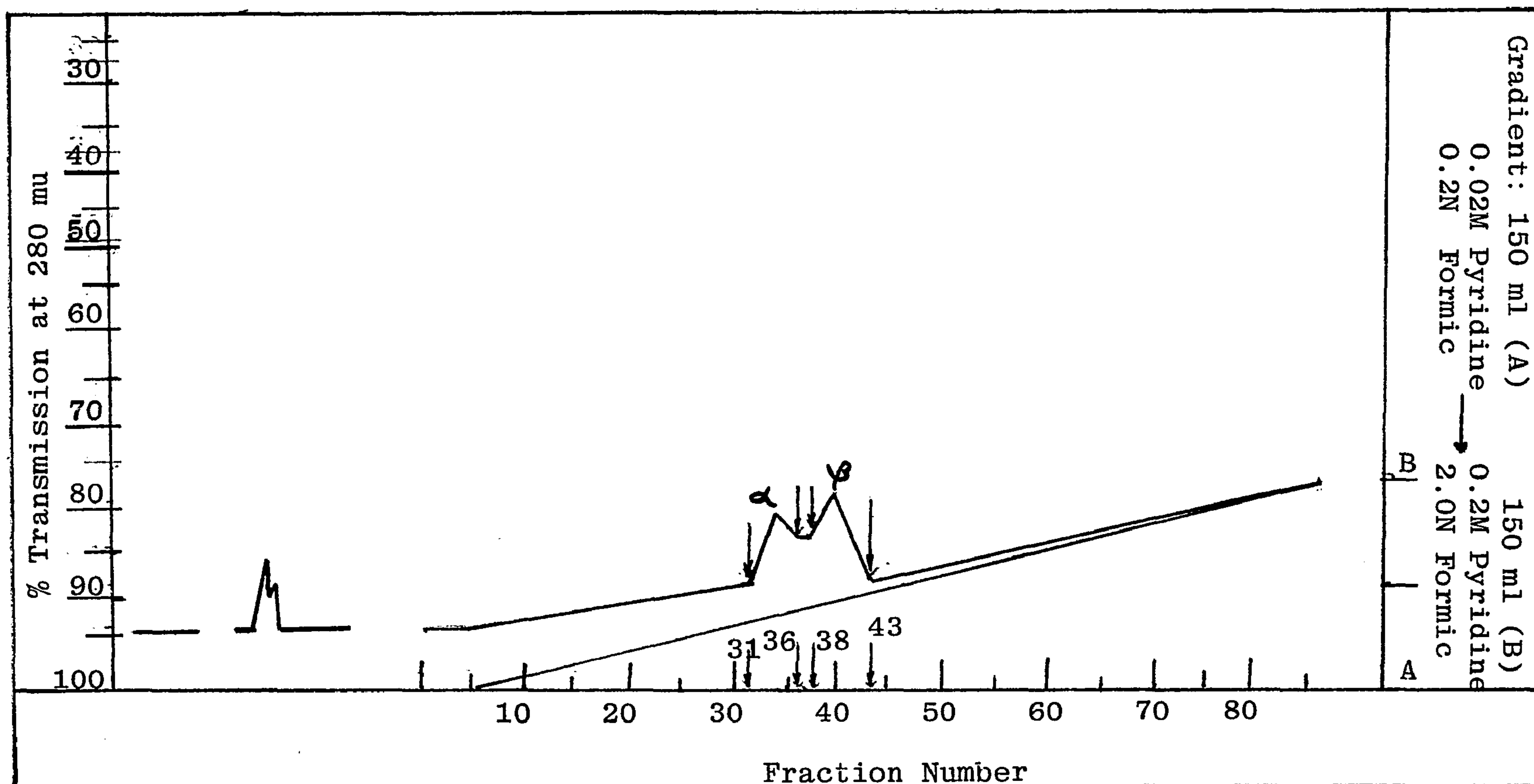

Figure 6 SEPARATION OF TOTAL RAT GLOBIN INTO α AND β CHAINS

PYRIDINE-FORMATE SYSTEM

15 mg protein applied to a 1x10 cm
carboxymethylcellulose column.

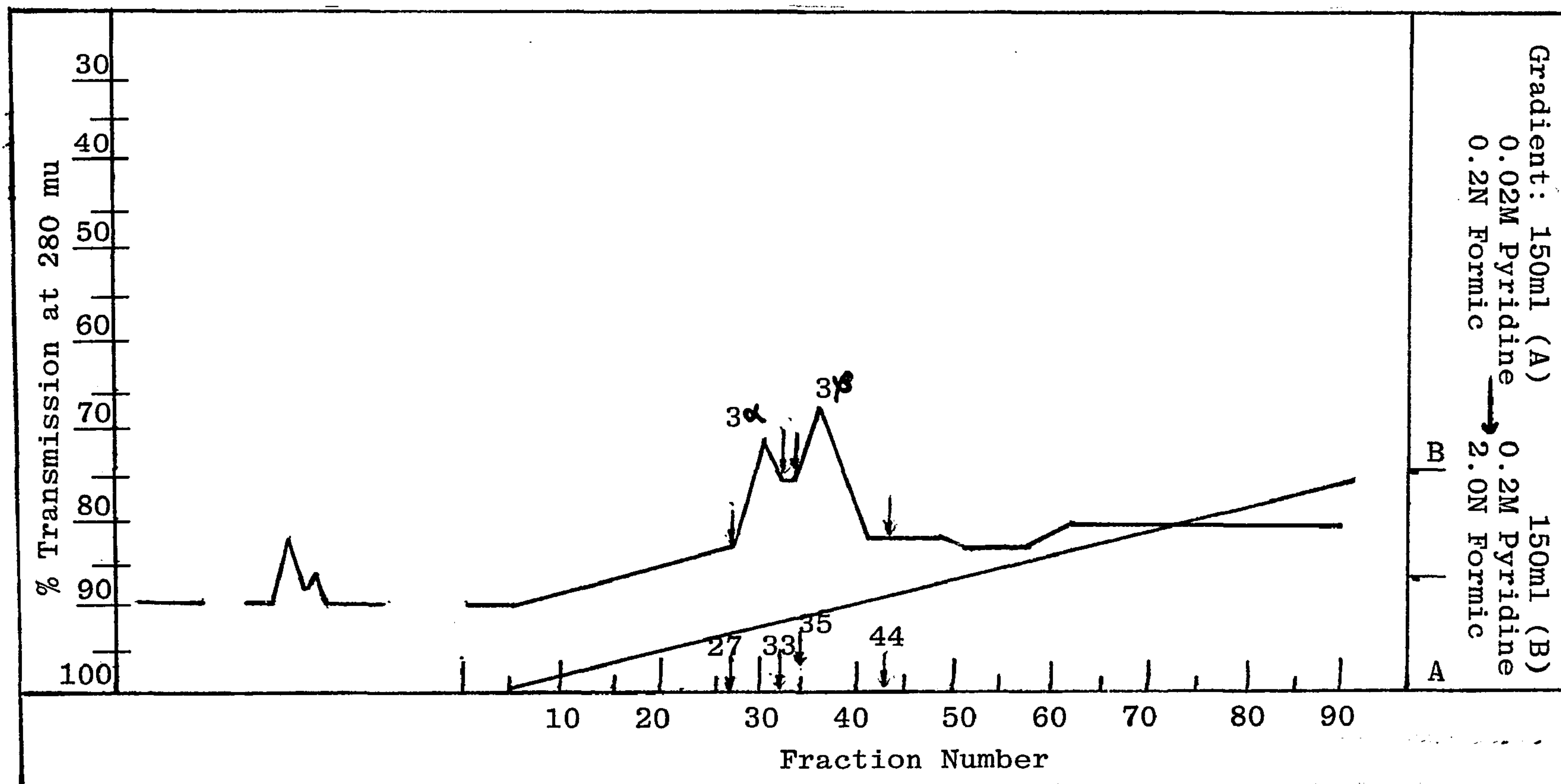

Figure 7

SEPARATION OF RAT GLOBIN #3 INTO α AND β CHAINS

PYRIDINE-FORMATE SYSTEM

28 mg protein applied to a 1x10 cm carboxymethylcellulose column.

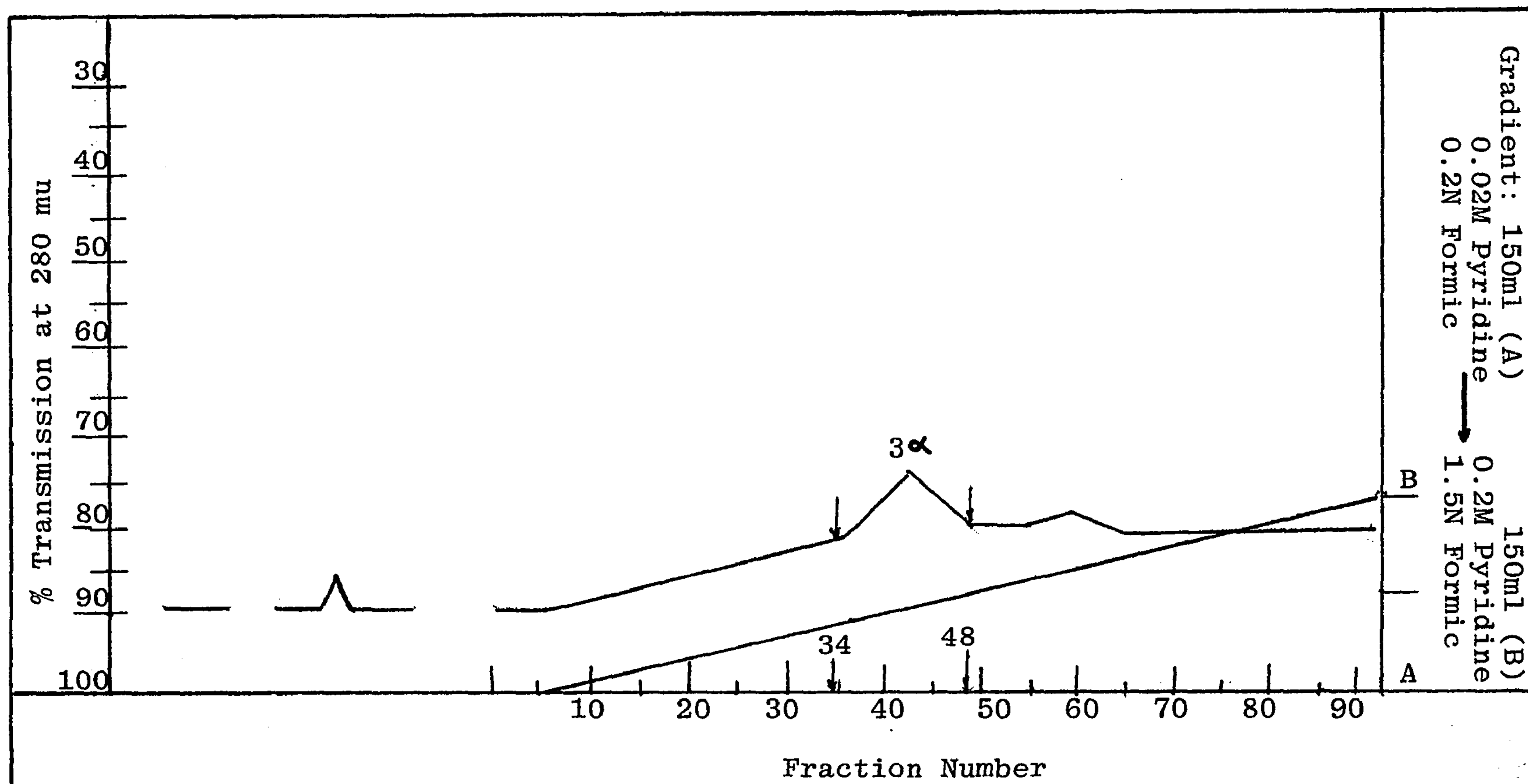

Figure 8 RECHROMATOGRAPHY OF RAT GLOBIN 3α CHAIN

PYRIDINE-FORMATE SYSTEM

~30 mg protein applied to a 1x11 cm carboxymethylcellulose column.

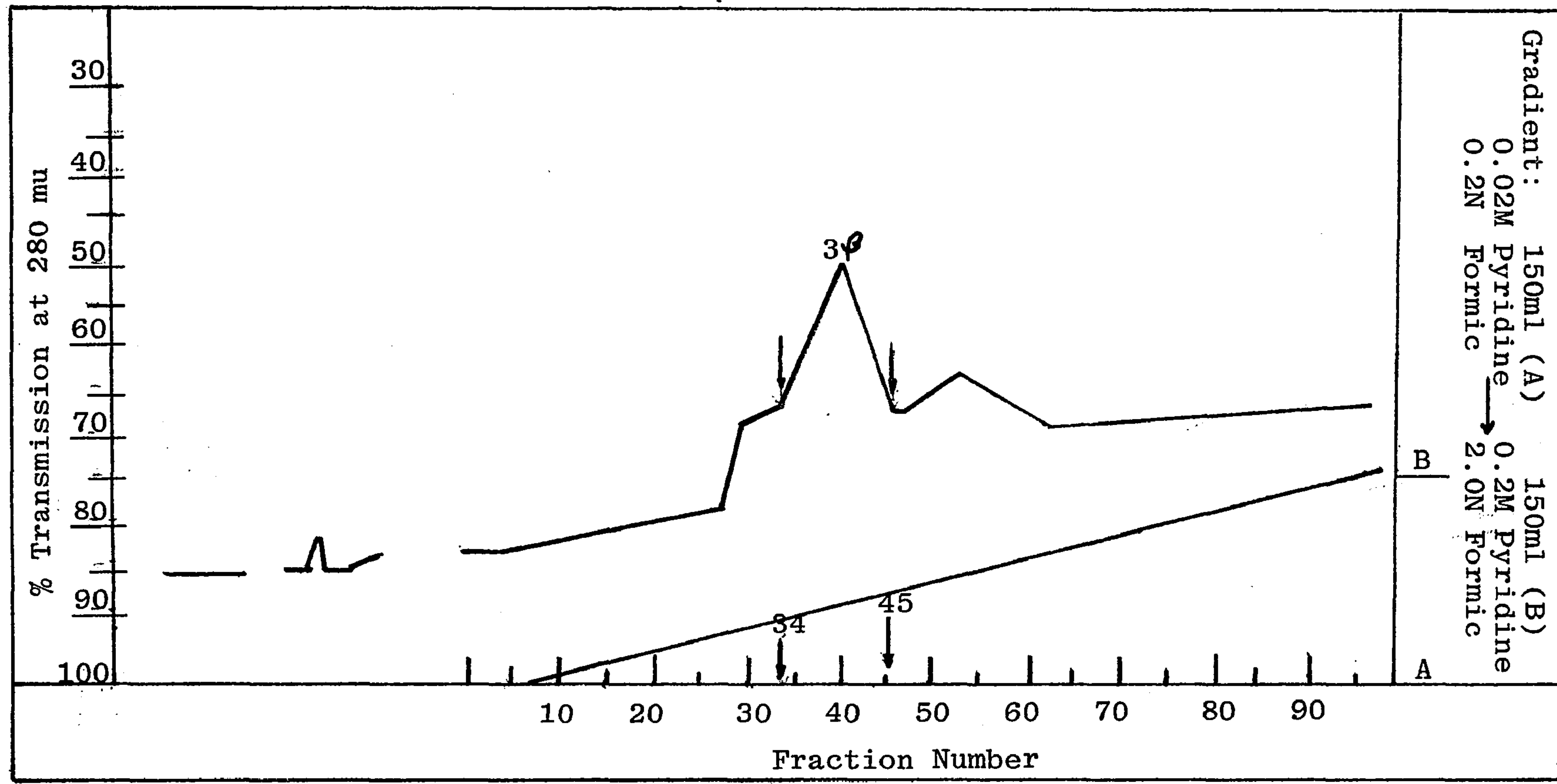

Figure 9 RECHROMATOGRAPHY OF RAT GLOBIN 3β CHAIN

PYRIDINE-FORMATE SYSTEM

~40 mg protein applied to a 1x12 cm carboxymethylcellulose column.

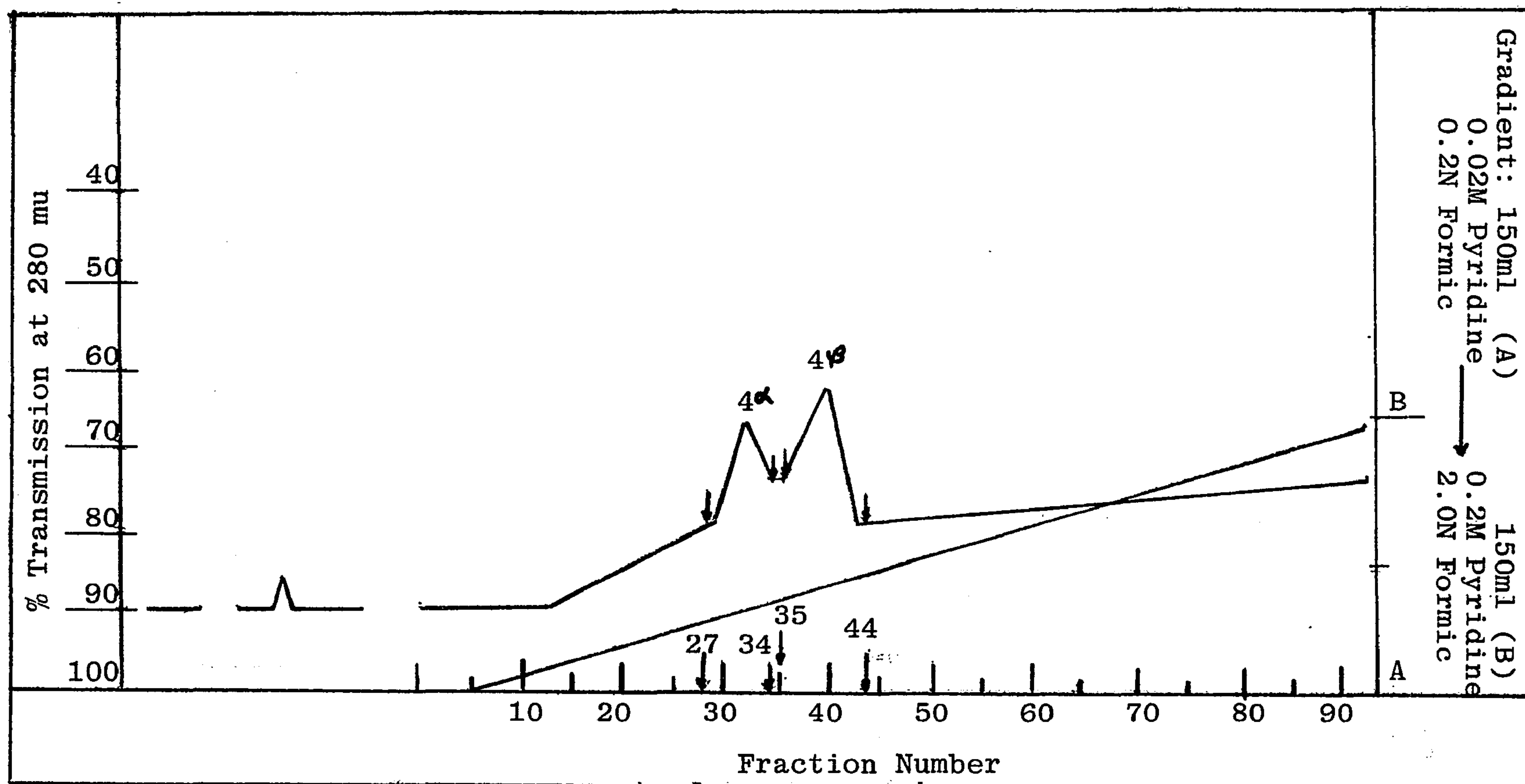

Figure 10

SEPARATION OF RAT GLOBIN #4 INTO α AND β CHAINS

PYRIDINE-FORMATE SYSTEM

23 mg protein applied to a 1x12 cm carboxymethylcellulose column.

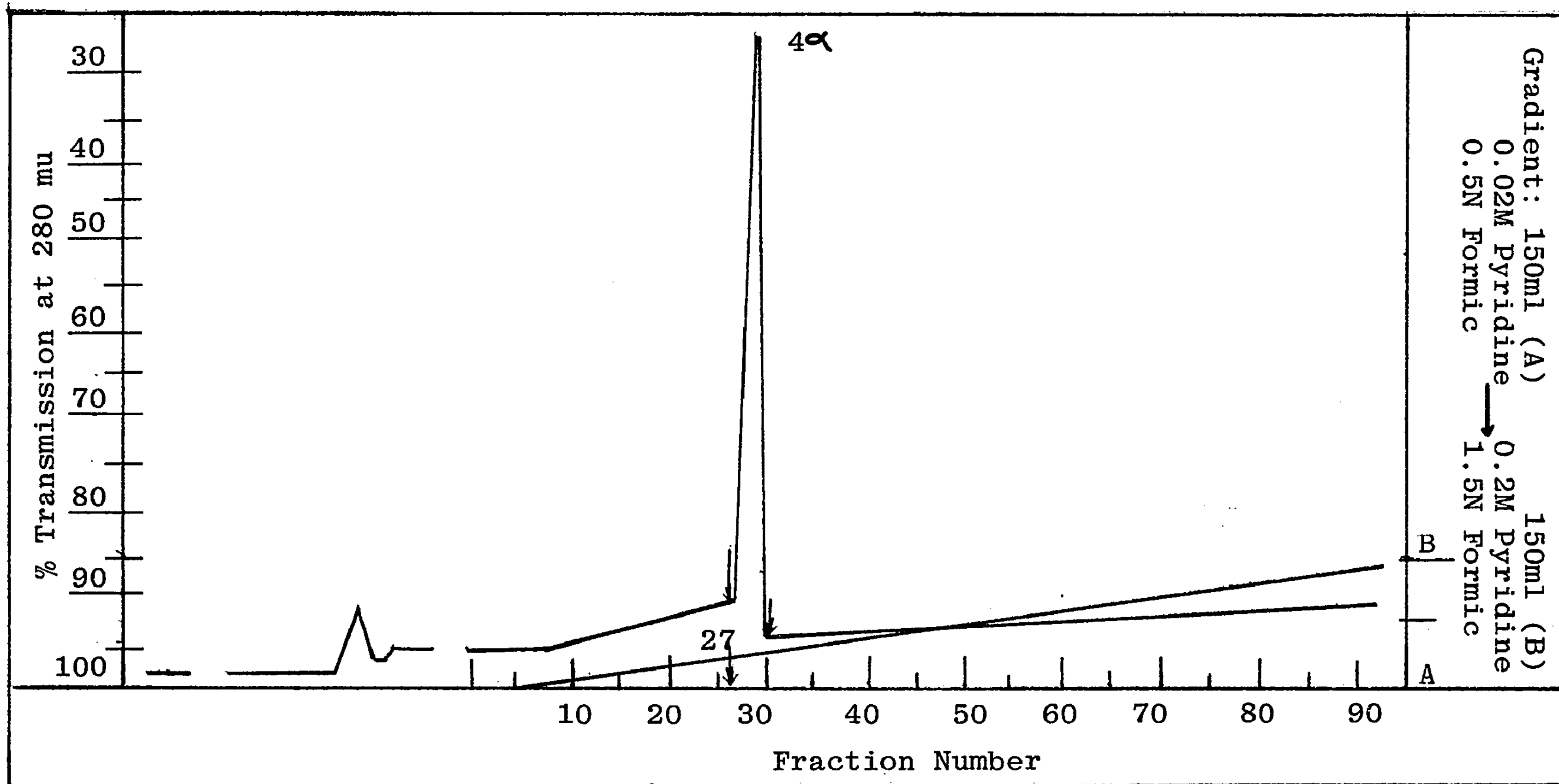

Figure 11 RECHROMATOGRAPHY OF RAT GLOBIN 4α CHAIN

PYRIDINE-FORMATE SYSTEM

~64 mg protein applied to a 1x20 cm carboxymethylcellulose column.

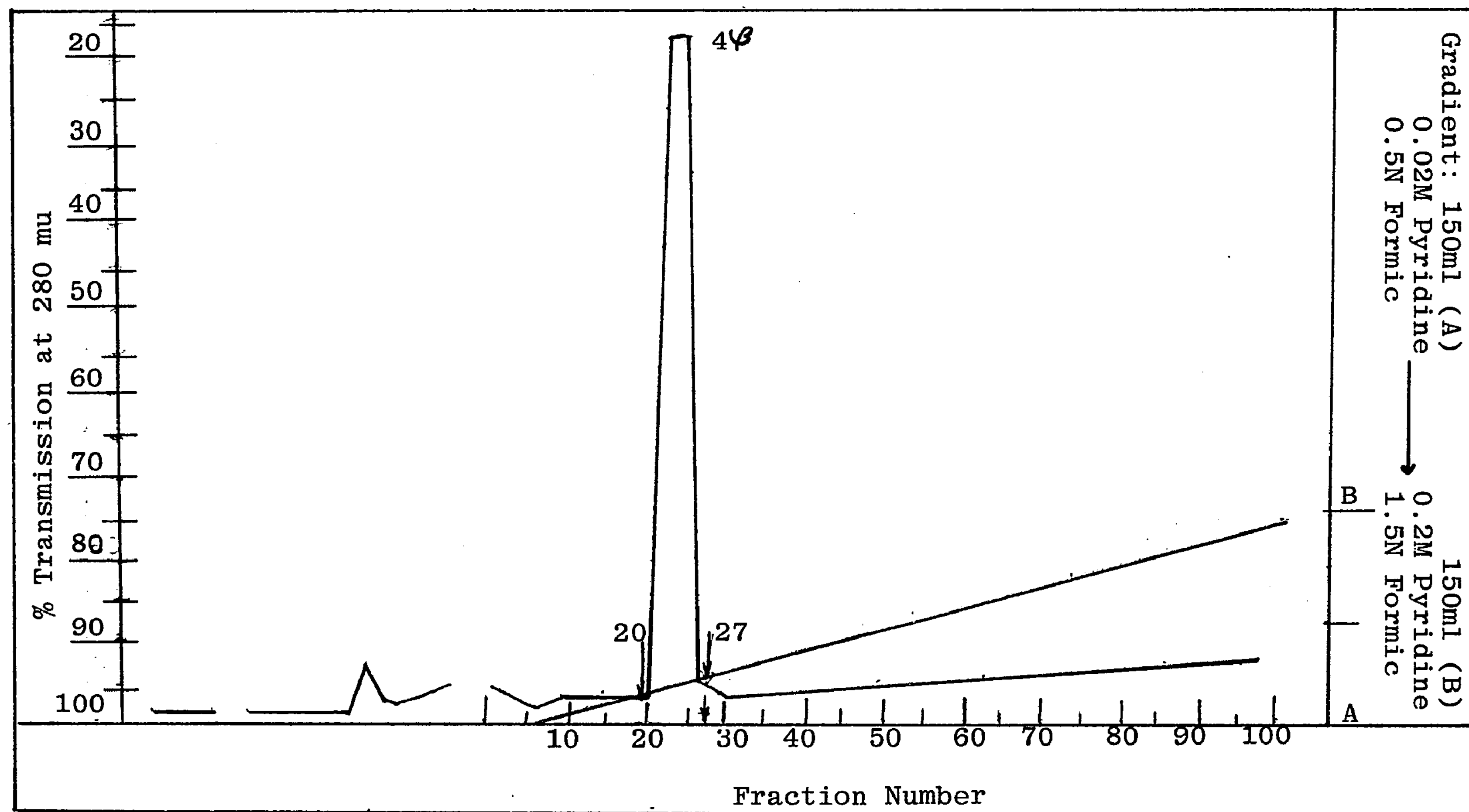

Figure 12

RECHROMATOGRAPHY OF RAT GLOBIN 4β CHAIN

PYRIDINE-FORMATE SYSTEM

~64 mg protein applied to a 1x20 cm carboxymethylcellulose column.

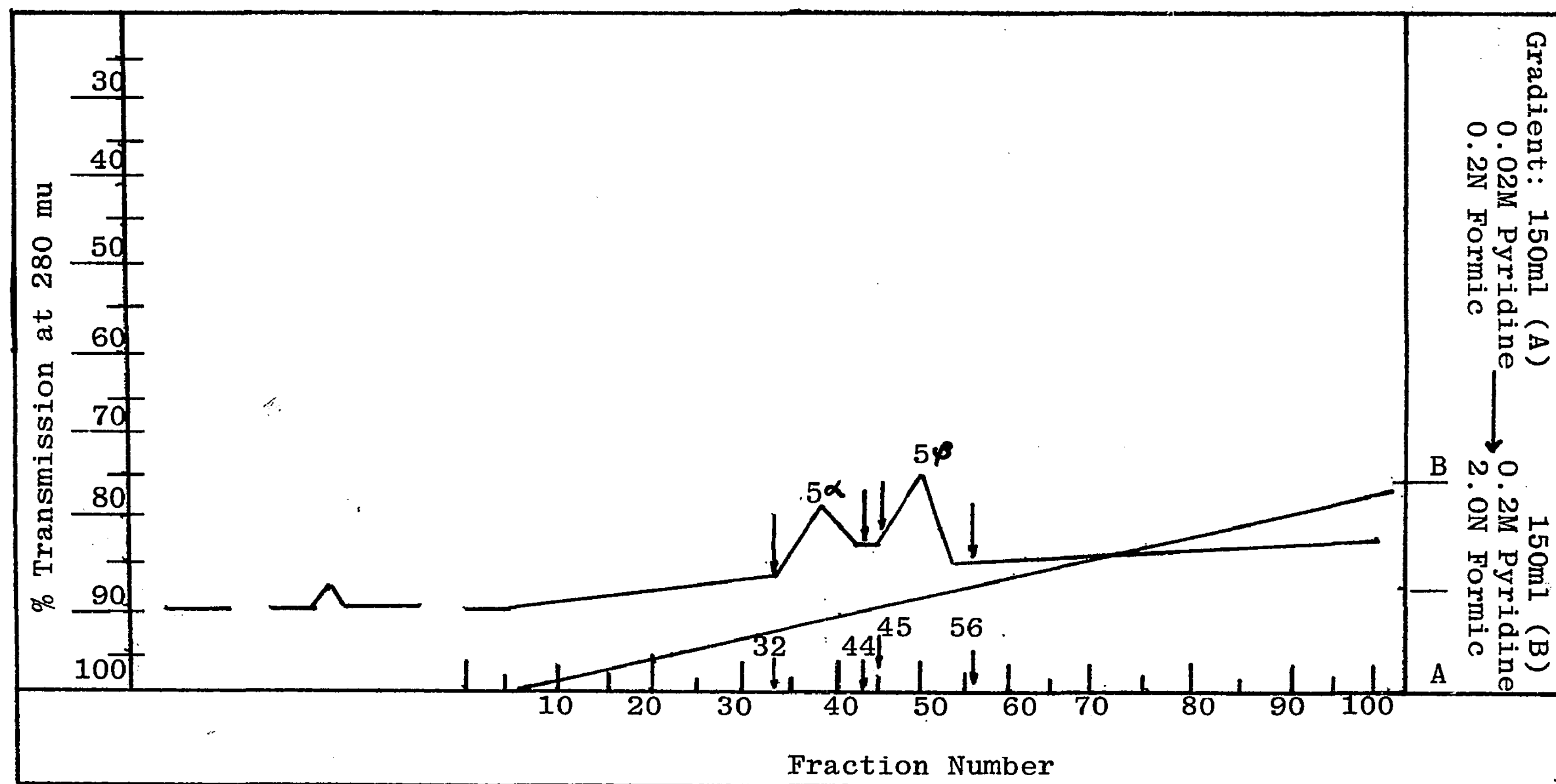

Figure 13

SEPARATION OF RAT GLOBIN #5 INTO α AND β CHAINS

PYRIDINE-FORMATE SYSTEM

16 mg protein applied to a 1x13 cm carboxymethylcellulose column.

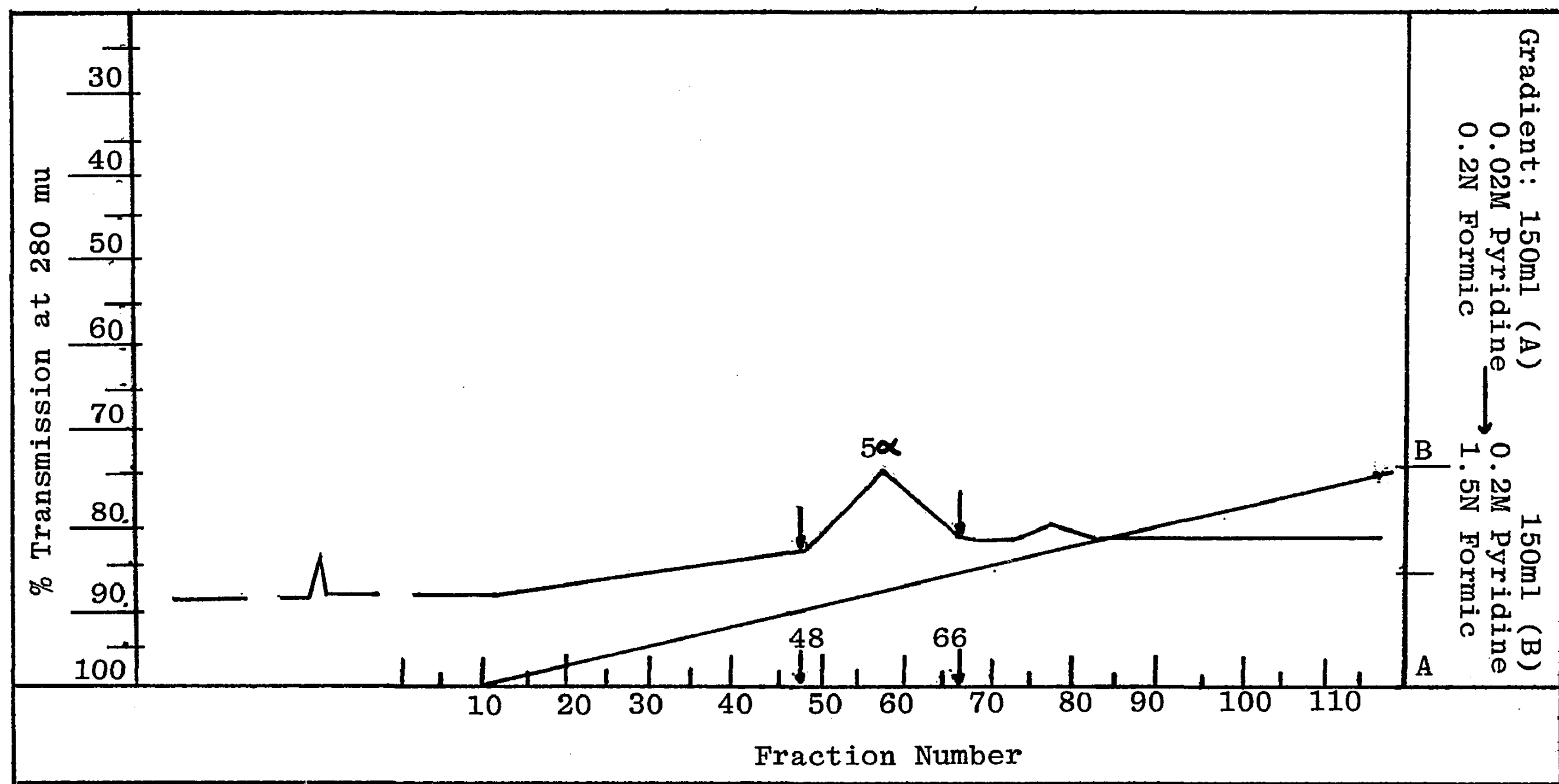

Figure 14

RECHROMATOGRAPHY OF RAT GLOBIN #5α CHAIN

PYRIDINE-FORMATE SYSTEM

~24 mg protein applied to 1x15 cm
carboxymethylcellulose column.

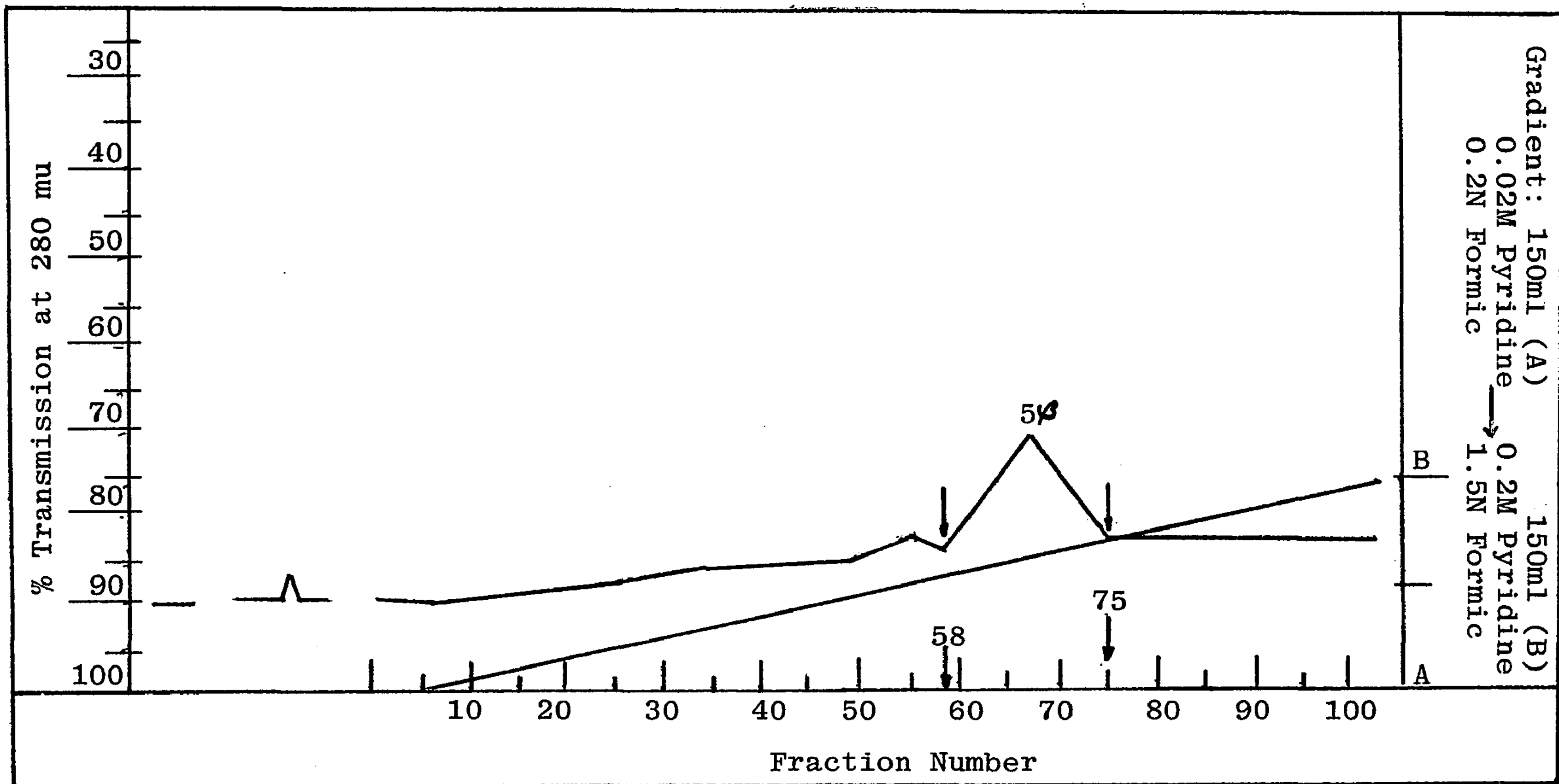

Figure 15

RECHROMATOGRAPHY OF RAT GLOBIN #5β CHAIN

PYRIDINE-FORMATE SYSTEM

~24 mg protein applied to a 1x15 cm carboxymethylcellulose column.

values represent an average of the "72 hour" determinations. The values for threonine, serine and tyrosine are determined by extrapolation back to zero time in order to correct for their partial destruction during acid hydrolysis. Tryptophan and cysteine which are completely destroyed during acid hydrolysis are determined by the methods of Lui (21,22) and Hirs (23), respectively. Cysteine yields cysteic acid after performic acid oxidation and hydrolysis. Other amino acids, not detected by acid hydrolysis, are asparagine and glutamine. These two amino acids are deamidated and recovered as aspartic and glutamic acids, respectively. Methionine is also difficult to recover after acid hydrolysis but can be quantitatively recovered as methionine sulphone after performic acid oxidation (23).

The total number of amino acids for each of the three α chains was found to be 141. The total for 3β and 4β was 146. These total are the same as those reported for other mammalian systems (31). The total for 5β, 147 amino acids, is one more than generally accepted for the β chain but its significance is in doubt.

From the protein hydrolyzate data, differences among α and β chains can be identified (Table IA and Table IIA). The 3α and 3β were found to contain 10 histidines each. This was one more histidine than contained in either the α and β chains of fractions #4 and #5 (Table IA and Table IIA). Fractions 4α, 4β, 5α and 5β all contain only 9 histidines per chain. While all three α chains have

TABLE I AMINO ACID COMPOSITION OF THE α CHAINS

Amino Acids	3α Average[a]			4α Average[a]			5α Average[a]		
Tryptophan	0.9[b]	(n.d.)	1	1.0[b]	(n.d.)	1	0.9[b]	(n.d.)	1
Lysine	12.4	(0.13)	12	12.3	(0.19)	12	11.9	(0.14)	12
Histidine	10.1	(0.26)	10	9.2	(0.26)	9	9.4	(0.18)	9
Arginine	3.0[d]	(0.00)	3	3.0[d]	(0.12)	3	3.1[d]	(0.06)	3
Aspartic	12.9	(0.32)	13	14.1	(0.50)	14	14.3	(0.03)	14
Threonine	8.3[c]	(n.d.)	8	9.8[c]	(n.d.)	10	8.4[c]	(n.d.)	8
Serine	9.6[c]	(n.d.)	10	8.8[c]	(n.d.)	9	10.2[c]	(n.d.)	10
Glutamic	7.2	(0.50)	7	6.7	(0.13)	7	7.4	(0.07)	7
Proline	5.9	(0.20)	6	5.7	(0.38)	6	6.0	(0.22)	6
Glycine	10.7	(0.28)	11	10.8	(0.31)	11	11.8	(0.16)	12
Alanine	17.6	(0.48)	18	14.8	(0.60)	15	17.8	(0.26)	18
Cysteine	1.9[e]	(n.d.)	2	2.0[e]	(n.d.)	2	2.0[e]	(n.d.)	2
Valine	10.9[d]	(n.d.)	11	11.9[d]	(0.24)	12	10.4[d]	(0.14)	10
Methionine	2.0[b,e]	(n.d.)	2	1.8[b,e]	(n.d.)	2	2.3[b,e]	(n.d.)	2
Isoleucine	2.9[d]	(0.10)	3	3.3[d]	(0.17)	3	3.0[d]	(0.09)	3
Leucine	14.0	(0.16)	14	15.0	(0.35)	15	14.4	(0.28)	14
Tyrosine	3.0[c]	(n.d.)	3	3.1[c]	(n.d.)	3	2.6[c]	(n.d.)	3
Phenyl-alanine	7.0	(0.08)	7	6.9	(0.21)	7	6.9	(0.12)	7
Total	140.3		141	140.2		141	142.8		141

[a] Average of protein hydrolysis with 6N HCl at 110°C for 24, 48 and 72 hours (20).

[b] Value from hydrolysis with 4N methane sulfonic acid + 0.2% 3 - (2-Aminoethyl) indole at 115°C for 24 hours (21,22).

TABLE I AMINO ACID COMPOSITION OF THE α CHAINS (Cont'd)

Legend

[c] Extrapolation back to "zero" time.

[d] 72 hour 6N HCl hydrolysis value.

[e] Value from 24 hour 6N HCl hydrolysis after performic acid oxidation (23). Cysteine determined as cysteic acid. Methionine determined as methionine sulphone.

n.d. = not determined

Values in parentheses ( ) represent average deviations.

TABLE IA DIFFERENCES IN THE AMINO ACID COMPOSITIONS OF THE α CHAINS (from Table I)

Amino Acid	3α	4α	5α	Difference
Histidine	10	9	9	3α has one more histidine than 4α or 5α .
Aspartic	13	14	14	3α has one less aspartic acid than 4α or 5α.
Threonine	8	10	8	4α has two more threonines than 3α or 5α .
Serine	10	9	10	4α has one less serine than 3α or 5α .
Glycine	11	11	12	5α has one more glycine than 3α or 4α .
Alanine	18	15	18	4α has three less alanines than 3α or 5α .
Valine	11	12	10	4α has one more valine than 3α and two more valines than 5α .
Leucine	14	15	14	4α has one more leucine than 3α or 5α .

TABLE II AMINO ACID COMPOSITION OF THE β CHAINS

Amino Acids	3β Average[a]			4β Average[a]			5β Average[a]		
Tryptophan	2.0[b]	(n.d.)	2	1.9[b]	(n.d.)	2	1.9[b]	(n.d.)	2
Lysine	11.7	(0.18)	12	11.7	(0.39)	12	12.4	(0.40)	12
Histidine	9.9	(0.24)	10	9.3	(0.37)	9	9.2	(0.02)	9
Arginine	3.0[d]	(0.03)	3	3.1[d]	(0.06)	3	3.1[d]	(0.12)	3
Aspartic	16.9	(0.22)	17	17.2	(0.60)	17	16.9	(0.43)	17
Threonine	4.7[c]	(n.d.)	5	7.9[c]	(n.d.)	8	5.5[c]	(n.d.)	5
Serine	6.6[c]	(n.d.)	7	6.8[c]	(n.d.)	7	8.4[c]	(n.d.)	8
Glutamic	8.7	(0.11)	9	8.0	(0.44)	8	8.8	(0.28)	9
Proline	5.0	(0.12)	5	6.3	(0.50)	6	5.7	(0.20)	6
Glycine	13.7	(0.10)	14	12.9	(0.22)	13	14.2	(0.13)	14
Alanine	16.1	(0.11)	16	17.0	(0.49)	17	17.0	(0.44)	17
Cysteine	0.8[e]	(n.d.)	1	0.9[e]	(n.d.)	1	1.3[e]	(n.d.)	1
Valine	12.7[d]	(0.46)	13	12.3[d]	(0.50)	12	13.5[d]	(n.d.)	13
Methionine	1.7[b,e]	(n.d.)	2	1.8[b,e]	(n.d.)	2	1.7[b,e]	(n.d.)	2
Isoleucine	3.3[d]	(0.33)	3	3.1[d]	(0.12)	3	3.2[d]	(0.19)	3
Leucine	16.8	(0.21)	17	15.9	(0.34)	16	16.4	(0.13)	16
Tyrosine	3.0[c]	(n.d.)	3	2.9[c]	(n.d.)	3	2.6[c]	(n.d.)	3
Phenyl-alanine	7.0	(0.01)	7	7.1	(0.03)	7	7.0	(0.00)	7
Total	143.8		146	146.1		146	148.8		147

[a] Average of protein hydrolysis with 6N HCl at 110°C for 24, 48 and 72 hours (20).

[b] Value from hydrolysis with 4N methane sulfonic acid + 0.2% 3-(2-Aminoethyl) indole at 115°C for 24 hours (21,22).

TABLE II AMINO ACID COMPOSITION OF THE β CHAINS (Cont'd)

Legend

[c] Extrapolation back to "zero" time.

[d] 72 hour 6N HCl hydrolysis value.

[e] Value from 24 hour 6N HCl hydrolysis after performic acid oxidation (23). Cysteine determined as cysteic acid. Methionine determined as methionine sulphone.

n.d. = not determined

Values in parentheses ( ) represent average deviations.

TABLE IIA — DIFFERENCES IN THE AMINO ACID COMPOSITIONS OF THE β CHAINS (from Table II)

Amino Acid	3β	4β	5β	Differences
Histidine	10	9	9	3β has one more histidine than 4β or 5β .
Glutamic	9	8	9	4β has one less glutamic acid than 3β or 5β .
Proline	5	6	6	3β has one less proline than 4β or 5β .
Serine	7	7	8	5β has one more serine than 3β or 4β .
Glycine	14	13	14	4β has one less glycine than 3β or 5β .
Valine	13	12	13	4β has one less valine than 3β or 5β .
Leucine	17	16	16	3β has one more leucine than 4β or 5β .
Alanine	16	17	17	3β has one less alanine than 4β or 5β .
Threonine	5	8	5	4β has three more threonines than 3β or 5β .

the same number of glutamic residues, 3α has one less aspartic acid residue than 4α and 5α . Similarly in the β chains, all three β chains have the same number of aspartic acids but 4β has one less glutamic acid residue than 3β and 5β . Since these values also include glutamic acid residues arising from glutamine (and aspartic residues from asparagine) the significance of these differences (and similarities) awaits an investigation into the amide components of each of the chains.

Other differences in the total amino acid composition of the α chains (Table IA) mainly revolve around changes in content of neutral and polar amino acids. Globin chain 4α has 10 threonines, two more than either 3α or 5α which have 8 threonines. It has one less serine (only 9) than 3α and 5α which have 10 serine residues each. There appear to be 12 glycines in 5α as opposed to the 11 such residues found in 3α and 4α . A significant difference can be seen in the alanines where there are 15 residues in 4α and 18 residues in 3α and 5α - a difference of 3 alanine residues. The valine amino acids vary according to the following pattern: 4α has one more valine than 3α which has one more valine than 5α . 4α then has two more valines than the 5α chain. Another difference among the α chains can be seen in the leucine values: 4α has one more residue than either 3α or 5α (Table IA).

These differences in the numbers of neutral and polar amino acids can also be found in the β chains (Tables

II & IIA). An interesting difference in the β 's involves the amino acid proline. It appears that 3β contains one less proline (5 residues) than 4β and 5β which contain 6 residues each. The 5β chain contains one more serine than the other two while 4β has one less glycine and one less valine than 3β and 5β. Also, 3β has one more leucine and one less alanine than either 4β or 5β. The only significantly large change in amino acid composition occurring among the β chains involves the residue threonine. The 4β chain appears to contain 8 threonines as compared to the five found in each of the 3β and 5β chains (Table IIA).

In summary, there are enough differences in overall amino acid composition among the α chains and among the β chains (Table IA & IIA) to suggest that 3α, 4α & 5α all differ from each other as do the 3β, 4β and 5β chains. In order to confirm this, tryptic peptide map comparisons of each of the α and β chains were made and amino acid analyses of the selected tryptic peptides were performed.

Trypsin Digestion; Two-Dimensional Peptide Mapping

The rat α and β chains under investigation were found to be digested readily by trypsin, leaving a minimum of insoluble core material after digestion was terminated. Trypsin acts at the site of lysine and arginine residues. Cleaving at the carbonyl function donated by the residue, trypsin releases peptides that contain lysine or arginine

as the C-terminal residues. For this reason, when determining the amino acid compositions of the tryptic peptides, the molar proportion ratios were always based on the assumption that the lysine (or arginine) present represented 1 lysine (or arginine) residue per mole.

The peptides released from each chain were resolved by two-dimensional chromatography. It was found that descending chromatography followed by high voltage paper electrophoresis gave the best separation of the peptides in each "fingerprint" map (Figures 16 and 17). The number of peptides expected after trypsin digestion is one more than the combined number of lysine and arginine in the protein. For rat α and β chains 16 peptides were predicted. In Figure 16 it can be seen that in the 3α and 4α peptide maps there are 15 spots and in the 5α map, a total of 17 peptides is seen. The fact that rat α chains, like other mammalian α chains contain arginine as a C-terminal residue (31) accounts for the one less peptide found. The two 5α peptides enclosed by dotted lines and labeled with the letters "a" and "b" represent spots that did not appear consistently on all 5α maps. They were, however, included on the 5α map. Amino acid analysis of these spots could reveal their significance. The β maps (Figure 17) revealed 16 peptides for each chain.

The peptide spots were first located by spraying a map with a 0.5% solution of ninhydrin. The same maps were then dipped into Ehrlich's reagent to reveal tryptophan-

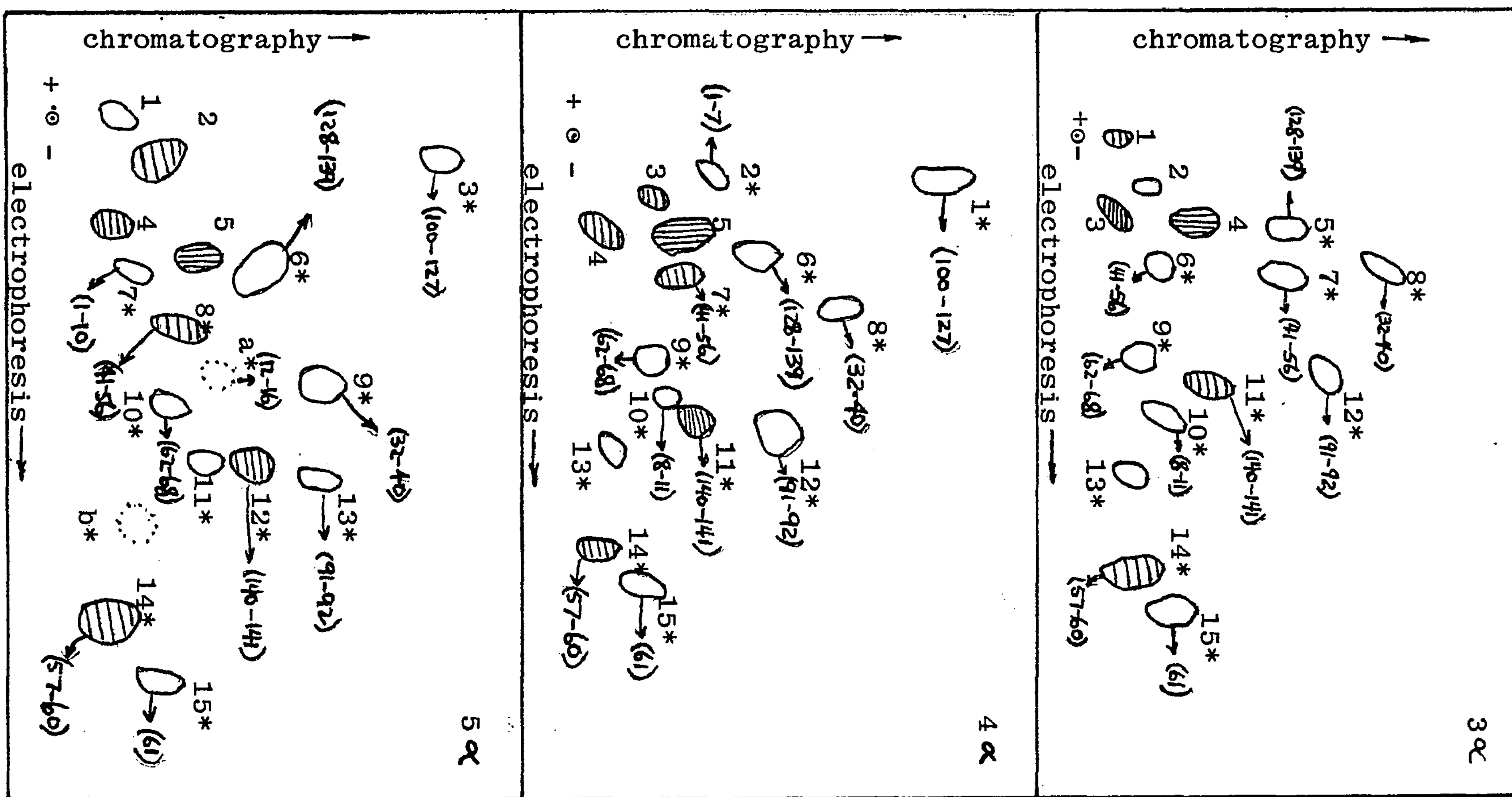

Figure 16 TRYPTIC PEPTIDE MAPS OF THE α CHAINS

Two-Dimensional Chromatography (see experimental for conditions).
(horizontally hatched spot) = Positive Pauly reaction for histidine and/or tyrosine.
(vertically hatched spot) = Positive Ehrlich reaction for tryptophan.
* = Peptide analyzed for amino acid content.
Numbers in ( ) indicate that portion of the α chain sequence the peptide represents.

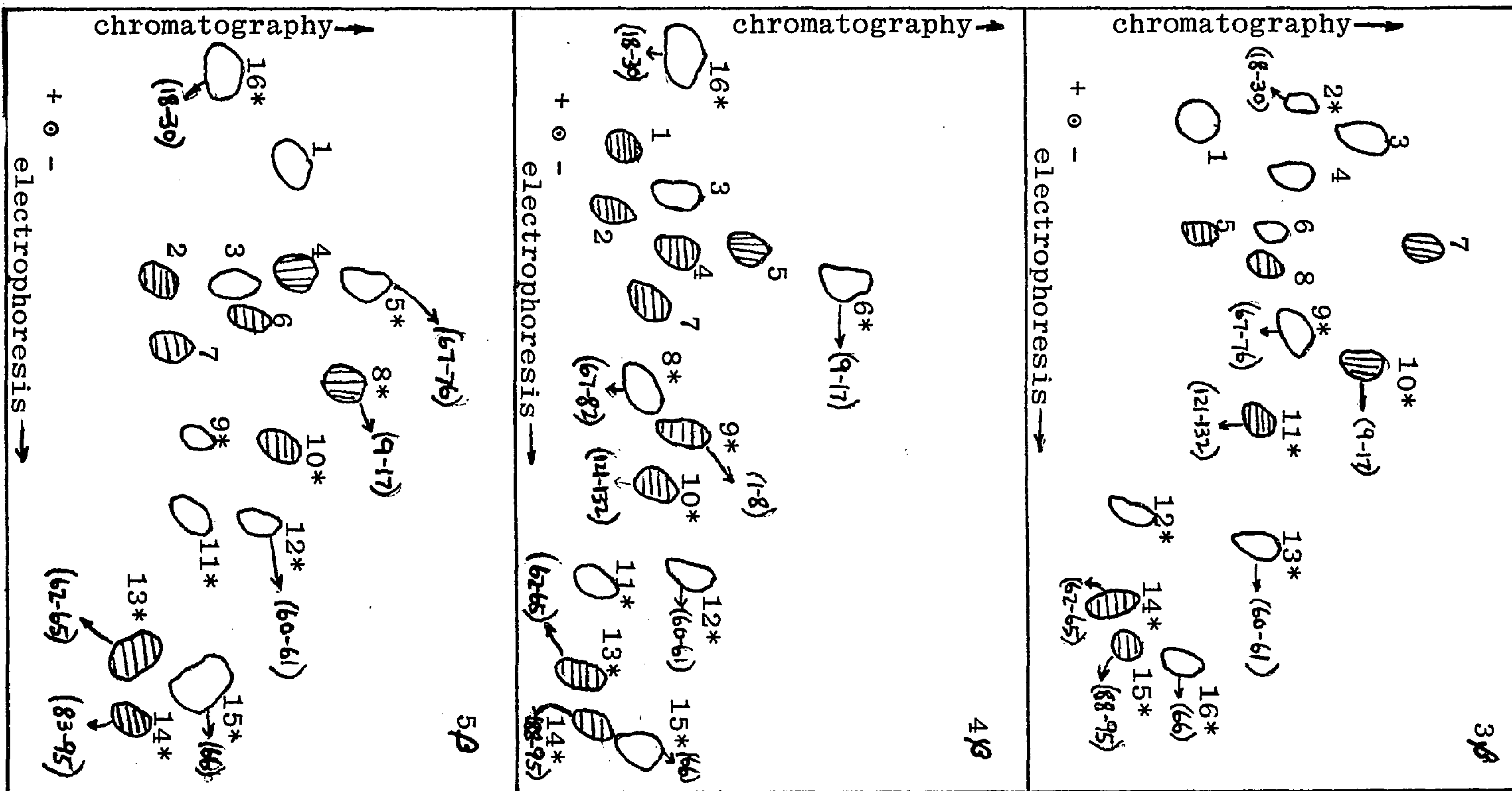

Figure 17 TRYPTIC PEPTIDE MAPS OF THE β CHAINS

Two-Dimensional Chromatography (see experimental for conditions).
(horizontally hatched spot) = Positive Pauly reaction for histidine and/or tyrosine.
(vertically hatched spot) = Positive Ehrlich reaction for tryptophan.
* = Peptide analyzed for amino acid content.
Numbers in ( ) indicate that portion of the β chain sequence the peptide represents.

containing spots. Since this particular preparation of Ehrlich's reagent contained HCl (see experimental), the purple color due to the ninhydrin was bleached out before the faint purple of a positive Ehrlich was seen. In Figure 16, it can be seen that one peptide in each of the 3α, 4α and 5α maps (peptides 3α4, 4α5 and 5α5) gave a positive Ehrlich reaction. This is in agreement with the tryptophan determination data (Table I) which shows one tryptophan per α chain. For the β chains, two peptides gave a positive Ehrlich reaction in the 3β and 5β maps (peptides 3β7, 3β10, and 5β4, 5β8) but only one peptide in the 4β map (peptide 4β5) gave such a positive reaction (Figure 17). The data in Table II indicates that there are two tryptophans per β chain. This means, in the case of the 4β map, that either peptide 4β5 contains both tryptophan residues or that the other tryptophan containing peptide is not revealed by the Ehrlich reagent. The Ehrlich reagent, requiring a longer time to develop than the ninhydrin reagent (33), might not give a positive response to a tryptophan-bearing peptide present in limited quantity thus favoring the latter of the two above possibilities.

A second peptide map was used to detect histidine and tyrosine containing peptides with the Pauly reagent. In each of the α peptide maps, five tryptic peptides gave a positive Pauly reaction. They were (from Figure 16): 3α1, 3α3, 3α4, 3α11, 3α14; 4α3, 4α4, 4α7, 4α11, 4α14; 5α2, 5α4, 5α8, 5α12, 5α14. As was discovered later,

some of these peptides contained more than one histidine, histidine and tyrosine, or tyrosine alone. Of the peptides examined, there were those that did not give a positive Pauly reaction but contained histidine after analysis. These were: 3α7, 3α6; 4α1; 5α3. The latter two also contained tyrosine (Tables IV, V, VI). In the case of the $\gamma\beta$ chains (Figure 17), five tryptic peptides in the 3$\gamma\beta$ map (peptides 3$\gamma\beta$5, 3$\gamma\beta$8, 3$\gamma\beta$11, 3$\gamma\beta$14 and 3$\gamma\beta$15), eight tryptic peptides in the 4$\gamma\beta$ map (peptides 4$\gamma\beta$1, 4$\gamma\beta$2, 4$\gamma\beta$4, 4$\gamma\beta$7, 4$\gamma\beta$9, 4$\gamma\beta$10, 4$\gamma\beta$13, 4$\gamma\beta$14) and six tryptic peptides in the 5$\gamma\beta$ map (peptides 5$\gamma\beta$2, 5$\gamma\beta$6, 5$\gamma\beta$7, 5$\gamma\beta$10, 5$\gamma\beta$13, 5$\gamma\beta$14) gave positive Pauly tests. Peptide 4$\gamma\beta$8 was found to contain histidine and tyrosine (Table VIII) even though it did not give a positive Pauly reaction. The Pauly test results were used more for peptide orientation and for comparison among maps rather than for a quantitation of histidine and/or tyrosine per chain. Hence, if the 4$\gamma\beta$ tryptic peptide map shows more positive Pauly spots than either the 3$\gamma\beta$ or the 5$\gamma\beta$ maps, it does not necessarily contradict the results seen in Table II with respect to the total number of histidines and tyrosines in the 3$\gamma\beta$, 4$\gamma\beta$ and 5$\gamma\beta$ chains.

Comparison of Tryptic Peptide Maps

An examination of the 3α, 4α and 5α tryptic peptide maps (Figure 16) shows that many of the peptide spots are common to all three chains. The same is true for the 3$\gamma\beta$, 4$\gamma\beta$ and 5$\gamma\beta$ tryptic maps (Figure 17). Common peptides are those peptides located in one position on one map that can

be found in the same position or at least in the same vicinity in another map. If two peptides display the same mobilities after two-dimensional chromatography, then they are expected to posess the same overall net charges and net polarities, indicating the same overall amino acid composition. After inspection of Figure 16, a group of common and unique α peptides can be assembled (Table IIIa). A summary of this table shows that there are ten tryptic peptides common to all three α chains, two tryptic peptides common to 3α and 4α alone, two common to 3α and 5α alone, and two tryptic peptides common to 4α and 5α alone. There are five peptides - one each from 3α and 4α and three peptides from 5α - that do not have matching counterparts in the other maps and are therefore unique to that particular α chain. From this data one is tempted to conclude that there are at least three different α chains in rat hemoglobin. Supporting evidence from amino acid composition studies will be used to confirm this conclusion.

In the same way, the common and unique β tryptic peptides can be identified from Figure 17 (Table IIIb). There are six peptides common to all three β chains, three peptides common to just 3β and 4β, four peptides common only to 3β and 5β, and five peptides common to 4β and 5β. Altogether, there are six unique peptides - three from the 3β map, two from the 4β map and one from the 5β map - that do not have counterparts in the other β maps. Again, a conclusion that there are at least three different

TABLE IIIa TRYPTIC PEPTIDE COMPARISONS BY MAP POSITIONING AFTER TWO-DIMENSIONAL CHROMATOGRAPHY-α CHAINS. (from Figure 16).

Tryptic Peptides Common to 3α, 4α and 5α

3α	4α	5α
*3α15	*4α15	*5α15
*3α14	*4α14	*5α14
*3α13	*4α13	*5α b
*3α12	*4α12	*5α13
*3α11	*4α11	*5α12
*3α10	*4α10	*5α11
*3α 9	*4α 9	*5α10
*3α 5	*4α 6	*5α 6
3α 4	4α 5	5α 5
3α 3	4α 4	5α 4

Tryptic Peptides Common to 3α and 4α Only

3α	4α
*3α 8	*4α 8
3α 2	4α 3

Tryptic Peptides Common to 3α and 5α Only

3α	5α
*3α 6	*5α 7
3α 1	5α 1

Tryptic Peptides Common to 4α and 5α Only

4α	5α
*4α 7	*5α 8
*4α 1	*5α 3

Tryptic Peptides Unique Only to 3α, 4α and 5α

*3α 7
*4α 2
*5α 9
5α 2
*5α a

* Indicates those tryptic peptides that have been analyzed for amino acid content (Tables IV, V and VI).

TABLE IIIb   TRYPTIC PEPTIDE COMPARISONS BY MAP POSITIONING AFTER TWO-DIMENSIONAL CHROMATOGRAPHY - β CHAINS. (from Figure 17).

Tryptic Peptides Common to 3β, 4β and 5β

*3β 16	*4β 15	*5β 15
*3β 15	*4β 14	*5β 14
*3β 14	*4β 13	*5β 13
*3β 13	*4β 12	*5β 12
3β 6	4β 4	5β 3
3β 5	4β 2	5β 2

Tryptic Peptides Common to 3β and 4β Only

*3β 12	*4β 11
3β 4	4β 3
3β 1	4β 1

Tryptic Peptides Common to 3β and 5β Only

*3β 11	*5β 10
*3β 10	*5β 8
3β 8	5β 6
3β 3	5β 1

Tryptic Peptides Common to 4β and 5β Only

*4β 16	*5β 16
*4β 8	*5β 9
4β 7	5β 7
*4β 6	*5β 5
4β 5	5β 4

Tryptic Peptides Unique Only to 3β, 4β and 5β

*3β 9
3β 7
*3β 2
*4β 10
*4β 9
*5β 11

* Indicates those tryptic peptides that have been analyzed for amino acid content (Tables VII, VIII and IX).

β chains in rat hemoglobin can be drawn from this evidence, and can be confirmed from the amino acid compositions of the peptides.

Amino Acid Compositions of Selected α and β Chain Peptides; Sequence Homologies of These Peptides.

Eleven tryptic peptides from the 3 α chain, twelve from the 4α chain and eleven peptides from the 5α chain were analyzed for their amino acid content (Tables IV, V, VI). Four to six maps of each chain were prepared and sprayed very lightly with a 0.05% ninhydrin solution (see experimental). As soon as the spots became barely visible, they were cut out and stored in glass vials at $-20^{0}C$ until needed. In this way, minimal destruction of the N - terminal amino acid of each peptide was achieved.

Tryptic peptides from the β chains were also analyzed for amino acid content. Nine 3β tryptic peptides, ten 4β tryptic peptides and seven 5β tryptic peptides were eluted and hydrolyzed (Tables VII, VIII, IX).

Elution of Peptides

At first, the peptides were eluted from the chromatography paper with 6N HCl and then hydrolyzed for 24 hours. Later, it was found that better results were obtained if peptides were eluted with distilled water, lyophilized and then hydrolyzed with 6N HCl. Eluting the peptides with 6N HCl sometimes causes destruction of the amino acids in the peptides. This was found to be the case with several peptides - 4α11, 5α13, 5αb, 5β11, 5β10, 5β9.

TABLE IV AMINO ACID COMPOSITION OF THE 3α CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	3α15	3α14	3α13	3α12	3α11	3α10	3α9	3α8	3α7	3α6	3α5
Lysine	1.00	1.00	1.00	0.02	0.37	1.00	1.00	1.00	1.00	1.00	1.00
Histidine		1.17	0.23		0.12	0.15		0.25	0.87	1.03	0.05
Arginine				1.00	1.00	0.07			0.15		
Aspartic	0.02		0.66	0.05	0.13	1.04	0.85	0.06	1.11	1.18	1.71
Threonine			0.52		0.03	0.72	0.07	1.72	0.93	0.85	0.14
Serine			1.32		0.18	0.43	0.20		2.06	2.63	0.23
Glutamic	0.14		2.27	0.22	0.38	0.42	0.27	0.39	1.18	1.22	0.09
Proline			1.90					0.61	0.95	0.65	0.55
Glycine	0.12	1.11	1.42	0.19	0.34	0.58	0.29	0.32	1.63	1.01	0.23
Alanine	0.04	0.78			0.35	0.23	1.86	1.82	1.11	1.02	2.48
Cysteine											
Valine			0.41		0.17	0.12	0.62		1.80	2.53	1.51
Methionine								0.83			
Isoleucine			0.39			0.92	0.20		0.94	1.29	0.24
Leucine	0.02		0.34	0.77	0.01	0.04	0.60		0.19		0.93
Tyrosine									0.61	0.71	
Phenylalanine			0.15			0.06		1.77	1.16	1.29	0.64

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE V AMINO ACID COMPOSITION OF THE 4α CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	4α15	4α14	4α13	4α12	4α11[b]	4α10	4α9	4α8	4α7	4α6	4α2	4α1[b]
Lysine	1.00	1.00	1.00	0.38	0.07	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Histidine		1.08	0.17		0.02	0.10		0.14	0.95		0.20	+++
Arginine				1.00	1.00	0.04						
Aspartic	0.04	0.03	0.41	0.23	0.09	1.06	0.76	0.32	1.20	1.68	2.20	0.64
Threonine	0.04		0.70	0.24	0.03	0.76	0.07	1.77	0.82	0.10	0.14	3.20
Serine	0.13	0.05	1.37	0.62	0.09	0.47	0.16	1.04	2.59	0.22	1.18	1.50
Glutamic	0.27	0.14	0.87	0.52	0.11	0.41	0.11	0.73	1.11	0.19	0.32	1.01
Proline								1.09	0.88	0.39		1.94
Glycine	0.04	1.14	0.93	0.49	0.18	0.49	0.34	1.72	1.25	0.31	0.47	1.34
Alanine	0.14	0.87	0.39	0.27	0.10	0.23	1.74	2.78	1.06	2.29	1.29	3.76
Cysteine												
Valine							0.57		1.80	1.52	0.96	0.47
Methionine								1.05				1.09
Isoleucine						0.92			1.00			0.45
Leucine	0.04			0.96	0.04	0.08	0.60	0.30	0.26	0.81	1.27	0.84
Tyrosine					0.52				0.49			0.30
Phenylalanine								1.58	0.91	0.54	0.07	2.76

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

[b] Peptide from a 4α chain that was isolated from an individual rat specimen.

+++ Indicates amino acid is present but is unresolved.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE VI AMINO ACID COMPOSITION OF THE 5α CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	5α15	5α14	5α12	5α11	5α10	5α9	5α8	5α7	5α6	5α3	5αa
Lysine	1.00	1.00	0.11	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Histidine	0.07	1.00				0.05	0.93	0.22		+++	0.13
Arginine			1.00								
Aspartic	0.14	0.11			0.80	0.15	1.34	1.57	1.59	2.24	0.83
Threonine	0.08	0.05				2.03	0.78	0.50	0.02	1.52	0.33
Serine	0.37	0.27	0.13		0.09	0.47	2.70	1.14	0.12	4.37	0.10
Glutamic	0.28	0.15	0.03		0.01	0.29	1.09	0.49	0.06	4.15	
Proline						1.15	0.92		0.43	1.03	
Glycine	0.48	1.33	0.08		0.29	0.50	1.21	0.76	0.18	4.80	0.90
Alanine	0.14	0.94	0.09		1.69	2.17	1.08	0.75	2.38	2.92	0.60
Cysteine											
Valine					0.55		1.94	0.47	1.42	1.49	
Methionine						0.69					
Isoleucine		0.07					0.95	0.54		1.19	
Leucine		0.08			0.56	0.08	0.35	0.63	0.96	1.98	0.26
Tyrosine			0.46				0.58			0.44	
Phenylalanine						1.85	0.91	0.13	0.45	0.80	

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

+++ Indicates amino acid is present but is unresolved.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE VII AMINO ACID COMPOSITION OF THE 3β CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	$3\beta 16$	$3\beta 15$	$3\beta 14$	$3\beta 13$	$3\beta 12$	$3\beta 11$	$3\beta 10$	$3\beta 9$	$3\beta 2$
Lysine	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Histidine	0.15	0.77	1.19		0.19	2.40		0.29	0.04
Arginine	0.04	0.19		0.10	0.14	0.18			1.00
Aspartic	0.23	0.41	0.04	0.09	0.51	0.31	0.87	1.79	0.68
Threonine	0.20	0.27	0.04	0.09	0.45	0.17	0.36	0.14	0.08
Serine	0.54	0.81	0.09	0.02	0.58	0.67	0.41	0.60	
Glutamic	0.38	0.61	0.06	0.11	1.40	0.50	0.48	0.20	0.47
Proline									0.57
Glycine	0.46	0.93	1.15		1.01	0.73	1.98	0.97	2.99
Alanine	0.20	0.50	0.78	0.09	0.67	3.18	1.52	0.95	0.96
Cysteine									
Valine	0.08			1.14	0.51	2.02	1.11	0.66	2.21
Methionine									
Isoleucine				0.05	0.29	0.26		0.71	
Leucine	0.14	0.13		0.06	0.60	0.93	1.11	0.63	1.06
Tyrosine						0.75		0.08	
Phenylalanine				0.09	0.36	0.53	0.16	0.71	

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE VIII AMINO ACID COMPOSITION OF THE 4β CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	4β16	4β15	4β14	4β13	4β12	4β11	4β10	4β 9	4β 8	4β 6
Lysine	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Histidine	0.33	0.06	1.00	1.10		0.19	0.79	0.96	1.18	
Arginine	1.00	0.15								
Aspartic	3.50	0.29	1.00	0.12		0.21	0.22	0.93	1.76	0.81
Threonine		0.24	0.47							0.39
Serine	0.27	0.58	1.20	0.22		0.53	1.39	0.98	1.51	1.30
Glutamic	1.89	0.89	1.10	0.32	0.08	0.86	1.12	0.34	0.94	1.34
Proline	1.41									
Glycine	3.10	0.70	0.34	1.10		0.91	2.01	3.30	3.20	0.85
Alanine	1.59	0.39	1.03	0.81	0.18	0.32	3.20	1.64	1.33	1.47
Cysteine										
Valine	2.02				1.20		1.34	0.68	0.78	2.00
Methionine										
Isoleucine									0.42	
Leucine	1.34	0.14	0.55			0.17	0.54	0.81	0.74	1.20
Tyrosine										
Phenylalanine							0.81		0.83	

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE IX AMINO ACID COMPOSITION OF THE 5β CHAIN TRYPTIC PEPTIDES[a]

Amino Acid	5β16	5β15	5β14	5β13	5β12	5β8	5β5
Lysine	0.08	1.00	1.00	1.00	1.00	1.00	1.00
Histidine			0.99	1.25		0.19	0.09
Arginine	1.00						
Aspartic	3.33	0.15	0.68	0.07		1.04	2.81
Threonine	0.09	0.13	0.61	0.09		0.23	
Serine	0.24	0.39	1.25	0.21		1.12	0.24
Glutamic	1.35	0.34	1.26	0.19		1.07	0.24
Proline	1.09						
Glycine	3.34	0.32	1.69	1.69		1.75	1.53
Alanine	1.22	0.20	1.57	0.97		1.54	1.58
Cysteine							
Valine	1.96		0.35		1.17	0.89	0.77
Methionine							
Isoleucine						0.32	0.69
Leucine	1.18	0.23	0.46	0.08		0.89	1.08
Tyrosine							
Phenylalanine							0.70

[a] Hydrolysis with 6N HCl at 110°C for 24 hours.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

TABLE X AMINO ACID COMPOSITION OF 4α CHAIN TRYPTIC PEPTIDES ISOLATED FROM AN INDIVIDUAL RAT SPECIMEN[a]

Amino Acid	4α14	4α13	4α12	4α11	4α 1
Lysine	1.00	1.00	0.09	0.07	1.00
Histidine	1.01	0.22		0.02	+++
Arginine			1.00	1.00	
Aspartic	0.05	0.41	0.25	0.09	0.64
Threonine	0.03	0.69	0.03	0.03	3.20
Serine	0.05	1.12	0.16	0.09	1.50
Glutamic	0.08	0.06	0.20	0.11	1.01
Proline					1.94
Glycine	1.19	1.03	0.26	0.18	1.34
Alanine	0.88	0.28	0.13	0.10	3.76
Cysteine					
Valine					0.47
Methionine					1.09
Isoleucine					0.45
Leucine	0.02	0.25	0.70	0.04	0.84
Tyrosine				0.52	0.30
Phenylalanine					2.76

[a] Hydrolysis wih 6N HCl at 110°C for 24 hours.

+++ Indicates amino acid is present but is unresolved.

The results are expressed in molar proportions of the amino acids to lysine or arginine.

In the case of the 4α11 peptide, it had first been eluted from the chromatography paper with acid and upon hydrolysis, it appeared to contain no amino acids. New maps of the 4α tryptic digest were made. The 4α11 peptide was then cut out, eluted with distilled water and then hydrolyzed with 6N HCl. This time it was found that the peptide did contain amino acids (Table X). As a test for reproducibility, several other 4α peptides were also re-analyzed for amino acid content (Table X). The composition of these peptides (peptides 4α14, 4α13, 4α12) was found to be the same as those 4α peptides reported in Table V.

In the following sections, amino acid composition comparisons of the tryptic peptides will be presented. Sequence alignments of the peptides will also be included using other known mammalian chains as models. "The Amino Acid Composition Comparisons" (Tables XI - XXXII) contain the values for the overall content of the peptides in question as presented originally in Tables IV - IX. These values are being presented in this fashion so that comparisons of amino acid content can be made easier. An amino acid was judged to be a part of a peptide if it had a molar proportion value of 0.40 or greater. Values of 0.35 were allowed for the amino acids valine, isoleucine and tyrosine due to poor recovery after hydrolysis.

Sequence Homology

Once the total amino acid composition of a peptide had been determined, its sequence was asigned by hom-

ology. For many of the peptides sequencing by homology was straightforward. The amino acid composition of these peptides were similar to peptides found in other species. They lent themselves to unambiguous sequence alignments. However, a few peptides contained amino acids quite unlike those found by homology in other species. Decisions as to their place along the sequence had to be made arbitrarily. Some of the guidelines used to help in such cases came from observing amino acid exchanges in other species of homologous peptides. In general, threonine and serine could be used interchangeably as could alanine and glycine or aspartic acid and glutamic acid. Sometimes alanine and glycine are seen as amino acid exchanges with serine and threonine. Histidine and glutamine, proline and glutamine, proline and histidine, proline and threonine are other sets of amino acid exchanges that commonly occurred. Some residue positions along the α and β chains show that they can be occupied by a variety of different amino acids and so assignment of an amino acid to that position in the rat chain was again arbitrary. It should be noted that in the sequence of the rat α and β chains, glutamic acid and aspartic acid residues are written as "GLX" and "ASX", respectively. This indicates that no distinction between the acid or the amide form of these amino acids has been made.

Once the total amino acid composition of a peptide has been determined, its sequence was assigned by homology. Residues placed by homology (31) are shown by

bracketing ( ) the amino acids. Periods between amino acids indicate that the amino acid to the left could be positioned with confidence by homology. Hyphens between amino acid residues indicate that the amino acid to the left has been positively identified. In addition, doubled amino acids (i.e., $\begin{matrix}\text{ser}\\\text{thr}\end{matrix}$ or $\begin{matrix}\text{ala}\\\text{gly}\end{matrix}$) indicate that either residue can be placed at this point in the sequence by homology. Blanks at positions along the sequence indicate that a residue could not be identified. Some amino acids appeared to be part of the peptide composition but, after sequencing, could not be placed by homology. These "extra amino acids" were probably present in the chromatography paper as background amino acid contamination and were eluted along with the peptide being studied. After hydrolysis, these amino acids would freely mix with those released from the peptide, increasing the relative amounts of amino acids detected. Amino acid analysis of arbitrary segments of chromatography paper that have undergone two-dimensional chromatography and ninhydrin spraying and that have never given a positive ninhydrin reaction revealed a constant background presence of such amino acids as serine, glutamic acid, glycine, alanine and sometimes even leucine. These "extra amino acids" were probably not picked up in the ampoules used for acid hydrolysis due to the fact that they were presoaked in a mixture of nitric and sulfuric acids before use (see experimental).

Human, dog, mouse strain C57Bl and rabbit chain sequences are from Dayhoff, 1972 (31). Mouse strain BALB/c

sequence is from Popp (12,13). Rat$^{I}\alpha$ and Rat$^{II}\alpha$ sequences are those assigned to the major α and minor α chains, respectively, of the rat by Ranney, et al (6). Rat$^{I}\alpha$* is the sequence assigned to the major α chain of the rat by Chua, et al (8).

<u>Sequence Homology of the α Chains - Residues 57 - 60</u>

Generally, tryptic peptides believed to be common to all three α chains by map location (Table IIIa) revealed that they also contained the same amino acid compositions. Common peptides 3α14, 4α14, 5α14 (Figure 16, Table XI) all contain the same four amino acids - lysine, histidine, glycine, and alanine. Sequencing by homology places these amino acids in positions 57 - 60 on the α chain.

Species	57 58 59 60
Human α	Gly-Hist-Gly-Lys
Dog α	Ala-Hist-Gly-Lys
Mouse (C57BL) α	Gly-Hist-Gly-Lys
Rabbit α	Ala-Hist-Gly-Lys
Rat$^{I}\alpha$	Ala-Hist-Gly-Lys
Rat$^{II}\alpha$	Ala-Hist-Gly-Lys
Rat$^{I}\alpha$ *	Ala-Hist-Gly-Lys
3α14	(Ala Hist Gly)Lys
4α14	(Ala Hist Gly)Lys
5α14	(Ala Hist Gly)Lys

<u>Sequence Homology of the α Chains - Residues 140 - 141</u>

In Table XII, common peptides 4α11 and 5α12 were found to contain the amino acids arginine and tyrosine. A composition study of peptide 3α11 showed that it

contained only arginine. Even though the amino acid tyrosine cannot be detected after acid hydrolysis (due to its probable destruction by the ninhydrin reagent and/or by acid hydrolysis), its presence is indicated by the Pauly reagent (Figure 16). Then by homology, the amino acids in these peptides - 3α11, 4α11, 5α12 - represent the last two amino acids (positions 140 and 141) in the α chain sequence.

Species	140 141
Human α	Tyr-Arg
Dog α	Tyr-Arg
Mouse (C57BL) α	Tyr-Arg
Rabbit α I	Tyr-Arg
Rat α II	Tyr-Arg
Rat α I	Tyr-Arg
Rat α *	Tyr-Arg
3α11	(Tyr)Arg
4α11	(Tyr)Arg
5α12	(Tyr)Arg

Sequence Homology of the α Chains - Residues 62 - 68

Residues 62-68 on the α chain sequence were located in peptides 3α9, 4α9 and 5α10 (Table XIII). All three peptides after acid hydrolysis were found to contain the same six amino acids : Lys, 1-2 Ala, Asx, Val, Leu. Homology studies indicate that there should be seven amino acids in these peptides. It appears that one of the amino acids was destroyed during acid hydrolysis. Although position 63 on the chains is left blank (see next page), it

is just as possible for residues in positions 65 or 67 to be the missing ones.

Species	62	63	64	65	66	67	68
Human α	Val-	Ala-	Asp-	Ala-	Leu-	Thr-	Asn-
Dog α	Val-	Ala-	Asp-	Ala-	Leu-	Thr-	Thr-
Mouse (C57BL)α	Val-	Ala-	Asp-	Ala-	Leu-	Ala-	Asn-
Rabbit α	Val-	Ser-	Gln-	Ala-	Leu-	Thr-	Lys
Rat α I	Val-	Ala-	Asx-	Ala-	Leu-	Ala-	Lys
Rat α II	Val-	Ala-	Asp-	Ala-	Leu-	Ala-	Lys
Rat α I	Val-	Ala-	Asp-	Ala-	Leu-	Ala-	Lys
3α 9	(Val		Asx	Ala	Leu	Ala)	Lys
4α 9	(Val		Asx	Ala	Leu	Ala)	Lys
5α10	(Val		Asx	Ala	Leu	Ala)	Lys

Sequence Homology of the α Chains - Residues 93-99 or 128-139

Another set of peptides which demonstrate similar chromatographic properties and therefore contain the same amino acid composition are peptides 3α5, 4α6, 5α6 (Table XIV). These peptides were found to contain the same ten amino acids : Lys, 1-2 Asp, Pro, 2 Ala, 1-2 Val, Leu, Phe. When trying to evaluate what part of the α chain sequence these peptides represented, two different sequence segments turned up as possible choices. The first sequence homology possible represents position 93-99. This sequence calls for seven amino acids (see next page). This would mean that at least three amino acids - two alanine and one leucine - would be in excess. From experience with other peptides, it is very unusual to find the amino acid leucine as an "extra amino acid".

Species	93 94 95 96 97 98 99
Human α	Val-Asp-Pro-Val-Asn-Phe-Lys
Dog α	Val-Asp-Pro-Val-Asn-Phe-Lys
Mouse (C57BL)α	Val-Asp-Pro-Val-Asn-Phe-Lys
Rabbit α	Val-Asp-Pro-Val-Asn-Phe-Lys
Rat α^{I}	Val-Asp-Pro-Val-Asn-Phe-Lys
Rat α^{II}	(Val-Asp Val-Asn-Phe)
Rat α^{I} *	Val-Asp-Pro-Val-Asn-Phe-Lys
3α5	(Val Asx Pro Val Asx Phe)Lys
4α6	(Val Asx Pro Val Asx Phe)Lys
5α6	(Val Asx Pro Val Asx Phe)Lys

The second sequence homology proposed for these peptides represents residues 128-139. This encompasses twelve amino acid residues leaving two amino acids to be accounted for in these peptides.

Species	128 129 130 131 132 133 134
Human α	Phe-Leu-Ala-Ser-Val-Ser-Thr-
Dog α^{I}	Phe-Phe-Ala-Ala-Val-Ser-Thr-
Dog α^{II}	Phe-Phe-Thr-Ala-Val-Ser-Thr-
Mouse (C57BL)α	Phe-Leu-Ala-Ser-Val-Ser-Thr-
Rabbit	Phe-Leu-Ala-Asn-Val-Ser-Thr-
Rat α^{I}	Phe-Leu-Ala-Ser-Val-Ser-Thr-
Rat α^{II}	(Phe.Leu.Ala.Ser.Val.Ser.Thr.
Rat α^{I} *	Phe(Leu-Gly-Asn-Met-Ser-Ala.
3α5	(Phe Leu Ala Asx Val Ala Pro
4α6	(Phe Leu Ala Asx Val Ala Pro
5α6	(Phe Leu Ala Asx Val Ala Pro

Continued

Species	135 136 137 138 139
Human α	-Val-Leu-Thr-Ser-Lys
Dog α^{I}	-Val-Leu-Thr-Ser-Lys
Dog α^{II}	-Val-Leu-Thr-Ser-Lys
Mouse (C57BL)α	-Val-Leu-Thr-Ser-Lys
Rabbitα	-Val-Leu-Thr-Ser-Lys
Rat α^{I}	-Val-Leu-Thr-Ser-Lys
Rat α^{II}	.Val.Leu.Thr.Ser)Lys
Rat α^{I} *	.Val.Leu)Thr-Ser-Lys
3 α 5	Val Asx)Lys
4 α 6	Val Asx)Lys
5 α 6	Val Asx)Lys

The amino acids that appear to be missing are serine and leucine. Serine destruction can occur during acid hydrolysis, accounting for its absence from these peptides (i.e. 3 α 5, 4 α 6, 5 α 6). It is proposed that position 135 is occupied by a valine and 136 should be a leucine. Since valine release during acid hydrolysis is slow, the release of a leucine residue coupled to the valine would also be slow. This effect of valine on neighboring amino acids is also reflected in the low proline value (Table XIV). Proline in this peptide is postulated to be at position 134 in this sequence.

In order to distinguish between these two possible sequences, other techniques of peptide investigation (i.e. sequential Edman degradation, N-terminal amino acid analysis, etc.) would need to be applied.

Sequence Homology of the α Chains - Residues 91-92

Peptides 3α12, 4α12 and 5α13 are listed as common peptides by map location (Table IIIa, Figure 16). On analysis, peptides 3α12 and 4α12 revealed two amino acids - arginine and leucine - while peptide 5α13 revealed no amino acids (Table XV). The amino acids in peptide 5α13 probably were destroyed during the elution step and subsequent acid hydrolysis. Peptides 3α12 and 4α12 correspond to position 91-92 on the α chain by homology.

Species	91 92
Human α	Leu-Arg
Dog α	Leu-Arg
Mouse (C57BL)α	Leu-Arg
Rabbit α I	Leu-Arg
Rat α II	Leu-Arg
Rat α I	Leu-Arg
Rat α *	Leu-Arg
3α12	(Leu)Arg
4α12	(Leu)Arg
5α13	

Sequence Homology of the α Chains - Residues 32-40

Tryptic peptides 3α8 and 4α8 have the same map locations (Figure 16), yet their total amino acid compositions appear to differ (Table XVI). Among the amino acids in peptide 4α8 are serine, glutamic acid, 1 to 2 glycines and 2 to 3 alanines. Of those amino acids in peptide 3α8, serine is not included, glutamic acid and glycine may or may not be present, and only 1 to 2 alanines are recorded. Sequencing these two peptides by homology places them at

position 32-40 - an eight amino acid segment. After sequencing, peptide 3α8 has two "extra amino acids" - glutamic and glycine - while peptide 4α8 has four to five "extra amino acids" - serine, glutamic, 1 to 2 glycines and an alanine. In both cases, this represents background that cannot be accounted for.

Species	32	33	34	35	36	37	38	39	40
Human α	Met-	Phe-	Leu-	Ser-	Phe-	Pro-	Thr-	Thr-	Lys
Dog α	Thr-	Phe-	Gln-	Ser-	Phe-	Pro-	Thr-	Thr-	Lys
Mouse (C57BL) α	Met-	Phe-	Ala-	Ser-	Phe-	Pro-	Thr-	Thr-	Lys
Rabbit α	Met-	Phe-	Leu-	Gly-	Phe-	Pro-	Thr-	Thr-	Lys
Rat α^{I}	Met-	Phe-	Ala-	Ala-	Phe-	Pro-	Thr-	Thr-	Lys
Rat α^{II}	Met-	Phe-	Ala-	Ala-	Phe-	Pro-	Thr-	Thr-	Lys
Rat α^{I} *	Met-	Phe-	Ala-	Ala-	Phe-	Pro-	Thr-	Thr-	Lys
3α8	(Met	Phe	Ala	Ala	Phe	Pro	Thr	Thr)	Lys
4α8	(Met	Phe	Glx Ala	Ser Ala	Phe	Pro	Thr	Thr)	Lys
5α9	(Met	Phe	Gly Ala	Ser Ala	Phe	Pro	Thr	Thr)	Lys

When peptide 5α9, which was thought to be a unique peptide to the 5α map (Table IIIa, Figure 16), was investigated, it unexpectedly matched peptides 3α8 and 4α8 after sequence homology. But, in overall amino acid composition it differed slightly from peptide 4α8 having no glutamic acid and only one glycine (Table XVI). It too had "extra amino acids" - serine and glycine - which could not be accounted for by homology.

<u>Sequence Homology of the α Chains - Residue 61</u>

Peptides 3α15, 4α15 and 5α15 appear in the

same positions on their respective maps (Table IIIa, Figure 16). In Table XVII it can be seen that 3α15 and 4α15 both contain only the amino acid lysine, while peptide 5α15 contains lysine and the amino acid glycine. Since glycine is a non-polar neutral amino acid, it would not necessarily result in a peptide with the composition (N) - glycine - lysine - (C) behaving differently on two-dimensional chromatography from a peptide containing just lysine. On the 3α and 4α chains these peptides represent residue 61 which, by homology, has always been a lysine.

Species	61
Human α	Lys
Dog α	Lys
Mouse (C57BL)α	Lys
Rabbit α	Lys
Rat α^{I}	Lys
Rat α^{II}	Lys
Rat α^{I} *	Lys
3α15	Lys
4α15	Lys
5α15	(gly)Lys

In the case of the 5α chain, peptide 5α15 might represent residues 60 and 61 ((N)-gly-lys-(C)) or 61 and 62 ((N)-lys-gly-(C)). From the homology results on page 73, residue 62 is reported to be a valine leaving the arrangement gly-lys (60-61). This means that somewhere between residues 1-59 in the 5α chain an amino acid could be missing from the sequence and its loss compensated for by this peptide. The results to be presented next will add

support to this suggestion.

Sequence Homology of the α Chains - Residues 1-11

Species	1	2	3	4	5	6	7
Human α	Val-	Leu-	Ser-	Pro-	Ala-	Asp-	Lys-
Dog α	Val-	Leu-	Ser-	Pro-	Ala-	Asp-	Lys-
Mouse (C57BL)α	Val-	Leu-	Ser-	Gly-	Glu-	Asp-	Lys-
Rabbit α	Val-	Leu-	Ser-	Pro-	Ala-	Asp-	Lys-
Rat$^{I}\alpha$	Val-	Leu-	Ser-	Ala-	Asp-	Asp-	Lys-
Rat$^{II}\alpha$	Val-	Leu-	Ser-	Ala-	Ala-	Asp-	Lys-
Rat$^{I}\alpha$ *	Val-	Leu-	Ser-	Ala-	Asp-	Asp-	Lys-
3 α							
4 α 2	(Val	Leu	Ser	Gly Ala	Asx	Asx)	Lys
5 α 7	(Val	Leu	Ser	Gly Ala	Glx Asx	Asx	
OR							
5 α 7	(Val	Leu	Ser	Gly Ala	Glx Asx	Asx	Ser Thr

Species	8	9	10	11
Human α	-Thr-	Asn-	Val-	Lys
Dog α	-Thr-	Asn-	Ile-	Lys
Mouse (C57BL)α	-Ser-	Asn-	Ile-	Lys
Rabbit α	-Thr-	Asn-	Ile-	Lys
Rat$^{I}\alpha$	-Thr-	Asn-	Ile-	Lys
Rat$^{II}\alpha$	-Thr-	Asn-	Ile-	Lys
Rat$^{I}\alpha$ *	-Thr-	Asn-	Ile-	Lys
3 α 10	(Ser Thr	Asx	Ile)	Lys
4 α 10	(Ser Thr	Asx	Ile)	Lys
5 α 7	Ser Thr	Asx	Ile)	Lys
OR				
5 α 7	Asx	Ile)	Lys	

The composition of peptide 4α2 (Table XVIII) appears to represent the first seven amino acids (from the N-terminal end). This peptide could not be located in the same position on the 3α and 5α peptide maps. Peptides 3α10, 4α10 and 5α11 all appear in the same general location on their respective maps (Figure 16). The composi-

tion of 3α10 and 4α10 is the same and represents positions 8-11 on the sequence by homology (Table XVIII and page 79). On the other hand, peptide 5α11 only contains the amino acid lysine (Table XVIII) and yet it has the same mobility as the 3α10 and 4α10 peptides (Figure 16). There are two explanations possible:

1) This peptide only contains lysine and because of hydration moves like a peptide containing (Thr Asn Ile)Lys.

2) This peptide also contained the amino acids thr, asn, and ile but during acid hydrolysis these amino acids were selectively destroyed.

The first explanation appears to be the more reasonable one. The number of umoles of lysine/ml released by the peptides 3α10, 4α10 and 5α11 are almost the same - 0.037, 0.033 and 0.022 umoles/ml respectively. Clearly, this indicates that lysine was released from all three peptides in the same amounts. Why then would the amino acids threonine, asparagine, and isoleucine, if present in 5α11, be totally degraded while the lysine is left untouched? It seems more likely that this peptide does not contain these amino acids - thr, asn and ile.

Peptide 5α7 (Table XVIII) contains ten amino acids. Six of these appear to match positions 1-6 on the α chain by homology with other known α chain sequences and with peptide 4α2 (page 79). The other four amino acids coincide with positions 8-11 by homology with peptides 3α10 and 4α10, as well as with other known α chain se-

quences (page 79). This presents four possible sequences for the first ten amino acids on the 5α chain:

	1	2	3	4	5	6	7	8	9	10	11
1)	Val	Leu	Ser	Ala	Asx	Asx	Lys	Thr	Asx	Ile	Lys
2)	Val	Leu	Ser	Ala	Asx	Asx	Lys	Thr	Asx	Ile	
3)	Val	Leu	Ser	Ala	Asx	Asx		Thr	Asx	Ile	Lys
4)	Val	Leu	Ser	Ala	Asx	Asx	Thr	Asx	Ile	Lys	

The favored alignments are 3) and 4) for the following reasons:

The first possibility can occur only if the lysine at position 7 has not been cleaved by trypsin. Since the amino terminal end is usually not buried inside the molecule (32) but readily available to trypsin attack and since on 3α and 4α chains, trypsin has cleaved at this location as indicated, this sequence is the least plausible. In addition, when determining the molar ratios of the amino acids in peptide $5\alpha 11$, if the concentration of alanine (an acid-stable amino acid) is considered to be equal to 1.0, then the value for lysine is 1.3. If lysine is present at position 7 then the molar ratio of lysine in this peptide should be almost 2.0.

The second possibility can be ruled out using an argument similar to the first. Trypsin has not cleaved at position 7 but at position 11 and at the amino side of the lysine instead of the carbonyl side. Since trypsin does not cleave at isoleucine and since chymotryptic activity has been quenched in the trypsin by TPCK treatment, cleavage on

the amino acid of lysine is ruled out; although it is possible that in this instance fragmentation of the peptide has occurred during elution and that the amino acid lysine at position 11 was lost.

This then leaves choices 3) and 4) as the best arrangements. In case 3), the amino acid at position 7 is probably not lysine but some other acid-labile amino acid (i.e. ser or thr). Choice 4) places the amino acids thr, asn and ile and lys into positions 7-10 instead of 8-11. This puts the remainder of the sequence of the 5α chain out of alignment. From the discussion on page 78 in connection with peptide 5α15, the appearance of an extra glycine in this peptide would compensate for the missing amino acid in peptide 5α7. This insertion allows the remainder of the 5α chain from residues 60-141 to remain in alignment.

Therefore, it can be seen that within the first eleven amino acids of the 3α, 4α and 5α chains there are two significant differences. The first being that the sequence of amino acids in position 1-6 in the 3α chain is probably not the same as that in the 4α and 5α chains. Second, the 5α chain has either i) an amino acid residue at position 7 that is not a lysine as it is in the 4α chain (and most probably is in the 3α chain) but some other one; or ii) its sequence is shifted over by one amino acid due to the loss of a residue that is compensated for at a later position.

It is interesting to note at this point that in

Table IIIa it was proposed that the position of peptide 5α7 matched that of peptide 3α6 (see also Figure 16). From amino acid composition studies, it was found that peptides 5α7 and 3α6 did not contain the same amino acids and in fact did not even represent the same portions on the α chain sequence (pages 79 and 84). This is one of the few instances where two peptides possessing the same mobilities in two-dimensional chromatography cannot be identical in amino acid content.

Sequence Homology of the α Chains - Residues 41-56

Sequence positions 41-56 are represented by peptides 3α7, 3α6, 4α7 and 5α8 (page 84). From map considerations, it was first thought that only 4α7 and 5α8 would be the same peptides (Table IIIa, Figure 16). However, it turns out that peptides 4α7, 5α8 as well as 3α6 all have the same amino acid compositions and therefore probably represent the same sequence (Table XIX). Peptide 3α7, which has a completely different mobility from the other three (Figure 16) has a similar amino acid composition as peptides 3α6, 4α7 and 5α8, with the difference that it contains one more glycine and one less serine than the other peptides (Table XIX). This amino acid exchange most likely takes place at position 44 in the 3α sequence (see page 84). In this situation, the 3α chain at position 44 is occupied by both a serine and a glycine residue. The number of umoles/ml of lysine released by 4α7 and 5α8 is 0.032 and 0.038 respectively. The values for 3α6 and

and 3α7 are 0.014 and 0.017 umoles lysine/ml respectively and when added together they equal o.o31 umoles lysine/ml - almost the same value as found for 4α7 and 5α8. This implies that 3α6 represents approximately 50% of the 3α chain sequence 41-56 and 3α7 represents the other 50%.

Species	41	42	43	44	45	46	47	48
Human α	Thr-	Tyr-	Phe-	Pro-	Hist-	Phe-	Asp-	Leu-
Dog α	Thr-	Tyr-	Phe-	Pro-	Hist-	Phe-	Asp-	Leu-
Mouse (C57BL)α	Thr-	Tyr-	Phe-	Pro-	Hist-	Phe-	Asp-	Val-
Rabbit$^{I}\alpha$	Thr-	Tyr-	Phe-	Pro-	Hist-	Phe-	Asp-	Phe-
Rabbit$^{II}\alpha$	Thr-	Tyr-	Phe-	Pro-	Hist-	Phe-	Asp-	Leu-
Rat$^{I}\alpha$	Thr-	Tyr-	Phe-	Asn-	Hist-	Ile-	Asp-	Val-
Rat$^{II}\alpha$	Thr-	Tyr-	Phe-	Asn-	Hist-	Ile-	Asp-	Val-
Rat $^{I}\alpha$ *	Thr-	Tyr-	Phe-	Ser-	Hist-	Ile-	Asp-	Val-
3 α 6	(Thr	Tyr	Phe	Ser	Hist	Ile	Asx	Val
3 α 7	(Thr	Tyr	Phe	Gly	Hist	Ile	Asx	Val
4 α 7	(Thr	Tyr	Phe	Ser	Hist	Ile	Asx	Val
5 α 8	(Thr	Tyr	Phe	Ser	Hist	Ile	Asx	Val

Species	49	50	51	52	53	54	55	56
Human α	-Ser-	Hist-	Gly-	Ser-	Ala-	Gln-	Val-	Lys
Dog α	-Ser-	Pro-	Gly-	Ser-	Ala-	Gln-	Val-	Lys
Mouse (C57BL)α	-Ser(	Hist.	Gly.	Ser.	Ala.	Gln.	Val)	Lys
Rabbit$^{I}\alpha$	-Thr-	Hist-	Gly-	Ser-	Glu-	Gln-	Ile-	Lys
Rabbit$^{II}\alpha$	-Ser-	Hist-	Gly-	Ser-	Glu-	Gln-	Ile-	Lys
Rat$^{I}\alpha$	-Ser-	Pro-	Gly-	Ser-	Ala-	Gln-	Val-	Lys
Rat$^{II}\alpha$	-Ser-	Pro-	Gly-	Ser-	Ala-	Gln-	Val-	Lys
Rat$^{I}\alpha$ *	-Ser-	Pro-	Gly-	Ser-	Ala-	Gln-	Val-	Lys
3 α 6	Ser	Pro	Gly	Ser	Ala	Glx	Val)	Lys
3 α 7	Ser	Pro	Gly	Ser	Ala	Glx	Val)	Lys
4 α 7	Ser	Pro	Gly	Ser	Ala	Glx	Val)	Lys
5 α 8	Ser	Pro	Gly	Ser	Ala	Glx	Val)	Lys

The ratio of glycine to serine in position 44 of the 3α chain is estimated to be 0.5 : 0.5. This peculiarity in the 3α chain is another indication that the 3α chain has a different sequence when compared with the 4α

and 5 α chains.

Sequence Homology of the α Chains - Residues 12-16

Species	12 13 14 15 16
Human α	Ala-Ala-Trp-Gly-Lys
Dog α	Ser-Thr-Trp-Asp-Lys
Mouse (C57BL)α	Ala-Ala-Trp-Gly-Lys
Rabbit α	Thr-Ala-Trp-Glu-Lys
Rat[I] α	Asn(Ala)Trp-Gly-Lys
Rat[II] α	Asn- Trp-Gly-Lys
Rat[I] α *	Asn-Cys-Trp-Gly-Lys
3α	
4α	
5α a	(Asx Ala Gly)Lys

Peptide 5αa (Table XX) containing four amino acids appears to match sequence positions 12-16 by homology. Position 14 in the sequence is indicated to be a tryptophan, although this particular peptide does not give a positive Ehrlich reaction nor can tryptophan be detected after acid hydrolysis. On the other hand, peptide 5α5 does give a positive Ehrlich reaction (Figure 16) indicating tryptophan. Since there is only one tryptophan per α chain (Table I), these two peptides (5αa and 5α5) probably represent the same positions on the chain (positions 12-16). However, because they display different mobilities, their amino acid compositions are probably different, reflecting an amino acid exchange at one of the positions. This would mean that there are two amino acids sharing one position on the 5 α chain sequence. Examples of this situation have been cited in other species - rabbit and mouse (31,32) - and is encountered in the 3 α chain as mentioned above.

The 5αa peptide has no counterpart in the 3 α and 4α maps (Figure 16). This possible arrangement in the 5 α chain would indicate that the 5 α chain definitely has a different amino acid composition and hence a different amino acid sequence from the 3 α and 4 α chains.

<u>Sequence Homology of the α Chains - Residues 100-127</u>

Species	100	101	102	103	104	105	106	107	108
Human α	Leu-	Leu-	Ser-	Hist-	Cys-	Leu-	Leu-	Val-	Thr-
Dog α	Leu-	Leu-	Ser-	Hist-	Cys-	Leu-	Leu-	Val-	Thr-
Mouse (C57BL)α	(Leu.	Leu.	Ser.	Hist.	Cys.	Leu.	Leu)	Val-	Thr-
Rabbit α	Leu-	Leu-	Ser-	Hist-	Cys-	Leu-	Leu-	Val-	Thr-
Rat^{I}α	Leu-	Leu-	Ser-	Hist-	Cys-	Leu-			
Rat^{II}α							Leu-	Val	(Thr.
Rat^{I}α *	Phe-	Leu-	Ser-	Hist-	Cys-	Leu-	Leu-	Val-	Thr-
3 α									
4 α 1	(Phe	Leu	Ser					Val	Thr
5 α 3	(Phe	Leu	Ser			Leu		Val	Thr

Species	109	110	111	112	113	114	115	116	117
Human α	-Leu-	Ala-	Ala-	Hist-	Leu-	Pro-	Ala-	Glu-	Phe-
Dog α	-Leu-	Ala-	Cys-	Hist-	Hist-	Pro-	Thr-	Gly-	Phe-
Mouse(C57BL)α	-Leu-	Ala-	Ser-	Hist	(Hist-	Pro.	Ala-	Asp.	Phe)
Rabbitα	-Leu-	Ala-	Asn-	Hist-	Val-	Pro-	Ser-	Glu-	Phe-
Rat^{I}α									
Rat^{II}α	.Leu.	Ala.	Asx)	Hist	(Gly.	Pro.	Ser.	Asx.	Phe.
Rat^{I} α *	-Leu-	Ala-	Cys-	Hist-	Hist-	Pro-	Gly-	Asp-	Phe-
3 α									
4 α 1		Ala	Thr			Pro	Gly	Glx	Phe
5 α 3		Ala	Thr			Pro	Gly	Glx	Tyr

Species	118	119	120	121	122	123	124	125	126	127
Human α	-Thr-	Pro-	Ala-	Val-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Dog α	-Thr-	Pro-	Ala-	Val-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Mouse α	Thr-	Pro-	Ala-	Val-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Rabbitα	-Thr-	Pro-	Ala-	Val-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Rat^{I} α				Met-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Rat^{II}α	.Thr)	Pro-	Ala-	Met-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
Rat^{I}α *	-Thr-	Pro-	Ala-	Met-	Hist-	Ala-	Ser-	Leu-	Asp-	Lys
3 α										
4 α 1	Thr	Pro	Ala	Met		Ala	Ser	Ile	Asx)	Lys
5 α 3	Thr		Ala	Val		Ala	Ser	Ile	Asx)	Lys

Peptides 4α1 and 5α3 do not have a peptide counterpart on the 3α map (Figure 16). The amino acid composition of these peptides indicates that they contain almost the same amino acids (Table XXI) and that they represent the same positions along their respective α chain sequences - amino acid residues 100-127 . These two peptides represent one of the longest segments found in the α chain - 28 amino acids. Because these peptides were so long, fragmentation may have occurred during elution from the chromatography paper. After acid hydrolysis, not all of the amino acids in these peptides were recovered. Further, the four histidines predicted for this sequence segment did not resolve properly during amino acid analysis and therefore could not be determined accurately. Due to the poor results obtained with peptides 4α1 and 5α3, no comparison between them with respect to amino acid content or possible amino acid sequence can be made at this time. It should be noted that the absence of this peptide from the 3α map does not necessarily mean that it is not present in the 3α chain. This particular peptide in the 3α chain may not be soluble after trypsin digestion and remains in the insoluble "core". In the 4α and 5α chains, this peptide is soluble and can be recovered after two-dimensional chromatography. In order for the equivalent peptide in the 3α chain to be recovered, the "tryptic core" might have to be redigested with chymotrypsin to reduce the peptide to smaller soluble fragments for recovery on two-dimensional chromatograms.

Amino Acid Composition Comparison - Peptides 3α13 and 4α13

The last set of α tryptic peptides to be discussed are the 3α13 and 4α13 peptides. They have the same map locations on their respective maps (Figure 16) and should therefore have the same amino acid composition. The amino acid content of 3α13 and 4α13 (Table XXII) did not match. Peptide 4α13 was one of those 4α peptides that was re-isolated (Table X). Its amino acid content was found to be exactly the same in both cases. In addition, neither 3α13 nor 4α13 could be placed by homology into the α chain sequence. Evidently these two peptides are pieces of larger peptides that have become fragmented during chromatographic isolation.

α Chain Summary

The results presented here reveal the following differences in the α chains:

1) Tryptic peptide 4α2 representing the first seven amino acids on the 4α chain has no peptide counterpart in the same locations on the 3α and 5α maps (Figure 16, Table XVIII, page 79).

2) Peptide 3α10 and 4α10 represent amino acids 8-11 on their respective chains. No peptide counterpart has been located on the 5α tryptic map (Table XVIII, page 79).

3) Peptide 5α7 appears to represent the first ten amino acids on the 5α chain sequence (Table XVIII, page 79). The first six amino acids may match those on the 4α chain. The remaining four amino acids of this peptide cannot be arranged.

arranged like those of the 3α and 4α chains (page 79).

4) In peptide 3α7, position 44 appears to be occupied by a glycine, while in 4α7 and 5α8 position 44 is a serine (page 83-84). Peptide 3α6 shows a serine also to be present in position 44 on the 3α chain. This means that the 3α chain has two possible amino acids occupying position 44.

5) Peptide 5α15 contains a glycine as well as a lysine, while its peptide counterparts in the 3α and 4α maps only contain lysine (Table XVII).

6) Peptide 5αa (Table XX), having no counterparts in the 3α and 4α maps, may represent residues 12-16 (page 85) along with peptide 5α5. Because of differing mobilities, 5αa and 5α5 necessarily have different amino acid contents.

From this data, the conclusion can be confirmed that in the rat the 3α , 4α and 5α globin chains differ from one another in their amino acid compositions and in their amino acid sequences.

Sequence Homology of the β Chains - Residues 60-61

Turning to the β chains, an examination of common peptides 3β13, 4β12 and 5β12 (Figure 17) revealed that all three contained only two amino acids - valine and lysine - (Table XXIII). This corresponds to positions 60-61 on the β chain sequence.

Species	60 61
Human γ	Val-Lys
Dog γ	Val-Lys
Mouse (C57BL)γ	Val-Lys
Rabbit γ	Val-Lys
3γ13	(Val)Lys
4γ12	(Val)Lys
5γ12	(Val)Lys

Sequence Homology of the γ Chains - Residues 62-65

Species	62 63 64 65
Human γ	Ala-Hist-Gly-Lys
Dog γ	Ala-Hist-Gly-Lys
Mouse (C57BL)γ	Ala-Hist-Gly-Lys
Rabbit γ	Ala-Hist-Gly-Lys
3γ14	(Ala Hist Gly)Lys
4γ13	(Ala Hist Gly)Lys
5γ13	(Ala Hist Gly)Lys
OR	
5γ13	(Ala Hist Gly Gly)Lys

Tryptic peptides 3γ14, 4γ13 and 5γ13, thought to be the same peptides by map location (Table IIIb, Figure 17), after amino acid analysis revealed a slight difference in the 5γ peptide. From the data in Table XXIV, peptides 3γ14 and 4γ13 contain four amino acids - lysine, histidine, glycine and alanine - which corresponds to positions 62-65 by homology. Peptide 5γ13, on the other hand, appears to have two glycines instead of one (Table XXIV). If this extra glycine is indeed part of this peptide and is not due to background contamination, then the sequence for this peptide (5γ13) becomes (Ala Hist Gly Gly)Lys. There is a precedent set for this type of unusual arrangement in the Llama and Grey Kangaroo (31).

Sequence Homology of the γ^{β} Chains - Residue 66

Species	66
Human γ^{β}	Lys
Dog γ^{β}	Lys
Mouse (C57BL) γ^{β}	Lys
Rabbit γ^{β}	Lys
3 γ^{β} 16	Lys
4 γ^{β} 15	Lys
5 γ^{β} 15	Lys

Another unusual set of common peptides involves tryptic peptides 3 γ^{β}16, 4 γ^{β}15 and 5 γ^{β}15 (Figure 17). The amino acid composition of peptides 3 γ^{β}16 and 5 γ^{β}15 is lysine, serine, glutamic acid and glycine (Table XXV). Peptide 4 γ^{β}15 has the same four amino acids plus possibly an added alanine. When attempts were made to place these three peptides by homology, no sequence in the γ^{β} chain appeared to fit. It was soon realized that these three γ^{β} peptides had the same mobilities as the peptides 3 α15, 4 α15 and 5 α15 (Figure 16) and that they contained only the amino acid lysine (Table XVII). (It is also possible that 5 α15 contains a glycine as well.) It could follow then that γ^{β} peptides 3 γ^{β}16, 4 γ^{β}15 and 5 γ^{β}15 represent position <u>66</u> on the γ^{β} chain containing only lysine (see above). These analyses should be repeated in order to resolve the problem of the "extra amino acids" found with these three peptides.

Sequence Homology of the γ^{β} Chains - Residues 1-8

Species	1	2	3	4	5	6	7	8
Human γ^{β}	Val-	Hist-	Leu-	Thr-	Pro-	Glu-	Glu-	Lys
Dog γ^{β}	Val-	Hist-	Leu-	Thr-	Ala-	Glu-	Glu-	Lys
Mouse (C57BL) γ^{β}	Val-	Hist-	Leu-	Thr-	Asp-	Ala-	Glu-	Lys
Rabbit γ^{β}	Val-	Hist-	Leu-	Ser-	Ser-	Glu-	Glu-	Lys
3 γ^{β}								
4 γ^{β} 9	(Val	Hist	Leu	Ser	Ala	Ala	Asx)	Lys
5 γ^{β}								

In Table XXVI, the amino acid composition of the 4 γ^{β} 9 peptide represents the first eight amino acids from the N-terminal end of the 4 γ^{β} chain, by homology (see above). Its peptide counterparts on the 3 γ^{β} and 5 γ^{β} maps could not be located in the same position nor did they happen to be one of those peptides from the 3 γ^{β} and 5 γ^{β} maps analyzed for amino acids. Therefore, it can be assumed that the sequence of amino acid residues 1-8 on the 3 γ^{β} and 5 γ^{β} chains is different from that of the 4 γ^{β} chain.

Sequence Homology of the γ^{β} Chains - Residues 9-17

Another indication that the 4 γ^{β} chain is different from 3 γ^{β} and 5 γ^{β} can be seen in Table XXVII. From map comparisons (Table IIIb, Figure 17) peptide 3 γ^{β} 10 and 5 γ^{β} 8 appear to be the same peptides. There is no peptide in this position in the 4 γ^{β} map. However, peptide 4 γ^{β} 6 is the counterpart to these two peptides in the 4 γ^{β} system. Among its approximately ten amino acids, peptide 4 γ^{β} 6 contains two valines and one glycine while peptides 3 γ^{β} 10 and 5 γ^{β} 8 contain one valine and two glycines. Peptides 3 γ^{β} 10, 4 γ^{β} 6 and 5 γ^{β} 8 represent positions 9-17 in the γ^{β} sequence by homology.

Species	9	10	11	12	13	14	15	16	17
Human β	Ser-	Ala-	Val-	Thr-	Ala-	Leu-	Trp-	Gly-	Lys
Dog β	Ser-	Leu-	Val-	Ser-	Gly-	Leu-	Trp-	Gly-	Lys
Mouse (C57BL)β	Ala-	Ala-	Val-	Ser-	Gly-	Leu-	Trp-	Gly-	Lys
Mouse (BALB/c)β	Ser-	Ala-	Val-	Ser-	Cys-	Leu-	Trp-	Ala-	Lys
Rabbit β	Ser-	Ala-	Val-	Thr-	Ala-	Leu-	Trp-	Gly-	Lys
3β10	(Ser Ala	Ala	Val	Asx	Gly	Leu	Trp	Gly)	Lys
4β6	(Ser Ala	Ala	Val	Ser Asx	Val	Leu		Gly)	Lys
5β8	(Ser Ala	Ala	Val	Ser Asx	Gly	Leu	Trp	Gly)	Lys

In contrast to position 16 which is almost always occupied by a glycine in other species, position 13 demonstrates more variability with respect to the amino acid residue that can be placed there. For this reason the glycine-valine exchange is indicated to be at residue position 13 rather than position 16. Position 15 in 3β and 5β is assigned tryptophan by homology and by the fact that these two peptides (peptides 3β10 and 5β8) give a positive Ehrlich reaction (Figure 17). It is still possible for position 15 on the 4β chain to contain tryptophan. Only one peptide on the 4β map has given a positive Ehrlich reaction (peptide 4β5, Figure 17) and since there are two tryptophans in the 4β chain (Table II), from homology considerations alone the 4β6 peptide can be assigned a tryptophan at position 15.

In these three peptides, there appears to be an excess of the amino acids serine and glutamic acids. Further studies into the actual sequence of these peptides need to be performed in order to determine whether the

"extra amino acids" are in some way an actual part of the chain sequence or just background contamination.

Sequence Homology of the β Chains - Residues 18-30

Species	18	19	20	21	22	23	24	25	26	27	28	29	30
Human β	Val-	Asn-	Val-	Asp-	Glu-	Val-	Gly-	Gly-	-Glu-	Ala-	Leu-	Gly-	Arg
Dog β	Val-	Asn-	Val-	Asp-	Glu-	Val-	Gly-	Gly-	-Glu-	Ala-	Leu-	Gly-	Arg
Mouse (C57BL)β	Val-	Asn-	Ala-	Asp-	Glx-	Val-	Gly-	Gly-	-Glu-	Ala-	Leu-	Gly-	Arg
Mouse (BALB/c)β	Val-	Asn-	Pro-	Asp-	Glx-	Val-	Gly-	Gly-	-Glu-	Ala-	Leu-	Gly-	Arg
Rabbit β	Val-	Asn-	Val-	Glu-	Glu-	Val-	Gly-	Gly-	-Glu-	Ala-	Leu-	Gly-	Arg
3β 2	(Val	Asx	Pro			Val	Gly	Gly	Glx	Ala	Leu	Gly)	Arg
4β16	(Val	Asx	Pro	Asx	Glx Asx	Val	Gly	Gly	Glx	Ala	Leu	Gly)	Arg
5β16	(Val	Asx	Pro	Asx	Asx	Val	Gly	Gly	Glx	Ala	Leu	Gly)	Arg

The 3β chain can be shown to be different from the 4β and 5β chains from the following results. In Figure 17, it can be seen that peptides 4β16 and 5β16 have the same map positions. No peptide on the 3β map is seen in that position. In Table XXVIII, an amino acid composition comparison of peptides 4β16 and 5β16 reveals that they are almost identical in content and constitute residues 18-30 (see above). Except for the fact that 4β16 contains an extra glutamic acid residue and perhaps some aspartic acid and alanine as well, it and 5β16 correspond exactly in sequence. Also it was discovered that peptide 3β2 was the counterpart to these two peptides in the 3β

system. As can be seen from Figure 17, the map position of $3\beta 2$ is clearly different from that of $4\beta 16$ and $5\beta 16$. While all three peptides appear to have the same relative polarities, after high voltage electrophoresis, peptide $3\beta 2$ has remained at the origin position while $4\beta 16$ and $5\beta 16$ have migrated toward the anode. This indicates that peptides $4\beta 16$ and $5\beta 16$ have an overall negative net charge while peptide $3\beta 2$ is neutral. This is in agrrement with the findings reported in Table XXVIII. It can be seen that peptide $3\beta 2$ contains one aspartic residue while $4\beta 16$ and $5\beta 16$ both contain three to four aspartic acid residues. By homology, positions <u>21</u> and <u>22</u> are assigned the amino acid aspartic acid in 4β and 5β . Even though $3\beta 2$ lacks the necessary aspartic acids, no increase in other amino acids is seen to compensate for their loss. It is possible that the $3\beta 2$ peptide has two amino acids at positions <u>21</u> and <u>22</u> that are destroyed during peptide isolation or acid hydrolysis. It could also be that this peptide lacks two amino acids. This latter situation would put the 3β chain out of alignment and therefore is least favored. In any case, the map position of $3\beta 2$ is clearly different from that of $4\beta 16$ and $5\beta 16$ and definitely points to an alteration in the 3β chain sequence.

<u>Sequence Homology of the β Chains - Residues 67-82</u>

From map positioning (Table IIIb, Figure 17), it was thought that peptide $4\beta 8$ and $5\beta 9$ were common and therefore, would have the same amino acid content. Upon

analysis, 5β9 was found to contain no amino acids while 4β8 did (Table VIII). Similarly, peptide 4β6 and 5β5 (Table IIIb, Figure 17) were thought at first to be common peptides while it was discovered, after amino acid analysis, that they were different (Tables VIII and IX). It was found that 4β8 and 5β5, along with peptide 3β9, had much in common with each other. Table XXIX reveals that peptides 3β9 and 5β5 represent sequence positions 67-76, while peptide 4β8 may represent residues 67-82 (see below).

Species	67	68	69	70	71	72	73	74
Human β	Val-	Leu-	Gly-	Ala-	Phe-	Ser-	Asp-	Gly-
Dog β	Val-	Leu-	Asn-	Ser-	Phe-	Ser-	Asp-	Gly-
Mouse (C57BL)β	Val-	Ile	Thr-	Ala-	Phe-	Asp/Ser-	Ser/Asp-	Gly-
Mouse (BALB/c)β	Val-	Ile-	Thr-	Ala-	Phe-	Asp/Glu-	Ser-	Gly-
Rabbit β	Val-	Leu-	Ala-	Ala-	Phe-	Ser-	Glu-	Gly-
3β9	(Val	Ile	Asx	Ala	Phe	Ser	Asx	Gly
4β8	(Val	Ile	Asx	Ala	Phe	Ser	Asx	Gly
5β5	(Val	Ile	Asx	Ala	Phe	Asx	Asx	Gly

Species	75	76	77	78	79	80	81	82
Human β	-Leu-	Ala-	Hist-	Leu-	Asp-	Asp-	Leu-	Lys
Dog β	-Leu-	Lys	-Asn	-Leu-	Asp-	Asn-	Leu-	Lys
Mouse (C57BL)β	-Leu	(Asn	.Hist	.Leu	.Asp	.Asn)	Leu-	Lys
Mouse (BALB/c)β	-Leu	(Lys/Hist	.Asn	.Leu	.Asp	.Asn)	Leu-	Lys
Rabbit β	-Leu-	Ser	-Hist-	Leu-	Asp-	Asn-	Leu-	Lys
3β9	Leu)	Lys						
4β8	Leu	Ser	Hist	Gly		Glx	Gly)	Lys
5β5	Leu)	Lys						

At position 72 in 3β9 and 4β8 a serine may be present, while in 5β5 an aspartic acid residue is assigned. This amino acid replacement is also reflected in their different map positions. All three peptides move toward the

cathode, but peptide 5β5 does not move as far as peptides 3β9 and 4β8. This indicates a larger net negative charge on peptide 5β5. Peptide 5β4 is also more polar than peptides 3β9 and 4β8 which can be accounted for in a serine-aspartic exchange. On the other hand, peptides 3β9 and 4β8 also demonstrate different mobilities. Peptide 4β8 moves closer to the cathode than 3β9 but 3β9 is more polar than 4β8 (Figure 17). This is due to the fact that peptide 4β8 appears to be longer than either 3β9 or 5β5 by six amino acid residues. In peptide 4β8, position 76 is shown to be occupied by a serine instead of a lysine (page 96). The other two peptides - 3β9 and 5β5 - terminate at residue 76 and positions 77-82 are contained in another peptide. This arrangement is comparable to what occurs in the dog β chains and in one of the mouse β chains (strain BALB/c). The rat 4β chain arrangement follows the same pattern set by human, rabbit and one of the other mouse β chains (strain C57BL) where position 76 is not lysine but some other amino acid. Although, it is possible that there is a lysine residue at position 76 in 4β8 and that trypsin did not cleave it, the molar ratio values for the amino acids in this peptide do not favor such an explanation.

Assuming that 4β8 does have this diffrent arrangememt, the composition of residues 78-81 appear not to match those amino acid residues usually found in these positions by homology (page 96). In fact, one amino acid appears to be lacking at position 79. From these considera-

ations it could be argued that peptide 4β8 actually represents the following:

67	68	69	70	71	72	73	74	75	76	77
Val	Ile	Asx	Ala	Phe	Ser	Asx	Gly	Leu	Lys	Hist

In this instance, trypsin cleavage occurred at the histidine - which really belongs to the next peptide, positions 77-82. The two remaining glycines and one serine and one glutamic acid would just be considered as "extra amino acids" from background contamination. It should be noted that peptide 5β5 also reveals some "extra amino acids" - glycine and alanine. They are not present in significant quantities. Clearly further investigation into this peptide will be needed to resolve the question.

<u>Sequence Homology of the β Chains - Residues 83-95</u>

Species	83	84	85	86	87	88	89
Human β	Gly-	Thr-	Phe-	Ala-	Thr-	Leu-	Ser-
Dog β	Gly-	Thr-	Phe-	Ala-	Lys-	Leu-	Ser-
Mouse (C57BL)β	Gly-	Thr-	Phe-	Ala-	Ser-	Leu	(Ser.
Rabbit β	Gly-	Thr-	Phe-	Ala-	Lys-	Leu-	Ser-
3β15						(Ala	Ser
4β14						(Leu	Ser
5β14	(Gly	Thr	Gly	Ala	Val	Leu	Ser

Species	90	91	92	93	94	95
Human β	-Glu-	Leu-	Hist-	Cys-	Asp-	Lys
Dog β	-Glu-	Leu-	Hist-	Cys-	Asp-	Lys
Mouse (C57BL)β	.Glu)	Leu	(Hist.	Cys.	Asp)	Lys
Rabbit β	-Glu-	Leu-	Hist-	Cys-	Asp-	Lys
3β15	Glx	Gly	Hist		Asx)	Lys
4β14	Glx	Ala	Hist		Asx)	Lys
5β14	Glx	Ala	Hist		Asx)	Lys

A similarsituation is encountered in the 5β chain. Peptides 3β15, 4β14 and 5β14 appear to be common peptides by map location (Figure 17, Table IIIb). However,

upon analysis, peptides 3γβ15 and 4γβ14 appear to encompass sequence residues 88-95 while peptide 5γβ14 may encompass residues 83-95 (Table XXX and page 98). By Homology with human and mouse γβ chains, peptide 5γβ14 contains an amino acid other than lysine at position 87. In this case it is thought to be valine. This makes peptide 5γβ14 five amino acids longer than peptides 3γβ15 and 4γβ14 which match dog and rabbit γβ chains by homology. Residues 83-87 in the 3γβ and 4 γβ chains would appear as separate tryptic peptides due to the presence of a lysine at position 87.

In addition, peptide 3γβ15 appears to have an amino acid exchange at position 91 - a glycine for an alanine. It is further thought that at position 88, in peptide 3γβ15, there is an alanine instead of a leucine. It is possible that a leucine residue does occupy position 88 but as it is the N-terminal amino acid, it might be completely destroyed during map spraying with ninhydrin. From other results, however, it can be seen that no other peptide has had its N-terminal amino acid completely destroyed by the ninhydrin reagent. At position 93 in all three peptides, it is probable that cysteine, which is destroyed during acid hydrolysis, is also present.

It can be argued that if peptide 5γβ14 has more amino acids than 3γβ15 and 4γβ14 why doesn't it have a dif-

ferent map position? This can be partially explained from an electrophoretic viewpoint. Since these five added amino acids are essentially neutral amino acids, they would have no affect on the overall net charge of the peptide. However, one of the five added amino acids is a threonine - a polar amino acid - and it should have affected the mobility of this peptide during descending chromatography.

Sequence Homology of the $\gamma\beta$ Chains - Residues 121-132

Species	121	122	123	124	125	126
Human $\gamma\beta$	Glu-	Phe-	Thr-	Pro	-Pro-	Val-
Dog $\gamma\beta$	Glu-	Phe-	Thr-	Pro	-Gln-	Val-
Mouse (C57BL) $\gamma\beta$	Asp-	Phe-	Thr-	Pro	-Ala	(Ala.
Rabbit $\gamma\beta$	Glu-	Phe-	Thr-	Pro	-Gln-	Val-
3$\gamma\beta$11	(Glx	Phe	Val	Hist	Ala	Val
4$\gamma\beta$10	(Glx	Phe	Ser	Gly	Ala	Val
5$\gamma\beta$						

Species	127	128	129	130	131	132
Human $\gamma\beta$	-Gln	-Ala-	Ala-	Tyr-	Gln-	Lys
Dog $\gamma\beta$	-Gln	-Ala-	Ala-	Tyr-	Gln-	Lys
Mouse (C57BL) $\gamma\beta$	.Gln	.Ala)	Ala-	Phe-	Gln-	Lys
Rabbit $\gamma\beta$	-Gln	-Ala-	Ala-	Tyr-	Gln-	Lys
3$\gamma\beta$11	Hist	Ala	Ala	Tyr	Leu)	Lys
4$\gamma\beta$10	Hist	Ala	Ala	Gly	Leu)	Lys
5$\gamma\beta$						

From the total acid hydrolysis data of the total $\gamma\beta$ Chains (Table II), it is revealed that the 3$\gamma\beta$ chain has one more histidine residue than the 4$\gamma\beta$ or 5$\gamma\beta$ chain. This extra histidine has been tentatively located in peptide 3$\gamma\beta$11 (Table XXXI). Peptide 3$\gamma\beta$11 has two histidines, while peptide 4$\gamma\beta$10, the counterpart to 3$\gamma\beta$11 in the 4$\gamma\beta$ system, contains only one histidine. It should be noted here that peptides 3$\gamma\beta$11 and 4$\gamma\beta$10 have different mobilities

on their respective maps (Figure 17).

Peptide 3β11 also has two valines and one tyrosine, while 4β10 has one valine, two glycines, one serine and no tyrosines. Matching the sequence of these two peptides by homology to other species reveals that they include residues 121-132 (see page 100). Placement of the amino acids by homology, on the other hand, is not reliable because of the disparate composition of these two peptides as compared with analogous sequences in other species. Therefore, the placement of some of these amino acids is arbitrary. As a result, many of the amino acid exchanges that appear to be taking place may not necessarily be real. Some of the exchanges indicated on page 100 occur at position _123_, a valine for a serine; position _124_, a histidine for a glycine; and at position _131_, a tyrosine for a glycine. This latter exchange may not take place due to the possibility that at position _131_ in 4β10, a tyrosine residue is present but is completely destroyed upon acid hydrolysis. As mentioned earlier, tyrosine is acid-labile.

Amino Acid Composition Comparison - Peptides 3β12 and 4β11

Peptides 3β12 and 4β11 are presented in Table XXXII for comparison. As it can be seen, 3β12 and 4β11 do not contain the same amino acids even though they do appear to have the same map position (Figure 17). Peptides 3β12 and 4β11 could not be placed by homology into the β sequence. They seem to represent fragmented peptides.

$\gamma\beta$ Chain Summary

The $\gamma\beta$ chains revealed the following sets of differences:

1) The first eight amino acids in the 4$\gamma\beta$ chain may be different from those in the 3$\gamma\beta$ and 5$\gamma\beta$ chains (p. 92).

2) At position 13 in the 4$\gamma\beta$ chain, there is probably a valine while 3$\gamma\beta$ and 5$\gamma\beta$ have a glycine (p. 93).

3) Peptide 4$\gamma\beta$16 and 5$\gamma\beta$16 are different from 3$\gamma\beta$2 by map position. Peptide 3$\gamma\beta$2 appears to be missing the two aspartic acid residues that 4$\gamma\beta$16 and 5$\gamma\beta$16 have at position 21 and 22 (p. 94).

4) Peptide 5$\gamma\beta$13 may have an extra glycine at position 62-65 altering the 5$\gamma\beta$ chain sequence (p. 90).

5) The 3$\gamma\beta$ and 4$\gamma\beta$ chains have serine at position 72 while aspartic acid is indicated in the 5$\gamma\beta$ chain (p. 96). In addition, peptide 4$\gamma\beta$8 appears to contain six more amino acids as compared with counterpart peptides 3$\gamma\beta$9 and 5$\gamma\beta$5. It is possible that position 76 in the 4$\gamma\beta$ chain is some amino acid other than lysine - probably a serine.

6) Peptide 3$\gamma\beta$15 has alanine at position 88 and glycine at position 91, while peptides 4$\gamma\beta$14 and 5$\gamma\beta$14 have leucine at 88 and alanine at 91 (p. 98). Here also, peptide 5$\gamma\beta$14 appears to be longer by five amino acids than peptides 3$\gamma\beta$15 and 4$\gamma\beta$14 due to a valine-lysine exchange at position 87 in the 5$\gamma\beta$ chain.

7) As can be seen on page 100, the 3$\gamma\beta$ chain has a histidine at position 130 while the 4$\gamma\beta$ chain a serine.

There also seems to be a tyrosine-glycine exchange at position 123 and a valine-glycine exchange at position 124. Both exchanges occur between the 3 $\gamma\beta$ and 4 $\gamma\beta$ chains. No data could be presented on the 5 $\gamma\beta$ chain for this segment.

From the nature of the data presented on the $\gamma\beta$ chains, it can be concluded that in the rat, the 3 $\gamma\beta$ chain is different in amino acid composition and sequence from the 4 $\gamma\beta$ chain which in turn is also different in composition and amino acid sequence from the 5 $\gamma\beta$ chain. There are then at least three different $\gamma\beta$ chains in the rat.

DISCUSSION

Each individual rat has six different hemoglobins present in circulating red blood cells (1,6,7,8). The relative distribution of the hemoglobins is constant from one animal to another regardless of age, sex or strain (Figure 1). This heterogeneity is shown to be a result of differences in the primary structure of the globin chains. In this investigation, of the three rat hemoglobins studied at least two and probably even three different α chains were found, as well as, three different β chains. Comparisons of the total amino acid content of globin chains 3α, 4α and 5α (Table Ia) as well as 3β, 4β and 5β (Table IIa) revealed significant differences among the three α chains and among the three β chains. Two-dimensional peptide mapping techniques have also shown significant differences among the three α chain tryptic peptides (Figure 16, Table IIIa) and among the three β chain tryptic peptides (Figure 17, Table IIIb). Examination of Figure 18 reveals amino acid sequence differences among the 3α, 4α and 5α chains.

One major difference in the α chains involves an amino acid exchange somewhere in residues 1-7 of the 3α chain (Figure 18). In the 4α map, the peptide (in this instance $4\alpha2$) representing this portion of the α sequence (residues 1-7) could not be located in the same position on the 3α map (Figure 16), indicating that an amino acid exchange has occurred in the 3α peptide corresponding to

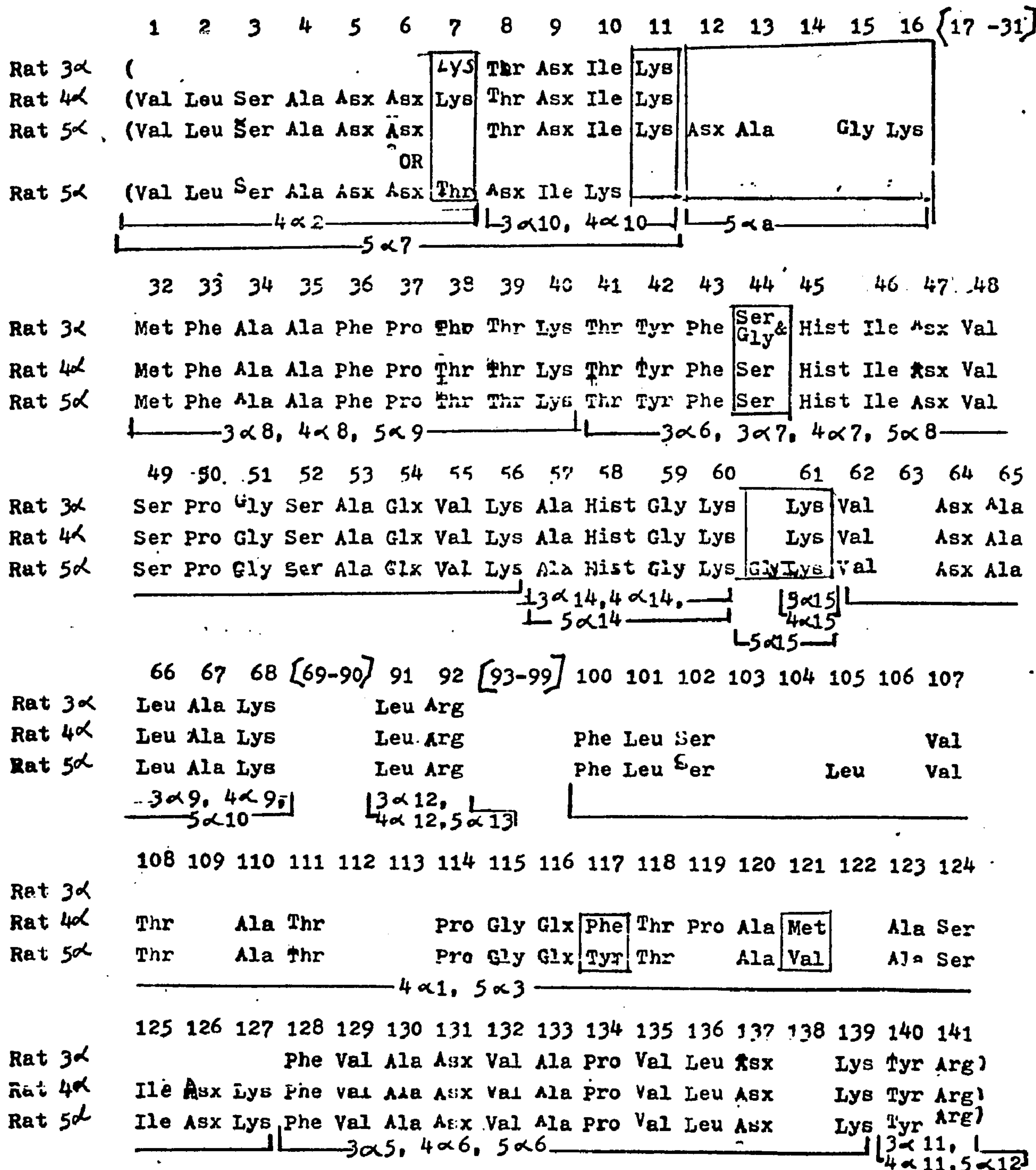

Figure 18 Sequence by Homology of the Rat 3α, 4α and 5α Chains

Boxed areas indicate residue differences among the three α chains.

Partitioned areas indicate the corresponding tryptic peptide.

Bracketed numbers indicate regions on the α chains not sequenced by homology.

residues 1-7.

In the 5 α chain, peptide 5α7 appears to represent the first ten amino acids on the 5 α chain sequence (Figure 18). The first six amino acids may match those on the 4 α chain. The remaining four amino acids of this peptide cannot be arranged like those of the 3 α and 4 α chains. Either position <u>7</u> or <u>11</u> in the 5 α chain does not contain the amino acid lysine which is present in the 3 α and 4 α sequences (see pages 79-83). Also, in the 5 α chain, residues 12-16 were located (Figure 18) but the peptide representing this sequence (peptide 5αa) had no counterpart in the 3 α and 4 α maps. This indicates that an amino acid exchange has occurred in the 3 α and 4 α peptides corresponding to residues 12-16 (see page 85).

Residue <u>44</u> in the 4 α and 5 α chains is glycine, while the 3 α chain is occupied at position <u>44</u> by a glycine and serine residue simultaneously (Figure 18). The presence of two amino acids at one position on the 3 α chain is an example of "heterozygosity at a single locus" and is discussed later in more detail.

In the 5 α chain, the appearance of a glycine residue (not present in the 3 α or 4 α chains) next to position <u>61</u> (Figure 18) compensates for the predicted loss of an amino acid (a lysine residue) at position <u>11</u> (or 7) in the 5 α sequence (see pages 78-83).

Position <u>117</u> and <u>121</u> in the 4 α and 5 α chains (Figure 18) are shown to be occupied by different amino

acids and are therefore areas of potential differences between the two chains. But, because of incomplete recovery of all the amino acids contained in the two peptides 4α1 and 5α3, representing residues 100-127, no reliable comparison between them with respect to amino acid content or possible amino acid sequence can be made at this time. In addition, the absence of a counterpart to these two peptides in the 3α map may indicate more amino acid differences between them and the 3α chain (pages 86-87).

The three β chains display more structural variations than the α chains. Among the first eight residues, an amino acid exchange has taken place in the 3β and 5β chains such that the peptides representing this part of the 3β and 5β sequence display a different mobility from the same peptide representing the 4β sequence (Figure 19). The following is a summary of the amino acid exchanges that have occurred among the three β chains (from Figure 19):

	13	21	22	65	72	76	87	88	91	123	124	130
3β	Gly	?	?	Lys	Ser	Lys	Lys*	Ala	Gly	Val	Hist	Tyr
4β	Val	Asx	Asx	Lys	Ser	Ser	Lys*	Leu	Ala	Ser	Gly	Gly
5β	Gly	Asx	Asx	Gly	Asx	Lys	Val	Leu	Ala			

(? Amino acids in this position cannot be identified. They are not aspartic acids.)

(* Lysine assumed at these positions.)

Of those portions of the β chains that were investigated and sequenced by homology, at least twelve different amino acid exchanges have taken place, as indicated above and in

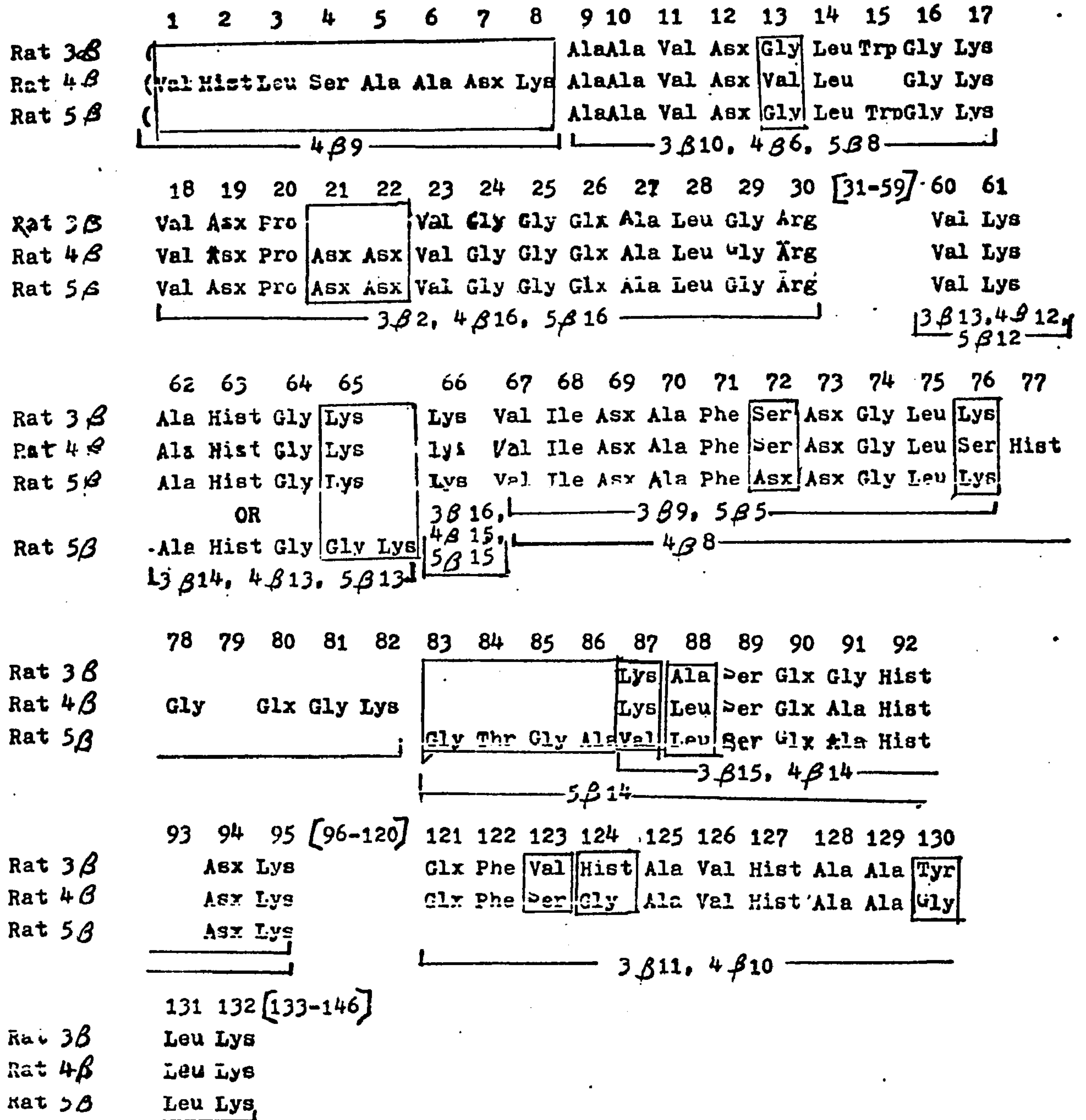

Figure 19 Sequence by Homology of the Rat 3β, 4β and 5β Chains

Boxed areas indicate residue differences among the three β chains.

Partitioned areas indicate the corresponding tryptic peptide.

[ ] Bracketed numbers indicate regions on the β chains not sequenced by homology.

Figure 19, among the 3, 4 and 5 β chains.

As previously mentioned, other investigators have demonstrated the presence of six rat hemoglobins in circulating erythrocytes. Ranney, et al (6) isolated six rat hemoglobins from hemolysates of adult Wistar Rats by cellulose-acetate electrophoresis and preparatively by DEAE-cellulose chromatography. From a comparison of charge and relative quantities, their hemoglobin components appeared to correspond to those of Stein, et al (1) as follows:

Stein, et al (1)	Distribution	Ranney, et al (6)	Distribution
Fraction I	8.7%	Hemoglobin 1	7%
Fraction II	17.9%	Hemoglobin 2	18%
Fraction III	20.6%	Hemoglobin 3	20%
		Hemoglobin 4	2%
Fraction IV	35.3%	Hemoglobin 5	45%
Fraction V	10.7%	Hemoglobin 6	13%
Fraction VI	7.2%	Hemoglobin 6a	

Ranney's hemoglobin component 4 does not correspond to any of the hemoglobin fractions isolated by IEF. The preparative separation of the rat hemoglobin into components by DEAE-cellulose chromatography yielded nine peaks. Using cellulose-acetate electrophoresis, Ranney's group determined that the nine peaks corresponded to six hemoglobins. Globin from the "total rat globin" was then separated into total α chains and total β chains by a modification of the Dintzis (19) pyridine-formate procedure. The α and β

chain pools were then further fractionated by the urea-phosphate method of Clegg, et al (18). By these methods, Ranney, et al found two distinct α chains and three distinct β chains. The α and β subunit compositions of the six rat hemoglobins was assigned by using urea-starch gel electrophoresis.

Chua, et al (7,8) also separated rat hemoglobin into six components (components A-F) using DEAE-sephadex Chromatography. The purity of the fractions was confirmed by starch gel electrophoresis. All subsequent work was carried out on major component E. Globin chain separation was performed on component E using the method of Clegg, et al (18).

Despite the different methods used for the separation of rat hemoglobin components, all groups agree to there being six rat hemoglobins. The major rat hemoglobin component is Fraction IV by IEF, component 5 by DEAE-cellulose chromatography, component E by DEAE-sephadex. The major difference in the technique for isolation of the α and β chains lies in the fact that Ranney's group first performs a gross separation of the α chains from the β chains using a pyridine-formate system (Dintzis) followed by a further fractionation of the total α chains (and the total β chains) by a urea-phosphate system (Clegg). On the basis of two resolved α peaks, Ranney's group claims that there are only two α chains. Chain $^{I}\alpha$ is present as 70% and $^{II}\alpha$ is present as 30% of the total α chain. The fact that one of these

α peaks might represent two or more unresolved α chains was not ruled out. If a neutral-neutral or an acidic-acidic amino acid exchange had occurred, it would go undetected due to the fact that the two α chains would co-elute and appear as a single α peak. By isolating α chains from individual hemoglobin fractions, the work of Chua, et al (7, 8) and of this investigator avoids this possibility. Chua's group has studied only one α chain from one hemoglobin component, while a partial investigation of three of the six rat hemoglobin α and β chains has been presented in this report. Therefore, the existence of three different rat α chains, as proposed by this investigator, is a possibility that should be seriously considered.

While Ranney, et al (6) do not give total amino acid composition data, Chua, et al (8) does give an amino acid analysis of their major α globin - $^{I}\alpha$ *. A comparison of Chua's group's data with that appearing in Table I reveals the following differences in amino acid content:

Amino Acid	$^{I}\alpha$ *(8)	3α	4α	5α
Histidine	10	10	9	9
Aspartic	13	13	14	14
Threonine	8	8	10	8
Serine	10	10	9	10
Glycine	12	11	11	12
Alanine	16	18	15	18
Cysteine	3	2	2	2
Valine	9	11	12	10
Leucine	16	14	15	14

Significant differences among the four chains reported above lie in the values for alanine, valine and leucine. The value of 3 cysteines reported by Chua, et al (7,8) is

not confirmed by Ranney, et al (6), who reports only two cysteines, nor does it correspond with the value of two cysteines reported by this investigator (Table I) for the three α chains presented in this study.

Other discrepancies among the work of Ranney, et al, Chua, et al, and this investigator also appear in the assignments of amino acid residues to certain positions along the major and minor α chain sequences (position 44 page 84; position 100 page 86; positions 129, 131-134, 137 page 74, 75). These differences in the compositional studies of the major and minor rat α chains may be due to the fact that a different strain of Wistar rat was used by each investigating group. Chua, et al (7,8) used a highly inbred New Zealand strain, while Ranney, et al (6) did not report which strain they studied. The CFN strain of Wistar rat was studied in this investigation. Even though Stein, et al (1) had investigated more than seven different strains of rat and found that all contained six hemoglobin components in the same % distribution in circulating erythrocytes, this did not necessarily mean that a hemoglobin component in one strain could not contain a different amino acid composition and/or sequence from the same hemoglobin component in other strains. Amino acid exchanges occuring in either the α chain or the β chain or in both chains that leave the overall net charge of a hemoglobin component unchanged would not be detected by IEF, cellulose-acetate electrophoresis, or ion-exchange chromatography. Clearly, further

work on different strains is necessary in order to resolve this point.

The fact that the rat has multiple α chains and multiple β chains is now evident. From the number and nature of the amino acid exchanges found by this investigator along with those of Ranney, et al (6), the variations among the 3α, 4α and 5α chains are due to a basic structural difference in the α chains and are not the result of a posttranslational modification. The fact that there are at least twelve amino acid differences among the three β chains, one of which is a polar-acidic amino acid exchange, also argues against these differences in the β chains arising from posttranslational modification.

Of the sequence assigned so far, by homology, in the rat α chains, there are at least sixteen differences in amino acid composition from the major mouse α chain. Three of these differences are found in residues that are in the region of contact with the heme group - residues 44, 46 and 50 (32). Generally, they are occupied by the amino acids proline, phenylalanine and histidine, respectively, in mouse and as well as in human, dog and rabbit α chains. In rat 3α, 4α and 5α chains, these same residues are serine (and glycine in 3α), isoleucine and proline, respectively. Carp α chains (31) also show a similar peculiarity in these residues being occupied as follows: 44 - alanine; 46 - tryptophan; 50 - proline. Oxygen affinity studies performed by Ranney, et al (6) have shown that the functioning of the

rat hemoglobin molecule has not been affected by the unusual amino acid changes in the α chain sequence.

There are at least twenty amino acid changes between the mouse β chain and the three rat β chains examined so far. Of these, only two occur in the heme contact region. This region includes position 88 in the 3β chain and position 91 in the 3β, 4β and 5β chains. Both these positions are leucine residues, by homology; but in 3β, position 88 is an alanine and 91 a glycine. In 4β and 5β , position 91 is assigned an alanine. Since these are neutral-neutral exchanges, no untoward affects on the heme-globin association or the heme-oxygen affinity is expected.

Multiple hemoglobins can be the result of non-genetic or genetic determinants (25). Non-genetic determinants, which are not being considered here, include those that are environmentally related, in vitro, (i.e. enzymatic degradation and/or polymerization of hemoglobin during isolation) or chemical modification, in vivo, (i.e. acetylation of the hemoglobin polypeptide, attachment of hexose moieties). In gene-related hemoglobin heterogeneity, two factors can be considered:

I Structural Gene Determinants

II Differential Gene Activity Related to Development - a study of the sequential emergence and arrest of respective polypeptide chain types.

In animals, multiple hemoglobins may be due to the presence of multiple alleles of a single structural

gene or to the existence of more than one structural gene for a polypeptide chain type. In either case, gene duplication has occurred resulting in the production of identical proteins by more than a single gene locus. This evolutionary event is followed by sequential mutation, accounting for two or more loci that produce polypeptide chains differing by one or several amino acids, while maintaining sufficient identity for the polypeptides to be assigned as α or β.

Heterogeneity caused by multiple hemoglobins in animals may be due to the existence of more than one structural gene for a polypeptide chain type as in some cases in man, where the α chain is coded by at least two pairs of nonallelic genes (9). Schroeder and Huisman (32) have produced evidence that multiple non-allelic γ chain genes can explain the heterogeneity and variable expression seen in human fetal hemoglobin. The occurrence of hemoglobin heterogeneity in the rabbit (10), mouse (14), goat (16) and horse (3) was found to be due to structural differences in the α chains arising from gene duplication. Allelism at the β chain locus with more than one amino acid replacement has been reported in rabbit and goat (32,9).

In the case of the rat α and β chains, it is proposed, then, that there are at least three different structural genes coding for the three different α chains and three different genes for the three different β chains. That there might be more structural genes involved in rat

hemoglobin synthesis awaits investigation of the amino acid composition and sequence of the α and β chains of hemoglobins I, II and VI.

Sequential analyses have shown that hemoglobin polypeptide chains from several mammalian species may be present in an electrophoretically and chromatographically inseparable mixture (32). Certain positions of the α or β chains are occupied by more than one amino acid. For instance, fractional residues of amino acids are present in more than one position of the α chains of the hemoglobin from the rabbit, horse, goat and certain strains of mice. This situation is also encountered in the rat 3 α chain at position 44 (Figure 18) which is reported to be occupied by both a glycine and a serine residue in a 50/50 ratio. The genetic heterogeneity can be caused by a heterozygosity at a single α (or β) chain structural locus or by the presence of two closely linked non-allelic genes. Example:

2 homozygous alleles coding for an α (or β) chain - aa, bb, cc.

2 heterozygous alleles coding for an α (or β) chain - ab, dc, bc.

In the case of the rat 3 α chain, it is proposed that this heterozygosity at a single locus is a result of there being a single point mutation on one of the 3α alleles:

$$(AG^{U}_{C} \text{ (ser) }) \longrightarrow (GG^{U}_{C} \text{ (gly) })$$

If equal amounts of the gene for 3α 44-ser and the gene for 3α 44-gly are being transcribed and translated, then two

different 3α polypeptide chains are being produced; one-half of which contain serine at position 44 and the other half with glycine at 44. In the mouse, strain BALB/c, an α chain is reported by Popp (12, 13) with position 68 occupied simultaneously by serine and threonine residues, also in a 50/50 ratio.

Although heterozygosity at a single locus is usually rejected by most investigators because the strains of the animal used are highly inbred, it cannot be ruled out in this case with the rat, in favor of the other alternative - presence of two closely linked non-allelic genes - for the following reason. All of the data presented represents results from α and β chains that have been isolated from hemoglobin pooled from 10-12 rats. It is conceivable that half these rats had hemoglobin genes coding for α chains with serine at position 44 and the others glycine. When the pooled hemoglobin was separated into α and β chains, the purified 3α chain actually represented a heterogeneous mixture of 3α chains; half of which contained serine and the other half glycine at position 44. If this were the case, then a 3α chain isolated from one individual rat should contain only one of these amino acids at position 44 (i.e. be homozygous for the 3α gene). To test this, hemoglobin from two individual rats (not pooled) was separated into the six fractions using IEF. Fraction III (as well as fraction IV and V) were then separated into α and β chains and subjected to tryptic digestion and map-

ping. In both cases, the α chain from fraction III revealed that tryptic peptides 3α6 and 3α7, containing position 44 in the sequence, were always present. This indicates that heterozygosity at a single α chain locus in an individual rat was responsible for there being two amino acids at position 44 in the 3α chain. This genetic heterogeneity, then, is not a result of some rats having genes coding for $3\alpha^{44\text{-ser}}$, only; while others have genes coding for $3\alpha^{44\text{-gly}}$, only.

In addition to the 3α chain, the 3β, 4α, 4β, 5α and 5β chains from individual rat specimens were also remapped and like the 3α chain, these maps matched the tryptic maps of α and β chains isolated from pooled rat hemoglobin. Thus, any conclusions reached about the α and β chains from pooled rat hemoglobin could be applied to hemoglobin α and β chains from individual rats of this strain.

With cell differentiation and development, during erythropoesis, the chain composition of hemoglobin differs due to changes in gene activity. In the study of sequential emergence and arrest of polypeptide chain types, Stein, et al (1), using in vivo "pulse" labeling of marrow erythroid cells with serum-bound ^{59}Fe to follow incorporation of the isotope into each of the six hemoglobins in circulating red cells, have concluded that erythroid cells, at different stages of their maturation, synthesize different hemoglobins and to a different extent. In the youngest erythroid cells, hemoglobin fraction V is the major hemoglobin syn-

thesized and in the oldest, fraction IV is the major hemoglobin synthesized. More recently, the work of C.K. Yeh (personal communication) has corroborated the conclusions of Stein, et al (1). By separating maturing erythroid cells of the rat bone marrow into young and old populations and examining their hemoglobin components via IEF, Yeh has found that the youngest erythroid cell does indeed contain only hemoglobin fraction V (of the red cell hemoglobins) and that in the older cells, hemoglobin fraction IV is the last hemoglobin to be synthesized. Clearly, with cell differentiation and development, the chain composition of the hemoglobins differs due to changes in gene activity.

The coordination of the synthesis of six different hemoglobins in the rat raises the following questions that remain unanswered:

1) During erythropoesis in the rat does the synthesis of each rat hemoglobin occur in the same cell, or is each hemoglobin produced in separate cell lines?

2) If each maturing erythroid cell synthesizes all six rat hemoglobins, is only one type of hemoglobin synthesized at a particular time or are several hemoglobins made at the same time at different rates? Without information on the primary structure of the other three rat hemoglobins - i.e fractions I, II and VI - it would be difficult to speculate on how the combinations of different α and β chains might take place.

3) A follow-up to the above question raises the possi-

bility that the gene for a particular hemoglobin might be polycistronic for α and β chains - i.e. the α and β chains would be coded for on one messenger - and hence translated simultaneously.

Present technology must advance before evidence can become available to solve these problems. In any case, evidence for repression and derepression is indicated whether hemoglobin is made in one cell or whether several are being synthesized simultaneously.

The data has implied that the heterogeneity of the α and β chains of rat hemoglobin fractions III, IV and V are the result of the presence of more than one structural gene. The available information, however, is at present insufficient to offer adequate insight into the genetic mechanism by which the production of these three types of α chains and of these three types of β chains is initiated and repressed. Continued research into all six hemoglobins of the rat, may perhaps, offer insight into the complex mechanisms responsible for production of different types of hemoglobins with respect to component α and β chains.

APPENDIX

TABLES XI - XXXII

TABLE XI Amino Acid Composition Comparison of α Chain Peptides - 3α14, 4α14, 5α14

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α14	4α14	5α14
Lys	1.00	1.00	1.00
Hist	1.17	1.08	1.00
Arg			
Asx		0.03	0.11
Thr			0.05
Ser		0.05	0.27
Glx		0.14	0.15
Pro			
Gly	1.11	1.14	1.33
Ala	0.78	0.87	0.94
Cys			
Val			
Met			
Ile			0.07
Leu			0.08
Tyr			
Phe			
	Lys Hist Gly Ala	Lys Hist Gly Ala	Lys Hist Gly Ala

TABLE XII Amino Acid Composition Comparison of α Chain Peptides - 3α11, 4α11, 5α12

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α11	4α11	5α12
Lys	0.37	0.07	0.11
Hist	0.12	0.02	
Arg	1.00	1.00	1.00
Asx	0.13	0.09	
Thr	0.03	0.03	
Ser	0.18	0.09	0.13
Glx	0.38	0.11	0.03
Gly	0.34	0.18	0.08
Ala	0.35	0.10	0.09
Cys			
Val	0.17		
Met			
Ile			
Leu	0.01	0.04	
Tyr	+	0.52	0.46
Phe			
	Arg + Tyr	Arg Tyr	Arg Tyr

+ Tyr: Tyrosine is present in peptide because of a positive Pauly Reaction (Figure 16).

TABLE XIII Amino Acid Composition Comparison of α Chain Peptides - 3α9, 4α9, 5α10

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α9	4α9	5α10
Lys	1.00	1.00	1.00
Hist			
Arg			
Asx	0.85	0.76	0.80
Thr	0.07	0.07	
Ser	0.20	0.16	0.09
Glx	0.27	0.11	0.01
Pro			
Gly	0.29	0.34	0.29
Ala	1.86	1.74	1.69
Cys			
Val	0.62	0.57	0.55
Met			
Ile	0.20		
Leu	0.60	0.60	0.56
Tyr			
Phe			
	Lys Asx 1-2 Ala Val Leu	Lys Asx 1-2 Ala Val Leu	Lys Asx 1-2 Ala Val Leu

TABLE XIV Amino Acid Composition Comparison of α Chain Peptides - 3α5, 4α6, 5α6

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α5	4α6	5α6
Lys	1.00	1.00	1.00
Hist	0.05		
Arg			
Asx	1.71	1.68	1.59
Thr	0.14	0.10	0.02
Ser	0.23	0.22	0.12
Glx	0.09	0.19	0.06
Pro	0.55	0.39	0.43
Gly	0.23	0.31	0.18
Ala	2.48	2.29	2.38
Cys			
Val	1.51	1.52	1.42
Met			
Ile	0.24		
Leu	0.93	0.81	0.96
Tyr			
Phe	0.64	0.54	0.45
	Lys 1-2 Asx Pro 2-3 Ala 1-2 Val Leu Phe	Lys 1-2 Asx Pro 2 Ala 1-2 Val Leu Phe	Lys 1-2 Asx Pro 2 Ala 1-2 Val Leu Phe

TABLE XV — Amino Acid Composition Comparison of α Chain Peptides - 3α12, 4α12

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α12	4α12
Lys	0.02	0.09
Hist		
Arg	1.00	1.00
Asx	0.05	0.25
Thr		0.03
Ser		0.16
Glx	0.22	0.20
Pro		
Gly	0.19	0.26
Ala		0.13
Cys		
Val		
Met		
Ile		
Leu	0.77	0.70
Tyr		
Phe		
	Arg Leu	Arg Leu

TABLE XVI Amino Acid Composition Comparison

Of α Chain Peptides - 3α8, 4α8, 5α9

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α8	4α8	5α9
Lys	1.00	1.00	1.00
Hist	0.25	0.14	0.05
Arg			
Asx	0.06	0.32	0.15
Thr	1.72	1.77	2.03
Ser		1.04	0.47
Glx	0.39	0.73	0.29
Pro	0.61	1.09	1.15
Gly	0.32	1.72	0.50
Ala	1.82	2.78	2.17
Cys			
Val			
Met	0.83	1.05	0.69
Ile			
Leu		0.30	0.08
Tyr			
Phe	1.77	1.58	1.85

3α8	4α8	5α9
Lys	Lys	Lys
1-2 Thr	1-2 Thr	2 Thr
	Ser	Ser
glx?	Glx	
Pro	Pro	Pro
	1-2 Gly	Gly
1-2 Ala	2-3 Ala	2 Ala
Met	Met	Met
1-2 Phe	1-2 Phe	1-2 Phe

TABLE XVII Amino Acid Composition Comparison

Of α Chain Peptides - 3α15, 4α15, 5α15

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acids	3α15	4α15	5α15
Lys	1.00	1.00	1.00
Hist			0.07
Arg			
Asx	0.02	0.04	0.14
Thr		0.04	0.08
Ser		0.13	0.37
Glx	0.14	0.27	0.28
Pro			
Gly	0.12	0.04	0.48
Ala	0.04	0.14	0.14
Cys			
Val			
Met			
Ile			
Leu	0.02	0.04	
Tyr			
Phe			
	Lys	Lys	Lys Gly

TABLE XVIII — Amino Acid Composition Comparison of α Chain Peptides - 4α2, 5α7; 3α10, 4α10, 5α11

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	4α2	5α7
Lys	1.00	1.00
Hist	0.20	0.22
Arg		
Asx	2.20	1.57
Thr	0.14	0.50
Ser	1.18	1.14
Glx	0.32	0.49
Pro		
Gly	0.47	0.76
Ala	1.29	0.75
Cys		
Val	0.96	0.47
Met		
Ile		0.54
Leu	1.27	0.63
Tyr		
Phe	0.07	0.13
	Lys	Lys
	2 Asx	1-2 Asx
		Thr
	Ser	Ser
	glu?	Glu
	gly?	Gly
	Ala	Ala
	Val	Val
		Ile
	Leu	Leu

Amino Acid	3α10	4α10	5α11
Lys	1.00	1.00	1.00
Hist	0.15	0.10	
Arg	0.07	0.04	
Asx	1.04	1.06	
Thr	0.72	0.76	
Ser	0.43	0.47	
Glx	0.42	0.41	
Pro			
Gly	0.58	0.49	
Ala	0.23	0.23	
Cys			
Val	0.12		
Met			
Ile	0.92	0.92	
Leu	0.04	0.08	
Tyr			
Phe	0.06		
	Lys	Lys	Lys
	Asx	Asx	
	Thr	Thr	
	ser?	ser?	
	glu?	glu?	
	gly?	gly?	
	Ile	Ile	

TABLE XIX Amino Acid Composition Comparison of α Chain Peptides - 3α6, 4α7, 5α8, 3α7

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α6	3α7	4α7	5α8
Lys	1.00	1.00	1.00	1.00
Hist	1.03	0.87	0.95	0.93
Arg		0.15		
Asx	1.18	1.11	1.20	1.34
Thr	0.85	0.93	0.82	0.78
Ser	2.63	2.06	2.59	2.70
Glx	1.22	1.18	1.11	1.09
Pro	0.65	0.95	0.88	0.92
Gly	1.01	1.63	1.25	1.21
Ala	1.02	1.11	1.06	1.08
Cys				
Val	2.53	1.80	1.80	1.94
Met				
Ile	1.29	0.94	1.00	0.95
Leu		0.19	0.26	0.35
Tyr	0.71	0.61	0.49	0.58
Phe	1.29	1.16	0.91	0.91

3α6	3α7	4α7	5α8
Lys	Lys	Lys	Lys
Hist	Hist	Hist	Hist
Asx	Asx	Asx	Asx
Thr	Thr	Thr	Thr
2-3 Ser	2 Ser	2-3 Ser	2-3 Ser
Glx	Glx	Glx	Glx
Pro	Pro	Pro	Pro
Gly	1-2 Gly	Gly	Gly
Ala	Ala	Ala	Ala
2-3 Val	2 Val	2 Val	2 Val
Ile	Ile	Ile	Ile
Tyr	Tyr	Tyr	Tyr
Phe	Phe	Phe	Phe

TABLE XX Amino Acid Composition of α Chain Peptide - 5αa

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α	4α	5αa
Lys			1.00
Hist			0.13
Arg			
Asx			0.83
Thr			0.33
Ser			0.10
Glx			
Pro			
Gly			0.90
Ala			0.60
Cys			
Val			
Met			
Ile			
Leu			0.26
Tyr			
Phe			
	Peptide not located	Peptide not located	Lys Asx Gly Ala

TABLE XXI Amino Acid Composition Comparison of α Chain Peptides - 4α1, 5α3

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α	4α 1	5α 3
Lys		1.00	1.00
Hist		+++	+++
Arg			
Asx		0.64	2.24
Thr		3.20	1.52
Ser		1.50	4.37
Glx		1.01	4.15
Pro		1.94	1.03
Gly		1.34	4.80
Ala		3.76	2.92
Cys			
Val		0.47	1.49
Met		1.09	
Ile		0.45	1.19
Leu		0.84	1.98
Tyr		0.30	0.44
Phe		2.76	0.80

3α	4α 1	5α 3
Peptide not located	Lys	Lys
	+++ Hist	+++ Hist
	Asx	2 Asx
	3 Thr	1-2 Thr
	1-2 Ser	4 Ser
	Glx	4 Glx
	2 Pro	Pro
	Gly	4-5 Gly
	3-4 Ala	3 Ala
	Val	1-2 Val
	Met	
	Ile	Ile
	Leu	2 Leu
	tyr?	Tyr
	2-3 Phe	Phe

+++ Indicates Histidine present in significant amounts but is unresolved.

TABLE XXII Amino Acid Composition Comparison of α Chain Peptides - 3α13, 4α13

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3α13	4α13
Lys	1.00	1.00
Hist	0.23	0.17
Arg		
Asx	0.66	0.41
Thr	0.52	0.70
Ser	1.32	1.37
Glx	2.27	0.87
Pro		
Gly	1.90	0.93
Ala	1.42	0.39
Cys		
Val	0.41	
Met		
Ile	0.39	
Leu	0.34	
Tyr		
Phe	0.15	
	Lys Asx Thr Ser 2 Glx 2 Gly 1-2 Ala Val Ile leu?	Lys Asx Thr Ser Glx Gly ala?

TABLE XXIII Amino Acid Composition Comparison of β Chain Peptides - 3β13, 4β12, 5β12

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β13	4β12	5β12
Lys	1.00	1.00	1.00
Hist			
Arg	0.10		
Asx	0.09		
Thr	0.09		
Ser	0.02		
Glx	0.11	0.08	
Pro			
Gly			
Ala	0.09	0.18	
Cys			
Val	1.14	1.20	1.17
Met			
Ile	0.05		
Leu	0.06		
Tyr			
Phe	0.09		
	Lys Val	Lys Val	Lys Val

TABLE XXIV Amino Acid Composition Comparison of β Chain Peptides - 3β14, 4β13, 5β13

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β14	4β13	5β13
Lys	1.00	1.00	1.00
Hist	1.19	1.10	1.25
Arg			
Asx	0.04	0.12	0.07
Thr	0.04		0.09
Ser	0.09	0.22	0.21
Glx	0.06	0.32	0.19
Pro			
Gly	1.15	1.10	1.69
Ala	0.78	0.81	0.97
Cys			
Val			
Ile			
Leu			0.08
Tyr			
Phe			
	Lys Hist Gly Ala	Lys Hist Gly Ala	Lys Hist 1-2 Gly Ala

TABLE XXV Amino Acid Composition Comparison of β Chain Peptides - 3β16, 4β15, 5β15

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β16	4β15	5β15
Lys	1.00	1.00	1.00
Hist	0.15	0.06	
Arg	0.04	0.15	
Asx	0.23	0.29	0.15
Thr	0.20	0.24	0.13
Ser	0.54	0.58	0.39
Glx	0.38	0.89	0.34
Pro			
Gly	0.46	0.70	0.32
Ala	0.20	0.39	0.20
Cys			
Val	0.08		
Met			
Ile			
Leu	0.14	0.14	0.23
Tyr			
Phe			
	Lys Ser Glx Gly	Lys Ser Glx Gly Ala	Lys Ser Glx Gly

TABLE XXVI Amino Acid Composition of β Chain Peptide - 4β9

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β	4β 9	5β
Lys		1.00	
Hist		0.96	
Arg			
Asx		0.93	
Thr			
Ser		0.98	
Glx		0.34	
Pro			
Gly		3.30	
Ala		1.64	
Cys			
Val		0.68	
Met			
Ile			
Leu		0.81	
Tyr			
Phe			
	Peptide not located	Lys Hist Asx Ser 3 Gly 1-2 Ala Val Leu	Peptide not located

TABLE XXVI Amino Acid Composition of β Chain Peptide - 4β9

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β	4β9	5β
Lys		1.00	
Hist		0.96	
Arg			
Asx		0.93	
Thr			
Ser		0.98	
Glx		0.34	
Pro			
Gly		3.30	
Ala		1.64	
Cys			
Val		0.68	
Met			
Ile			
Leu		0.81	
Tyr			
Phe			
	Peptide not located	Lys Hist Asx Ser 3 Gly 1-2 Ala Val Leu	Peptide not located

TABLE XXVII Amino Acid Composition Comparison of β Chain Peptides - 3β10, 4β6, 5β8

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β10	4β6	5β8
Lys	1.00	1.00	1.00
Hist			0.19
Arg			
Asx	0.87	0.81	1.04
Thr	0.36	0.39	0.23
Ser	0.41	1.30	1.12
Glx	0.48	1.34	1.07
Pro			
Gly	1.98	0.85	1.75
Ala	1.52	1.47	1.54
Cys			
Val	1.11	2.00	0.89
Met			
Ile			0.32
Leu	1.11	1.20	0.89
Tyr			
Phe	0.16		
Trp[a]	+		+
	Lys	Lys	Lys
	Asx	Asx	Asx
	ser?	Ser	Ser
	Glx	Glx	Glx
	2 Gly	Gly	1-2 Gly
	1-2 Ala	1-2 Ala	1-2 Ala
	Val	2 Val	Val
	Leu	Leu	Leu
	+ Trp		+ Trp

[a]Detected by Ehrlich reagent on tryptic peptide map.

[+]Peptide gave a positive stain for tryptophan.

TABLE XXVIII Amino Acid Composition Comparison of β Chain Peptides - 3β2, 4β16, 5β16

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β2	4β16	5β16
Lys		0.02	0.08
Hist	0.04	0.33	
Arg	1.00	1.00	1.00
Asx	0.68	3.50	3.33
Thr	0.08		0.09
Ser		0.27	0.24
Glx	0.47	1.89	1.35
Pro	0.57	1.41	1.09
Gly	2.99	3.10	3.34
Ala	0.96	1.59	1.22
Cys			
Val	2.21	2.02	1.96
Ile			
Leu	1.06	1.34	1.18
Tyr			
Phe			
	Arg	Arg	Arg
	Asx	3-4 Asx	3 Asx
	Glx	1-2 Glx	Glx
	Pro	Pro	Pro
	3 Gly	3 Gly	3 Gly
	Ala	1-2 Ala	Ala
	2 Val	2 Val	2 Val
	Leu	Leu	Leu

TABLE XXIX Amino Acid Composition Comparison of β Chain Peptides - 3β9, 4β8, 5β5

The results are expressed in molar proportions of the amino acids to lysine (or Arginine).

Amino Acid	3β9	4β8	5β5
Lys	1.00	1.00	1.00
Hist	0.29	1.18	0.09
Arg			
Asx	1.79	1.76	2.81
Thr	0.14		
Ser	0.60	1.51	0.24
Glx	0.20	0.94	0.24
Pro			
Gly	0.97	3.20	1.53
Ala	0.95	1.33	1.58
Cys			
Val	0.66	0.78	0.77
Met			
Ile	0.71	0.42	0.69
Leu	0.63	0.74	1.08
Tyr	0.08		
Phe	0.71	0.83	0.70

3β9	4β8	5β5
Lys	Lys	Lys
	Hist	
1-2 Asx	1-2 Asx	2-3 Asx
Ser	1-2 Ser	
	Glx	
Gly	3 Gly	1-2 Gly
Ala	Ala	1-2 Ala
Val	Val	Val
Ile	Ile	Ile
Leu	Leu	Leu
Phe	Phe	Phe

TABLE XXX Amino Acid Composition Comparison of β Chain Peptides - 3β15, 4β14, 5β14

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β15	4β14	5β14
Lys	1.00	1.00	1.00
Hist	0.77	1.00	0.99
Arg	0.19		
Asx	0.41	1.00	0.68
Thr	0.27	0.47	0.61
Ser	0.81	1.20	1.25
Glx	0.61	1.10	1.26
Pro			
Gly	0.93	0.34	1.69
Ala	0.50	1.03	1.57
Cys			
Val			0.35
Met			
Ile			
Leu	0.13	0.55	0.46
Tyr			
Phe			

3β15	4β14	5β14
Lys	Lys	Lys
Hist	Hist	Hist
Asx	Asx	Asx
	thr?	Thr
Ser	Ser	Ser
Glx	Glx	Glx
Gly	gly?	1-2 Gly
Ala	Ala	1-2 Ala
		Val
	Leu	Leu

TABLE XXXI Amino Acid Composition Comparison of β Chain Peptides - 3β11, 4β10

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β11	4β10
Lys	1.00	1.00
Hist	2.40	0.79
Arg	0.18	
Asx	0.31	0.22
Thr	0.17	
Ser	0.67	1.39
Glx	0.50	1.12
Pro		
Gly	0.73	2.01
Ala	3.18	3.20
Cys		
Val	2.02	1.34
Met		
Ile	0.26	
Leu	0.93	0.54
Tyr	0.75	
Phe	0.53	0.81
	Lys	Lys
	2 Hist	Hist
	asp?	
	Ser	Ser
	Glx	Glx
	Gly	2 Gly
	3 Ala	3 Ala
	2 Val	Val
	Leu	Leu
	Tyr	
	Phe	Phe

TABLE XXXII Amino Acid Composition Comparison of β Chain Peptides - 3β12, 4β11

The results are expressed in molar proportions of the amino acids to lysine (or arginine).

Amino Acid	3β12	4β11
Lys	1.00	1.00
Hist	0.19	0.19
Arg	0.14	
Asx	0.51	0.21
Thr	0.45	
Ser	0.58	0.53
Glx	1.40	0.86
Pro		
Gly	1.01	0.91
Ala	0.67	0.32
Cys		
Val	0.51	
Met		
Ile	0.29	
Leu	0.60	0.17
Tyr		
Phe	0.36	

	3β12	4β11
	Lys	Lys
	Asx	
	Thr	
	Ser	Ser
1-2	Glx	Glx
	Gly	Gly
	Ala	ala?
	Val	
	ile?	
	Leu	
	phe?	

REFERENCES

1. Stein, S., Cherion, M.G. and Mazur, A.
J. Biol. Chem. 246 5287 (1971)

2. Huisman, T.H.J., Brandt, G. and Wilson, J.B.
J. Biol. Chem. 243 3675 (1967)

3. Kilmartin, J.V. and Clegg, J.B.
Nature 213 269 (1967)

4. Popp, R.A.
Fed. Proc. 24 1252 (1965)

5. Schneck, A.G. and Schroeder, W.A.
J. Amer. Chem. Soc. 83 1472 (1961)

6. Ranney, H.M., Garrick, L.M., Sharma, V.S. and McDonald, M.J.
Biochem. J. 149 245 (1975)

7. Chua, C.G. and Carrell, R.W.
Biochim. Biophys. Acta 326 328 (1974)

8. Chua, C.G., Carrell, R.W. and Howard, B.H.
Biochem. J. 149 259 (1975)

9. Kitchen, H.
Hemoglobin: Comparative Molecular Biology Models for the Study of Disease Annal. of N.Y. Acad. Sci. Kitchen, H. and Boyer, S. (eds) Vol. 241 (1974)

10. Hunter, T. and Munro, A.
Nature 223 1270 (1969)

11. Garrick, M.D., Hafner, R., Bricker, J. and Garrick, L.
cf - 9 p. 436

12. Popp, R.
Biochim. Biophys. Acta 303 52 (1973)

13. Popp, R. and Bailiff, E.
Biochim. Biophys. Acta 303 61 (1973)

14. Popp, R.A.
J. Mol. Biol. 27 9 (1967)

15. Clegg, J.B.
cf - 9 p. 61

16. Huisman, T.H.J., Wilson, J.B. and Adams, H.R.
Arch. Biochem. Biophys. 121 528 (1967)

17. Rossi-Fannelli, A., Antonini, E. and Caputo, A.
Biochim. Biophys. Acta 30 608 (1958)

18. Clegg, J.B., Naughton, M.A. and Weatherall, D.J.
J. Mol. Biol. 19 91 (1966)

19. Dintzis, H.
Proc. Natl. Acad. Sci. U.S.A. 47 247 (1961)

20. Moore, S. and Stein, W.H.
Methods in Enzymology S.P. Colowick and N.O. Kaplan (eds) Acad. Press N.Y. Vol. VI p. 819 (1963)

21. Lui, T.-Y.
Methods in Enzymology S.P. Colowick and N.O. Kaplan (eds) Acad. Press N.Y. Vol. XXV p. 45 (1971)

22. Lui, T.-Y. and Chang, Y.H.
J. Biol. Chem. 246 2842 (1971)

23. Hirs, C.H.W.
Methods in Enzymology S.P. Colowick and N.O. Kaplan (eds) Acad. Press N.Y. Vol. XI p. 59 (1967)

24. Guidotti, G., Hill, R.J. and Konigsberg, W.
J. Biol. Chem. 237 2184 (1962)

25. Bennet, J.C.
cf - 23 p. 330

26. Baglioni, C. and Ingram, V.M.
Biochim. Biophys. Acta 48 253 (1961)

27. Ingram, V.M.
Biochim. Biophys. Acta 28 539 (1958)

28. Katz, A.M., Dreyer, W.J. and Anfinsen, C.B.
J. Biol. Chem. 234 2897 (1959)

29. Dhounldt, J.L., Cartigny, B. and Farrieux, J.P.
Clin. Chem. Acta 50 297 (1974)

30. Smith, I.
Nature 171 43 (1953)

31. Dayhoff, M.O.
Atlas of Protein Sequence and Structure Vol. 5 (1972)

32. Huisman, H.J. and Schroeder, W.A.
New Aspects of the Structure, Function and Synthesis of Hemoglobins Chemical Rubber Co. (1971)

33. Dawson, R.M.C., Elliott, D.C., Elliott, W.H. and Jones, C.M.
Data for Biochemical Research Oxford Univ. Press (1969)