THE ORIGIN, PARTIAL STRUCTURE, AND PROPERTIES OF HEMOGLOBIN A_{Ic}

Thesis by

Walter Richard Holmquist

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1966

(Submitted October 1, 1965)

DEDICATION

To My Parents

THE SECRET SITS

We dance around in a ring and suppose,
But the Secret sits in the middle and knows.

-- Robert Frost

ACKNOWLEDGMENTS

I express my thanks to Dr. Schroeder for his patience in guiding my research on hemoglobin A_{Ic}, especially during those few discouraging moments when "nothing seemed to work." It is also through his silent efforts that I have (of necessity) absorbed some of the philosophy that what is worth doing is frequently difficult.

I should also like to thank Dr. Keith Shillington at Washington and Lee University for first making chemistry the exciting field it has remained for me to this day.

Joan and Roger Shelton, Barbara Robberson, Barbara Olson, Jean Cormick, John Cua and Bill Fenninger have--quite aside from their technical assistance--provided that camaraderie which goes so far towards making a laboratory a happy place to work.

To the Institute, and to Dr. Waser in particular, I owe one of the more rewarding periods of my stay at Caltech--the teaching of freshman chemistry during my first graduate year.

Financial support for this research has been provided by the Institute, by Du Pont Chemical Company and by the American citizen through his tax support of the National Science Foundation, the National Institutes of Health, and the United States Public Health Service.

Blood samples for the research have been donated by Dr. Walter Schroeder, Dr. Donald Babin, Dr. Robert Leif, and Hyland Laboratories in Los Angeles. The medical proficiency of Dr.'s Perry, Armstrong, and Matossian has made the drawing of some of these samples a relative pleasure. The high reticulocyte count blood samples which were used

in studying the biosynthesis of A_{Ic} were obtained through the cooperation of Dr. Ken Williams at Children's Hospital, Los Angeles and Dr. David Chernof at the University of California at Los Angeles Medical Center. I especially thank Dr. Georg Philipps, now at Washington University in St. Louis, for instructing me in the practical aspects of *in vitro* biosynthetic systems.

ABSTRACT

Hemoglobin A_{Ic} is a minor hemoglobin component of normal adult humans and comprises approximately 5% of the total hemoglobin. This thesis describes the in vitro biosynthesis of A_{Ic} and the chemical properties and structure of this component and its degradation products.

A_{Ic} was found to be biosynthesized concurrently with hemoglobin A, which is the major component (80-90%) in normal adult hemoglobin. A_{Ic} is thus a normal constituent of all red cells, irrespective of their physiological age.

Chemically, A_{Ic} is the condensation product--a Schiff base--between one molecule of hemoglobin A and one molecule of a ketone or aldehyde R=O. The point of linkage of R=O to hemoglobin A is the N-terminus of one of the two β chains. Other than this, no difference has been found between the primary amino acid sequence of the α and β chains of hemoglobin A_{Ic} and A.

The ketone or aldehyde R=O has a molecular weight of approximately 281, is not an aromatic aldehyde, is not a steroid, and does not contain phosphorus, carbohydrates or amino acids. Esterified nonketo acyl groups of less than five carbon atoms are absent. It is probable, but not certain, that R=O does not contain nitrogen and that it does not contain any free carboxyl groups.

Methods are described for isolating A_{Ic} as well as smaller peptides which contain the blocking group R such as R=val-his. The use and construction of a large (10 x 100-cm) chromatographic column for the preparative isolation of minor hemoglobin components is described. This column

is rather unique and should find applications in the isolation of other biological compounds. Certain other isolation procedures which are described in the thesis, such as the purification of peptides on cellulose phosphate, have not been previously described.

Finally several miscellaneous theoretical aspects of the thesis are of interest. These include the theory of self-hybridization experiments, a theoretical treatment of protein biosynthesis which includes a method for correcting the experimental data for nonconstant ribosomal activity and for interconversions of the protein of interest to other proteins during the biosynthesis. A practical method for calculating ionization constants from paper electrophoretic data is also given.

Photographic materials on pp. 25, and 46 are essential and will not reproduce clearly on Xerox copies. Photographic copies should be ordered.

ORGANIZATION OF THESIS

The thesis is divided into seven major divisions. Part I serves as a general introduction to the thesis problem with particular emphasis on historical perspective, Part II describes the chemical characterization of hemoglobin A_{Ic}, and Part III describes its biosynthesis. Part IV summarizes all the currently (October, 1965) known facts about A_{Ic}. Part V is a set of appendices to the thesis proper, and Part VII is the bibliography. Results of the experiments are to be found at the end of each individual section, although several summaries have been placed at strategic locations throughout the thesis.

Footnotes have been placed at the bottom of the page to which they apply. These footnotes can be ignored in a first reading of the thesis without a significant loss in comprehension, but they are essential for anyone desiring to repeat the experiments.

Part VI is the candidate's propositions. The bibliography for each proposition is located at the end of that proposition and not in the bibliography (Part VII) for the thesis proper.

All errors that are given are probable errors (P.E.) unless otherwise stated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure | Page

Part I

INTRODUCTION

I: Historical Background of Thesis Problem

I: HISTORICAL BACKGROUND OF THESIS PROBLEM

I A: Definition of Hemoglobin
I B: Major and Minor Hemoglobin Components
I C: Hemoglobin A_{Ic}
I D: Purpose of Thesis

I A: DEFINITION OF HEMOGLOBIN

Hemoglobin is the generic term which is applied to the family of proteins that are obtained from the erythrocytes of many biological species.* Hemoglobins have the following characteristics in common: a molecular weight of approximately 64,000 or an integral submultiple thereof, the ability to bind molecular oxygen reversibly, and a red color which results from their similar absorption spectra. Each hemoglobin molecule is composed of an aggregate of polypeptide chains whose molecular weights are about 16,000;** each chain contains a single heme-bound iron atom. However, the chains may be alike or different in their primary amino acid sequence. For example, the human hemoglobins designated*** A, H, and F have the structures $\alpha_2\ \beta_2$, β_4, and $\alpha_2\ \gamma_2$, respectively; in these formulas α, β, and γ each represents a single polypeptide chain which differs from the other chains in its primary amino acid sequence.

---

*The hemoglobins of humans, primates, horses, donkeys, cattle, sheep, goats, rabbits, hares, mice, birds, cetaceans, and fish have been studied. A review of these hemoglobins is given by Schroeder and Jones (131). Because of their medical significance and relative availability, the human hemoglobins have been the most thoroughly studied; an excellent review of these hemoglobins is given by Huehns and Shooter (61). Both reviews contain extensive references to the original papers.

**In the literature each chain is sometimes referred to as a <u>subunit</u>. The term <u>subunit</u> is also used to denote various aggregates of the individual chains. The context in which the word is used defines the meaning intended.

***Schroeder and Jones, and Huehns and Shooter, <u>loc. cit.</u>, give very complete correlations between hemoglobin designations (<u>i.e.</u>, nomenclature) and molecular structure.

Other similarities or differences may exist among the hemoglobins of the same or different species; however, the above criteria provide a rapid preliminary classification of a protein--a classification which is useful for further characterization.

I B: MAJOR AND MINOR HEMOGLOBIN COMPONENTS

If the hemoglobin from any single species is subjected to any one of a variety of physico-chemical procedures, it is frequently separated into two or more protein components. Each of these components possesses the common characteristics described above and may therefore be considered a hemoglobin. These procedures include electrophoresis (21, 53, 61, 84, 107, 125), chromatography (1, 64, 65, 77, 78, 125), and selective denaturation (82, 92, 135). Typically (125, and Fig. 16, p. 183, this thesis) one or more of these proteins will represent 30%-90% of the total hemoglobin; such proteins are for convenience called major hemoglobin components. Hemoglobin proteins which represent less than 30% of the total hemoglobin are called minor components.

I B1: Major Components

Most adult humans have only a single major hemoglobin component, which is designated* hemoglobin A or A_{II}, and which has the subunit

---

*No standard nomenclature exists for designating all the known hemoglobins. The reasons for this state of affairs have been summarized by Schroeder and Jones (131). The only sure guide to the identity of a hemoglobin is an understanding of the methods by which it was isolated and characterized. In this thesis the nomenclature developed by Schroeder and his associates (1, 22, 76, 125) will be used because it lends itself particularly well to the problems under study. Partial correlations between different systems of nomenclature are given in References 61, 125, and 131. Conventions do exist, however, for express-

structure (cf. p. 3) $\alpha_2 \beta_2$. The primary amino acid sequence of both the α and β chain is known in detail (15, 16, 61, 80, 81, 126, 131). The biosynthesis of the α chain is under the control of a gene segregating independently from the gene which controls the biosynthesis of the β chain.* Hemoglobin A_{II} comprises 80-90% of the total hemoglobin in adults.

The newborn human infant also has but a single major hemoglobin component, which has the subunit structure $\alpha_2 \gamma_2$, which is called fetal or umbilical cord blood hemoglobin, and which is designated hemoglobin F or F_{II} (cf. p. 3). The primary amino acid sequence of the γ chain is known (61, 129, 131) and differs from that of the β chain of A_{II}. The α chains of F_{II} and A_{II} are identical. The gene responsible for the synthesis of the γ chain is probably closely linked to the β-chain gene.* Hemoglobin F_{II} comprises 75-90% (1; 13; 61, Fig. 5) of the total hemoglobin in the fetus and newborn child. Shortly after birth the synthesis of γ chains ceases and the synthesis of the β chain is augmented (61, Fig. 5); the net effect is a replacement of hemoglobin F_{II} by hemoglobin A_{II}. Because the erythrocyte containing F_{II} has a greater affinity for oxygen than the adult erythrocyte containing A_{II} (109), it may be that fetal hemoglobin is important in insuring the

---

ing the structure of a hemoglobin in terms of its subunits (3, 131), and these conventions have been adopted for this thesis: indeed, they have already been used (cf. p. 3). The writer has attempted to include sufficient descriptive material in the text to make the meaning of the nomenclature self-evident, at least with respect to the topic under discussion.

*Reviews of the genetic evidence on which this statement is based can be found in References 61 and 131.

transference of adequate quantities of oxygen across the placenta from the mother's blood to the developing embryo.

Major hemoglobin components other than A_{II} and F_{II} occur infrequently and are termed abnormal* hemoglobins. The primary amino acid sequences of the polypeptide chains of all the normal and abnormal hemoglobins bear striking resemblances to one another. These resemblances persist even in comparisons between the hemoglobins of humans and those of lower phylogenetic species. Because of these resemblances it has been postulated that the various polypeptide chains have all arisen by evolution via point mutations of a single gene (145). For example, hemoglobins S (61, 73, 112, 131), C (61, 68, 69, 131), E (61, 70, 131), and D_{Punjab} (6, 8, 61, 131) each differ from hemoglobin A_{II} by only a single amino acid residue in the β chain.

The presence or absence of a major hemoglobin component in an individual can exert good, bad, or relatively slight effects on the health of that individual. Individuals who have the gene for hemoglobin S, C, or D_{Punjab} have an increased resistance to malaria (61, 103), a clearly beneficial effect. On the other hand, persons homozygous for the hemoglobin S gene exhibit a severe pathological hemolytic anemia which is called sickle cell disease. Yet, homozygosity for the hemoglobin C or

---

*An abnormal hemoglobin is defined as a hemoglobin which is found in a small minority of the earth's population. Huisman (66) lists 80 reports of abnormal hemoglobins, and Huehns and Shooter (61) list 48. In certain geographical regions, an abnormal hemoglobin may be found in a high proportion, even a majority, of the population of that region. The gene responsible for hemoglobin S (see text), for example, though relatively rare in the United States, is found in 40% (61) of the population of certain geographical regions of East Africa.

D gene results in only a mild, nonpathological anemia. This diversity of clinical manifestations* is somewhat surprising in view of the fact that these hemoglobins differ from each other by only 2-4 amino acid residues (1-2 for each β chain) out of a total of 574! Murayama (105, 106) has proposed a molecular mechanism for explaining the pathological manifestations of hemoglobin S in sickle cell disease.

An individual can have more than one major hemoglobin component provided he carries the genes required for the synthesis of the polypeptide chains of the hemoglobins (79, 123).

Although the above discussion of major hemoglobin components is by no means exhaustive, it has attempted to show that the origin, molecular structure, and function of these hemoglobins is relatively well, though not completely, understood. This situation is to be contrasted with lack of knowledge about the minor components, one of which, hemoglobin A_{Ic}, is the subject of this thesis.

I B2: Minor Components

The significance of the minor hemoglobin components of both men and other animals is less well understood than that of the major components, primarily because less attention has been devoted to their study. This lack of attention is in part a consequence of difficulties in obtaining sufficient amounts of purified minor components.

Whereas most adult humans have only a single major component, A_{II} (cf. p. 4), there exists evidence for ten minor components. In the order of their chromatographic mobilities on IRC-50 cation exchange

---

*Huehns and Shooter (61) review the clinical aspects of the abnormal hemoglobins in considerable detail.

resin (cf. p. 23, footnote) these are: A_{Ia_1} (most rapid mobility), A_{Ia_2}, A_{Ia_3} (86);* A_{Ib} (125); F_{II} (1; 131, p. 122, Fig. 1); A_{Ic} (102, 125); A_{II}·GSSG (65, 67, 104); A_4 (85); Hb-α^A (20, 32);** and A_{III} (125).*** In the following paragraphs, each of the above hemoglobins will be described briefly.

Under many experimental conditions hemoglobins A_{Ia_1}, A_{Ia_2}, and A_{Ia_3} are not resolved; the mixture of unresolved components is equivalent to Schnek and Schroeder's (125) A_{Ia}. If A_{Ia} is not resolved from A_{Ib}, the mixture of these two chromatographic fractions is referred to as A_{Ia+b} (125). A_{Ia+b} is identical to Kunkel's and Wallenius' (84) A_3. The amino acid compositions of A_{Ia} and A_{Ib} appear to be identical with that of the major component A_{II} (56, 76), though there is a margin of uncertainty in this identification because of a slight contamination of A_{Ia} and A_{Ib} with nonheme proteins that could not be removed by the chromatographic methods which were used. The results of Rosa and Labie (122) suggest that the β chains of A_{Ia+b} differ in some manner from those of A_{II}, though the facts just stated imply that the difference is probably not one of amino acid composition. A_{Ia} and A_{Ib} each comprise

---

*A_{Ia_1} is not given a designation by Laurent et al. (86); the latter investigators designate A_{Ia_2} and A_{Ia_3} as A_{Io} and A_{Ia} respectively. The latter hemoglobin is not to be confused with Schnek's and Schroeder's (125) A_{Ia} which is a mixture of A_{Ia_1}, A_{Ia_2}, and A_{Ia_3}.

**Unfortunately, Hb-α^A has also been designated A_4 (20); it is not to be confused with the hemoglobin A_4 of Rosa and Labie (85).

***A_{III} is identical to Kunkel's and Wallenius' (84) A_2.

approximately 1% of the total adult hemoglobin.

F_{II} is fetal hemoglobin, which has been discussed earlier (pp. 5, 6). Though it is the major component of the fetus, it comprises at most 0.4% of adult hemoglobin (12).

Hemoglobin A_{Ic} comprises 4% - 7% of the total adult hemoglobin and is the subject of this thesis. It is probably identical to component A_I^c which was isolated by Huisman and Meyering (64). It is sometimes incorrectly identified with Kunkel's A_3 (84) by some authors. Decreased percentages of A_{Ic} are found in persons having a hemolytic anemia (59).

Hemoglobin A_{II}·GSSG is a complex of oxidized glutathione (GSSG) with the major adult hemoglobin, A_{II}. The GSSG is attached to the β chains. A_{II}·GSSG is probably not present <u>in vivo</u>. It may be an artifact of the isolation procedures (67).

The nature of hemoglobin A_4 is puzzling. In short term <u>in vivo</u> and <u>in vitro</u> incubations with radioactive Fe^{59} or C^{14}-glycine, the specific activity of A_4 is almost twice as high as the activities of A_I* and A_{II} (85). This implies that hemoglobin A_4 is being synthesized at a more rapid rate than the latter two hemoglobins. As the duration of the incubation is increased to several days** the specific activity of A_4 falls off until at 34 days it is roughly equal to that of A_I and A_{II}.

---

*A_I is the designation which was given by Schroeder and his associates (cf. footnote, pp. 4, 5) to the unresolved mixture of minor hemoglobins A_{Ia+b}, F_{II}, and A_{Ic}.

**The short term incubations were done with both rabbits and with anemic human patients. The long term experiments which involved the decrease in specific activity of A_4 were done on rabbits only; extrapolation to humans may not be valid.

A_4 does not appear to be a precursor of A_{II} at least in in vitro experiments (85). No differences between the primary structures of A_4 and A_{II} could be demonstrated (85).

Hb-α^A is a hemoglobin which has the structure α_2 at physiological pH (20). The α chain is identical to that of A_{II}. It is probable, but not certain, that Hb-α^A exists in vivo (20, 32).

Hemoglobin A_{III} is identical to Kunkel's A_2 (84). Its molecular structure is α_2 δ_2. The α chain is identical to that of A_{II}. There are 10 known differences between the primary amino acid sequence of the β chain of A_{II} and the δ chain of A_{III} (61, Table 1; 131, Table 1). The gene responsible for the synthesis of the δ chain is closely linked to the β-chain gene (60). A_{III} comprises about 2.5% of the total hemoglobin in the adult. This percentage is increased or decreased in certain hemoglobinopathies (131, p. 124). This fact, together with the ease by which A_{III} can be isolated as a pure component by electrophoresis (84), make A_{III} a useful clinical diagnostic aid.

Huisman and Norton (67) have shown that the "in vivo aging of red blood cells does not result in the formation of additional minor fractions nor in changes in the quantities of the normally occurring minor hemoglobin components, as demonstrable by Amberlite IRC-50 chromatography."* A point which is particularly relevant to this thesis is the fact that even in vitro storage of red blood cells at 4°C for 49 days or at 25°C for 7 days results in no increase or decrease in the quantity of A_{Ic} initially present in the cells. Such in vitro aging does, however,

---

*Huisman and Norton's results contradict those of earlier investigations (63, 83, 93).

result in slight increases in the quantity of A_{Ia+b} (67). These facts clearly demonstrate that A_{Ic} is not an artifact of the method which is used to isolate it.

Three minor components have been found in human fetal hemoglobin (1, 76). These have been designated f_3, f_4, and F_I. Hemoglobin F_I has been shown to be a monoacetylated derivative of the major fetal hemoglobin, A_{II} (128). F_I has the molecular structure $\alpha_2 \gamma \gamma^{N\text{-acetyl}}$; it comprises 10% of the total fetal hemoglobin. Hemoglobins f_3 and f_4 together comprise roughly 4% (76) of fetal hemoglobin. Nothing is known about their structure.

Detailed studies of the minor components of individuals having abnormal major hemoglobins are virtually nonexistent, though Jones (76) has shown that up to seven minor components may be associated with hemoglobins S and C.

Still fewer studies exist with respect to the minor hemoglobin components of animals. Lingrel and Borsook have found that some of the minor components of rabbits are intermediates in the synthesis of the major rabbit component (89). None of the minor components found in humans has been shown to play an analogous role.

I B3: <u>Commentary on the Human Minor Components</u>

The above examples demonstrate the diverse biological roles which have been ascribed to the minor components, and suggest that their further study, with respect to structure, function, and evolution, would be fruitful. It was with this viewpoint in mind that the study of hemoglobin A_{Ic}, which is that adult human hemoglobin minor component present

in the greatest relative amount, viz. 4% - 7% of the total adult hemoglobin, was begun; and indeed, the results which have been obtained suggest that A_{Ic} has a role in biological function which differs from those ascribed to some of the minor components discussed in the preceding section.

IC: HEMOGLOBIN A_{Ic}

Hemoglobin A_{Ic} was discovered by Morrison and Cook (102) and Allen et al. (1). When the latter investigators chromatographed a sample of human hemoglobin on the cation exchange resin IRC-50 they observed a hemoglobin component which comprised 5.3% of the total hemoglobin in the sample, and which at neutral values of pH was eluted from the chromatographic column more rapidly than the major component A_{II}. This more rapidly moving component was designated hemoglobin A_{Ic}. Because IRC-50 is a carboxylic type resin it followed from these observations that the structure of hemoglobin A_{Ic} was such as to make it more negatively charged at neutral pH values than hemoglobin A_{II}.

The next advance toward an understanding of the structure of A_{Ic} was made by Jones (76) who reasoned in the following manner: Since the structure of A_{II} was representable in the form $\alpha_2\beta_2$ where α and β were distinct single polypeptide molecules (cf. p. 4), might not A_{Ic} have a similar structure? He was able to confirm his hypothesis experimentally by showing that the molecular weights of A_{Ic} and A_{II} were the same (75). Further, the total amino acid composition, i.e. the number and kinds of amino acids which were present in the hemoglobin irrespective of their

sequential arrangement, was to within the experimental error the same for A_{Ic} and A_{II}. Using Sanger's technique (124) for determining the N-terminal amino acid which was present in a polypeptide, he was able to show that in the A_{Ic} molecule there were at least two and possibly three polypeptide chains which had N-termini that were reactive toward dinitrofluorobenzene. As in A_{II} these chains appeared to be alike in pairs, because two of them contained the same amino terminal sequence val-leu (120) as the α chain of A_{II}. Only one of the possibly two remaining chains appeared to contain the same amino terminal tripeptide val-his-leu (121) as the β chain of A_{II}. These results suggested to him that the chemical difference(s) responsible for the chromatographic behavior of A_{Ic} was located in the β-like chain of A_{Ic}. He obtained additional support for this interpretation by showing that the α chain of radioactive S_{II}* would exchange with the α-like chain of A_{Ic} under suitable experimental conditions.**

Jones (76) found no difference between the ultraviolet spectra of A_{Ic} and A_{II} between 245 mμ and 320 mμ.

At the time the experiments which are reported in this thesis were begun, the structure of A_{Ic} that was most consistent with the data out-

---

*S_{II} is the major component of persons with sickle cell anemia (76). The molecular structure of S_{II} is $\alpha_2 \beta_2^S$ in which the α chain is the same as that in A_{II} and the primary amino acid sequence of the β^S chain differs from that of the β chain of A_{II} (61, 131).

**Exchange of subunit chains between two hemoglobin molecules does not establish the identity of these chains; it does, however, suggest that such identity may in fact exist.

lined above was $\alpha_2 \beta \beta^{A_{Ic}}$, where $\beta^{A_{Ic}}$ differed from the β chain of A_{II} in that the N-terminal amino acid of the former was blocked by some chemical group R whose nature was not known.*

ID: PURPOSE OF INVESTIGATION

The goals toward which the thesis work has been directed have been twofold: first, to characterize A_{Ic} chemically as fully as possible; and second, to obtain information about the biosynthetic origin of A_{Ic} in the erythrocyte and the functional role this minor hemoglobin component plays therein.

The chemical characterization of A_{Ic} has proceeded along three lines of investigation: confirmation of the subunit structure, $\alpha_2 \beta \beta^{A_{Ic}}$, which was proposed in the closing paragraph of the preceding section; determination of the primary amino acid sequence of the α and β chains of A_{Ic}; and chemical characterization of the group R which appears to be blocking the N-terminus of one of the β chains.

The isolation of the blocking group R has proved to be difficult. The larger portion of the candidate's time has been expended in develop-

---

*Jones himself (76) proposed the structure $\alpha_2 X_2$ for A_{Ic}; here the polypeptide chain X was assumed to be similar to, but not identical to the β chain of A_{II}. However, the fact that only one N-terminal val-his-leu peptide per A_{Ic} molecule was obtained by Sanger's technique (cf. p. 13) favors the formula given in the text. No hemoglobins with the unusual structure of hemoglobin F_I, $\alpha_2 \gamma \gamma^{N\text{-acetyl}}$ (cf. p. 11), were known at the time of Jones' work. Those hemoglobins which had been studied all had two α-like and β-like chains each. The structure found for F_I was totally unexpected and has given considerable impetus to a more searching examination of various aspects of hemoglobin chemistry.

ing suitable methods for its isolation, and it is only in the last year that this search has been brought to a successful conclusion. Because of these factors, the chemical characterization of R is not yet complete. Nevertheless, characterization is sufficiently advanced so that the directions which additional efforts should take are now well marked out.

Part II

CHEMICAL CHARACTERIZATION OF HEMOGLOBIN A_{Ic}

II: ISOLATION PROCEDURES
STEPWISE DEGRADATION OF HEMOGLOBIN A_{Ic} TO R-VAL-HIS

II A: INTRODUCTION

The aim of the procedures to be described in this section has been to obtain a purified compound which contained the blocking group R and which was sufficiently small to allow its complete chemical structure, and hence that of R, to be determined. In particular, because hemoglobin A_{Ic} and A_{II} have similar structures (cf. Part I), and because the molecular weight of A_{II} is approximately 64,000, it was clear that masswise the percentage of R in the "sufficiently small compound" would have to be of the order of 300 times the percentage of R in A_{Ic} if the structure determination were not to become prohibitively difficult.

A priori there were two ways of obtaining this enrichment. The first was to cleave R chemically from hemoglobin A_{Ic} and to purify it. The second was to degrade hemoglobin A_{Ic} into progressively smaller peptides. At each step of the degradation that peptide which contained R would be isolated, purified, and further degraded to still smaller peptides. In this manner a peptide which contained R eventually would be obtained that was small enough to make its complete structure determination possible. It would seem that the first method, being more direct, would be the method of choice. However, several factors dictated that the second be followed. At the time this investigation was begun, nothing was known about the chemical structure of R. Thus even had we known how to cleave R from A_{Ic}, which we did not, there was really no rational basis on which to purify it: in fact, there was no method even to detect R by itself. On the other hand, if A_{Ic} and A_{II} differed in the manner which has been indicated at the top of p. 14, R was attached

to the N-terminus of the β chain. Methods were known by which the N-terminal octapeptide val-his-leu-thr-pro-glu-glu-lys of the β chain of A_{II} could be obtained and purified (127). Further, Hill et al. (54) had shown that this octapeptide, which is designated* βT-1, could be further degraded with papain to the N-terminal dipeptide val-his. It seemed reasonable to expect that the hypothesized corresponding octapeptide R-val-his-leu-thr-pro-glu-glu-lys (designated $\beta^{A_{Ic}}$T-1 for conciseness) from $\beta^{A_{Ic}}$ and the dipeptide R-val-his could be obtained by similar methods. Equally important was the fact that though we had no method of detecting R itself, it was probable that the peptide to which R was attached could be detected with ninhydrin or histidine spray reagents. Once detected, this peptide could be distinguished from the peptides of the normal β chain by the altered electrophoretic behavior which was expected of the blocked peptide because of the chromatographic behavior of A_{Ic} on IRC-50 (cf. p. 12). In this manner progress of the enrichment of R could be followed simply by obtaining the amino acid composition of the peptide to which R was attached. Finally, the confirmation of the hypothesized gross structure** of A_{Ic} and the determination of the

---

*Throughout this thesis we shall designate peptides according to the nomenclature given by Baglioni (7). In βT-1, the β indicates the hemoglobin chain from which the peptide was obtained, the T indicates that this peptide resulted from hydrolysis of the chain with trypsin, and the -1 indicates that it is the N-terminal tryptic peptide. βT-2 would designate the tryptic peptide adjacent to βT-1 in the hemoglobin chain, etc.

**By the gross structure of a protein is meant a description of the number and type of subunits from which it is composed. Here, a subunit will be defined as a chemically definite molecule that is composed of single polypeptide chains. The chains of which a subunit is composed may be alike or different. Two subunits will be said to be of the same type if they are identical with respect to certain arbitrary criteria

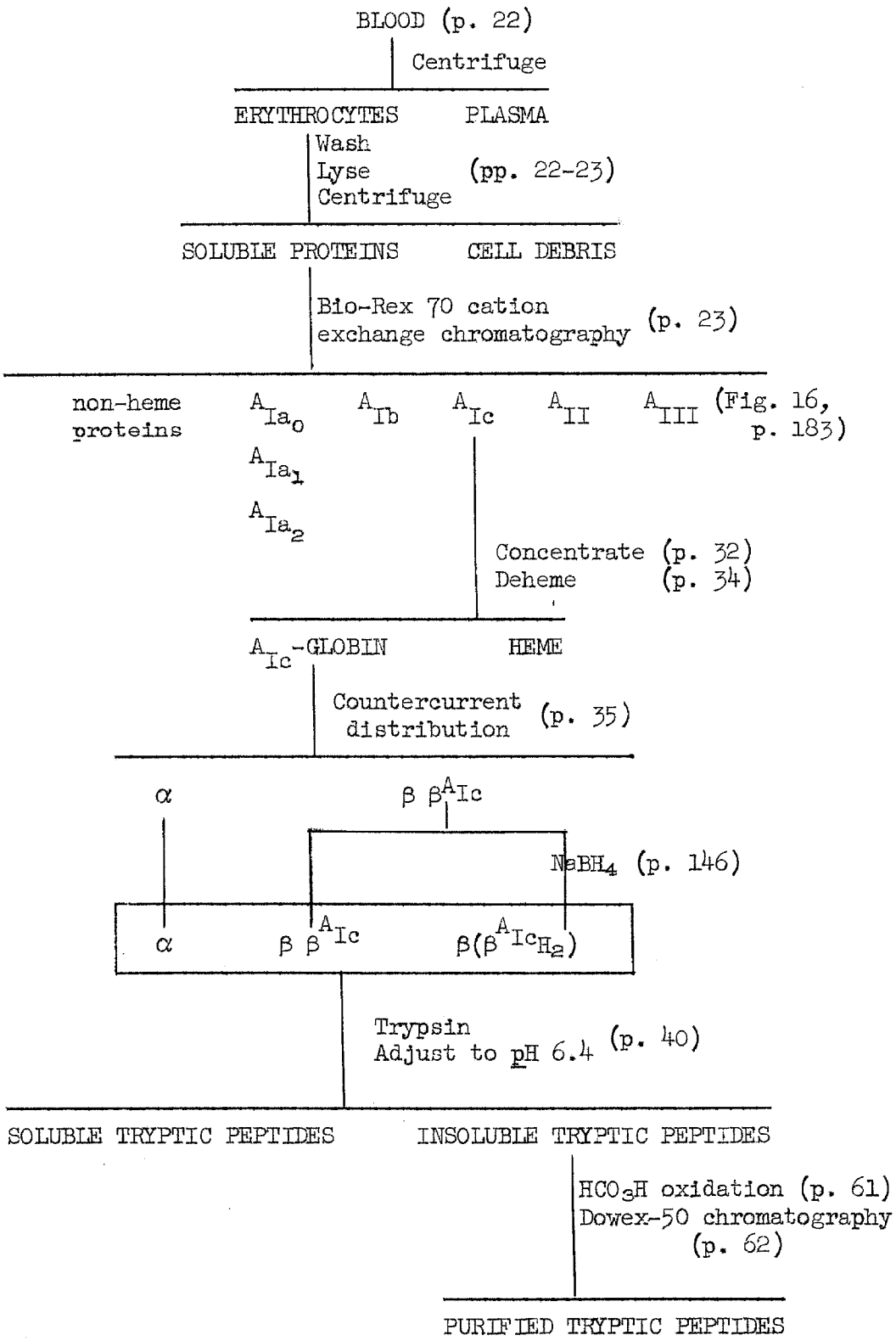
primary amino acid sequences of the α and β chains of A_{Ic} would in any case have required the isolation of peptides obtained from the degradation of hemoglobin A_{Ic}. By making the procedures which were necessary to purify R overlap as much as possible with those required to isolate the peptides of the α and β chains, the total labor could be reduced.

A summary of the isolation procedures, which are discussed in detail in subsequent sections, is given in the flow sheet on pp. 20a and 21.

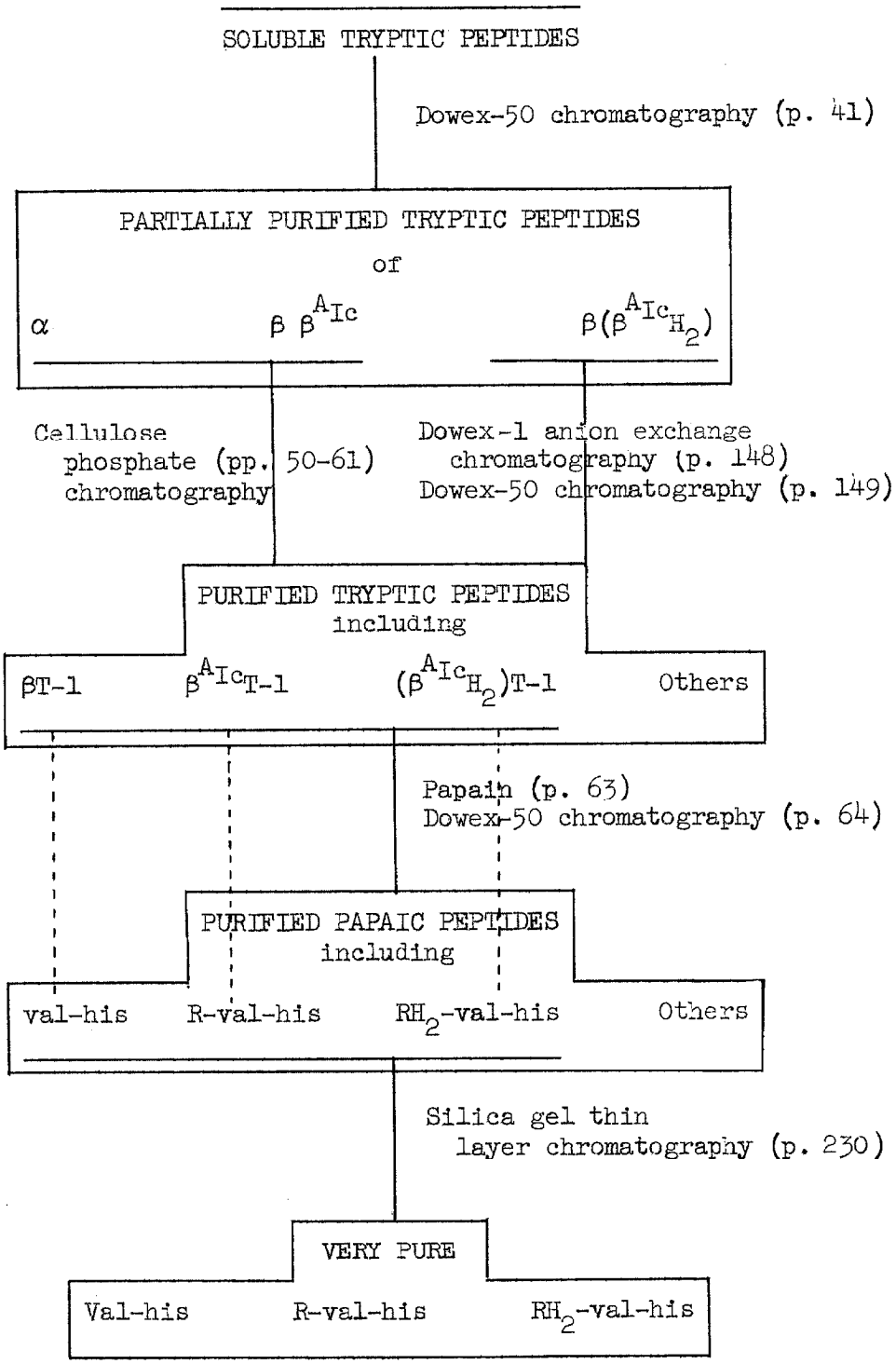
---

that are selected by the experimentalist. These criteria become progressively more stringent the more highly refined the structure determination becomes until ultimately they coincide with the usual definition of chemical identity between two compounds.

ISOLATION FLOW SHEET



ISOLATION FLOW SHEET CONTINUED



II B: PREPARATION OF UNFRACTIONATED HEMOGLOBIN LYSATE

All blood donors were adult, presumably normal, males. Sources of all blood were recorded so that any unexpected abnormalities could be traced down. No such abnormalities were found in any of the blood samples drawn. For small scale experiments, the purpose of which was to establish the experimental conditions to be used in later preparative work, 50 ml of venous blood were collected in Alsever's* solution. The donors for such small scale isolations were Walter Schroeder, Robert Leif, and the writer; hemoglobins from the blood of these donors were known from earlier experiments to behave normally and thus served as controls for other blood samples. For preparative experiments the red cells from two pints of whole blood, which was collected in citrate and from which the plasma had been removed, were obtained from Hyland Laboratories in Los Angeles. For both the small scale and preparative isolations the blood was worked up the same day that it was drawn from the donor.

If the plasma had not previously been removed, the blood was centifuged for 5 min at 2,000 g, and the plasma was removed with a suction pipet. The centrifuged cells were washed four times with three times their volume of NKM,** lysed at 8°C for 30 min by shaking them with 1.3

---

*Manufactured by Delta Laboratories, Inglewood, California. Alsever's solution is an anticoagulant which contains sodium citrate, citric acid, sodium chloride and dextrose in sterile water. The exact percentage of each of these components is not known, because Delta Laboratories would not give out this information.

**The composition of NKM is given in Table XX, p. 174.

volumes of distilled water and 0.4 volumes of toluene, and then were centrifuged for 1 hr at 15,000 g. The hemoglobin layer was transferred to a clean centrifuge tube and was again centrifuged for 1 hr at 15,000 g to remove the last traces of particulate material. The hemoglobin was converted to carbonmonoxyhemoglobin by gently swirling the aqueous solution at 0°C under a stream of carbon monoxide. The carbonmonoxyhemoglobin was dialyzed three times at 2°C for 8 hr against 10 times its volume of the buffer to be used in the chromatography of the lysate. The dialysate from 50 ml of whole blood contained 5 gm of hemoglobin in a total volume of 50 ml.

II C: BIO-REX 70 FRACTIONATION OF LYSATE INTO PURIFIED A_{1c}

II C1: Introduction

The separation of human hemoglobin lysates into purified major and minor hemoglobin components by cation exchange chromatography on Bio-Rex 70* was accomplished with the system of sodium phosphate-potassium cyanide chromatographic buffers which were developed by Allen et al. (1), Clegg and Schroeder (22), and Jones (76). For analytical purposes, columns 1 x 35 cm were used. For preparative isolations, columns 3.5 x 35 cm or 10 x 100 cm were more suitable. These columns had a capacity of 0.1, 2.5, and 50 gm of hemoglobin respectively. The use of the analytical column has been described in detail in the references just cited. The use of the 3.5 x 35-cm column, with specific reference to

---

*Bio-Rex 70 (Bio-Rad Laboratories, Richmond, California) is prepared by the manufacturer from Duolite CS-101 and is equivalent to the IRC-50 used by Allen et al. (1). These resins perform similarly.

the isolation of hemoglobin A_{Ic}, has been described by Holmquist and Schroeder (57). However, the large 10 x 100-cm column has not previously been described. The need for this large column became evident when it was realized that the structure determination of A_{Ic} was more difficult than anticipated and that a successful attack on the structure would require considerable amounts of A_{Ic}. Because of its large size, special techniques were required both for its construction and use. A detailed description of this column follows.

II C2: Description of the 10 x 100-cm Chromatographic Column

A photograph of this column in operation is shown on the frontispiece of this thesis. In Fig. 1, p. 25, the details of the construction of the column are shown. Figure 2, p. 26, shows the accessories which were needed for the operation of the column. Both figures should be referred to during the following description.

The upper (A) and lower (B) halves of the column, made of 4" i.d. Pyrex tubing which was sealed onto Pyrex pipe joints, were separated by a Koroseal gasket* (C) and were held together by two aluminum pipe flanges** (D). Each half of the column was jacketed (E) so that it could be warmed or cooled by circulating water. The lower flange was permanently bolted to a tripod (F) which supported the assembled column. The lower ends of the tripod were attached to a pressboard base (G) which

---

*Gasket, interface, for flange (see following footnote), Koroseal, Type R-3, Braun Catalogue No. 88164, Los Angeles, Calif.

**Flange, Pyrex pipe to Pyrex pipe, 4" size, Style 1, aluminum, with inserts, Braun Catalogue No. 87140-1.

LEGEND TO FIG. 1 AND FIG. 2

A: Top half-column, 4" Pyrex pipe and flange
B: Bottom half-column, 4" Pyrex pipe and flange
C: Koroseal gasket
D: Aluminum pipe flange with asbestos insert, six nuts and bolts
E: Water jacket, inlet; the water layer was 6 mm thick
F: Steel tripod
G: Pressboard base, triangular
H: Swivel wheels
I: Rubber covered metal ring support
J: Turnbuckles
K: Coarse frit funnel
L: Ball joint 18/9
M: Socket joint 65/40
N: Adapter; ball joint 65/40, socket joint 18/9
O: Pressure clamp with spring
P: Teflon stopcock with 18/9 ball joint to fit N
Q: 56-liter polyethylene carboy (aspirator bottle type)
R: Tygon tubing
S: Adapter for eluting lower half-column; 4" Pyrex pipe flange and 18/9 socket joint
T: Adapter for eluting upper half-column; 4" Pyrex pipe flange, coarse frit funnel, and 18/9 ball joint
U: Spigot
V: Wooden pressure collar for holding in rubber stopper
W: Tank of nitrogen
X: Small centrifugal pump
Y: Concave bottom of carboy not under pressure
Z: Tygon tubing

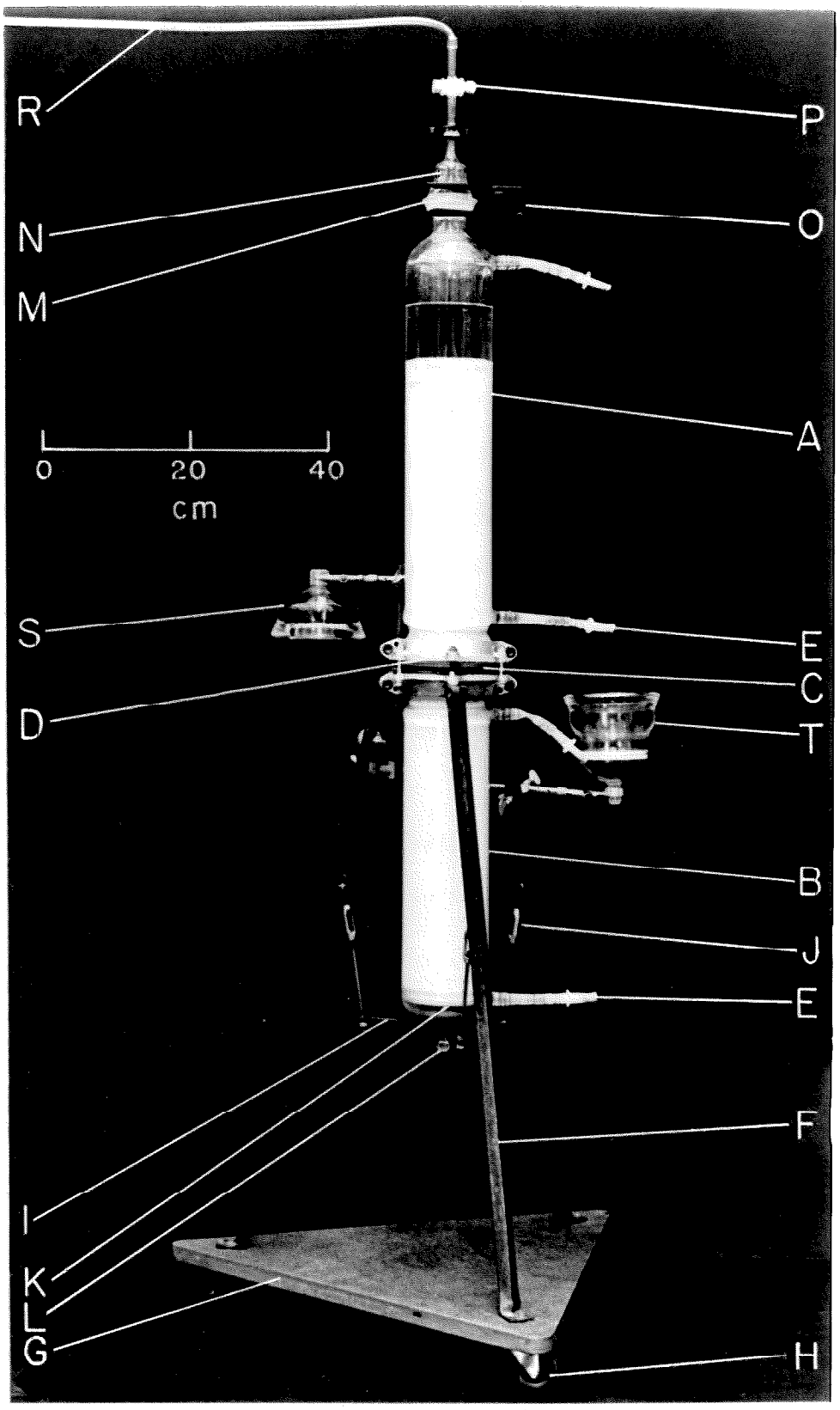


Fig. 1: 10 x 100-cm chromatographic column

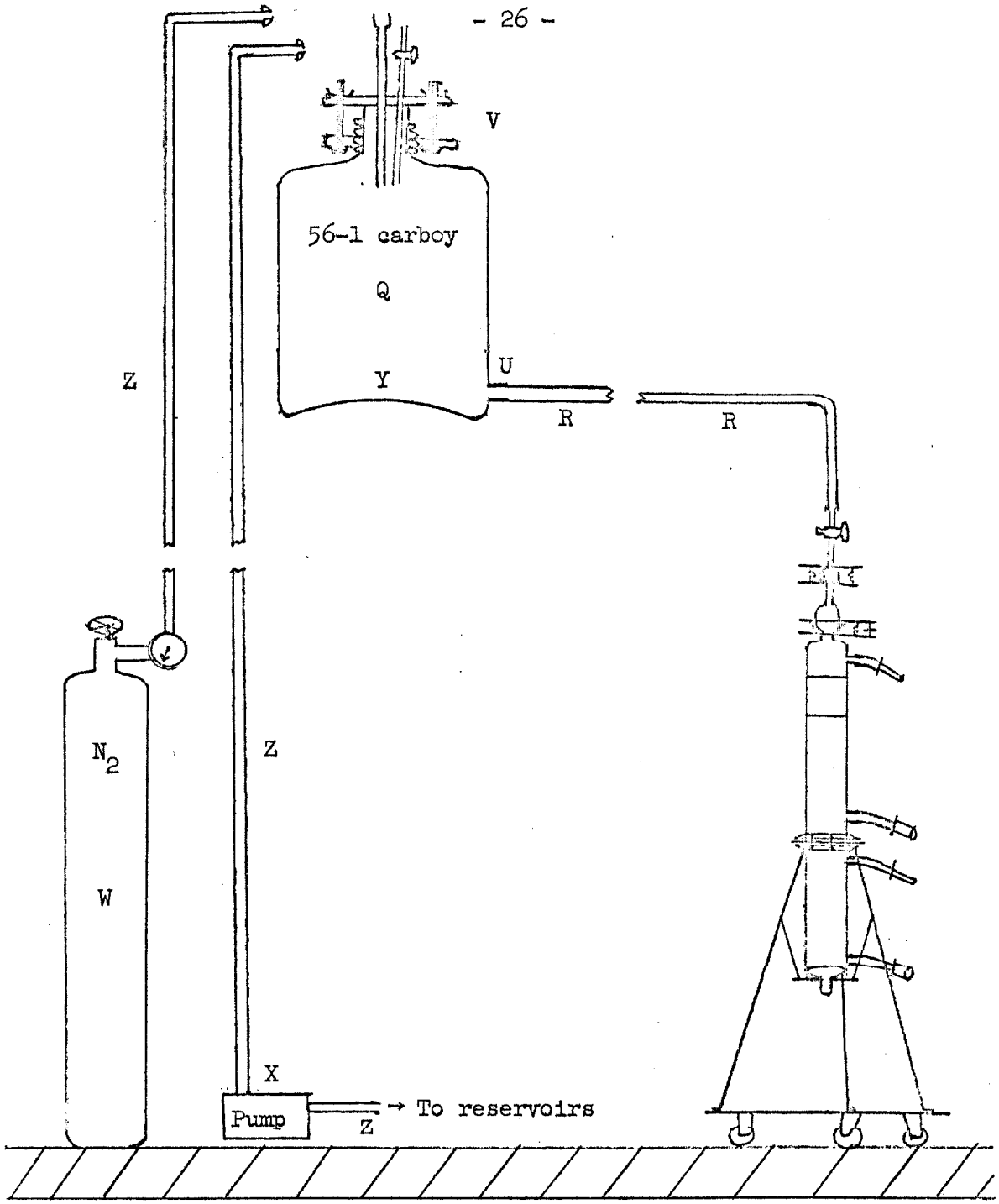


Fig. 2: Accessories to the 10 x 100-cm chromatographic column.

was supported on three wheels (H). The rubber-covered metal ring (I), the height of which could be adjusted by three turnbuckles (J), provided additional support for the column; it also prevented undesirable lateral motion of the column when the latter was transported from place to place.

The lower end of the bottom half-column was permanently sealed to a coarse frit funnel (K) which supported the resin and which terminated in an 18/9 ball joint (L). The upper end of the top half-column terminated in a 65/40 socket joint (M). An adapter (N), which was held to the socket joint (M) by the pressure clamp (O), connected the column to the Teflon stopcock (P) that led to the buffer reservoir (Q) (Fig. 2) by way of the Tygon tubing (R). Adapters (S) and (T) were provided for eluting the hemoglobin from the bottom and top half-column respectively.*

Three 56-liter screw-cap polyethylene carboys,** one of which (Q)*** had a side spigot (U) at the bottom, were required to operate the column. The carboy with the spigot was permanently attached to a shelf of the cold room. This carboy was placed as high as possible above the bottom of the chromatographic column in order to take advantage of the hydrostatic pressure which was thereby established. The

---

*The use of these adapters is described later in the text. In Fig. 1 they are supported by clamps for illustrative purposes. Normally, when they were not in use, these adapters were stored in a cabinet.

**Carboy, 13 gallon, with handgrips, polyethylene, Nalge No. 2210, Aloe Catalogue No. V17452, Aloe Scientific Co., Los Angeles, Calif.

***Bottle, aspirator, 13 gallon capacity, with tubulation, polyethylene, Nalge No. 2302, Aloe Scientific Co., Los Angeles, Calif.

mouth of this carboy was equipped with a rubber stopper that was held in place by a wooden collar (V) which prevented the stopper from popping out when the carboy was under pressure. Two wide-diameter glass tubes, one of which was equipped with a Teflon stopcock, fitted tightly into the stopper. The tube with the stopcock was used to release pressure from the carboy. The other tube served to fill the empty carboy with buffer, or to connect the filled carboy to a tank of nitrogen (W) during the operation of the column. The piece of Tygon tubing (R) that terminated in the Teflon stopcock (P) (Fig. 1) was attached with a band clamp to the spigot (U) at the bottom of the carboy. The remaining two carboys were used as buffer reservoirs; while one was cooling, fresh buffer was made up in the other. These two carboys sat on separate dollies for ease of movement from the laboratory to the cold room. A small centrifugal pump (X) was used to transfer buffer from a reservoir to the elevated carboy.

II C3: Packing the Chromatographic Columns with Bio-Rex 70 Resin

Bio-Rex 70* cation exchange resin, as received from the manufacturer and without further purification, was suspended in ten times its volume of distilled water and allowed to settle for 30 min by gravity. The fine particles which remained suspended in the supernatant were decanted and discarded. This process was repeated three times. The resin was then suspended in twice its volume of the buffer which was to be used in equilibrating the column prior to chromatography and was

---

*Purchased from Bio-Rad Laboratories, Richmond, California. The actual wet mesh size, as purchased, was 200-325 mesh. Cf. footnote, p. 23.

adjusted to the pH of the equilibrating buffer by adding the appropriate acid to the sodium form of the resin (for example, the pH was adjusted with orthophosphoric acid in the case of a pH 7 sodium phosphate buffer). The column was poured from this suspension as a single section: this was accomplished by keeping the column continuously filled to its top with suspended resin. The resin was allowed to settle into the column by gravity; it was not necessary or desirable to use pressure or to vibrate the column during the packing. After twenty column volumes* of the equilibrating buffer had passed through the column at the temperature to be used later in the actual chromatography of the hemoglobin, the column was ready for use.

II C4: Separation of Hemoglobin Lysate into Hemoglobin A_{Ic} and A_{II}

For fractionation of less than 2.5 gm of hemoglobin see References 1, 22, 57, and 76. The description which follows applies to the isolation of the A_{Ic} that is present in the lysate from a pint of whole blood.

The 10 x 100 cm-column was packed with resin as described in the preceding section and was equilibrated at 2°C with 160 liters of developer #6.** This amount of buffer was conveniently made up 56 liters at a time by draining an empty carboy, weighing the solid reagents directly into the carboy, and filling it to the 56-liter mark.

---

*By one column volume is meant the volume of the resin bed, including the interstitial volume.

**The composition of developer #6 is 18.40 gm of $NaH_2PO_4 \cdot H_2O$, 4.70 gm of Na_2HPO_4, and 2.60 gm of KCN per 4 liters of distilled water.

It was not necessary to wait for one reagent to dissolve before adding the next. Solution was achieved by mechanical stirring: a motor-driven Monel shaft that was equipped with a two-inch propeller was used to stir the salts into solution. Complete solution required less than an hour.

After the column had been equilibrated, the buffer from the top of the column was removed until the top of the Bio-Rex 70 resin bed was exposed; if the top of the resin had collected a surface scum from settled impurities,* it was removed to expose a fresh surface. Approximately 500 ml of the hemoglobin dialysate (cf. p. 23) containing 50 gm of carbonmonoxyhemoglobin in developer #6 was poured from a 1-liter beaker onto the resin bed and was stirred into the resin to a depth of 5 cm to give an even boundary between the stirred and unstirred portions of the resin. That portion of the dialysate which remained above the resin was allowed to flow into the column by gravity. When the resin surface was again exposed, the sides of the column were rinsed down with developer #6 at 2°C. The sides were rinsed three times; a small amount of developer was used each time, though the amount was sufficient to completely cover the top of the resin bed. The top of the column was carefully filled with developer #6 so that the top of the resin bed was not disturbed.

---

*These impurities came from the reagent grade chemicals which were used in preparing the buffer. They are not a reflection of the lack of purity of these chemicals, but simply of the large amounts of buffer which were used in equilibrating the column. The impurities could have been removed by filtering the buffer prior to use, but it was simpler, and just as effective, to use the column itself as the filter.

The column was developed at 2°C with developer #6 at a flow rate of approximately 2 liters/hr. This rate of flow required that one-half atmosphere of nitrogen pressure be applied to the buffer in the reservoir. The first proteins to emerge in the effluent were the "nonheme" proteins, and Hb-A_{Ia} and Hb-A_{Ib}: Because we were interested only in the isolation of A_{Ic}, the effluent from the column during the emergence of the preceding proteins was run directly to the cold room drain where it was diluted by a continuously running stream of tap water. The running tap water was a safety measure the purpose of which was to dilute the hydrogen cyanide that was present in the buffer. Approximately 36 hr after column development had been started, the front of hemoglobin A_{Ic} reached the fritted disc (K); this process was readily observed visually. At this point the development of the chromatogram was stopped, and the top half-column (A) was lifted away from the lower half and was set aside. The dimensions of the column had been chosen so that only A_{Ic} was present in the bottom half-column (B). Occasionally, because the column had been slightly overloaded, small amounts of A_{II} from the top half of the column were found in the resin topping the lower half of the column. If this occurred the A_{II} was easily scooped out with a spoon. Chromatographic zones A_{Ic} and A_{II} were usually well separated by a region of resin several centimeters in length.

The lower half-column (B), which was still attached to its support system (D, F, G, I, H, J), was taken from the cold room to the laboratory. After the adapter S had been attached to the lower half-column, the column was warmed to 38°C by circulating water through the jacket

(E). Because of the large diameter of the resin bed, this process took two to three hours. The rise in temperature was observed on a thermometer which was inserted along the axis of the resin bed. After the column had reached 38°C, developer #6, at room temperature, was used to elute the A_{Ic}. Approximately 2 gm of A_{Ic} was eluted in a total volume of 2 liters. The eluate was collected in a large bottle that was chilled in a bucket of ice. The solution of A_{Ic} was concentrated at 2°C through collodion membrane bags* to a final volume of 5 to 10 ml.

Hemoglobin A_{II} was eluted from the top half of the column in the same manner with the adapter (T). In order to remove the last traces of A_{II} from the resin, it was necessary to raise the column temperature to 50°C.

After the hemoglobin had been eluted from the resin, the resin was removed from the column by up-ending the column over a rectangular plastic reservoir of roughly 40-liter capacity** and by washing the resin into the reservoir with a stream of distilled water. This procedure was best carried out by two people--one holding the column, the other controlling the stream of water. The clean resin was suspended in water, and the column was repoured with this suspension. The column was then ready for re-equilibration.

---

*Membrane Filter: type, collodion bags CB; porosity, less than 5 mμ, size 8 ml; Membranfiltergesellschaft, Gottingen, Germany.

**Tank, rectangular, polyethylene, 10 gallon capacity, with cover, Nalge No. 11000, Aloe Scientific Company, Los Angeles, Calif.

II C5: Safety Considerations in the Fractionation of Hemoglobin Lysates

The use of large amounts of cyanide-containing buffers in the operation of the 10 x 100-cm column requires that special care be exercised at all times by the operator. At the neutral pH values of these buffers, all the cyanide is present as HCN in solution. The release of the hydrogen cyanide present in one 56-liter carboy into a cold room three meters on a side would produce an immediately lethal concentration of gas. In actual practice, where all containers are kept capped, the concentration of HCN in the air during the operation of the column seldom exceeds 2 ppm, a level which is safe for work over long periods without the need for a gas mask. However it is simply good safety practice to have a positive pressure type mask (canister type masks are not effective against high concentrations of HCN) ready at hand in the immediate working area and to be familiar with the relevant first aid principles in case of an accident. The hydrocyanic acid gas detector manufactured by Mine Safety Appliance Company* is a moderately priced, simply operated instrument which we have used to monitor the HCN concentration in working areas.

A second safety consideration involves the use of polyethylene carboys under pressure. When under half an atmosphere of pressure the bottom (Y) (Fig. 2) of these carboys rounds out (it changes from concave to convex in contour), and this tends to overturn the carboy.

---

*Address: 201 N. Braddock Ave., Pittsburgh 8, Pa.; detector, hydrocyanic acid gas, complete with leather case and set of 12 detector tubes, No. DH73493.

This tendency to overturn can be prevented by judicious arrangement of the equipment. In any case the carboy should not be allowed to present any danger to those working in the area of its use.

We have used the 10 x 100-cm column almost continuously over a six-month period without mishaps of any sort with respect to its operation.

II D: REMOVAL OF HEME FROM HEMOGLOBIN A_{Ic} TO OBTAIN A_{Ic}-GLOBIN (4, 129)

Between 5 and 10 ml of a solution containing 2 gm of hemoglobin A_{Ic} in developer #6 was dialyzed against three changes of distilled water at 2°C. The dialyzed solution was added dropwise to 1.5 liters of 0.01 F HCl in acetone. The acetone was kept at -15°C by immersing the reaction flask in a dry ice - methyl cellosolve bath. The mixture was stirred continuously for one-half hour and was then centrifuged at 0°C. The precipitated globin was washed several times with cold acetone and finally with ether. It was then vacuum dried at room temperature. When a more purified sample of globin was needed, the ether-washed globin was dialyzed against several changes of water at 2°C and was freeze-dried. The quality of the A_{Ic}-globin thus obtained varied from run to run; in some cases it was white in color and dissolved in aqueous solvents to give a very pale yellow solution; in others it was tan in color and gave a virtually opaque dark brown solution. However, regardless of the globin's color, it was suitable for use in the counter-current distribution procedure to be described next.*

---

*This statement is not meant to imply that it is not necessary to remove the heme from the globin. The procedure which has just been

II E: SEPARATION OF A_{Ic}-GLOBIN INTO α AND β $\beta^{A_{Ic}}$ SUBUNITS

A_{Ic}-globin was separated into α and β $\beta^{A_{Ic}}$ subunits by two phase countercurrent distribution at room temperature as described by Hill et al. (53). Initially it was hoped that the $\beta^{A_{Ic}}$ subunit could also be separated from the β subunit by this method, but in practice this did not prove possible.

II E1: Experimental

The countercurrent apparatus consisted of 18 glass-stoppered 125-ml graduated cylindrical separatory funnels which were equipped with a stopcock at their bottom end.* These cylinders were attached by buret clamps to a common axle which was motor driven at 35 rpm.

The two phase system was prepared by mixing 2-butanol,** 0.5 F HOAc, and 10% v/v aqueous dichloroacetic acid*** in the volume ratios 42.51:52.26:5.22 respectively. The sample, consisting of 1 to 2 gm of globin, was dissolved overnight in a total volume of 50 ml of a

---

described fails to remove only a very small fraction of the heme originally present in the hemoglobin. However, that small fraction which is not removed is intensely colored.

*This stopcock provided a convenient means of transferring the lower phase. A thin coat of silicone grease was applied to each stopcock to prevent it from freezing.

**The 2-butanol should have an optical absorbancy at 280 mμ of less than 0.150. We found by experience that reagent grade 2-butanol purchased from J. T. Baker Co. satisfied this criterion; we have heard that the same is true of the Merck product. 2-Butanol purchased from several other sources had to be purified by distillation or charcoal filtration before it was suitable for use.

***Commercial preparations of dichloroacetic acid were purified by distillation at 45 mm Hg and 113°C.

solution containing 25 ml each of the upper and lower phases.* After the sample had dissolved, it was transferred to the first mixing cylinder, and the phases were allowed to separate. Each phase was then adjusted to exactly 50 ml for a total volume of 100 ml. The first-cylinder was stoppered, and the motor was turned on for two minutes, i.e., seventy inversions. The motor was turned off, and after the phases had separated the lower phase of the first cylinder was transferred to the second cylinder.** To the first and second cylinders were added 50 ml of the lower and 50 ml of the upper phase respectively. The cylinders were stoppered, and the cycle was repeated until the distribution had been completed.

The degree of separation achieved was determined by mixing 0.5 ml of each phase from each cylinder with 1.5 ml of methanol and reading the optical absorbancy at 280 mμ against the appropriate reagent blank. From the data obtained, the distribution coefficients of the α and β $\beta^{A_{Ic}}$ subunits between the organic and aqueous phase could be calculated.*** If we were not interested in determining the distribu-

---

*The dissolved sample varied in color from a very pale yellow to a very dark opaque brown depending on the amount of residual heme in the starting material. Good separations of the subunits were obtained in either case.

**If the phases were slow to separate, three expedients helped hasten the process: centrifuging the solution, or cooling it to 10°C, or warming it to 60°C. The presence of material having the appearance of fuzz at the interface between the two phases usually was an indication that the settling would be slow; removal of this material with a syringe before continuing the distribution reduced the time necessary for subsequent separations.

***The distribution coefficient is moderately sensitive to extraneous components, such as heme, which may be present in the starting material

tion coefficients, 4 ml of methanol were added to the 100 ml of solution in each cylinder; the cylinders were stoppered and inverted a few times to give a single phase. One milliliter of this phase was mixed with 1 ml of methanol, and the absorbancy of the mixture was determined as before. The mg* of protein in each cylinder was plotted against cylinder number, and a decision was made as to which cylinders would be pooled (Fig. 3). For example, in Fig. 3, p. 38, cylinders 1-11 were pooled to obtain the $\beta\ \beta^{A_{Ic}}$ subunits, and cylinders 13-18 to obtain the α subunits of the 1-gm sample; cylinder 12 was discarded. For the 2-gm sample, the $\beta\ \beta^{A_{Ic}}$ subunits and α subunits respectively were obtained by pooling cylinders 1-13, and 16-18; cylinders 14 and 15 were discarded.**

The pooled solutions were extracted once with their volume of cyclohexane.*** The aqueous phase was saved. The cyclohexane phase was then washed three times with one-tenth its volume of water,**** and

---

and can therefore only be used for identification purposes under well-defined conditions. Normally we did not calculate it; when we did, it was only to check that we were "in the right ball park."

*The number of mg in each cylinder is given approximately by the relation mg = AV, where A and V are the absorbancy and volume respectively of the solution in the cylinder.

**The following was helpful in determining which cylinders to pool. That cylinder to the right and left of which half the total number of mg lay was located, and depending on the exact shape of the distribution curve this cylinder and a certain number on either side of it were discarded. Usually no more than one or two cylinders needed to be discarded.

***The purpose of this step was to remove much of the 2-butanol; a considerable reduction in the total volume of the solution was thereby achieved.

****Any protein which may have been extracted into the cyclohexane was recovered in this step.

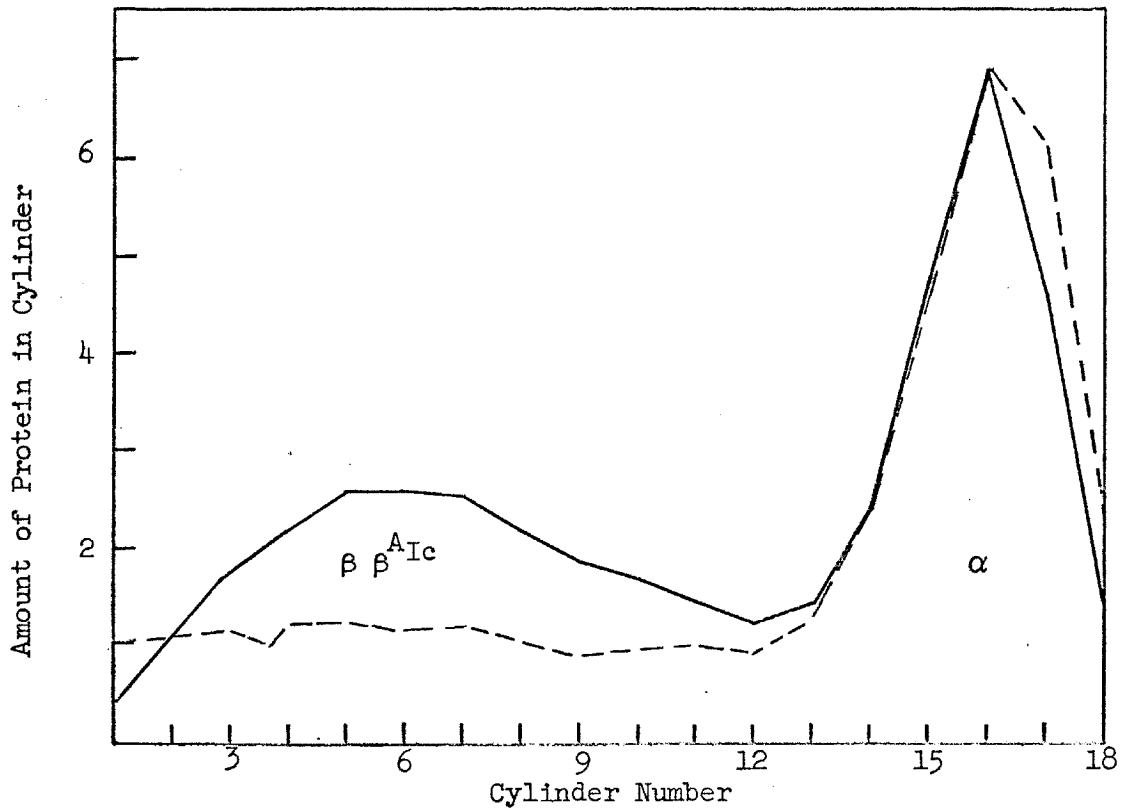


Fig. 3: Countercurrent distribution of A_{Ic}-globin into α and β $\beta^{A_{Ic}}$ subunits; ——— 1 gm of globin, – – – 2 gm of globin. The amount of protein in each cylinder has been normalized to cylinder no. 16 which was assigned the arbitrary value 6.9.

the washings were added to the aqueous phase that had been saved. The combined aqueous phases were dialyzed against three changes of 0.1 F HOAc at room temperature and then against three changes of water. If the volume of the aqueous phases was very large, it was convenient to concentrate it to a smaller volume by evaporation in a rotary evaporator before proceeding to the dialysis; if this was done, three drops of n-octanol were added to the solution to prevent foaming during the evaporation. Finally, the aqueous dialysate was lyophilized to dryness, and the purified subunits were stored in a vacuum desiccator.

II E2: Results

Typical distribution curves are shown in Fig. 3.* Two curves are shown: in one we started with 1 gm of globin, in the other with 2 gm. The flattening of the β $\beta^{A_{Ic}}$ peak with increasing concentration of globin was observed in many instances. When the cylinders were pooled as described in the footnote, p. 37, the purity of the subunits was shown by amino acid analysis and Sanger's end group method (cf. Table IV, p. 77) to be better than 90%; this was true for both the separations that are shown in Fig. 3. The reason for the flattening at high protein concentrations was not investigated since it did not seem to have an adverse effect on the purity of the subunits. Approximately 85% of each chain was recovered.

The failure of the β subunit to separate from the $\beta^{A_{Ic}}$ subunit

---

*Figure 3 should be interpreted in a semiquantitative manner only, for the method by which it was calculated (cf. pp. 37, 39--especially the footnotes thereto) was meant to give approximate results only. In practice we found this method to be a reliable guide for isolating subunits of high purity.

confirmed the similarity between these two chains which was suspected from Jones' (76) and our (Table VI, p. 99a) findings that the total amino acid composition of A_{Ic} was the same as that of A_{II}, and that the β chains of both A_{Ic} and A_{II} had the same N-terminal amino acid sequence (cf. pp. 12, 13, and Table IV, p. 77).

II F: HYDROLYSIS OF GLOBIN SUBUNITS INTO PEPTIDES BY TRYPSIN

To facilitate the determination of the primary amino acid sequence of the α- and β $\beta^{A_{Ic}}$-globin subunits from A_{Ic} and to obtain the N-terminal octapeptide $\beta^{A_{Ic}}$T-1 (cf. p. 19), which contained the blocking group R, we hydrolyzed the above subunits with trypsin (6).

A 1.7 gm sample of α- or β $\beta^{A_{Ic}}$-globin was dissolved under nitrogen in 64 ml of water in a 150-ml beaker. The solution was a turbid straw yellow and had a pH of 5.7. The solution was titrated to pH 8.0 with 2 F NaOH. As the pH passed through the pH 6 to pH 7 region, a curd-like white precipitate formed: this precipitate redissolved as the pH approached its final value of 8.0. At 0, 6 and 12 hr, 8.5 mg of trypsin* in 0.300 ml of 0.001 F HCl were added. A Radiometer** pH-Stat kept the pH constant at 8.0 during the hydrolysis. The hydrolysis was stopped after 24 hr by lowering the pH to 6.4 with 1 F HCl. At this pH a precipitate was obtained. The mixture was centrifuged at room temperature, the precipitate was washed three times with its volume of

---

*Worthington Biochemicals: TRL 6227, salt free, twice recrystallized, and freeze-dried.

**Radiometer, Copenhagen, Denmark; Titrator, Type TT1b; Titrigraph, Type SBR 2c.

pH 6.4 pyridine-acetic acid buffer which was 0.62 F in pyridine and 0.035 F in acetic acid and the washings were added to the supernatant. The washed precipitate contained the tryptic peptides which were insoluble at pH 6.4 and was stored frozen. The supernatant, containing the soluble tryptic peptides, was lyophilized to near dryness, and the mixture of nearly dry peptides was dissolved in a total volume of 25 ml of 0.2 F pyridine-4.90 F acetic acid pH 3.1 buffer (127). After the pH of this solution had been adjusted to 2.2 with 6 F HCl, the peptides were separated from each other on Dowex-50 as described in the following section.

II G: PRELIMINARY SEPARATION OF THE TRYPTIC PEPTIDES ON DOWEX-50

II G1: Introduction

The tryptic peptides from the α or β chains of hemoglobin A_{II} had been separated by Schroeder et al. (127) by column chromatography on the cation exchange resin Dowex-50 X2 into several zones, each of which contained between one and four individual peptides. The individual, purified peptides from these zones were obtained by further chromatography on the anion exchange resin Dowex-1 X2. Because of the many similarities between hemoglobin A_{Ic} and A_{II} the decision was made to attempt the purification of the tryptic peptides from the α and β $\beta^{A_{Ic}}$ subunits of A_{Ic} in the same manner.

The major goal of our experiments was to isolate the N-terminal octapeptide $\beta^{A_{Ic}}$T-1 which contained the blocking group R (cf. pp. 18, 19). Now the basis of peptide separations on Dowex-50 and Dowex-1 is differences between the ionic charges on the different peptides, though

other factors also influence the chromatography (127). Because A_{Ic} appeared to be more negatively charged than A_{II}, as evidenced by the former's chromatographic behavior on IRC-50 (cf. p. 12; and Fig. 16, p.183), there was promise of being able to isolate $\beta^{A_{Ic}}$T-1 directly by these methods.

In this section the separation of the soluble tryptic peptides from β $\beta^{A_{Ic}}$-globin on Dowex-50 X2 is described. The peptides from the α chain of A_{Ic} were separated in exactly the same manner.

II G2: Experimental

A 3.5 x 100-cm Dowex-50 X2 cation exchange column* was equilibrated at 38°C with 1,000 ml of 0.2 F pyridine-4.90 F acetic acid pH 3.1 buffer. The solution of tryptic peptides, in 25 ml pH 2.2 buffer (cf. top p. 41), was driven into the column with air pressure. The vessel which had contained the peptides was rinsed six times with 1 ml of pH 2.2 buffer (prepared from the pH 3.1 buffer by adding HCl); these rinses were used to wash down the sides of the column above the resin bed. These rinses were followed by three 5-ml rinses with pH 5.1 buffer. Each rinse was driven into the resin bed before applying the next.

The column was developed at 38°C at a flow rate of 150 ml/hr**

---

*Purified (99) 200-400 mesh resin which, when wet and in the sodium form, had passed through a 60 mesh wire screen, and from which the fines had been removed by settling the resin in water, was washed consecutively with 1 F NaOH, H_2O, 2 F HCl, H_2O, and pyridine, as described by Schroeder et al. (127). The washed resin was suspended in twice its volume of pH 3.1 buffer, and the column was poured in five 20-cm sections with this suspension.

**This flow was maintained by a Milton Roy Mini-Pump, model CH (Philadelphia, Pennsylvania), which was set for 100% stroke.

with the gradient system* described by Schroeder et al. (127). Fifteen-milliliter fractions were collected. At fraction 600 the gradient was discontinued, and development was continued for 80 fractions with 8.5 F pyridine-3.17 F acetic acid pH 5.6 buffer. At this point the chromatogram was stopped.

The locations of the tryptic peptides in the chromatographic fractions were determined by hydrolyzing 0.100-ml aliquots from the even numbered fractions with 1 ml of 2.5 F NaOH for 1.5 hr at 100°C. The samples were contained in clean polyallomer or polypropylene** centrifuge tubes during the hydrolysis. At the end of the 1.5-hr period, the hydrolysates were neutralized with 1 ml of 30% HOAc in water (v/v) and were analyzed on the Technicon Auto-Analyzer*** by monitoring the color produced on reacting the neutralized sample with ninhydrin. A plot of the chromatogram is shown in Fig. 4, p. 44.

Each chromatographic zone was evaporated in a rotary evaporator to a few milliliters, was diluted with water, and was lyophilized to dryness. At times the zones were evaporated to dryness under a stream of air.****

---

*The gradient device consisted of three 8-liter aspirator bottles which were connected in series at their base by Tygon tubing. The bottle nearest the pump was filled with 4.6 liters pH 3.1 buffer, and each of the remaining two bottles were filled with 4.6 liters 2.0 F pyridine-2.50 F acetic acid pH 5.0 buffer (127).

**Beckman Instruments, Spinco Division, Palo Alto, California.

***Technicon Instruments Corporation, Chauncey, New York. See Ref. 132 for details.

****The temperature of zone 5 should not be allowed to exceed 25°C during evaporation: a higher temperature leads to the destruction of

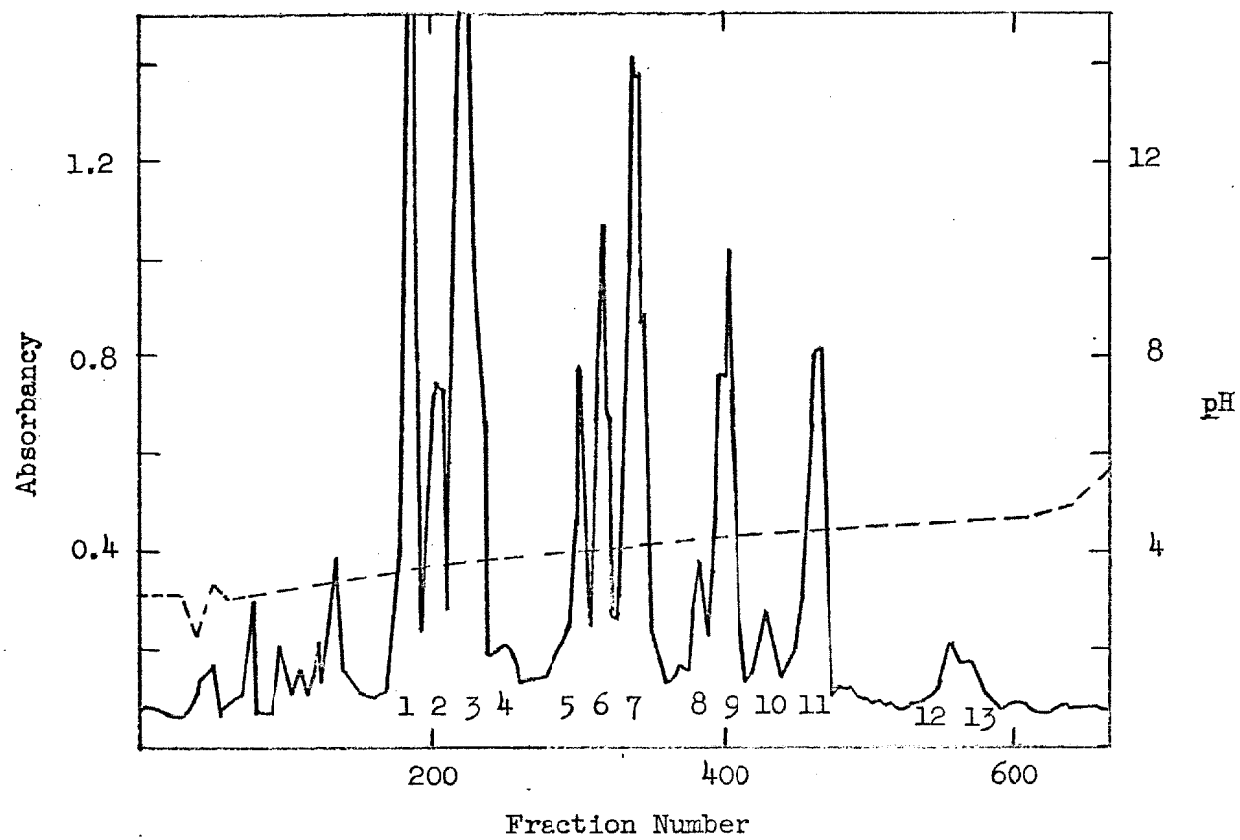


Fig. 4: Chromatographic separation of the soluble tryptic peptides from 1.7 gm of β β^{A}Ic-globin on Dowex-50 X2; ——— optical absorbancy (ninhydrin), – – – pH. The zones have been numbered consecutively from left to right. See text for discussion.

In order to locate the zone containing the N-terminal octapeptide $\beta^{A_{Ic}}T-1$, an aliquot of each of the zones in Fig. 4 was electrophoresed for 3/4 hr at 3 kilovolts on a 70 cm-length of Whatman 3MM paper. The electrophoretic buffer was pH 6.4 pyridine acetic acid 0.62 F in pyridine and 0.035 F in acetic acid (72). The sample was applied 35 cm from either end of the paper. An equal amount of each zone was paper chromatographed for 12 hr by ascending chromatography on Whatman 3MM paper with a developer composed of isoamyl alcohol, pyridine, and water in the volume ratios 7:7:6 respectively (7, 144).

II G3: Results: Detection of $\beta^{A_{Ic}}T-1$

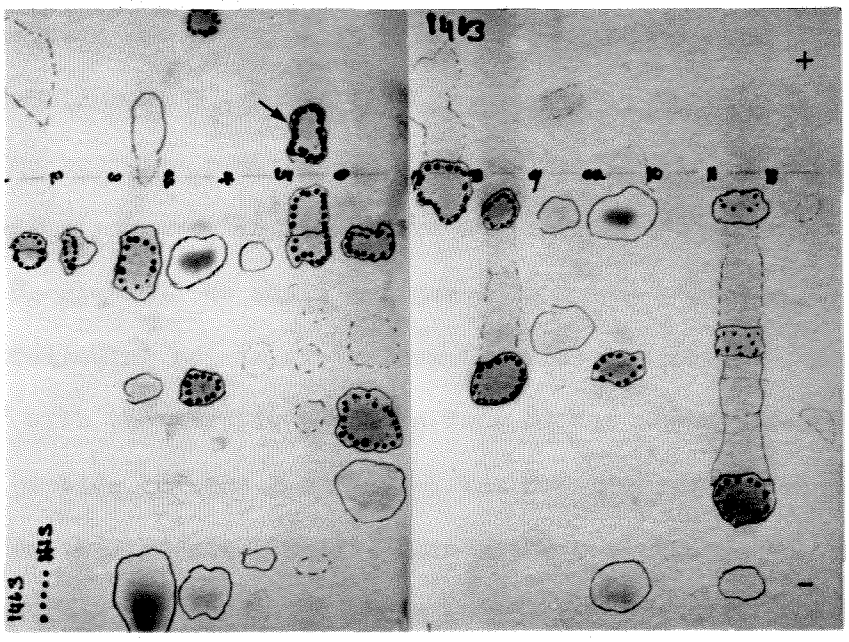
The chromatogram of the soluble tryptic peptides from 1.7 gm of $\beta\ \beta^{A_{Ic}}$-globin is shown in Fig. 4, p. 44. This chromatogram showed no obvious differences when compared with a chromatogram of the soluble tryptic peptides from the β chains of A_{II}. This fact was not surprising in view of the many similarities which exist between A_{Ic} and A_{II}.

The electrophoretic and paper chromatographic patterns of the zones from a Dowex-50 chromatogram of the typtic peptides of $\beta\ \beta^{A_{Ic}}$-globin are shown in Fig. 5, p. 46. The papers have been stained with ninhydrin and diazotized sulfanilic acid:* the latter reagent gives

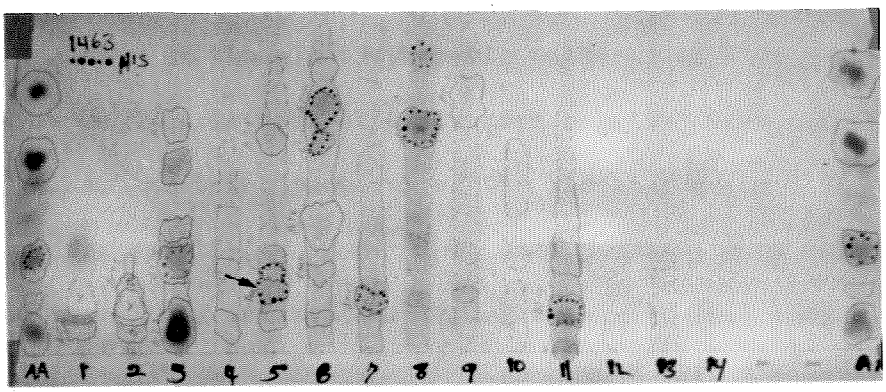
---

the blocked N-terminal octapeptide $\beta^{A_{Ic}}T-1$ (cf. Section II K, p. 67). The remaining zones were evaporated at 40°C.

*The dry paper chromatogram or electropherogram was sprayed with 0.5% ninhydrin in acetone; after the acetone had evaporated, the paper was held at 80°C for 5-10 min, or until maximum color development had occurred. The test for histidine was made after the peptides had been located with ninhydrin: The paper was heavily sprayed with an aqueous solution 0.5% in sulfanilic acid and 0.35% in $NaNO_2$. After the paper had partially dried, it was sprayed with 10% Na_2CO_3 in water. Histidine containing peptides showed up as orange spots; free histidine gave a brownish-orange color. The ninhydrin spots were bleached by the sulfanilic acid-$NaNO_2$ spray.



Electrophoretic Pattern of Zones (Fig. 4)
from Dowex-50 Chromatography



Paper Chromatography of Zones (Fig. 4)
from Dowex-50 Chromatography

Fig. 5: Electrophoretic and paper chromatographic patterns from zones (Fig. 4) from Dowex-50 chromatography of the soluble tryptic peptides of β $\beta^{A_{Ic}}$; ———, ninhydrin-positive; ....., histidine-positive; aa = amino acid standard containing leu, his, lys, glu. Zone numbers are placed along line of application of sample. Arrow points (→) to $\beta^{A_{Ic}}$T-1.

orange colored spots with peptides containing histidine. The only difference between the patterns in Fig. 5 and those obtained from the tryptic peptides of β_2-globin is the presence in Fig. 5, at zone 5, of a histidine positive peptide of low anodic mobility. This peptide is designated by an arrow in the Figure and is absent from tryptic hydrolysates of β_2-globin. We shall designate it $\beta^{A_{Ic}}$T-1 (cf. p. 19).

The following points are relevant. Previous chromatographic experience* with the tryptic peptides from β_2 had shown that the histidine positive peptide in zone 7 of Fig. 5 was β T-1, which has the structure val-his-leu-thr-pro-glu-glu-lys. Since the new peptide found in zone 5 was eluted from the Dowex-50 column before the peptides in zone 7, this peptide must have a relatively more negative net ionic charge than β T-1. This fact is in agreement with both the chromatographic behavior of A_{Ic} on IRC-50 and with the slight anodic mobility of the new peptide. However, a glance at the electrophoretic patterns of Fig. 5 suggests that the new peptide is just slightly more negative than leucine at pH 6.4.** Thus, whatever property of the blocking group R is responsible for the greater negativity of $\beta^{A_{Ic}}$T-1, that property does not seem to correspond to a fully ionized carboxyl group. The fact that $\beta^{A_{Ic}}$T-1 stained positively for histidine is in agreement with the postulated structure of this peptide: R-val-his-leu-thr-pro-glu-glu-lys (cf. p. 19). However, the positive reaction of this peptide with ninhydrin was

---

*Cf. Schroeder et al. (127): zone 10 of Fig. 4 in that reference corresponds to zone 7 of Fig. 4, p. 44, this thesis.

**Glutamic acid lies outside of the field of view in the electrophoretic patterns shown in Fig. 5. Its position corresponds to that of lysine, but in the anodic direction.

surprising, unexpected, and therefore important: a primary amino group is essential for a positive ninhydrin reaction. If the N-terminus of $\beta^{A_{Ic}}T$-1 is blocked by a group R as hypothesized, how can this peptide react with ninhydrin?! We shall return to this point much later* (pp. 67, 142).

II G4: Discussion

Figure 5 shows that there were 16-18 peptides which were applied to the Dowex-50 column.** These peptides were separated into ten major zones which contained an average of three peptides each. This represents a 6-fold increase in the purity of each individual peptide with respect to the starting mixture. The average yield of each peptide, based on the weight of the lyophilized zones was 81%.

However, the above yield is misleading, because there was a moderate to large contamination of each chromatographic zone by unidentified solid material from the Dowex-50 resin, which is apparently soluble to some extent in the pyridine-acetic acid developers. Actual yields were often nearer 40%-60%. The problem of contamination by resin material was particularly troublesome in the case of the 3.5 x 100-cm chromato-

---

*The importance of this fact became evident only in retrospect after much additional data, which were consistent with it, had been accumulated. These data will be presented as the thesis is developed.

**From the amino acid analysis of $\beta\ \beta^{A_{Ic}}T$-1 (cf. p. 99a), and the specificity of trypsin for lysyl and arginyl peptide bonds, at most 16 of these peptides are tryptic peptides, even if contamination of the soluble tryptic peptides by minor amounts of the tryptic peptides insoluble at pH 6.4 is allowed for. The remaining peptides are nonspecific cleavage products arising during the various procedures. Quantitatively, their amount is small relative to the amount of tryptic peptides.

graphic column (cf. p. 42). Because of the large amounts of buffer which were needed to develop this column, the peptides in the effluent were present in large volumes of buffer; and these large volumes carried with them correspondingly large amounts of resin material. For example, the peptide $\beta^{A_{Ic}}$T-1 from zone 5 of Fig. 4 was present in the residue from 300 ml of buffer.

A second source of contamination was impurities possibly present in the buffers themselves. Although distilled pyridine and reagent grade acetic acid were used in preparing the buffers, the presence of trace amounts of impurities was not excluded. When large volumes of such buffers are involved, the total amount of impurities is appreciable.*

Usually, even large amounts of nonprotein contaminants do not seriously interfere with subsequent purification and characterization of a peptide, for such contaminants seldom attack the peptide in such a way as to destroy its chemical identity. Further, in most instances, these contaminants are removed in subsequent purification steps; and even if they are not, the structure of the peptide can usually be es-

---

*As an illustration, roughly equivalent to the actual chromatographic conditions, the residue from 300 ml of pyridine-acetic acid buffer which is 0.6 F in pyridine, and which was prepared from 99.99% pure distilled pyridine (This degree of purity is entirely hypothetical; actual distillates are probably less pure.), contains 18 μm of impurities from the pyridine alone. This amount is about equal to the μm of a peptide present in one of the chromatographic zones in Fig. 4. Even if the impurities are volatile, their concentration could be large enough during the evaporation to effect chemical changes on other compounds which might be present; this would be particularly true in the presence of acetic acid, in which case acid catalysis would be possible. This catalysis would be even more enhanced in the presence of any type of chromatographic resin because of the large surface area that is presented to the reactants by the resin.

tablished by the use of reagents, such as ninhydrin, which have a specificity for certain aspects of the peptide structure, but not for the contaminants.

In the subsequent sections it will be demonstrated that the structure of $\beta^{A_{Ic}}$T-1, viz. R-val-his-leu-thr-pro-glu-glu-lys, is such as to make it extremely susceptible to certain types of contaminants, even as mild a "contaminant" as water. Here we merely want to establish the fact that the presence in zone 5 of Fig. 4 of contaminants from the Dowex-50 chromatography made the purification of $\beta^{A_{Ic}}$T-1 difficult.

II H: FINAL PURIFICATION OF THE TRYPTIC PEPTIDES SOLUBLE AT pH 6.4

II H1: Introduction

Schroeder et al. (127) and Babin et al. (6) had shown that the mixture of tryptic peptides in the chromatographic zones which had been isolated from Dowex-50 X2 columns could be separated into purified individual peptides on the quaternary amine anion exchange resin Dowex-1 X2. However, when the attempt was made to purify the aberrant N-terminal octapeptide $\beta^{A_{Ic}}$T-1 from the other peptides in zone 5 of Fig. 4 by this method, it was quantitatively converted into βT-1, the normal octapeptide; i.e., the blocking group R had been removed from $\beta^{A_{Ic}}$T-1 during the Dowex-1 chromatography.* As has been stated earlier

*The loss of the blocking group R was demonstrated by the facts that only a single peptide having the amino acid composition (val, his, leu, thr, pro, glu_2, lys) was isolated from the Dowex-1 chromatogram, and that on electrophoresis at pH 6.4 this peptide had a cathodic, rather than anodic, mobility. This mobility corresponded to that of βT-1, the histidine positive peptide from zone 7 of Fig. 4 (cf. Fig. 5). That the latter peptide was in fact βT-1 was shown by its subsequent isolation and structure determination.

(p. 19), it was absolutely essential that we were able to isolate R still attached to the peptide. Consequently, we were forced to consider other methods of purifying this peptide.

The simplest explanation for the loss of R from $\beta^{A}I_{c}T$-1 was that this peptide was not stable under the conditions of Dowex-1 chromatography: viz: in the presence of a quaternary amine anion exchange resin and of the basic pyridine-collidine buffer (6) with which the Dowex-1 chromatogram was developed. Because of this, we limited our search for a better chromatographic system to cation exchange resins which could be developed with pyridine-acetic acid buffers at acidic pH's.

Attempts to resolve a mixture of amino acids on IRC-50, a carboxylic acid resin, soon demonstrated that this resin was not likely to have the resolving power which would be necessary to separate a mixture of peptides.

Canfield and Anfinsen (19) had successfully separated peptides from egg white lysozyme on cellulose phosphate. When modified as described below, this chromatographic system was capable of separating $\beta^{A}I_{c}T$-1 in better than 90% yield from the other peptides which were present in zone 5 of Fig. 4, without the loss of the blocking group R.

At this point, because the laboratory had not had prior experience with cellulose phosphate chromatography, the decision was made to attempt to purify many of the tryptic peptides from the Dowex-50 chromatogram on this material in order to determine its capabilities. This attempt was successful.

II H2: Chromatography on Cellulose Phosphate

Preparation of Column

Cellulose phosphate* was sized into a 250-325 mesh fraction by dry sifting. This fraction was suspended in an equal volume of water, filtered onto a Buchner funnel, and allowed to drain under slight suction. It was washed successively with 100 ml of 1 F NaOH, 200 ml of H_2O, 100 ml of 1 F HCl, and 300 ml of H_2O. The washed material was suspended in twice its volume of water and allowed to settle. The fine particles were decanted. The suspension and decantation of fines was repeated a second time. Finally the settled cellulose phosphate was suspended in three times its volume of buffer B, Table I, and a 0.9 x 100-cm chromatographic column was poured in a single section with this suspension.**

TABLE I

Buffer Compositions for Cellulose Phosphate Chromatography

Buffer	pH	Ionic Strength†	Ml Pyridine/l	Ml HOAc/l
A	3.93	0.025	2.00	8.42
B	3.93	0.05	4.05	15.00
C	3.93	0.10	8.10	30.00
D	4.55	0.10	10.06	12.80
E	5.50	0.19	42.00	12.97
F	6.12	0.10	83.00	5.60
G	6.12	0.20	166.00	11.38
H	6.12	0.45	390.00	25.70

†Calculated from pK values.

*Whatman Cellulose Phosphate Powder P70.

**In buffer B, cellulose phosphate had a fibrous texture. This gave the suspension a tendency to form agglomerates of cellulose phosphate along the length of the column during the pouring. The formation of these agglomerates was easily avoided by thinning the suspension with additional buffer B.

Equilibration and Loading of Column

The column was equilibrated at 28°C with 250 ml of buffer B. The mixture of peptides* from a given Dowex-50 chromatographic zone was dissolved in 3 ml of buffer A and driven into the column with air pressure. The sample container and the sides of the column were rinsed three times with 1 ml of buffer B; each rinse was driven into the column before the next was applied.

Development of Column

With the exception of zone 5 of Fig. 4, the peptides in a given zone from Dowex-50 were separated as follows. For zone 5, a slightly modified system of higher resolving power was used because we wanted to obtain $\beta^{A_{Ic}}{}_{T-1}$ in the highest purity possible. This modified system is described in the next section.

The column was developed at 28°C with a two-vessel gradient device (15) in which the ratio of the areas of the mixer and reservoir was 2:1 respectively. The mixer and reservoir were filled with 232 ml and 116 ml of buffer B and C (Table I) respectively. A flow of 50 ml/hr was maintained (cf. second footnote, p. 42) and 4.4-ml fractions were collected. At the 80th fraction, any buffer which remained in the gradient device was removed with a suction line, and the mixer and reservoir were refilled with 440 ml and 220 ml of buffer C and G re-

---

*A mixture containing up to 50 μm of each of several peptides can be chromatographed on a 0.9 x 100-cm cellulose phosphate column. The actual capacity of the column may be larger than this. We did not have sufficient material available to determine the maximum capacity. For chromatographing quantities less than 15 μm, a column 0.6 x 60 cm is more convenient to use.

spectively. At fraction 220 the gradient was discontinued, and development was continued with 300 ml buffer H. The chromatogram was finished at fraction 260.

Modified Development of Column for Purifying $\beta^{A_{Ic}}T$-1

The sample (zone 5, Fig. 4) was loaded in the usual manner with buffer A onto a column which had been equilibrated with buffer D. The ratio of the cross-sectional areas of the mixer and reservoir was 1:2. The mixer and reservoir were filled with 366 ml and 732 ml of buffer D and E respectively. Gradient development with buffers D and E was continued until fraction 220. At this point the gradient was discontinued, and development was continued to fraction 260 with buffer H. Flow rate and fraction size were the same as for the normal procedure.

Re-equilibration of Column

The column was re-equilibrated with buffer B immediately after a chromatogram was finished, because on prolonged standing in buffer H, the cellulose phosphate tended to decompose. This was evidenced by an irreversible shrinkage of the column as well as by increased pressure within the column during subsequent chromatograms. Even with this precaution, a column eventually developed sufficient resistance to flow so that it was necessary to discard it and pour a new column with freshly prepared cellulose phosphate. It is likely that the useful life of a column could be extended by operating it at reduced temperatures.

Results

Normal Gradient

Figure 6, p. 56, summarizes the chromatographic behavior on cellulose phosphate of several tryptic peptides from the α and β $\beta^{A_{Ic}}$ chains of A_{Ic}. This figure was composed from 18 individual chromatograms. Because of slight variations between chromatograms the actual position of emergence of a particular peptide may vary somewhat from the position shown in Fig. 6. A correlation between the length and ionic charge of the peptide and its position of emergence is given on p. 55a. With some exceptions, the longer more negative peptides emerge prior to the shorter more positive ones. More particularly, peptides which contain more than 6 amino acid residues and an ionic charge no greater than 2 emerge prior to pH 4.7 irrespective of the sign of the ionic charge. Peptides less than 7 amino acid residues in length and which bear a net positive charge emerge after pH 4.7. Some peptides which contain histidine appear to be anomalously retarded, αT-9, αT-4, βT-7, for example. βT-7 is an extreme case: the yield of this peptide from the cellulose phosphate column was 0%; however, we continued to use the column on which we had chromatographed βT-7 to purify other peptides, and were very much surprised when several chromatograms later βT-7 was found in the eluate. The excessive retardation of this peptide may be due to the combination of its short length, large positive charge, and the presence of histidine.

The correlation of peptide structure and chromatographic behavior shown in Fig. 6 and Table II is very similar, but not identical, to that found by Schroeder et al. (127) for Dowex-50 chromatography. Even

TABLE II

Correlation between Peptide Length, Charge, and Chromatographic Mobility on Cellulose Phosphate

Peptide[c]	L[a]	Q[b]	Peptide[c]	L[a]	Q[b]
βT-5	19	-2			
βT-13	12	0			
βT-3	13	-2	αT-1, αT-11	7	0
			αT-5	9	1
βT-9	16	$-\frac{1}{2}$	αT-9	29	$-1\frac{1}{2}$
βT-14	12	$1\frac{1}{2}$			
βT-4	10	1	αT-14	2	1
βT-2	9	1			
			αT-4	15	$-1\frac{1}{2}$
βT-6	2	1	αT-2	4	1
βT-8	1	1	αT-8	1	1
			αT-10	2	1
βT-15	2	$\frac{1}{2}$			
βT-7	4	$1\frac{1}{2}$			

[a]L is the number of amino acid residues in the peptide.

[b]Q is an ionic charge assigned to the peptide and was calculated by assigning a value of -1 to glutamic acid and aspartic acids, +1 to lysine and arginine, $+\frac{1}{2}$ to histidine, and 0 to all other amino acids.

[c]*The peptides are designated according to the footnote on p. 19 of* this thesis. Their amino acid compositions, from which Q was calculated, can be found on pp. 194 and 195. The peptides are arranged on this page, from top to bottom, in descending order of their chromatographic mobilities on cellulose phosphate.

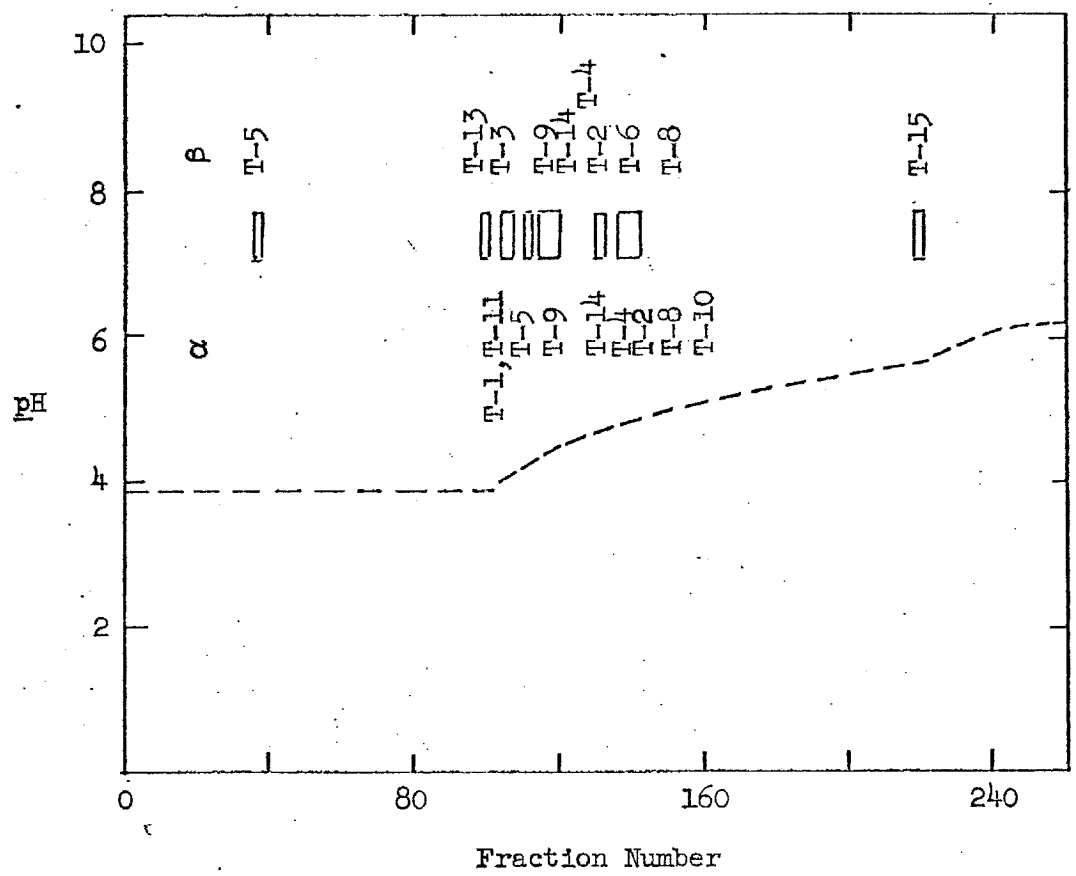


Fig. 6: Cellulose phosphate chromatography of tryptic peptides from the α and β chains of A_{Ic}; – – – pH of effluent fractions; the boxes show the approximate position of emergence of the peptides listed above and below them. See text for discussion.

the anomalies occur for the same peptides. This similarity suggests that peptides which are difficultly separable on Dowex-50 by the method of Schroeder et al. might be purified on that resin by rechromatography with the appropriate modifications of the buffers of Table I. This has indeed been found to be the case (cf. pp. 64-67, and p. 149).

The usefulness of cellulose phosphate chromatography, as described above, can be measured by the purity of the peptides which were obtained as well as by the yields of these peptides. Tables VII and VIII (cf. pp. 101a-103) show that the purity of the peptides was generally satisfactory, although for some peptides, such as αT-5, it is probable that by slightly modifying the gradient a purer peptide could be obtained. However, the yields of the purified peptides were less satisfactory. The average yield was 53%, but the yield for different peptides was highly variable: βT-9 and βT-7, for example were obtained in yields of 92% and 0% respectively. Generally the better yields were obtained with acidic peptides, and the poorer yields with basic peptides.

Two major difficulties encountered with cellulose phosphate chromatography were double zoning and increased resistance to flow in the column as it was repeatedly used. The latter effect has already been discussed (p. 54). The former was manifested by two zones, one in the position shown in Fig. 6, and the other displaced from it by several tens of fractions. The practical effect of double zoning was to lead to contamination of one zone by another when the zones were otherwise well separated. For example, βT-13 and βT-9 have been found in the vicinity of βT-5, and vice versa. Double zoning has been observed by

others on Dowex-1 (127); however, it has been our experience that the problem is more troublesome in the case of cellulose phosphate.*

A noteworthy advantage of cellulose chromatography is the lack of contamination of the effluent by cellulose phosphate itself. In fact, resin contaminants from Dowex-50 chromatography (cf. p. 48-50) were to a large extent removed by subsequent chromatography on cellulose phosphate.

Modified Gradient

The purification of $\beta^{A_{Ic}}$T-1 from zone 5 of a Dowex-50 chromatogram is shown in Fig. 7, p. 59. The zone labeled βT-1 at fraction 82 is not usually present in zone 5. Here, it arose by the loss of R from $\beta^{A_{Ic}}$T-1 <u>prior</u> to chromatography, because cellulose phosphate chromatography of a known amount of $\beta^{A_{Ic}}$T-1 gave a recovery of 92%. βT-9 is a normal constituent of zone 6 (cf. Fig. 4, p. 44) and frequently overlaps into zone 5.

When zone 5 from the tryptic peptides of 1.7 gm of β $\beta^{A_{Ic}}$-globin was chromatographed with the modified gradient, 20 μm of purified $\beta^{A_{Ic}}$T-1 was found in the effluent. However, after the effluent had been concentrated at room temperature on a rotary evaporator and evaporated to dryness with a stream of air, the dry residue contained <u>no</u> $\beta^{A_{Ic}}$T-1; rather, it contained 20 μm of βT-1, the normal N-terminal

---

*There is no reason to believe that the difficulties which have been discussed in this paragraph are insurmountable. Our investigation of the properties of cellulose phosphate has been limited to the chromatographic systems which were described on previous pages. Variations in buffer compositions, temperature, etc., might reduce or eliminate some or all of these difficulties.

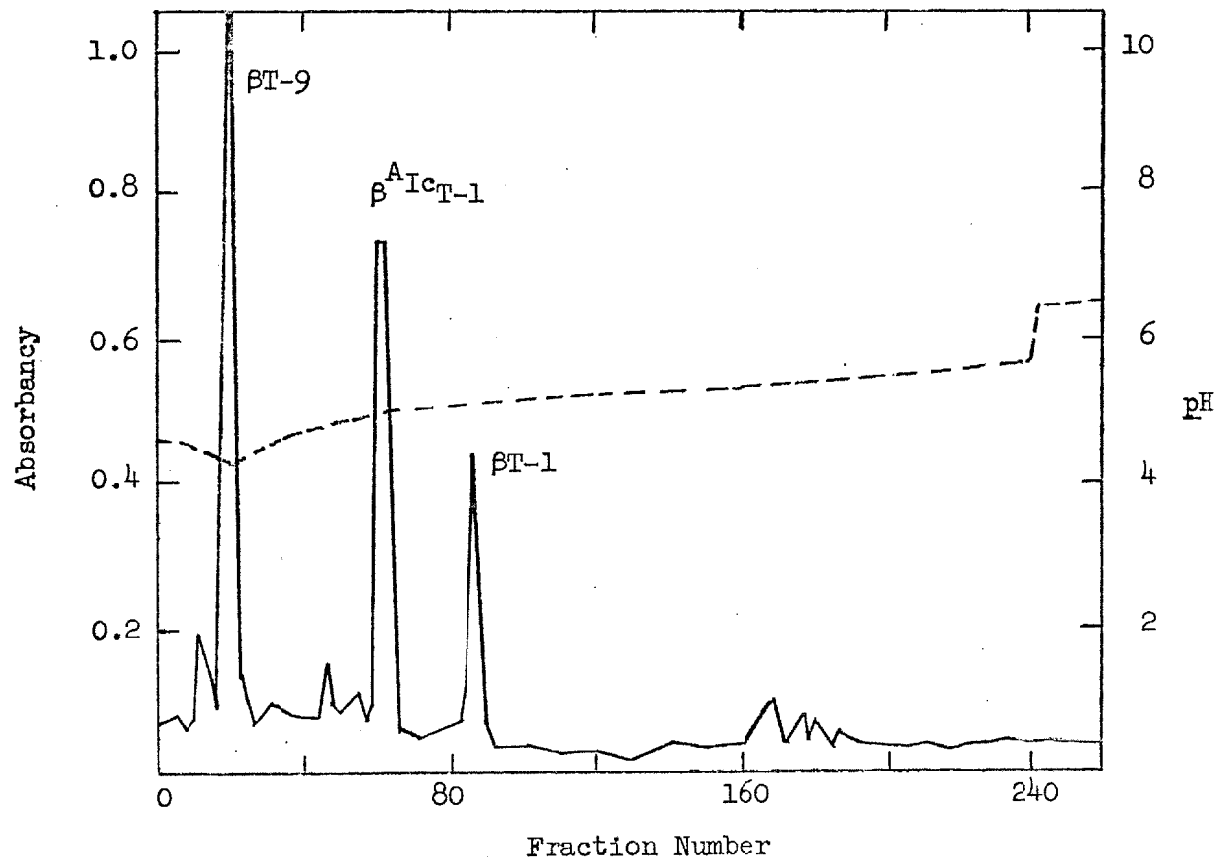


Fig. 7: Separation of the tryptic peptides from zone 5 of Fig. 4, p. 44, on cellulose phosphate with modified gradient; ——— absorbancy (ninhydrin), – – – pH of effluent fractions. The βT-1 in this chromatogram arose by the loss of R from $\beta^{A}Ic_{T-1}$ prior to chromatography; normally it is not present in this zone. See text for discussion.

octapeptide: that is, the blocking group R had been quantitatively removed from $\beta^{A_{Ic}}T-1$.

This catastrophe was the end result of six months of experimental effort. It was totally unexpected, because careful experiments with smaller amounts of material on 0.9 x 100-cm and 1.8 x 100-cm Dowex-50 columns, and on 0.6 x 60-cm and 0.9 x 100-cm cellulose phosphate columns respectively, had resulted in the successful isolation of $\beta^{A_{Ic}}T-1$. The procedures which have been described in the preceding sections differed from the small scale experiments only in their increased scale, and in particular the 4-fold increase in scale which resulted from Dowex-50 chromatography on a 3.5 x 100-cm column. We believe the cause of the loss of the blocking group was due to excessive resin or buffer contaminants from the large scale Dowex-50 chromatography (cf. pp. 48-51). When the effluent, which contained purified $\beta^{A_{Ic}}T-1$, from the cellulose phosphate chromatography was evaporated to dryness with a stream of air (cf. p. 58), the evaporating solution underwent a series of color changes from colorless, to pale green, to blue-green, to brown, to very dark brown. The dry residue was a dark brown and dissolved in water to give a very dark brown solution. This sequence of color changes had never been observed in smaller scale experiments.

The misfortune which has just been described is the major reason the complete structure of the blocking group R is not known today. The chemical characterizations which we have accomplished, and which are described later in the thesis, were made possible by small scale isolations of $\beta^{A_{Ic}}T-1$.

This section concludes the description of the final purification of the soluble tryptic peptides. Before describing the degradation of $\beta^{A_{Ic}}$T-1 into smaller peptide fragments, we shall describe the isolation and purification of the tryptic peptides which were insoluble at pH 6.4 (cf. p. 41, top).

II I: PURIFICATION OF THE TRYPTIC PEPTIDES INSOLUBLE AT pH 6.4

The peptides from the β chains of A_{II} which are insoluble at pH 6.4 were known to be βT-10,11 and βT-12* since earlier workers had purified these peptides by first oxidizing the cysteine residues which they contain to cysteic acid residues (55), and then separating the mixture of the two peptides on Dowex-50 (37). Because it turned out that the insoluble tryptic peptides from β $\beta^{A_{Ic}}$ were identical to those from A_{II}, we were able to purify these peptides successfully in the same manner.

II I1: Performic Acid Oxidation

Approximately 160 mg of the insoluble peptides from the tryptic hydrolysis of β $\beta^{A_{Ic}}$ (cf. p. 40, bottom) were dissolved in a mixture of 2.5 ml of formic acid and 0.5 ml of methanol in a 25-ml Erlenmeyer flask. A mixture of 0.125 ml of 30% H_2O_2 and 2.38 ml of 99% formic acid were allowed to stand in a test tube at room temperature for 2 hr. At the end of this time both the Erlenmeyer flask and the test tube were cooled to -10°C. The contents of the cooled test tube were poured into the flask which was then kept at -10°C for 3 additional hours.

---

*For nomenclature cf. footnote p. 19.

At the end of this time the contents of the flask were poured into 200 ml of water and were lyophilized to dryness. Fifty milliliters of water were added to the dry residue, and the solution was again lyophilized to dryness.

II I2: Isolation of Purified βT-10,11 and βT-12 on Dowex-50

The mixture of peptides obtained from the performic acid oxidation were purified by the method of Goldstein *et al.* (37). A 0.9 x 22-cm Dowex-50 X2 cation exchange column was equilibrated at 50°C with 90 ml of 1.07 F pyridine-0.50 F acetic acid *pH* 5.52 buffer.

The column was loaded with 16 mg of the mixture of performic oxidized peptides which had been dissolved in 1.0 ml of the 1.07 F *pH* 5.52 buffer and adjusted to *pH* 1.5* with HCl.

The column was developed for 60 fractions with the 1.07 F pyridine *pH* 5.52 buffer. The flow rate was 31 ml/hr, and 2-ml fractions were collected. At fraction 60, development with the 1.07 F buffer was stopped, and development was continued to fraction 140 with 2.00 F pyridine-1.00 F acetic acid *pH* 5.52 buffer.

II I3: Results

Peptide βT-10,11 emerged in the effluent between fractions 10 and 20. βT-12 emerged between fractions 90 and 100. Goldstein *et al.* (37) presents these results graphically.

---

*If the *pH* was greater than *pH* 2, there was much precipitation.

II J: DEGRADATION OF βT-1 AND $\beta^{A_{Ic}}$T-1 WITH PAPAIN: ISOLATION OF VAL-HIS AND R-VAL-HIS

II J1: Introduction (cf. p. 19)

The N-terminal octapeptide βT-1 of the β chain of A_{II} is known to have the amino acid sequence val-his-leu-thr-pro-glu-glu-lys. Hill *et al.* (54) had hydrolyzed this octapeptide with papain; among the peptide fragments from the papain hydrolysis they found the dipeptide val-his. If an analogous degradation of the N-terminal octapeptide $\beta^{A_{Ic}}$T-1 of the $\beta^{A_{Ic}}$ chain of A_{Ic} could be achieved--and this appeared likely in view of the fact (cf. Table VIII, p. 102a) that $\beta^{A_{Ic}}$T-1 and βT-1 had the same amino acid composition--the resulting dipeptide R-val-his might be small enough to permit the determination of its complete structure by infrared spectroscopy, nuclear magnetic resonance spectroscopy, mass spectroscopy, and modern chemical methods of classical structure analysis.

II J2: Papain Hydrolysis of βT-1 and $\beta^{A_{Ic}}$T-1

To 0.4-12.0 μm of the octapeptide in 0.100 ml of H_2O in a small test tube were added 0.050 ml of 0.1 F 2,3-dimercaptopropanol, 0.010 ml of 0.2 F *p*H 4.25 NaAc-HOAc* buffer, and 0.040 ml of papain as a 15 mg/ml aqueous solution. The hydrolysis was allowed to proceed at 38°C for 24 hr at which time the contents of the test tube were evaporated to dryness with a stream of air and stored until needed.

---

*Prepared by dissolving 1.64 gm of NaAc (or 2.70 gm of $NaAc \cdot 3H_2O$) and 2.63 ml of HOAc in 100 ml of H_2O.

**Purchased from Nutritional Biochemicals Corporation, Cleveland, Ohio.

II J3: Electrophoretic Detection of Val-his and R-val-his in the Papain Hydrolysate

The papain hydrolysates from βT-1 and $\beta^{A_{Ic}}$T-1 were examined electrophoretically at pH 6.4 as described on p. 45. The electrophoretic patterns are shown in Fig. 8, which is taken from Reference 57. Preliminary identification of the peptides indicated as R-val-his and val-his was made by electrophoretic and paper chromatographic comparison with authentic samples* of val-his and acetyl-val-his. Final identification was made after the peptides from the papain hydrolysate had been purified on Dowex-50 X8.

II J4: Isolation of R-val-his

The peptides from the papain hydrolysis of $\beta^{A_{Ic}}$T-1 were separated from each other by a modification of the method devised by Goldstein et al. (38). These investigators, using the buffers of Konigsberg and Hill (80), succeeded in separating a mixture of the peptides in Fig. 8 by two chromatograms on Dowex-50 X2. By using Dowex-50 X8, and buffers of lower ionic strength, we have resolved these peptides on a single chromatogram.**

---

*The synthesis of val-his and acetyl-val-his is given in Appendix II, p. 201.

**The difference between Dowex-50 X2 and Dowex-50 X8 polystyrene sulfonate type cation exchange resins is that the former is 2% cross-linked and the latter 8%. Discussions with Richard T. Jones had suggested that peptides difficult to separate on the former resin were sometimes readily separable on the latter. We wanted to be able to separate a mixture of R-val-his from val-his, if it should have proved necessary. The similar structures of these two peptides suggested that they might be difficult to separate, and it is for this reason that Dowex-50 X8 was chosen. The same consideration was responsible for the choice of low ionic strength buffers. At high ionic strengths, the interaction between the resin and peptides is reduced. By increasing the

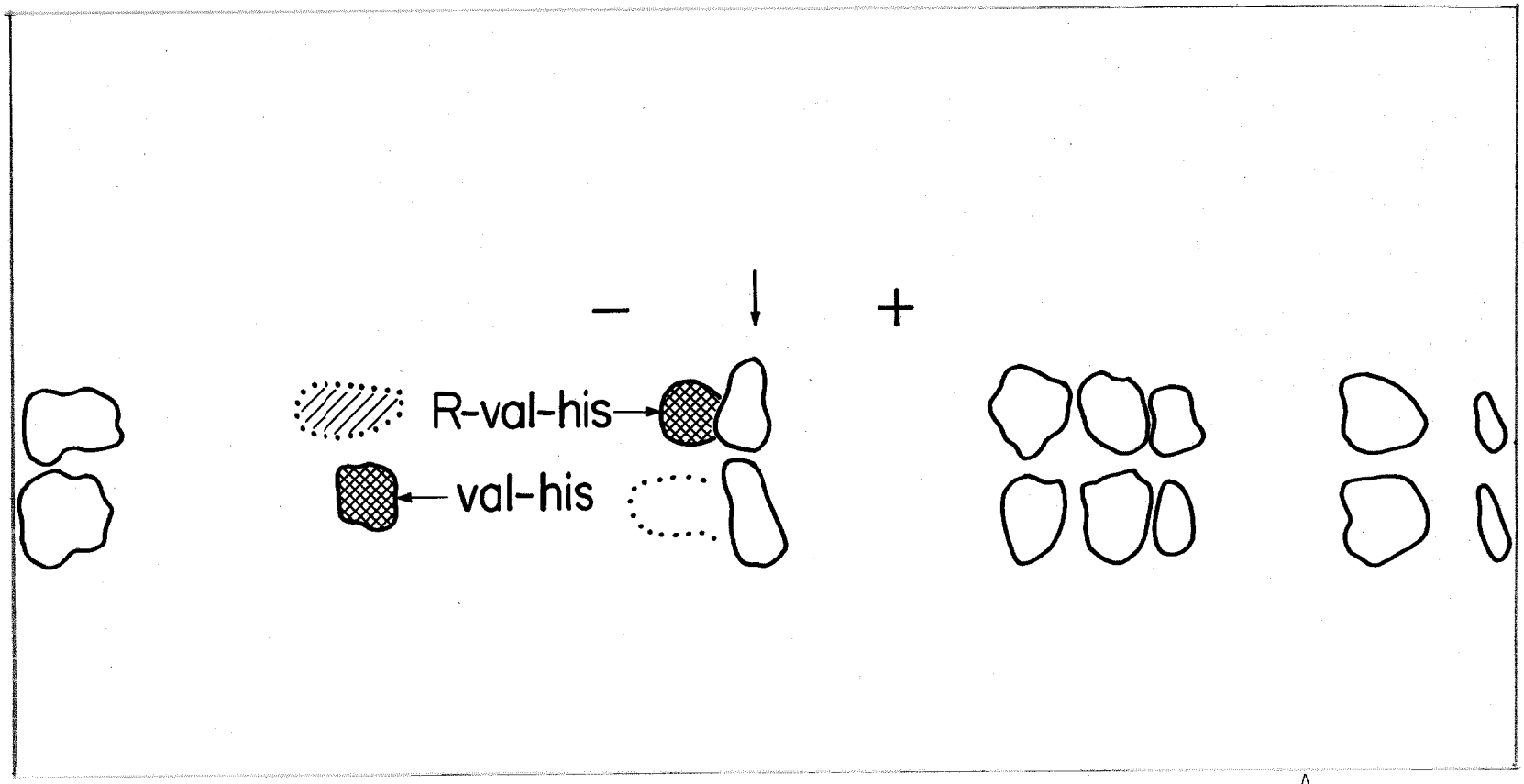


Fig. 8: Paper electrophoretic patterns at pH 6.4 of papain hydrolysates of $\beta^{A_{Ic}}{}_{T-1}$ (upper) and βT-1 (lower). The arrow shows the line of application. Solid and dotted lines outline spots that gave a strong and very faint color, respectively, with ninhydrin. The crosshatched spots reacted very strongly for histidine, and the singly hatched spot very weakly for histidine.

A 0.6 x 45-cm Dowex-50 X8 cation exchange column* was equilibrated at 38°C with 200 ml pH 3.28 buffer (Table III). The mixture of peptides

TABLE III

Buffer Compositions for Dowex-50 X8 Chromatography

Buffer	pH	Ionic Strength†	Ml Pyridine/l	Ml HOAc/l
A	2.41	0.004	0.00	62.5
B	3.28	0.05	4.03	61.80
C	4.21	0.05	4.59	10.40
D	5.28	0.05	9.42	4.20
E	5.28	0.10	18.85	7.80
F	5.64	0.30	686.	181.

†Calculated from pK values

from the papain hydrolysis of 12 μm** of $\beta^{A_{Ic}}$T-1 was dissolved in 0.4 ml of buffer A and was driven into the column with air pressure. This was followed by three 0.2-ml rinses with buffer B. The column was developed stepwise: development with buffer B was continued to fraction

---

interaction by lowering the ionic strength of the chromatographic buffers, increased resolution between peptides is frequently obtained. The weaker buffers also result in lesser contamination of the isolated peptides with any trace impurities which may have been present in the reagent grade chemicals from which the buffers were made.

*Purified (99) resin was prepared from Amberlite CG-120, Type II (finer than 200 mesh) resin. The purified resin was sized into 25-30 μ particles by hydraulic flotation (49, 100). The 25-30 μ fraction was washed consecutively with 1 F NaOH, H_2O, 2 F HCl, H_2O, and pyridine (127), and was then suspended in twice its volume of pH 3.28 buffer (Table III). The column was poured in five 10-cm sections with this suspension. The first three sections were settled by pumping pH 3.28 buffer through the column at 15 ml/hr; the latter two sections were settled at 10 ml/hr.

**The capacity of the column is probably considerably larger than this.

100; at fraction 100, buffer B was replaced with buffer C; at fraction 200, C was replaced with D; at fraction 300, D with E; and at fraction 400, E with F. The chromatogram was finished at fraction 500. A flow rate of 10 ml/hr was maintained, and 1-ml fractions were collected.

Figure 9, p. 68, shows the separations that were obtained with the above system. Val-his emerges from the column after lysine (38), so if a mixture of R-val-his and val-his had been present, it would have been readily resolved by the chromatography. The separation between glu-lys and R-val-his was approximately five fractions; amino acid analysis of these two peptides showed that neither was contaminated by the other. The yield of R-val-his, based on the amount of $\beta^{A_{Ic}}$T-1 which was taken for papain hydrolysis, was 28%.

II K: THE LABILITY OF THE BLOCKING GROUP R

In this section we wish to draw together several miscellaneous observations which were made during the course of the isolation procedures and which demonstrate the lability of the bond between the blocking group R and the N-terminus of the $\beta^{A_{Ic}}$ polypeptide chain. Several of these observations have been described in previous sections and are repeated here for emphasis.

First, each of the compounds $\beta^{A_{Ic}}$T-1 and R-val-his reacted positively with ninhydrin at 80°C; however, at room temperature only a weak positive reaction with ninhydrin occurred. Second, when $\beta^{A_{Ic}}$T-1 was chromatographed on the quaternary amine anion exchange resin Dowex-1, the blocking group R was quantitatively cleaved from the N-terminal valine of this octapeptide. Third, when $\beta^{A_{Ic}}$T-1 was obtained from the

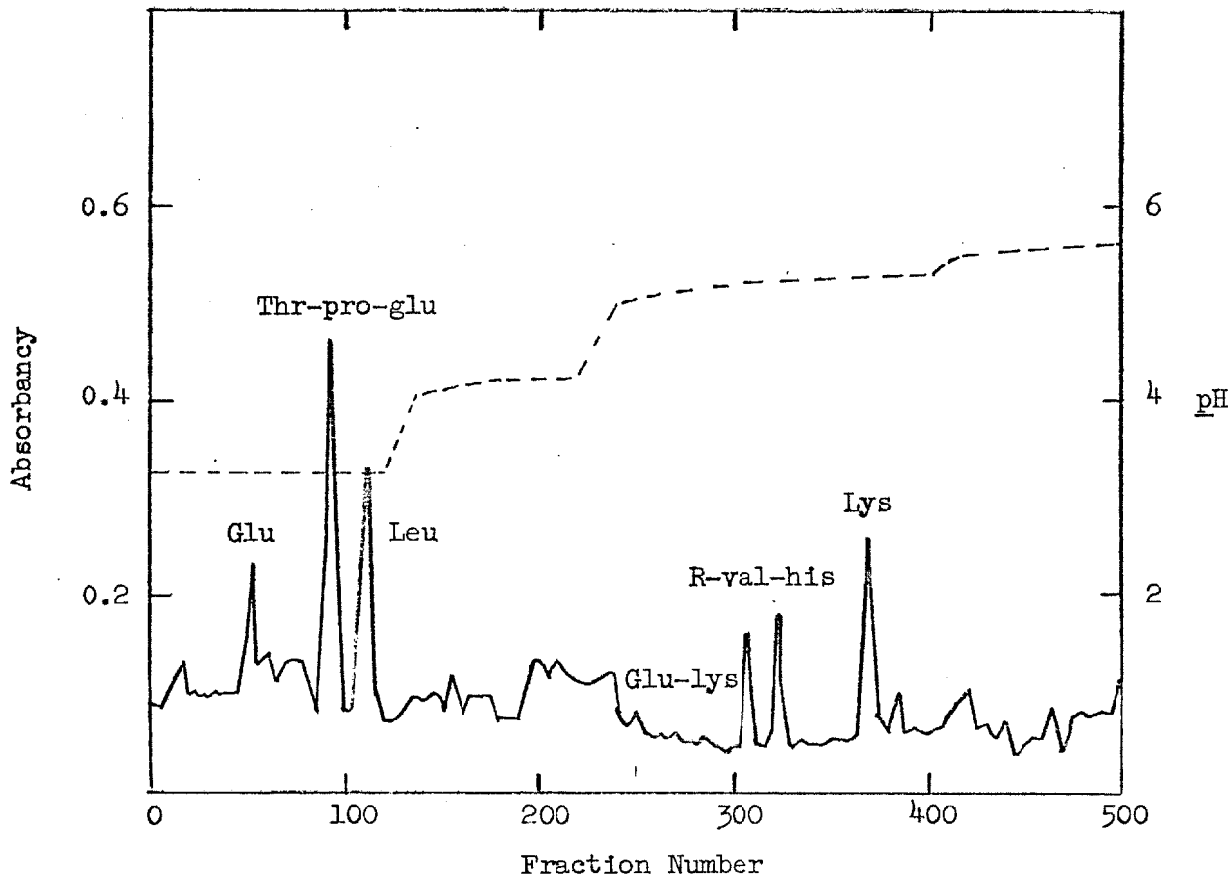


Fig. 9: Separation of peptides from papain hydrolysis of $\beta^{A}Ic_{T}$-1 on Dowex-50 X8; ——— Absorbancy (ninhydrin), – – – pH of effluent fractions. See text for discussion.

evaporation of large quantities of buffers, the blocking group R was cleaved. Fourth, three experiments were done in which 0.2 μm R-val-his was heated at 38°C for 16 hr with 1 ml of water, 1 ml of 0.0001 F (pH 10.08) NaOH, and 1 ml of pH 8.92 pyridine-collidine-acetic acid buffer (127),* respectively. After the three solutions had been evaporated to dryness at 38°C with a stream of air, analysis of the residues by electrophoresis showed that in the first two cases the blocking group R was quantitatively cleaved from R-val-his to give val-his; in the third case R-val-his was unaltered.** Fifth, in the course of chemically characterizing $\beta^{A_{Ic}}$T-1 and R-val-his we took many aliquots of these compounds from the test tube which contained the purified compound. This was accomplished by dissolving the dry compounds in a tenth of a milliliter or so of water, removing the aliquot, and evaporating the remaining solution to dryness at 25°C. Each time an aliquot was taken in the above manner, some of the $\beta^{A_{Ic}}$T-1 or R-val-his which remained in the test tube was converted into βT-1 or val-his respectively.

It is interesting that the degree of cleavage of R in the above cases appears to increase as the molecular weight of the peptide to which R was attached decreases: there was very little, if any, loss of R from hemoglobin A_{Ic} or β $\beta^{A_{Ic}}$-globin; on the other hand, loss of R from $\beta^{A_{Ic}}$T-1 and R-val-his was frequently the factor which limited the yields of these compounds.

---

*This buffer was the same one with which the Dowex-1 chromatogram had been developed.

**The failure of the pyridine-collidine-acetic acid buffer alone to cleave the R group implies that Dowex-1 resin itself, in the presence of this buffer, is the major cause for the loss of the R group in Dowex-1 chromatography.

It is also interesting that the peptides which contain R are relatively stable below pH 5 regardless of their molecular weight.

The relevance of the above facts for the structure of R will be discussed in Section VII, p. 142. From the point of view of the isolation procedures, the practical result of the lability of R is to reduce the overall yield of R-val-his.

II L: OVERALL YIELD OF R-VAL-HIS FROM THE ISOLATION PROCEDURES

The yield of hemoglobin in A_{Ic} from Bio-Rex 70 chromatography of whole blood lysates was about 95%. Dehemed A_{Ic}-globin was obtained from hemoglobin A_{Ic} in 98% yield. Countercurrent distribution of A_{Ic}-globin resulted in an 85% yield of $\beta\ \beta^{A_{Ic}}$-globin. $\beta^{A_{Ic}}$T-1 was isolated in 81% yield by Dowex-50 X2 chromatography of the tryptic hydrolysate of $\beta\ \beta^{A_{Ic}}$-globin. When $\beta^{A_{Ic}}$T-1 was purified on cellulose phosphate, 92% of the peptide was recovered. Chromatography on Dowex-50 X8 of the papain hydrolysate of $\beta^{A_{Ic}}$T-1 returned R-val-his in 28% yield. Thus, the overall yield of R-val-his was about 17%. More concretely, approximately 7 mg of purified R-val-his were isolated from three pints of whole blood. The blocked dipeptide was pure in the sense that it was free from other amino acids and peptides; however it was still contaminated with about 20% (cf. Table XIV, p. 139) of resin and/or buffer impurities. Seven months were required to isolate these seven milligrams of R-val-his.

The overall yield of 17% for R-val-his probably does not include losses which are caused by the lability of the blocking group R. Losses from the latter source will further reduce the overall yield.

Finally, it is worth mentioning the relative merits of large and small scale isolations. Wherever column chromatography was involved, it was more convenient to carry out the isolation on a large rather than a small scale, because the time required per chromatogram was about the same in either case.* However, because of the limited size of the countercurrent distribution apparatus, many small scale distributions had to be run in order to obtain large amounts of $\beta\ \beta^{A}Ic$ subunits. This factor tended to offset the advantages which were gained by the large scale chromatographic procedures. It therefore would not be worthwhile at present to increase the scales of isolation over those which have already been described. Rather, efforts should be directed towards developing a column chromatographic method for isolating the $\beta\ \beta^{A}Ic$ subunit. In this manner full advantage could be taken of the gains which result from large scale operations.

---

*This was only approximately true. It required more time and effort to physically manipulate the equipment which was used in the larger scale isolations. In any case, the amount of product which was obtained per unit time was less than proportional to the scale of the isolation.

III: THE GROSS STRUCTURE OF A_{Ic}

III A: Introduction

III B: Determination of N-Terminal Amino Acids of A_{Ic} and $\beta\ \beta^{A_{Ic}}$

III C: Self-hybridization of A_{Ic}

III D: Starch Gel Electrophoresis of A_{Ic}

III A: INTRODUCTION

Four experiments were done to confirm and refine the gross structure* $\alpha_2\ \beta\ \beta^{A_{Ic}}$ which was proposed for hemoglobin A_{Ic} at the top of p. 14. In the first experiment, Jones' (76) work with dinitrofluorobenzene (cf. p. 13) was repeated to determine the number and types of polypeptide chains which were present in A_{Ic}. In the second, the subunits of A_{Ic} were dissociated from each other and recombined to form a hemoglobin which chromatographed like A_{II} on IRC-50: this experiment confirmed the presence of the normal--*i.e.*, unblocked--β chain in A_{Ic}. In the third, the globin from A_{Ic} was electrophoresed at *pH* 1.8 in an attempt to detect charge differences between the polypeptide chains of A_{Ic} and A_{II}. In the fourth, A_{Ic}-globin was separated into two polypeptide fractions by countercurrent distribution, and each fraction was analyzed for the number and types of subunits present, for its amino acid composition, and its amino acid sequence.

Only the first three experiments will be described in this section. The fourth is sufficiently complex to require a section of its own.

III B: DETERMINATION OF THE N-TERMINAL AMINO ACIDS IN A_{Ic} AND IN $\beta\ \beta^{A_{Ic}}$

Sanger's (124) end-group method, in which 2,4-dinitrofluorobenzene is used to determine the N-terminal amino acid present in a polypeptide, was applied to hemoglobin A_{Ic}, A_{Ic}-globin, and to $\beta\ \beta^{A_{Ic}}$-globin to

---

*The term <u>gross structure</u> has been defined in the footnote to p. 19.

confirm the presence of two α chains and of at least one normal β chain. The experimental procedures for dinitrophenylation followed Rhinesmith's et al. (119, 120, 121) adaptation of the method of Levy and Li (87).

III B1: Experimental

Approximately 1.5 μm of A_{Ic}, A_{Ic}-globin, or β $\beta^{A_{Ic}}$-globin, which were prepared as described on pp. 23, 34, and 35, were suspended in 10 ml of water at 40°C. The solution was titrated to pH 9.0 with 0.2 F NaOH, 0.15 ml of 2,4-dinitrofluorobenzene was added, and the reaction was allowed to proceed for 2 hr. The pH was kept constant at 9.0 by the automatic addition of 0.2 F NaOH. At the end of 2 hr, the reaction was stopped by bringing the solution to pH 1.0 with 6 F HCl. At pH 1.0 the dinitrophenylated globin precipitated. The precipitate was isolated by centrifugation, washed twice with 0.1 F HCl at 4°C, rinsed three times with cold acetone, three times with ether and was vacuum dried at room temperature overnight.

When A_{Ic} hemoglobin was the starting material, after the dinitrophenylated hemoglobin had been precipitated at pH 1.0 and washed twice with 0.1 F HCl at 4°C, the heme was removed by suspending the dinitrophenylated hemoglobin in acetone which was 0.01 F in HCl. The mixture was shaken continuously at room temperature for 1 hr, centrifuged, and the precipitated dinitrophenylated globin was washed with acetone and ether and dried as described above.

The dried, weighed dinitrophenylated globin was suspended in 10 ml of 6 F doubly glass-distilled HCl, was allowed to stand for 1 hr to insure thorough wetting, and was then hydrolyzed under reflux for 15

min* in an oil bath set at 240°C.** The hydrolysate was rinsed into a 60-ml separatory funnel with four 1-ml portions of 0.06 F HCl and was extracted with four 25-ml portions of ether. The combined ether extracts were washed with three 10-ml portions of 0.06 F HCl in a 125-ml separatory funnel. The first washing was usually a light yellow and was added to the aqueous phase from the ether extraction. The other washings, if colorless, were discarded. The ether extract was evaporated under a stream of air; the residue contained the dinitrophenylated α chain N-terminal dipeptide DNP-val-leu and small amounts of DNP-val from the N-termini of the starting material.

The aqueous phase from the ether extraction was extracted with six 25-ml portions of ethyl acetate. The combined ethyl acetate extracts were washed with six 20-ml portions of 0.06 F HCl; all washings were discarded. The residue which was obtained on evaporating the ethyl acetate extracts contained the dinitrophenylated β chain N-terminal di- and tripeptides diDNP-val-his and diDNP-val-his-leu.

Any longer N-terminal dinitrophenylated peptides from the 15-min hydrolysis of the β subunits were not extractable by the above proce-

---

*Rhinesmith et al. (119, 120, 121) have shown that in 15-min hydrolyses, N-terminal DNP-val-leu is quantitatively cleaved from the α chain, and diDNP-val-his and diDNP-val-his-leu are partially cleaved from the N-terminus of the β chain. Small amounts of DNP-val are also present from the further hydrolysis of the above compounds. On prolonged hydrolysis for 22 hr N-terminal DNP-valine is quantitatively cleaved from both α and β chains. By combining the data from a 15-min and 22-hr hydrolysis it is possible to calculate the number and types of α and β chains that are present in the starting material.

**The high temperature was used to bring the solution to its boiling point rapidly. In this manner, the time of hydrolysis was accurately controlled.

dure but remained in the aqueous phase from the ethyl acetate extract.* The longer peptides were degraded to N-terminal DNP-valine by hydrolyzing the aqueous phase at 100°C for an additional 22 hr. The extraction procedure that was described above was repeated to recover the N-terminal DNP-valine from the 22-hr hydrolysis.

The residues obtained from the evaporation of the organic phases were chromatographed on silicic acid-Celite columns as described by Green and Kay (41) and Rhinesmith et al. (120, 121, 125). The N-terminal dinitrophenylated peptides and amino acids in the residues were identified by their chromatographic behavior on these columns and were estimated by their ultraviolet and visible spectrum in the manner described by the above investigators.

III B2: Results

The results of the above determination are shown in Table IV, p. 77. These results confirm the earlier work of Jones (76), and support the subunit structure $\alpha_2\ \beta\ \beta^{A_{Ic}}$ that is proposed by us in which the N-terminus of the $\beta^{A_{Ic}}$ chain is unreactive towards 2,4-dinitrofluorobenzene because this N-terminus is blocked by some group R.

III C: SELF-HYBRIDIZATION OF HEMOGLOBIN A_{Ic}

By the self-hybridization of a protein we mean a dissociation of that protein into its constituent subunits that is followed by a recombination of those subunits to form a different protein(s), chemically

*These longer peptides arise from the fact that the hydrolysis of the dinitrophenylated β chains to the N-terminal dinitrophenylated di- and tripeptides is incomplete after 15 min. See first footnote on preceding page.

TABLE IV

Amount of DNP-Peptide Isolated[a]

Starting Compound	n[b]	From α Subunits	From β Subunits
		val-leu + val	val-his-leu + val-his
DNP-A_{II}	3	100 ± 9	100 ± 9
DNP-A_{Ic}	4	128 ± 18	58 ± 8[c]
DNP-β $\beta^{A_{Ic}}$	1	4[d]	33[c,e]

[a]The amounts of peptides which were experimentally obtained from DNP-A_{II}-globin were assigned the arbitrary value 100. The amounts of peptides found for DNP-A_{Ic} and DNP-β $\beta^{A_{Ic}}$ are expressed relative to those found for DNP-A_{II}. DNP means 2,4-dinitrophenylated.

[b]n is the number of independent determinations made.

[c]The value to be expected if the N-terminus of one of the two β subunits is blocked is 50.

[d]The small amount of α subunits which are present confirms the purity of the starting protein.

[e]This value is somewhat low for some nonobvious reason; it is the result from a single determination, and it is likely that the low value is simply a reflection of a poorer than normal run.

distinct from the first. In particular, the experiments in this section were designed to show that the self-hybridization of hemoglobin A_{Ic}, $\alpha_2\ \beta\ \beta^{A_{Ic}}$, would result in the formation of hemoglobin A_{II}, $\alpha_2\ \beta_2$. The formation of the latter protein would confirm the presence of the β subunit, as distinguished from the $\beta^{A_{Ic}}$ subunit, in hemoglobin A_{Ic}.*

II C1: Theoretical

If a sample of A_{Ic} that is contaminated with a fraction P of A_{II}** dissociates under appropriate experimental conditions via

$$(1 - P)\ \alpha_2\ \beta\ \beta^{A_{Ic}} + P\ \alpha_2\ \beta_2 \longrightarrow 2\alpha + (1 + P)\ \beta + (1 - P)\ \beta^{A_{Ic}}\ ,$$

and if the recombination of the monomeric subunits is random, then if the experimental conditions are altered so as to cause these subunits to reassociate into hemoglobins, each of which contains two α and two

---

*The possibility of self-hybridizing A_{Ic} was suggested by the similarities between the gross structures of A_{Ic} and hemoglobin F_I (cf. pp. 11, 14), and a comment by Dr. Alexander Miller that the latter might be expected to self-hybridize via

$$\underset{F_I}{2\alpha_2\ \gamma\ \gamma^{N\text{-Acetyl}}} \rightleftharpoons \underset{F_{II}}{\alpha_2\ \gamma_2 + \alpha_2\ \gamma_2^{N\text{-Acetyl}}}$$

under suitable conditions. To the best of our knowledge this suggestion of Dr. Miller with respect to F_I has never been followed up.

**It is difficult to be sure that a sample of A_{Ic}, which has been isolated by IRC-50 chromatography is absolutely free from A_{II}. The reason for this is that A_{Ic}, when chromatographed on this resin, exhibits a long chromatographic "tail" (cf. Fig. 10, p. 85) which extends into the A_{II} region. When the chromatographic column is warmed in order to elute A_{II}, the tail from the A_{Ic} is also eluted as a chromatographic

β-like (i.e., β or $\beta^{A_{Ic}}$ subunits)* the overall result of the dissociation of A_{Ic} and the reassociation of its subunits can be written

$$4(1 - P)\ \alpha_2\ \beta\ \beta^{A_{Ic}} + 4P\ \alpha_2\beta_2 \longrightarrow (1 + P)^2\alpha_2\ \beta_2 + (1 - P)^2\alpha_2\ \beta_2^{A_{Ic}} + 2(1 + P)(1 - P)\alpha_2\ \beta\ \beta^{A_{Ic}}.$$

---

peak in the position of A_{II}. The A_{Ic} which was used in the hybridizations did not contain any hemoglobin from this tail. Nevertheless, rechromatography of "tail-free" A_{Ic} on IRC-50 always produced a new tail which extended into the A_{II} region. It may or may not be that this tailing phenomenon is due to the lability of the blocking group R (cf. p. 67), and that the hemoglobin A_{Ic} which chromatographs like A_{II} is in fact A_{II} which has been formed from A_{Ic} by the loss of R. In any case, the contamination with this A_{II}-like material is quantitatively small: it is not, however, negligible, and its effects on the outcome of the hybridization must be allowed for. First, the identification of the hemoglobins in a self-hybridization experiment was made by their chromatographic mobility on IRC-50: any A_{II}-like material originally present in the A_{Ic} sample before self-hybridization would naturally be present after self-hybridization and could lead to the false conclusion that self-hybridization of the A_{Ic} had occurred when in fact it had not. Second, if the A_{II}-like material were in fact hemoglobin A_{II}, there would exist a statistical bias during the recombination of the subunits which favors the formation of more A_{II} by self-hybridization than would have been formed in the absence of the contaminating A_{II}. Both these effects have been allowed for in the derivation given in the text.

*The assumption of equal numbers of α- and β-like chains in each hemoglobin molecule is not necessary, but without it the equations become algebraically cumbersome. In interpreting the outcome of a self-hybridization experiment quantitatively, it is necessary to assume some sort of model. The ultimate justification of the model selected rests, of course, on experiment. Qualitatively, it suffices to state that if self-hybridization occurs or does not occur for one model, it must occur or not occur respectively, for all models--and hence for the correct one. This last statement is true even if the model chosen to interpret the results happens to be incorrect. An incorrect model may lead to incorrect quantitative interpretations of the experimental results, but it cannot lead to a false conclusion as to whether qualitatively self-hybridization did or did not occur.

If only a fraction f of the original sample of A_{Ic} undergoes the dissociation and reassociation reaction, then after reassociation the relative amounts of the different hemoglobins which are present are:

$$A_{Ic} = 4(1 - f)(1 - P) + 2f(1 + P)(1 - P)$$

$$A_{II} = 4(1 - f)P + f(1 + P)^2$$

$$\alpha_2 \beta_2^{A_{Ic}} = f(1 - P)^2$$

Defining

$$R_x = A_{II}/A_{Ic}$$

$$R_y = A_{II}/\alpha_2 \beta_2^{A_{Ic}} ,$$

where these ratios refer to the hemoglobins present in the reassociated sample, then by substituting the values of A_{Ic}, A_{II}, and $\alpha_2 \beta_2^{A_{Ic}}$ given above into these definitions and solving for f in terms of R_x and P, and for R_y in terms of f and P we get

$$f = \frac{4R_x - \frac{4P}{(1 - P)}}{(1 + 2R_x)(1 - P)} ,$$

$$R_y = 1 + \frac{4P}{(1 - P)^2 f} .$$

The overall result of the self-hybridization of A_{Ic} can be nicely summarized in the form of an equilibrium* reaction

---

*It should be emphasized that this equilibrium is a <u>hypothetical</u> one which may or may not exist in fact. Here it is simply a convenient, formal way of expressing the data. The equilibrium reaction which actually exists under the conditions or dissociation is almost certainly <u>not</u> the one given above but either

$$2\alpha_2\ \beta\ \beta^{A_{Ic}} \rightleftharpoons \alpha_2\ \beta_2 + \alpha_2\ \beta_2^{A_{Ic}}\ .$$

The equilibrium constant for this reaction under the conditions of dissociation is then

$$K_e = \frac{(f/2)(f/2)}{[2(1 - f) + f]^2} = \frac{f^2}{4(2 - f)^2}\ .$$

Under the assumption of random recombination of subunits the maximum possible value of K_e is 0.25, which corresponds to 100% self-hybridization of the A_{Ic}.

$$\alpha_2\ \beta\ \beta^{A_{Ic}} \rightleftharpoons 2\alpha + \beta + \beta^{A_{Ic}}$$

with an equilibrium constant

$$K_e' = \frac{2f^3}{(1 - f)}\ ,$$

or

$$\alpha_2\ \beta\ \beta^{A_{Ic}} \rightleftharpoons \alpha\ \beta + \alpha\ \beta^{A_{Ic}}$$

with an equilibrium constant

$$K_e'' = \frac{f^2}{(1 - f)}\ ,$$

or perhaps both. That equilibria of the above nature exist is fairly well established. Schroeder and Jones (131) have reviewed the evidence for such equilibria. The f in the above expressions is the same as that on p. 80. One advantage of the hypothetical equilibrium given at the top of this page over those given in this footnote is that it shows at a glance the products to be expected in the reassociated sample, and the magnitude of K_e (p. 81) permits one to calculate the relative amounts of these products.

Summarizing, if the ratio of A_{II} to A_{Ic} in the reassociated sample is determined experimentally, and the purity of the starting sample is known, then the fraction of self-hybridization f, the equilibrium constants for the various modes of dissociation, and R_y can be calculated. If R_y can be determined experimentally, it can be compared with the calculated value of R_y as a consistency check.

A measure of the maximum error to be expected in f, R_y^{calc}, K_e, K_e', and K_e'' can be obtained by differentiating the expressions for these quantities. Explicit equations for these errors are given in Appendix II, p. 213. The important point here is that the relative error of K_e is always approximately three times the relative error in f, irrespective of the numerial value of f. On the other hand, the relative errors of K_e' and K_e'' become infinite as the fraction of self-hybridization approaches 100%. This superiority of K_e over K_e' and K_e'' shows up nicely in the data to be presented later in Table V, and is among the reasons the hypothetical, but nonexistent, equilibrium given at the top of p. 81 was chosen over other modes of representing the data.

III C2: Experimental

Determination of P

The contamination of the carbonmonoxy-A_{Ic} which was to be used in the self-hybridization experiment by material which chromatographed like hemoglobin A_{II} was determined by chromatographing the carbonmonoxy-A_{Ic} with developer #5* at 6°C on a 0.6 x 35-cm IRC-50 column as described by Clegg and Schroeder (22).

---

*Developer #5 (22) was prepared by dissolving 16.56 gm of $NaH_2PO_4 \cdot H_2O$, 7.10 gm of Na_2PO_4, and 2.60 of gm KCN in 4 liters of water. The pH of this buffer was 6.85.

Self-Hybridization

Hemoglobin can be dissociated into its constitutent subunits in both acidic and basic solutions, and in neutral solutions at high and low ionic strengths (Schroeder and Jones, Ref. 131, have reviewed the known methods of dissociating hemoglobins). In the following experiments, both acidic and basic solutions were used to induce dissociation.

For subunit hybridization under acidic conditions 15 mg of carbonmonoxy-A_{Ic} in 0.200 of ml water was added to 0.800 ml of pH 3.63 or pH 4.66 HOAc-NaOH buffer at an ionic strength of 0.10.* The mixture was dialyzed at 2°C for 24 hr against 2 liters of the HOAc-NaOH buffer. This was followed by dialysis at 2°C against pH 7.61 tris-HCl buffer for 8 hr, then by dialysis at 25°C for 30 min against fresh pH 7.61 buffer.** At this point the dialysis sac was transferred to developer

---

*The pH 3.63 buffer was prepared by diluting 8.20 gm of NaAc and 47.9 ml of HOAc to 1 liter. The pH 4.66 buffer was prepared in like manner with 8.20 gm of NaAc and 5.0 ml of HOAc. Boiled water which had been cooled to room temperature was used in preparing these buffers. Boiling the water minimized the amount of oxygen in it from dissolved air. It was conceivable that small amounts of oxygen might adversely affect the hemoglobin molecule under the conditions of hybridizations (74).

**To 2038 ml of freshly boiled H_2O at 98°C were added 34.8 gm of tris(hydroxymethyl)aminomethane, 47.92 gm of NaCl, and 18.5 ml of concentrated HCl. After the mixture had cooled to room temperature, 6.46 gm of cysteine hydrochloride which had been dissolved in 10 ml of 2 F KOH and 10 ml of 2 F NaOH was added. The sodium and potassium ion concentrations of this buffer were equal to their concentrations in chromatographic developer #5 (cf. footnote, p. 82). The purpose of the cysteine was to reduce any ferrihemoglobin which had been formed during the hybridization back to ferrohemoglobin: ferrihemoglobin, and particularly ferrihemoglobin H (β_4), was known to denature (i.e., precipitate) more readily than the corresponding carbonmonoxyferrohemoglobin (71). The pH (7.61) of the tris-HCl buffer and the concentration of cysteine were chosen so that even had all the hemoglobin in the hybridi-

#5 (cf. footnote, p. 82) at 2°C. This developer had been previously bubbled with carbon monoxide. Dialysis at 2°C in developer #5 was continued for 24 hr with three changes of buffer. The sample was removed from the dialysis sac, any precipitate was centrifuged off, and the supernatant was chromatographed at 6°C with developer #5 on a 0.9 x 35-cm IRC-50 column (cf. bottom p. 82) to determine quantitatively the hemoglobins which were present.

For subunit hybridization under basic conditions the same procedure was used except that pH 11.24 diethylamine-HCl buffer* at an ionic strength of 0.10 was used to induce dissociation.

In both the acidic and basic self-hybridizations, the pH of the hemoglobin solution at the beginning and end of the 24-hr dissociation period (but prior to neutralization) was found to be the same, viz. that of the buffer which was used to induce the dissociation.

III C3: Results

A chromatogram which shows the purity of the A_{Ic} which was used for self-hybridization is shown in Fig. 10a, p. 85. The "tailing" of the trailing edge of the A_{Ic} peak is evident (cf. 2nd footnote, p. 78). There was approximately 7.4% (P = 0.074) of A_{II}-like material in this sample.

---

zation mixture been present as ferrihemoglobin, 90% of it would have been reduced to ferrohemoglobin in 23 min at 25°C (58). The neutral pH was of course also necessary to obtain the reassociation of the subunits.

*Prepared by diluting 17.87 ml of diethylamine and 8.55 ml of concentrated HCl to 1 liter with boiled, cooled water.

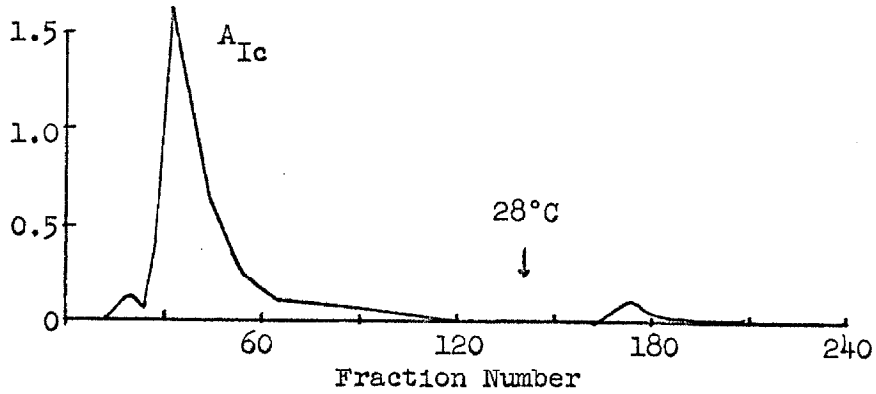


Fig. 10a: Chromatographic determination of P on 3.8 mg of A_{Ic} before self-hybridization.

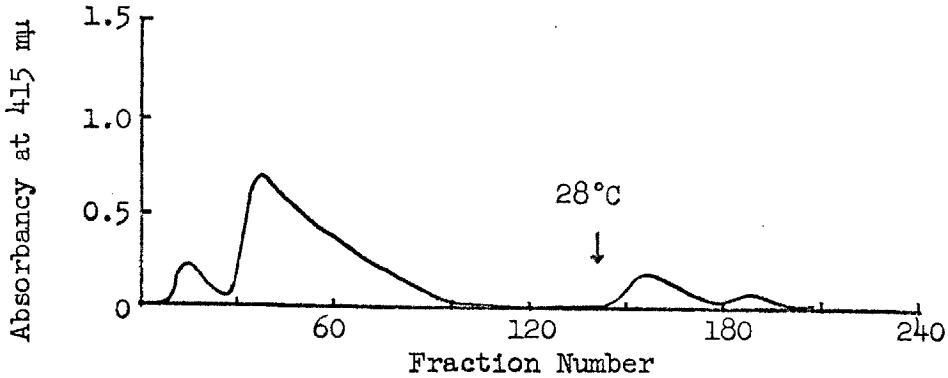


Fig. 10b: Chromatogram of 3.8 mg of A_{Ic} which has been self-hybridized at pH 3.63.

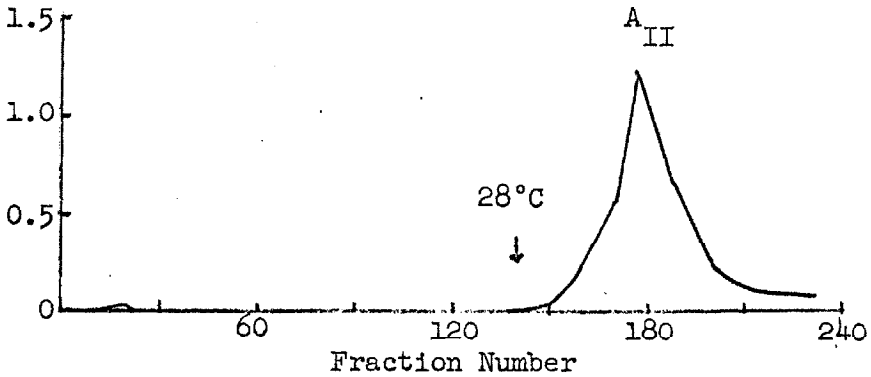


Fig. 10c: Chromatogram of 3.8 mg of A_{II} which has been self-hybridized at pH 3.63.

A chromatogram of self-hybridized A_{Ic} is shown in Fig. 10b. The increased percentage (relative to that in Fig. 10a) of A_{II}-like material and of a component of more rapid chromatographic mobility than A_{Ic} is evidence that self-hybridization of the A_{Ic} did, in fact, occur. Presumably, the increased percentage of A_{II}-like material was due to the formation of hemoglobin A_{II} by self-hybridization of A_{Ic} in the manner described on pp. 78-81. The component whose chromatographic mobility was greater than that of A_{Ic} was assumed to be $\alpha_2 \beta_2^{A_{Ic}}$.

A chromatogram of a control sample of A_{II}, which had been self-hybridized under the same conditions as the sample of A_{Ic} in Fig. 10b is shown in Fig. 10c. This control demonstrated that the self-hybridization procedure itself did not create any hemoglobin components with a more rapid chromatographic mobility than A_{II} from either the small amount of A_{II} which may have been present as a contaminant in the A_{Ic} sample, or from A_{II} which was formed by self-hybridization of A_{Ic}. Jandl and his collaborators (74) have shown that under acidic conditions in the presence of oxygen, hemoglobin A_{II} can give rise to hemoglobin components of rapid chromatographic mobility on IRC-50. Had these components been present in the reassociated sample of A_{Ic}, they could have caused falsely low determinations of the A_{II} formed by the self-hybridization, and falsely high determinations of the A_{Ic} and $\alpha_2 \beta_2^{A_{Ic}}$.

The total area under the peaks in Fig. 10b and Fig. 10c corresponds to 3.8 mg hemoglobin. Since 13 mg A_{Ic} was self-hybridized in each of the two figures, only 29% of the starting material was re-

covered. The remaining 71% represented precipitated, denatured hemoglobin which had been removed by centrifugation prior to chromatographic analysis. The fact that the same amount of precipitation was observed for both the self-hybridized A_{Ic} and the self-hybridized A_{II} control suggests that there was no preferential denaturation of the β and $\beta^{A_{Ic}}$ chains relative to each other: under these conditions, the theory which was developed on pp. 78-82 retains its validity irrespective of the amount of precipitation. The large amounts of precipitate that were observed at pH 3.63 were probably caused by the extremely acid pH at which these self-hybridizations were carried out.

A summary of the self-hybridization of hemoglobin A_{Ic} at acidic and basic pH values is given in Table V, p. 88. These data show that self-hybridization of carbonmonoxy-A_{Ic} went to approximately 48 $\pm$ 16% completion under both acidic and basic conditions: allowing for all types of experimental error, in no case was the fraction of self-hybridization less than 32% nor greater than 64%. About two-thirds of this $\pm$ 16% uncertainty was due to the error in the ratio R_x, and about one-third to the error in P: all of these errors were determined by the accuracy with which the area under a chromatographic peak could be measured ($\pm$ 5%).

III C4: Discussion

The purpose of this section is to discuss certain limitations of the data in Table V. An understanding of these limitations is necessary before intelligent inferences about the structure of A_{Ic} can be drawn therefrom.

TABLE V

Self-hybridization of Carbonmonoxyhemoglobin A_{Ic}

pH[a]	P[b]	R_x[c]	f[d]	R_y[e]	R_y^{theor}[f]	Precipitate?[g]
3.63	0.074	0.184	0.41	2.10	1.84	75%
4.66	0.145	0.338	0.47	5.44	2.69	24%
11.24	0.145	0.377	0.55	1.24	2.44	47%

Equilibrium Constants[h]

$K_e = 0.03\ (0.01 - 0.06)$

$K_e' = 0.25\ (0.06 - 0.89)$

$K_e'' = 0.50\ (0.17 - 1.33)$

[a]This is the pH at which hemoglobin was dissociated into subunits.

[b]P is the fraction of A_{II}-like material which was present in the A_{Ic} prior to self-hybridization. P was determined as in Fig. 10a.

[c]R_x is the experimentally determined ratio of A_{II}/A_{Ic} in the self-hybridized A_{Ic} after reassociation of the subunits. See Fig. 10b.

[d]Formally, during a self-hybridization, a fraction f of the starting A_{Ic} can be looked upon as having undergone the stoichiometric reaction given at the top of p. 79, whereas the remaining fraction (1 - f) of A_{Ic} can be considered to have undergone no change at all. f was calculated from R_x and P by the formula given on p. 80.

[e]This is the experimentally determined ratio of $A_{II}/\alpha_2\beta_\alpha^{A_{Ic}}$ in the self-hybridized A_{Ic} after reassociation of the subunits. (See Fig. 10b.)

[f]Calculated from P and f as per p. 80. R_y and R_y^{theor} should be equal.

[g]This is the percentage of the hemoglobin taken for self-hybridization which precipitated during the course of the experiment.

[h]Calculated from f with the formulas of p. 81. The most probable value is given first. The values in parentheses are values of K which are allowable within the experimental error.

The conclusion that hemoglobin A_{-I} has been formed by self-hybridization of A_{Ic} was based on increased amounts of a hemoglobin which had the same chromatographic behavior on IRC-50 as A_{II}, and of a rapidly moving component which was assumed to be $\alpha_2 \beta_2^{A_{Ic}}$. These assignments were reasonable in view of the more negative net ionic charge of the $\beta^{A_{Ic}}$ chain. Nevertheless, the positive identification of the newly formed hemoglobins as A_{-I} and $\alpha_2 \beta_2^{A_{Ic}}$ would require their isolation and a reasonably complete chemical characterization, particularly with respect to the presence or absence of the blocking group R. This additional characterization of the newly formed hemoglobins was not carried out because there was insufficient material available to make it even remotely feasible.

<u>A priori</u> it was possible that the formation of hemoglobin A_{II} in the self-hybridization experiments was not due to dissociation and reassociation of the individual polypeptide chains of A_{Ic}: A_{II} would also have been formed if the conditions of the experiments favored the removal of the blocking group R from the N-terminus of the $\beta^{A_{Ic}}$ chain <u>via</u>

$$\alpha_2 \beta \beta^{A_{Ic}} \longrightarrow R + \alpha_2 \beta_2 .$$

There are, however, two facts which argue strongly against this mode of formation of hemoglobin A_{II}. The first, is the fact that the blocking group R is more readily removed from the $\beta^{A_{Ic}}$ chain at basic <u>pH</u> (cf. pp. 67-70). If loss of R were the primary source of A_{II} in the self-hybridizations, the fraction self-hybridization f, calculated as described in Table V, would be expected to be significantly greater at

pH 11.24 than at pH 4.66 or 3.63. Table V shows that f was about the same at all three pH values. The second is the fact that the formation of A_{II} by the loss of R would not lead to any formation of $\alpha_2 \beta_2^{A_{Ic}}$: this fact would cause the ratio of $A_{II}/\alpha_2 \beta_2^{A_{Ic}}$ in the neutralized solution to be very large numerically relative to R_y^{theor}. Table V shows that this disproportion is not, in fact, observed: R_y and R_y^{theor} are actually of the same order of magnitude.

A quantitative defect in Table V is the inequality between R_y and R_y^{theor}. In principle they should be equal. In two of the experiments (pH 4.66 and pH 11.24), the difference between them is larger, though not greatly so, than the experimental error. The resolution between $\alpha_2 \beta_2^{A_{Ic}}$ and A_{Ic} is not good (cf. Fig. 10b). It may be that the error in the amount of $\alpha_2 \beta_\alpha^{A_{Ic}}$ is somewhat larger than we have assumed ($\pm$ 5%). By lowering the temperature of the IRC-50 chromatography, and/or changing the chromatographic developer, it is likely that $\alpha_2 \beta_2^{A_{Ic}}$ could be better resolved from A_{Ic}.*

The diethylamine buffer which was used for the basic self-hybridization was not satisfactory. It was chosen for its high buffering capacity at the pH of the hybridization. In practice the use of this buffer resulted in a large amount of precipitate and in a brownish colored final solution which did not look at all like undenatured hemo-

---

*Experience has shown that improved resolution can frequently be obtained without a change of buffer by lowering the temperature from 6°C to 2°C. Similarly, improved resolution has been obtained without a change of temperature (6°C) by developing the chromatogram with developer #4 (1, 76) rather than developer #5. The composition of developer #4 is the same as developer #5 (cf. p. 82) except that 5.68 gm of Na_2HPO_4, not 7.10 gm, is used.

globin. It is known that NH_3 (58) and NH_2OH (10) complex with the heme portion of hemoglobin. Perhaps diethylamine acts similarly. Phosphate and glycine are not particularly good buffers at pH 11.2. Pyrogallol might be satisfactory at this pH, for it does not contain any groupings which can obviously complex with the hemoglobin; it would have the additional advantage of removing residual oxygen from the solution (88).

III C5: Conclusions

Within the limitations which were discussed in the preceding section, the data in Table V support the gross structure, $\alpha_2\ \beta\ \beta^{A_{Ic}}$ which we have proposed for A_{Ic}. Approximately 12% of a sample of purified A_{Ic} was converted into roughly equimolar amounts of hemoglobin A_{II} and $\alpha_2\ \beta_2^{A_{Ic}}$. The formation of these hemoglobins could not have occurred unless A_{Ic} contained both a β and $\beta^{A_{Ic}}$ polypeptide chain.

If the dissociation of hemoglobin A_{Ic} by any (including those in the footnotes) of the mechanisms on pp. 78-81 were complete, and if the reassociation of the subunits were random, the fraction of self-hybridization f to be expected would be 1.00.* The experimental fact that f is considerably less than 1.00 implies that A_{Ic} is much more stable--either thermodynamically or kinetically--with respect to acid-base dissociation than either A_{II} or $\alpha_2\ \beta_2^{A_{Ic}}$. Jones and Schroeder (77) have shown that A_{II} is more stable than F_{II} in the above respect, and the data in Table V suggest A_{Ic} is still more stable than A_{II}.

---

*f was deliberately defined in such a way as to make this so.

An interesting aspect of the stability of A_{Ic} is the relevance of this stability to the picture of dynamic equilibrium of the hemoglobin molecule that Guidotti et al. (48) have discussed. According to their ideas a molecule with the gross structure $\alpha_2\ \beta\ \beta'$ which is in rapid equilibrium with the hybrids $\alpha_2\ \beta_2$ and $\alpha_2\ \beta_2'$* should not be isolable by column chromatography provided that the time required to separate the two hybrids is long relative to the time required to establish the equilibrium. The reason that a molecule such as $\alpha_2\ \beta\ \beta'$ is not separable under these conditions is simply that during the course of the chromatography it is converted into the hybrids. It is the two hybrids which are isolated. The dynamic equilibrium hypothesis has been used to explain the fact that in hybridizations between hemoglobins A_{II} and S_{II}, with the gross structures $\alpha_2\ \beta_2$ and $\alpha_2\ \beta_2^S$ respectively, no hemoglobin with the structure $\alpha_2\ \beta\ \beta^S$ has ever been isolated. Clearly, hemoglobin A_{Ic}, with the gross structure $\alpha_2\ \beta\ \beta^{A_{Ic}}$, fails to fit the dynamic equilibrium hypothesis, because A_{Ic} is easily isolable by IRC-50 chromatography. This anomaly becomes understandable if the stability of A_{Ic} is such as to preclude the establishment of a rapid equilibrium between itself and the hybrids A_{II} and $\alpha_2\ \beta_2^{A_{Ic}}$.

---

*The existence of such an equilibrium presupposes some mechanism whereby $\alpha_2\ \beta\ \beta'$ can be dissociated into subunits. This dissociation might be catalyzed, for example, by the large surface area of the resin in a chromatographic column--even at neutral pH values. The important point is not the position of the equilibrium, but that the equilibrium be rapidly established.

III D: STARCH GEL ELECTROPHORESIS OF A_{Ic}-GLOBIN

These experiments were designed to detect charge differences between the globin subunits of A_{Ic} and A_{II}.

III D1: Experimental

The method which was used followed Muller (104). The gel was prepared by stirring 63 gm of hydrolyzed starch into 200 ml pH 1.8 formic acid-sodium hydroxide buffer* at room temperature. This solution was added quickly, with constant stirring, to 300 ml of the same buffer which had been heated to 75°C. When the mixture had become an easy flowing viscous liquid it was poured into a lucite tray 317 mm in length and 6 mm deep and was covered with a plastic lid which was held in place by C-clamps.** The covered tray was placed in the refrigerator at 10°C for 1 hr until the starch had gelled.

Six-tenths milligram A_{Ic}- or A_{II}-globin, which was prepared as described on p. 34, was dissolved in 0.050 ml of the pH 1.8 formate buffer and placed into a rectangular slot** of the gel. The sample was

---

*This buffer was prepared by diluting 139 ml of 98% formic acid and 20 ml of 2 F NaOH to 2.5 liters. The exact amount of NaOH which must be used depends on the exact concentration of the formic acid and may vary slightly from 20 ml. This buffer had an ionic strength of 0.016.

**This tray was purchased from Mr. Otto Hiller, P.O. Box 1294, Madison, Wisconsin. The plastic lid which came with this apparatus had eight 0.1 x 0.6 x 1.2-cm rectangular protrusions which were at right angles to both the lid surface and the length of the tray. These protrusions were arranged linearly with approximately 2-mm spacing between each. Their purpose was to make a line of eight slots, for receiving the samples, in the center of the gel.

sealed into the slot with vaseline, and the gel was completely covered with Saran-wrap to prevent evaporation of the buffer from the gel during electrophoresis.

The globin was electrophoresed at room temperature (25°C) for 20 hr at 75 volts. The electrode compartments contained pH 1.8, ionic strength 0.068, formic acid-sodium hydroxide buffer.* Electrical contact between these compartments and the gel was made with clean white towels which had been saturated with the buffer in the electrode compartments. The breeze from a small fan was kept blowing across the gel surface during the electrophoresis: otherwise the electrophoresis generated sufficient heat which sometimes melted either the vaseline and/or the gel itself.

For staining, the gel was sliced into two sections with a thin piano wire. Prior to sectioning the gel, it was cooled in the refrigerator to give it a consistency suitable for easy mechanical handling. Each section was stained for 30 sec with 0.1% Buffalo Black in MeOH:H_2O:HOAc::50:50:1 v/v/v and was destained with several washes of MeOH:HOAc::9:1 v/v. For photographic or photometric purposes the destained sections could be made transparent by soaking them in benzyl alcohol.

III D2: Results

Starch gel electrophoresis at pH 1.8 revealed no differences between the globin subunits of A_{Tc} and A_{II}. In each case two well separated bands were observed. (The separation between the bands was about 0.5 cm.) The bands with the lesser and greater cathodic mobility re-

*111 ml of 98% HCO_2H, 34 ml of 2 F NaOH, and 855 ml of H_2O.

spectively, corresponded to the α and β chains of A_{II}, which have net positive ionic charges of 25 and 24 respectively at pH 1.8.* If the N-terminal amino acid of the $\beta^{A_{Ic}}$ chain is blocked in an amide type linkage, this chain should have a charge of + 22; a band corresponding to this charge was not observed.

III D3: Discussion

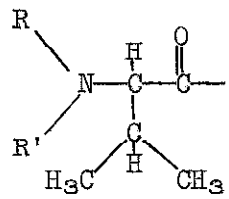
Huisman and Norton (67) have shown that in pH 8.1 Tris-EDTA-borate buffer hemoglobin A_{Ic} moves toward the anode barely perceptibly faster than A_{II}. Their results demonstrate the existence of a more negative charge at pH 8.1 on A_{Ic} than on A_{II}, but the difference in charge is extremely slight, much less than that to be expected if the more negative nature of A_{Ic} were due to an extra carboxylic acid group or its equivalent. Further, the identical electrophoretic behavior of A_{Ic}- and A_{II}-globin at pH 1.8 suggests that the chemical bond between the blocking group R and the N-terminal amino acid valine in the $\beta^{A_{Ic}}$ chain of A_{Ic} is such as to give the blocked amino terminal nitrogen of $\beta^{A_{Ic}}$ the same charge, at this pH, as the unblocked amino terminal nitrogen of β. These facts make the existence of an amide bond between R and N-terminal valine unlikely.

---

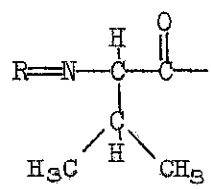
*These charges were obtained by assigning a charge of + 1 to each lysine, arginine, histidine, and N-terminal amino acid in the chain, and a charge of zero to all other amino acids. At pH 1.8 the carboxyl groups of the peptide chains probably have a slight negative charge. This negative charge works in such a way as to increase the difference between the charges of the α and β chains, but always the former remains more positive than the latter. It is interesting that the observed order of the cathodic mobilities of the α and β chain is opposite to that which would be expected from their ionic charges. This anomaly may be due to the greater number of hydroxy groups in the α chain as compared to the β. The interaction of these hydroxy groups with those of the starch may be sufficiently strong to account for the inverted order of mobilities.

The existence of practically no charge difference between the β and $\beta^{A_{Ic}}$ chain at pH 1.8 and pH 8.1 can be explained by assuming that the bond between R and the N-terminal valine is one of the following two types:*

a) R is bonded to the N-terminal amino group as a secondary or tertiary amine:



b) R is bonded to the N-terminal amino group as a Schiff base:



As we shall see (p. 142) the Schiff base linkage is the correct one.

---

*At the time these experiments were done we were not convinced, nor are we now, that starch gel electrophoresis of proteins, especially those of unknown structure, is sufficiently well understood to state with certitude that if charge differences did exist between β and $\beta^{A_{Ic}}$ separate bands would be observed in the gel. A major source of this doubt is the inverted order of mobilities for the α and β chains which was discussed in the footnote on p. 95. The two bond types which are proposed in the text are the result of retrospection in the light of experiments which preceded and followed the starch gel experiments. The latter experiments themselves are only corroborative.

IV: THE PRIMARY STRUCTURE OF A_{Ic}

IV A: Introduction

IV B: The Amino Acid Composition of A_{Ic} and $\beta\ \beta^{A_{Ic}}$

IV C: The Amino Acid Composition of the Tryptic Peptides of α_2 and $\beta\ \beta^{A_{Ic}}$

IV D: The Primary Amino Acid Sequence of the Tryptic Peptides of $\beta\ \beta^{A_{Ic}}$

IV A: INTRODUCTION

Although the major goal of the experiments which are described in this thesis was the isolation and characterization of the blocking group R of A_{Ic}, when time was available it was spent in refining the primary amino acid sequence of the polypeptide chains of A_{Ic}.

The gross structure of hemoglobin A_{Ic} (cf. Section III, pp. 72-96), $\alpha_2\ \beta\ \beta^{A_{Ic}}$, suggested that the α, β and $\beta^{A_{Ic}}$ chains of A_{Ic} were identical to the α and β chains respectively of A_{II}, with the exception of a blocking group R at the N-terminus of the $\beta^{A_{Ic}}$ polypeptide chain. It is the purpose of the experiments in this section to more firmly establish this hypothesized identity (except for R) between the polypeptide chains of A_{Ic} and A_{II}.

The primary amino acid sequence of the polypeptide subunits of A_{Ic} was determined in the following steps: The amino acid composition of A_{Ic}, irrespective of the sequence of the amino acids in the polypeptide chains, was determined. A_{Ic} was separated into α and $\beta\ \beta^{A_{Ic}}$ subunits by countercurrent distribution (pp. 35-40), and the amino acid composition of $\beta\ \beta^{A_{Ic}}$ was determined. The α and $\beta\ \beta^{A_{Ic}}$ subunits were hydrolyzed with trypsin (p. 40), the tryptic peptides were purified by column chromatography (pp. 41-62), and their amino acid compositions were determined. Finally, the sequence of amino acids in each tryptic peptide from $\beta\ \beta^{A_{Ic}}$ was determined by the Edman degradation procedure (6, 29, 129, 130). The order of the tryptic peptides in the α, β, and $\beta^{A_{Ic}}$ chains were assumed to be the same in A_{Ic} as the known order in A_{II} (16, 17, 61, 80, 81, 126, 131).

IV B: THE AMINO ACID COMPOSITION OF A_{Ic} AND $\beta\ \beta^{A_{Ic}}$

IV B1: Experimental (129)

Between 1-5 mg of hemoglobin A_{Ic} or of $\beta\ \beta^{A_{Ic}}$-globin was transferred with a small quantity of water to a weighed 13 x 100-mm Pyrex test tube. The water was removed by air evaporation at 40°C, and the tube was dried to constant weight at 110°C. Two milliliters of 6 F doubly glass-distilled HCl were added, and the test tube was evacuated with a water pump and sealed under this vacuum. The sealed tube was placed in an oven at 110°C for 22-72 hr. The hydrolysate was dried at 40°C under a stream of air and analyzed on the Spinco Model 120 Automatic Amino Acid Analyzer.

IV B2: Results

The results of these analyses are shown in Table VI, pp. 99a and 100. Tryptophan is not reported because it was destroyed under the conditions of the hydrolysis.

Within the experimental errors which are indicated in the footnotes to Table VI, the amino acid composition of A_{Ic} is identical to that of A_{II}; and that of $\beta\ \beta^{A_{Ic}}$ to that of β_2. It follows from conservation of mass that the α chains of A_{Ic} and A_{II} must have the same amino acid composition. In terms of residues, the maximum error in the amino acid composition for a particular amino acid is $\pm$ 0.8 residues per α, β, or $\beta^{A_{Ic}}$ chain, and is more typically of the order of one-half this magnitude.

TABLE VI

Amino Acid Composition[a] of Hb-A_{Ic} and $\beta\ \beta^{A_{Ic}}$-Globin[b]

		A_{Ic}[c]	A_{II}[e]	β_2[e]	$\beta\ \beta^{A_{Ic}}$[d]
lys		42.68	44	22	21.34
his		37.62	38	18	16.74
arg		11.76	12	6	5.88
asp		49.50	50	26	27.04
thr	(f)	33.60	32	14	14.00
ser	(f)	34.24	32	10	10.40
glu		32.32	32	22	22.67
pro		13.72	14	14	13.86
gly		40.40	40	26	25.74
ala		73.44	72	30	29.70
cys/2	(f)	5.22	6	4	3.04
val		62.62	62	36	33.84
met	(f)	5.82	6	2	1.82
leu		72.00	72	36	36.00
ileu		Not present	0	0	Not present
phe		28.80	30	16	15.84
tyr	(f)	13.44	12	4	3.68

[a] The figures in the Table are the number of amino acid residues in the compound heading each column.

[b] This sample of $\beta\ \beta^{A_{Ic}}$-globin had been reduced with $NaBH_4$ (cf. p. 146) prior to hydrolysis.

[c] These values were calculated as follows. The amino acid composition of A_{II}, from three independent analyses by Jones (76), is entered in Table VIII, p. 87 , of that reference. These analyses had been performed in the manner described in this thesis on p. 99, and were expressed in grams free amino acid per 100 grams A_{II}. Recoveries differed for the three analyses, so in order to make the entries among the analyses comparable each entry was expressed relative to the entry for leucine, which was given the arbitrary value of 10 for all three analyses, by multiplying the entry for each amino acid by 10/leu, where leu was the entry for leucine for that analysis. The average of the three relative values will be called the "normalized amino acid composition." The normalized amino acid composition of A_{Ic} was calculated in exactly the same manner from a total of seven analyses, three by Jones (76) and four by us. The ratio of the normalized amino acid composition of A_{Ic} to that of A_{II} was multiplied by the known number of residues (column 3 of

Table VI) for each amino acid: it is this product which is entered in column 2 of Table VI. For an individual analysis, errors due to weighing and the amino acid analysis were approximately 3% each. From these figures it can be calculated that the entries in column 2, Table VI, should be accurate to $\pm$ 2%.

d) The figures in this column were calculated analogously to those in column 2 (cf. preceding footnote), except that the normalized amino acid composition of β_2 was computed from the normalized amino acid composition of A_{II} and the known primary amino acid sequence of the α and β chain. The figures in this column are accurate to $\pm$ 5%. The error is larger than for column 2 because only a single determination of the composition of β $\beta^{A_{Ic}}$ was made.

e) The number of residues in this column are exact and were taken from Refs. 16, 17, 61, 80, 81, 126, and 131.

f) This amino acid is partially destroyed during the hydrolysis (129). Consequently the number of residues obtained for this amino acid is somewhat variable and deviates from the expected value in excess of the 2% and 5% experimental errors mentioned in footnotes c) and d) above. There is no reason to consider the deviations from the expected values of columns 3 and 4 significant.

IV B3: Conclusions

With the exception of a blocking group R on the N-terminus of the $\beta^{A_{Ic}}$ chain, the amino acid compositions of the polypeptide chains of A_{Ic} are identical to those of A_{II}.

It is unlikely that the blocking group R contains any oligopeptides or amino acids, for if it did, differences in the amino acid composition of the β $\beta^{A_{Ic}}$ and β_2 subunits would have been found.*

IV C: THE AMINO ACID COMPOSITION OF THE TRYPTIC PEPTIDES OF α_2 AND β $\beta^{A_{Ic}}$

IV C1: Experimental

The isolation of the purified tryptic peptides has been described on pp. 41-62 of this thesis. The amino acid analyses of these peptides were carried out exactly in the manner described in Section IV B1, p. 99, with the exception that the peptide was not weighed, for we were interested only in the relative, not absolute amounts of the amino acids which were present in the peptide.

IV C2: Results

The results of these analyses are presented in Tables VII and VIII. These Tables can be found on the following four pages. In interpreting these results it should be kept in mind that the presence of small non-integral amounts of methionine and tryptophan is significant of at

---

*This statement should be qualified somewhat: experimental errors would not have allowed us to detect with certainty a difference in amino acid composition of less than $\pm$ 2 residues for a given amino acid.

TABLE VII[a]

Amino Acid Composition[b] of the Tryptic Peptides of α^{AIc}

	αT-1	αT-2	αT-3	αT-4	αT-5	αT-6	αT-7	αT-8	αT-9
Yield %[c]	65	80	75	66	10	10	51	70	11
lys	0.97	1.00	1.00		1.06	0.95	1.10	1.00	1.07
his				0.98		1.78	0.93		2.52
arg				0.98					
$cySO_3H$									
asp	1.06	1.00			0.45	1.01			5.80
thr		0.92			1.88	0.88			0.99
ser	0.73				0.97	1.90			1.92
glu	0.16			2.96	0.27	1.07			0.12
pro	1.01				1.02	1.22			1.04
gly			1.06	3.08	0.26	1.18	1.94	0.12	0.23
ala	1.01		1.94	4.34	0.40	1.06	0.23	0.10	6.95
cys/2									
val	1.12	1.00		1.16	0.11	1.09			2.95
met					0.46				0.39
ileu									
leu	0.83			1.19	1.20	1.01			3.96
tyr				0.93		0.78			
phe	0.15				1.97	1.85			
try			0.14						

[a]This Table is concluded on the following page.

[b]The numbers in the Table are the number of amino acid residues relative to each other. Amounts less than 0.10 residue are not reported.

[c]Based on 10 μm α_2^{AIc} taken for tryptic hydrolysis.

TABLE VII

Amino Acid Composition of the Tryptic Peptides of $\alpha^{A_{Ic}}$

	αT-10	αT-11	αT-12 + αT-13[d]	αT-14
Yield %	39	43	-	100
lys		1.01	2	
his			3	
arg	1.00			1.04
$cySO_3H$				
asp		2.04	1	
thr			4	
ser		0.11	5	
glu			1	
pro		0.94	2	
gly				
ala			6	
cys/2			1	
val		2.02	4	
met				
ileu				
leu	1.00		9	
tyr				0.94
phe		0.96	2	
try				

[d]The number of residues in this column are only accurate to about ± 0.6 residues. They were calculated by subtracting the sum of the number of amino acid residues for the other $\alpha^{A_{Ic}}$ tryptic peptides from the known amino acid composition of the $\alpha^{A_{Ic}}$ chain (cf. pp. 99-101).

TABLE VIII[a]

Amino Acid Composition[b] of the Tryptic Peptides of β $\beta^{A_{Ic}}$

	βT-1	$\beta^{A_{Ic}}$T-1	βT-2	βT-3	βT-4	βT-5	βT-6	βT-7	βT-8
Yield %[c]	36	47	64	53	12	19	45	8	4
lys	1.05	1.03	0.94			0.94	1.02	1.06	1.00
his	1.00	0.86						0.83	
arg				1.00	0.98				
$cySO_3H$									
asp				2.03		2.76			
thr	0.96	0.96	0.88		0.96	0.93			
ser			0.80	0.15	0.12	1.59			
glu	2.13	2.00	0.13	2.08	1.04	1.01			
pro	1.02	1.00		0.13	1.04	1.82			
gly			1.06	2.99		2.06		1.15	
ala		0.15	2.05	1.06		0.97		0.97	
cys/2									
val	0.97	0.85	1.01	2.89	1.58	1.05	0.97		
met						0.47			
ileu									
leu	1.04	1.01	1.00	1.04	1.96	1.02			
tyr					0.90				
phe						2.60			
try			0.56		0.19				

[a]This Table is concluded on the following page.

[b]The numbers in the Table are the number of amino acid residues relative to each other. Amounts less than 0.10 residue are not reported.

[c]Based on 30 μm β $\beta^{A_{Ic}}$ taken for tryptic hydrolysis.

TABLE VIII

Amino Acid Composition of the Tryptic Peptides of β $\beta^{A_{Ic}}$

	βT-9	βT-10 + βT-11	βT-12	βT-13	βT-14	βT-15
Yield %	39	20	9	23	54	36
lys	1.04	1.13	1.03	1.02	1.03	
his	1.02	1.73	1.66		1.01	0.98
arg		0.84				
$cySO_3H$		0.83	1.06			
asp	2.95	2.90	1.05	1.02	0.98	
thr		1.77		0.92		
ser	0.89	1.14				
glu		1.90		3.04		
pro		1.16		2.04		
gly	2.01	1.35	2.04		1.01	
ala	2.00	1.33	1.23	1.94	3.89	
cys/2						
val	1.01	1.25	2.66	1.04	2.45	
met						
ileu						
leu	4.48	3.16	3.37		1.00	
tyr				0.89		1.02
phe	1.00	1.84	1.06	1.00		
try						

least one residue of these amino acids in the peptide, because these two amino acids are partially destroyed by oxidation during the hydrolysis: αT-3 and αT-9 provide examples. On the other hand the presence of up to 0.2 residues of serine, glycine, or alanine in an analysis is not usually significant because these amino acids are sometimes found (129) in hydrolysates of peptides which are known not to contain any serine, glycine, or alanine: αT-4 and αT-8, for example, probably contain four and zero residues of alanine respectively.

Subject to the above provisos, the amino acid composition of the tryptic peptides is unambiguously determined by the analyses given in Tables VII and VIII with the exception of the number of residues of aspartic acid in αT-5, histidine in αT-9, valine in βT-4, serine in βT-5, and leucine in βT-9 and βT-12. In these six cases there is an uncertainty of ± 1 residue. The uncertainty of ± 1 residue is far in excess of the accuracy of the amino acid analysis, which is about 3%. Deviation of the number of residues in a peptide by more than 5% from an integral value is an indication that the peptide is not pure. For the above six peptides, the impurities probably arise from inadequacies of the cellulose phosphate chromatography which was used to isolate and purify the individual tryptic peptides. These inadequacies have already been discussed at length on pp. 57 and 58. We shall assume, by analogy with A_{II}, that the correct numbers of residues for these peptides are 0, 3, 2, 2, 4, and 4 respectively.

IV C3: <u>Conclusions</u>

The recovery of approximately equal amounts of $\beta^{A_{Ic}}$T-1 and βT-1 (Table VIII) is additional support for the proposed structure of A_{Ic},

$\alpha_2\ \beta\ \beta^{A_{Ic}}$, in which one β and one $\beta^{A_{Ic}}$ polypeptide chain are present. Further, the fact that the amino acid compositions of $\beta^{A_{Ic}}$T-1 and βT-1 are identical supports the hypothesis that the two peptides differ only in the presence of a blocking group R at the N-terminus of the former. Also, R cannot contain any oligopeptide or amino acids unless (and this is unlikely) the overall amino acid composition of R is the same as that of βT-1, for otherwise the amino acid compositions of βT-1 and $\beta^{A_{Ic}}$T-1 would have differed.*

Finally, no peptides were found whose amino acid compositions differed from the peptides of the α and β chains of A_{II}.** Again this supports the hypothesis that R is the only difference which exists between A_{Ic} and A_{II}.

IV D: THE AMINO ACID SEQUENCE OF THE TRYPTIC PEPTIDES OF $\beta\ \beta^{A_{Ic}}$

The sequence of the amino acids in each of the purified tryptic peptides from the $\beta\ \beta^{A_{Ic}}$ subunit was determined by the Edman degradation*** (28) as described by Schroeder and his associates (6, 129, 130).

---

*This statement can be made with absolute certainty, for the number of residues of a particular amino acid in βT-1 and $\beta^{A_{Ic}}$T-1 was known to ± 0.1 residues.

**Tables VII and VIII are not in form which is convenient for comparison with the peptides from A_{II}. This comparison can better be made from Figs. 18, 19, pp. 194, 195, in which the data are compared directly with the known structure of the peptides of A_{II}.

***This method and its modifications are frequently referred to in the literature as the PTC (phenyl isothiocyanate) or PTH (phenylthiohydantoin) method. The principle of all these methods is the same, though the experimental details differ rather markedly among the various modifictions: When the adduct of phenyl isothiocyanate with the N-terminal

Because they were of slight interest, the amino acid sequences of the tryptic peptides from the α chains of A_{Ic} were not determined.

Table IX, p. 107, shows the amino acid sequence of the tryptic peptides from the β $\beta^{A_{Ic}}$ subunit. In most cases the complete sequence of the peptide could not be determined by the Edman method. There were two reasons (both expected) for this: 1) the yield of the PTH-amino acid from the peptide decreased progressively, for some peptides, at each cycle of the Edman procedure, and consequently the degradation could not be continued beyond a certain point; and 2) as the degradation continued, there was usually some carryover of the PTH-amino acid from the preceding cycle of the degradation. This carryover was cumulative and eventually sufficiently interfered with the identification of the PTH-amino acid last removed from the peptide to make further progress impossible. Because our primary interest was in the blocking group R, no attempt was made to carry the sequence determination beyond what could conveniently be learned from the Edman procedure alone.

As far as we went, no differences were found between the amino acid sequences of the tryptic peptides from β $\beta^{A_{Ic}}$ and the corresponding tryptic peptides of β_2.

---

amino group of a peptide or protein is subjected to acidic conditions, the 5-thiazolinone of the N-terminal amino acid is spontaneously cleaved from the adduct. The thiazolinone immediately rearranges to form the phenylthiohydantoin of the N-terminal amino acid. The phenylthiohydantoin of the N-terminal amino acid is then identified by paper chromatography. The process is then repeated on the peptide which resulted from ths loss of the N-terminal 5-thiazolinone. In this manner the amino acids can be removed from the peptide, starting at the N-terminal end, one at a time, until in principle the complete amino acid sequence of the peptide has been determined.

TABLE IX

Primary Amino Acid Sequence[a)] of the Tryptic Peptides of β $\beta^{A_{Ic}}$

$\beta^{A_{Ic}}$T-1:	R-val-his(leu)(thr,pro,glu)glu-lys[b)]
βT-1:	val-his-leu-thr(pro,glu)glu-lys
βT-2:	ser-ala-val-thr-ala-leu(try,gly)lys
βT-3:	val-asg-val(asp,glu,val,gly,gly,glu,ala,leu,gly)arg
βT-4:	leu-leu-val-val-tyr-pro(try,thr,glm)arg
βT-5:	phe-phe-glu(ser)phe-gly(asp,leu,ser,thr,pro,asp,ala,val,met, gly,asg,pro)lys
βT-6:	val-lys
βT-7:	(ala,his,gly)lys
βT-8:	lys
βT-9:	val-leu-gly-ala-phe-ser(asp,gly,leu,ala,his,leu,asp,asg, leu)lys
βT-10,11:	(gly,thr,phe,ala,thr,leu,ser,glu,leu,his,cys,asp,lys,leu,his, val,asp,pro,glu,asg,phe,arg)
βT-12:	(leu,leu,gly,asg,val,leu,val,cys,val,leu,ala,his,his,phe, gly)lys
βT-13:	glu(phe,thr,pro,pro,val,glm,ala,ala,tyr,glm)lys
βT-14:	val(val)ala-gly-val-ala(asg,ala,leu,ala,his)lys
βT-15:	tyr-his

a)The amino terminal end is on the left. The position of those amino acids joined by a hyphen is certain. The position of those amino acids in parentheses or separated by commas is uncertain, either because of an ambiguity in interpreting the results from the Edman degradation, or because the sequence of this part of the peptide was not determined; in these cases the amino acids have been placed in the same position that they occupy in the corresponding tryptic peptide from hemoglobin A_{II}. The position of C-terminal lysine or arginine is required by the specificity of trypsin.

b)The sequence of this peptide was determined by analysis of the peptide fragments from a papain hydrolysis (cf. pp. 63-68).

V: TESTS FOR SPECIFIC CHEMICAL GROUPINGS IN THE BLOCKING GROUP R

V A: INTRODUCTION

The evidence which has been presented up to this point has shown that hemoglobin A_{Ic} differs from hemoglobin A_{II} not in its primary amino acid sequence, but only in that the former contains a blocking group R at the N-terminal end of the $\beta^{A_{Ic}}$ chain of A_{Ic}. The sequence of the latter chain thus begins R-val.

Testing whether the blocking group R was itself, or contained, certain specific chemical groupings was among the more economical uses to which the small amounts of A_{Ic} available were put. The groupings which were searched for were determined by two criteria: 1) was it reasonable that R might be this group? and 2) were microanalytical methods available, or could they be developed, to assay for the group in question?

VA 1: Chemical Groups which Were Investigated

General considerations, arising from the known structures of other proteins, from certain facts about hemoglobins in general, from the principles of biochemistry, and from the available chemical evidence, led us to test for the following groups. Acyl, R = $-\overset{O}{\overset{\|}{C}}-R'$, and in particular acetyl; carboxy, R = $-CO_2^-$; carbamyl, R = $-\overset{O}{\overset{\|}{C}}-NH_2$; pyridoxal; phosphorus, both organic as in substituted phosphines, and inorganic as in pyrophosphate; carbohydrate, R = a sugar, or derivative thereof; and steroids. The considerations which led to a search for these groups are summarized in the following paragraphs.

The existence of the fetal hemoglobin minor component F_I, which has

the structure $\alpha_2\ \gamma\ \gamma^{N\text{-acetyl}}$ (cf. p. 11), and which has a similar chromatographic behavior to A_{Ic} in IRC-50 chromatography, suggested that A_{Ic} might have a similar structure: $\alpha_2\ \beta\ \beta^{N\text{-acetyl}}$.

Faurholt (30), Ferguson and Roughton (31), Stadie and O'Brien (136), Margari (91), Giustina et al. (33, 34, 35) and Milla et al. (98) have presented strong evidence for the existence of a compound, and possibly two different compounds, between CO_2 and hemoglobin. Some of these investigators have proposed that one of these compounds might be formed via

$$Hb\text{-}NH_2 + CO_2 \rightleftharpoons Hb\text{-}\overset{H}{N}\text{-}CO_2H \rightleftharpoons Hb\text{-}\overset{H}{N}\text{-}CO_2^- + H^+ .$$

Whatever the structure* of this compound between hemoglobin and CO_2 may be, the compound itself appears to have maximum stability at pH 10; above pH 12 or below pH 6, the compound cannot exist but is hydrolyzed to $CO_3^=$ or CO_2 respectively and hemoglobin. The formation or hydrolysis of the compound is almost instantaneous at all pH's. The stability of hemoglobin A_{Ic} itself between pH 3 and pH 11 (cf. p. 88) made it unlikely that the blocking group R was a carboxy group. Also the stability of both $\beta^{A_{Ic}}$T-1 and R-val-his below pH 6 and their instability above pH 6 (cf. p. 67) are characteristics that are the exact opposites of the hemoglobin-CO_2 compounds. Nevertheless, in view of the lack of knowledge of the actual structures of the hemoglobin-CO_2 compounds, it was decided to test for the carboxy group.

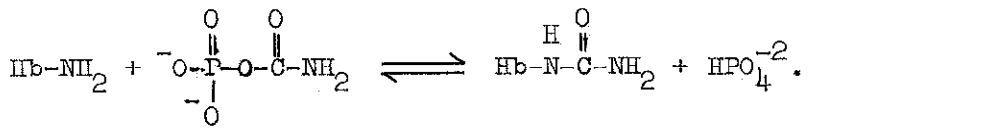
---

*The compounds between hemoglobin and CO_2 are so labile that they have never been isolated. It was this lability that first suggested to us that R might be $-CO_2^-$.

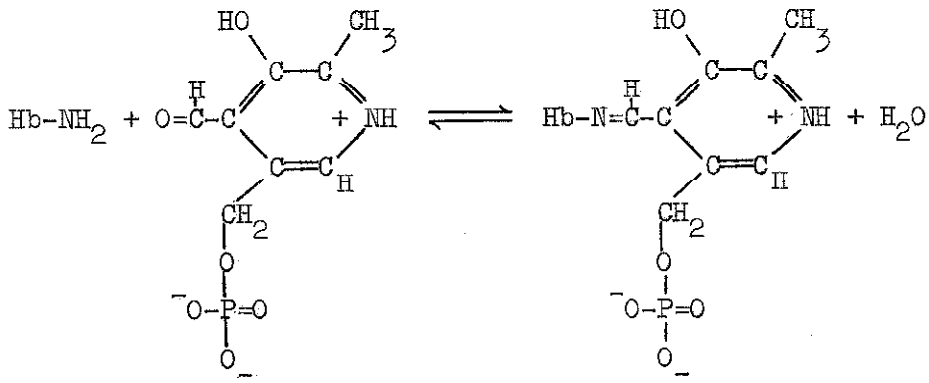
The carbamyl group could conceivably be introduced into hemoglobin by either the biological amidation of the hemoglobin-CO_2 compound,

$$\mathrm{Hb{-}\overset{H}{N}{-}CO_2H + NH_3 \rightleftharpoons Hb{-}\overset{H}{N}{-}\overset{O}{\overset{\|}{C}}{-}NH_2 + H_2O}$$

or by a transamidation with carbamyl phosphate,



Pyridoxal and pyridoxal phosphate are known to be very active in biological transamination reactions (14) <u>via</u> Schiff base intermediates. Such a Schiff base with hemoglobin might be formed as follows:



The relative stability at <u>p</u>H 5 of the Schiff bases between pyridoxal and some enzymes (14), and between pyridoxal and valine (97), and the instability of these Schiff bases at <u>p</u>H 8.4 are analogous to the stability of the blocking group R under acidic conditions and its instability at neutral or alkaline <u>p</u>H (cf. p. 67). Further, the <u>p</u>K_a of the nitrogen atom which is involved in the Schiff base bond is approximately 6.3 (14, 23, 97). A <u>p</u>K_a of this value would account for our failure to find

a charge difference between A_{Ic} and A_{II} during starch gel electrophoresis at pH 1.8 and pH 8.1 (cf. p. 93).

Phosphorus might be introduced into hemoglobin via pyridoxal phosphate (see preceding paragraph), or conceivably, by some type of rearrangement of the N-terminal aminoacyladenyl-RNA via which protein synthesis is known to proceed (131).

Several mucoproteins are known in which a protein is covalently bonded to a carbohydrate moiety (5). A priori there is no reason to suppose that hemoglobin could not be thus bonded.

Finally, two considerations led us to test for the presence of steroid compounds. First, in the bile acids, glycine or taurine is joined in an amide linkage to the carboxyl group of the cholesterol derivatives cholic acid, deoxycholic acid, or chenodeoxycholic acid. Might not the N-terminal amino group of one of the two β-like polypeptide chains of hemoglobin A_{Ic} be blocked by such an amide linkage with these acid derivatives of cholesterol? Second, both the adrenal cortical and gonadal hormones contain keto and/or aldehydo groupings which conceivably could be covalently bonded as a Schiff base to the N-terminus of one of the two β-like chains of A_{Ic}. The Schiff base between an N-terminal amino group of a polypeptide chain and a 3-keto-Δ^4-steroid such as cortisol or testosterone might be particularly favored because of the resonance stabilization of the imine bond of the Schiff base by the Δ^4-double bond.

VA 2: Considerations of Methodology

Tests for the above chemical grouping were carried out on either

A_{Ic} itself, or degradation products from it such as $\beta^{A_{Tc}}$T-1 or R-val-his. The isolation of these compounds has already been described in Section II (p. 17).

The presence of nonprotein impurities* in these compounds made it necessary to carry out the assays in a quantitative manner, so that any contribution from such impurities could be allowed for.

Wherever feasible, a given chemical group was assayed for by two independent methods. Standard curves covering the nano- to micromole range in which we expected the group to be present were prepared from substances of known composition. A test was considered negative unless the mole ratio of the group to the compound being assayed was, within reasonable limits of error, greater than or equal to unity.

V B: ASSAY FOR ACYL GROUPS

V B1: By Gas Chromatography

The acetyl group was assayed by the method of Ludowieg and Dorfman (90) as modified by Schroeder et al. (128). Sixty milligrams of vacuum dried hemoglobin were hydrolyzed with 0.5 ml of methanol which was 2 F in HCl in a sealed glass tube for 4 hr at 100°C. The volatile components (methyl esters, methanol, H_2O and HCl) of the hydrolysate were distilled between 40-50°C at 70 mm Hg into a glass receiving tube which was immersed in a dry ice-methanol bath. Fifty microliters of the distillate was immediately chromatographed at 50°C on a Perkin Elmer Model 154 Vapor Fractometer that was equipped with a polyethylene glycol type KX column.

---

*For the source of these impurities, see pp. 48-50, this thesis.

V B2: By Titration with $Ba(OH)_2$

Acyl groups in general were assayed by Bartley's procedure (11). The acyl groups which are detectable by this method are defined by the experimental procedure described below and include those whose vapor pressure at 50°C is comparable to that of acetic acid: formyl, acetyl, propionyl, and butyryl are among these. It is doubtful, though we have not ruled out the possibility, that acyl groups with more than four carbon atoms would be detected. Keto acids are not detected by the method, though a simple modification of the experimental procedure permits their detection.

Ten to twenty milligrams of vacuum dried hemoglobin was hydrolyzed with 0.5 ml of 2 F H_2SO_4 in a sealed tube at 100°C for 4 hr. The hydrolysate was centrifuged briefly, and 0.400 ml of the supernatant was transferred to a 1-ml test tube. Fifty microliters of 0.1 F $AgClO_4$ was added to the supernatant to precipitate any volatile halides, and 0.100 ml of 0.1 F 2,4-dinitrophenylhydrazine in 1 F H_2SO_4 was added to precipitate any volatile keto acids.* The mixture was centrifuged, and 0.500 ml of the pale straw yellow supernatant was placed on top of 3.7 gm of anhydrous sodium sulfate in the side arm of the diffusing apparatus.** The purpose of the sodium sulfate was to increase the vapor pressure of the volatile acids by decreasing the water activity. The ratio of liquid to sodium sulfate was critical if bumping was to be

---

*Omission of the $AgClO_4$ or 2,4-dinitrophenylhydrazine precipitations permits the detection of volatile halo-acids or volatile keto-acids respectively.

**A Thunberg tube which had been modified as described by Bartley (11).

avoided in subsequent steps. Exactly 1 ml of carbon dioxide-free water (prepared by boiling distilled water and allowing it to cool under CO_2-free conditions) was added to the receiving arm of the diffusing apparatus, and the entire system was evacuated with a water pump for 1 min, then sealed. The side arm, which contained the sample on the bed of sodium sulfate, was placed in a water bath at 50°C for 5 min in order to vaporize most of the water which was present in the sodium sulfate. This precaution helped prevent bumping in subsequent steps. At the end of this time the receiving arm was immersed in an ice bath at 0°C while the side arm containing the sample was kept at 50°C. After 20 min, the receiving arm was opened to the air and the contents were titrated under nitrogen with a micrometer syringe* which was filled with 0.01 F $Ba(OH)_2$. Exactly 0.050 ml of 0.1% phenol red** was used as an indicator. Except for the titration under nitrogen, and the use of CO_2-free

---

*The titration was conveniently carried out with a Radiometer Titrator (cf. footnote p. 40) set for manual operation. The contents of the receiving arm were magnetically stirred during the titration. Because the bottom of the Thunberg tubes were rounded it was necessary to make a rounded magnetic stirrer. This stirrer was made by passing a short section of thin iron wire through a glass capillary which had been bent into an arc whose radius of curvature was about the same as that of the Thunberg tube. The length of the capillary arc was just longer than that of the iron wire. After the wire had been centered in the arc, the ends of the glass capillary were sealed. This stirrer was tested prior to use by boiling it in a small volume of concentrated HCl. If the ends of the capillary had not been sealed correctly, the solution turned yellow from the dissolving iron.

**Prepared by dissolving 1 gm of phenol red in 28.4 ml of 0.1 F NaOH and diluting the mixture to 1 liter. The diluted solution was then titrated with NaOH to approximately pH 7.50. The exact pH was such that when 0.02 μm OH^- ion was added to a CO_2-free water blank, a distinct color change from yellow to pink was observed against a white background.

water to prepare the reagents, no special precautions were necessary to exclude atmospheric carbon dioxide.

V B3: Results and Conclusions

The results of these assays are shown in Table X, p. 117. The first 13 entries in this Table served to establish the validity of the two methods of assay. The final entry is the result from an analysis of 2 samples of A_{Ic}.

The two preparations from rabbit γ-globulin were from different sources. The average of the gas chromatographic and titrimetric determinations for the first and second preparations are 16.2 and 21.8 respectively. Of the difference of about 6 residues, only 2 can be attributed to experimental errors. The remaining 4 represent real differences in the acetyl content of the preparations, not a deficiency in the accuracy of the assays.

Hemoglobins A_{II} and F_{II} do not contain any acyl groups. The finding of 0.66 μm of volatile acids by the titrimetric procedure for these two compounds is probably an indication of some degradation of the hemoglobin during hydrolysis with 2 F H_2SO_4. These two values are to be considered as blanks which are to be subtracted from the determinations for A_{Ic} and F_{II} respectively. This subtraction has been carried out in that section of Table X just preceding the notes to the Table.

The gas chromatographic analyses of F_I and F_{II} were performed by John Cua. These analyses were the basis for the publication (128) by Schroeder et al. in which it was demonstrated that hemoglobin F_I was monoacetylated.

TABLE X

Assay for Acyl Groups

Compound	μm Taken[a]	Gas Chromatography	$Ba(OH)_2$
H_2O	-	-	0.01 ± .01 (3)[b]
HOAc[c]	0.10	-	0.11 ± .01 (2)
	0.50	-	0.55 ± .01 (2)
Acetyl glycine[d]	0.37	-	0.41 (1)
	0.94	-	0.88 (1)
	0.95	0.92 ± .01 (2)[b]	-
	1.90	2.01 ± .03 (2)	-
	3.77	3.73 ± .02 (2)	-
Rabbit γ-globulin[e]	1.00	17.9 ± .6 (4)	14.6 (1)
	1.00	23.3 ± .1 (2)	19.8 ± .5 (2)
F_{II}[e]	1.00	0 (9)	0.44 ± .08 (2)
F_{I}[e]	1.00	0.84 ± .05 (29)	1.74 (1)
A_{II}[e]	1.00	0 (5)	0.66 (1)
A_{Ic}[e]	1.00	0.20 ± .06 (2)	0.92 ± .08 (2)

μm acyl groups in F_I: 1.74 - 0.44 = 1.30 by titration
0.84 - 0 = 0.84 by gas chromatography

μm acyl groups in A_{Ic}: 0.92 - 0.66 = 0.26 by titration
0.20 - 0 = 0.20 by gas chromatography

[a]This is the amount of compound which was assayed.

[b]The numbers in parentheses are the number of independent assays which were made.

[c]Test of titration procedure only; not carried through diffusion steps.

[d]Carried through entire procedure.

[e]The actual amount of starting compound varied between 0.1 and 1.3 μm. For ease of comparison the results have been expressed on a common basis of 1.00 μm.

By combining the gas chromatographic and titrimetric data, we find that F_I contains 1.07 ± .15 μm acyl groups, and A_{Ic} 0.23 ± .02 μm acyl groups, per mole of hemoglobin. The conclusion is that nonketo acyl groups of less than 5 carbon atoms are absent from hemoglobin A_{Ic}.

V C: ASSAY FOR CARBOXY

The presence of a carboxy group in R-val-his was tested for indirectly by incubating synthetic L-valyl-L-histidine* with CO_2 and comparing the resulting compound electrophoretically with R-val-his and val-his.

V C1: Experimental

A 1 x 10-cm Pyrex test tube was flushed with CO_2 by allowing 0.2 gm of dry ice to evaporate at room temperature from the tube inside an evacuated desiccator. Exactly 0.200 ml of pH 7.60 sodium phosphate buffer which was 0.127 F in Na_2HPO_4 and 0.0194 F in NaH_2PO_4 was placed in the tube with 0.1 gm of dry ice. The tube was kept at room temperature until all the dry ice had evaporated. Approximately 0.3 μm L-val-L-his in 0.010 ml of water and 0.1 gm of dry ice were added to the tube, which was immediately sealed with a rubber stopper that was held in place mechanically.** The incubation was continued at 25°C for a total

---

*The synthesis of L-valyl-L-histidine is given in Appendix I, p. 201.

**Approximately 8 atm CO_2 pressure was developed in the sealed tube. The concentration of L-val-L-his in the tube was 0.0015 F. By assuming solution ideality, and Henry's law behavior for gaseous CO_2, the formalities of other species in the solution were estimated to be: Na^+, 0.26; $H_2PO_4^-$, 0.14; HCO_3^-, 0.12; and H_2CO_3, 0.27. The pH of the solution was approximately 5.6.

of 108 hr. Each 24 hr a fresh 0.1-gm piece of dry ice was added to the tube. At the end of the incubation, the contents of the tube were evaporated to dryness, and an aliquot was taken for electrophoretic comparison with R-val-his, and val-his. The electrophoresis was carried out as described earlier (p. 45).

V C2: Results

From Fig. 11, p. 120, it is clear that R-val-his cannot be formed by incubating val-his with carbon dioxide under the above conditions, because the compound that is formed during the incubation has an electrophoretic behavior which differs from that of either R-val-his or val-his. The tailing of the compound that is formed between val-his and CO_2 is probably due to the salts from the sodium phosphate buffer.

We conclude that R is not $-CO_2^-$.

V D: ASSAY FOR CARBAMYL

Stark and Smyth (137) used carbamylated proteins to determine the N-terminal amino acid present therein, The principle of the method is to convert the N-terminal amino group of the protein to the carbamylate by reacting it with KCNO at basic pH. On acidification the N-terminal amino acid carbamyl derivative cyclises, and the hydantoin of the N-terminal amino acid is spontaneously cleaved from the remainder of the peptide chain. The hydantoin is readily purified by chromatography and can be identified after hydrolysis to the corresponding amino acid, carbon dioxide, and ammonia.

The method can be used "in reverse" to determine whether an N-terminal carbamylate is present in a polypeptide chain: the treatment

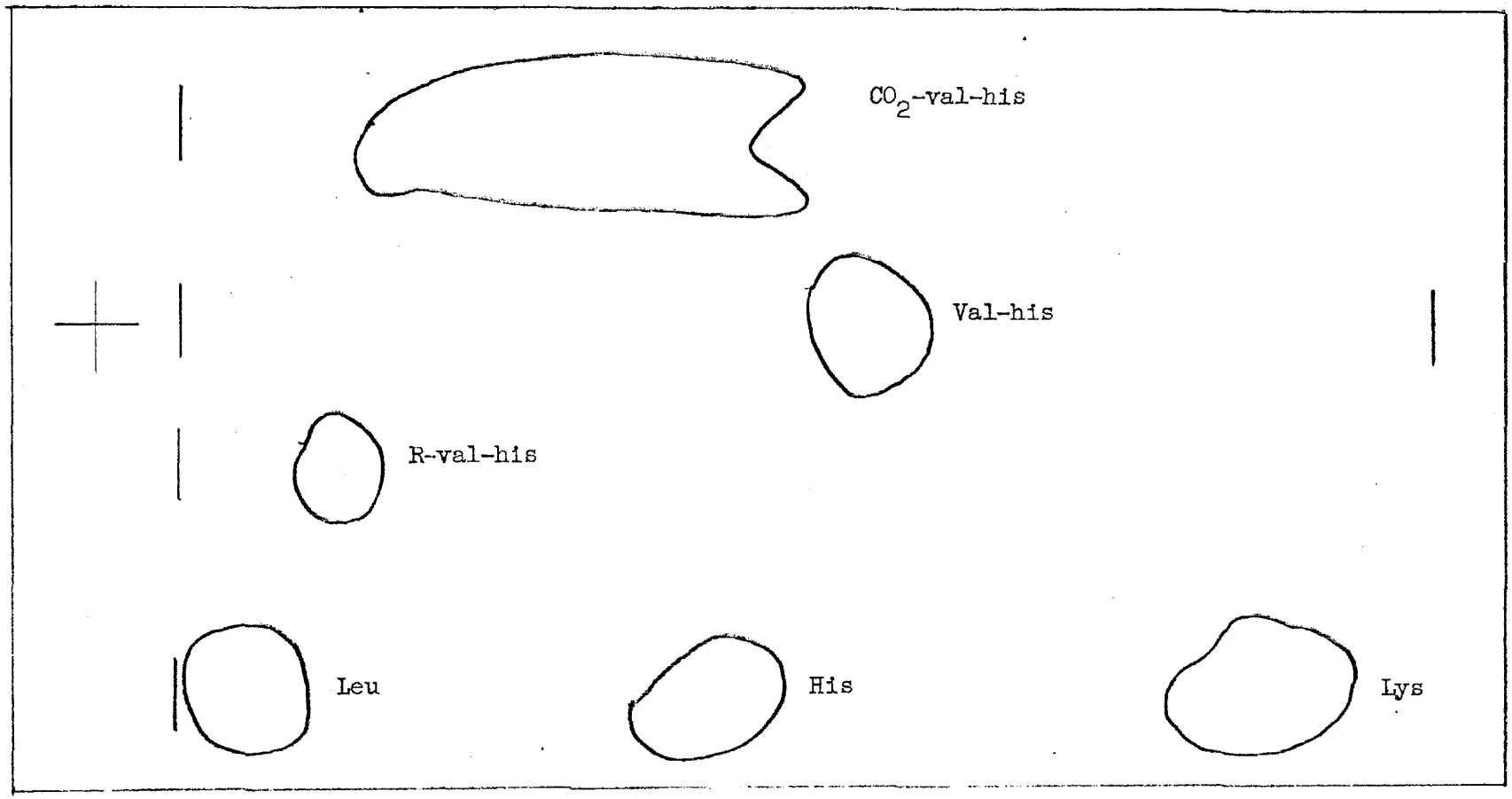


Fig. 11: Comparison of paper electrophoretic behavior at pH 6.4 of R-val-his, val-his, and val-his which has been incubated with CO_2.

with potassium cyanate is omitted; subsequent treatment is exactly the same as that described above. If the carbamylate is present, the corresponding hydantoin will be found; if it is absent, it will not.

The experimental results of this assay are given in Table XI.

TABLE XI

Assay for Carbamyl

Starting Peptide	Amount μm	N-Terminal Amino Acid Found	Amount μm
val-gly*	1.00	val	1.00
βT-1*	1.00	val	0.76
$\beta^{A_{Ic}}{}_{T-1}$**	1.00	none	0.00

*Reacted with KCNO; val-gly and βT-1 were used to check out method.

**Not reacted with KCNO.

Were the N-terminal amino group of $\beta^{A_{Ic}}{}_{T-1}$ blocked as the carbamylate, the Stark and Smyth procedure should have yielded the amino acid valine even though $\beta^{A_{Ic}}{}_{T-1}$ had not been reacted with KCNO. No valine was found for $\beta^{A_{Ic}}{}_{T-1}$. This conclusively demonstrates that the blocking group R is not $-\overset{O}{\overset{\|}{C}}-NH_2$, because had this group been present, the entry for βT-1 which was reacted with KCNO shows that it would have easily been detected.

V E: ASSAY FOR PYRIDOXAL

R-val-his was assayed indirectly for pyridoxal by synthesizing the Schiff base between pyridoxal and L-valine and comparing this base with R-val-his.

The synthesis followed Heyl (51). A mixture of 2 mg of L-valine and 3.48 mg of pyridoxal hydrochloride, which had been dissolved in 0.050 ml of 0.68 F KOH in MeOH, was allowed to stand 15 min at 25°C. This solution exhibited the deep yellow color that is characteristic (97) of the Schiff bases of aromatic aldehydes. When this Schiff base was diluted to a concentration equal to that of a comparison sample of R-val-his the yellow color was still pronounced. The comparison sample of R-val-his was colorless.

We conclude that R is not pyridoxal, a derivative thereof, or an aromatic aldehyde.

V F: ASSAY FOR PHOSPHORUS

The method of Nakamura (108) was used to determine total inorganic plus organic phosphorus. It was modified as follows to make it applicable to the small amount of material available.

V F1: Reagents

61 wt % $HClO_4$ in water

5% w/v $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ in water

Buffer: A mixture of 0.0200 gm of 1-amino-2-naphthol-4-sulfonic acid,* 1.20 gm of $NaHSO_3$, and 0.24 gm of Na_2SO_3 was dissolved in 10 ml of H_2O. The solution was stored in the refrigerator.

---

*The reagent grade product may range in color from white to a dull mauve depending on its exposure to light. It should be used as obtained because further purification is unnecessary as well as extremely difficult.

V F2: Experimental

The sample, containing 0-10 μ gm phosphorus, was dried in a 13 x 100-mm Pyrex test tube. Exactly 0.100 ml of 61% $HClO_4$ was added, and the tube was placed in a NaCl "sand bath" at 190°C-200°C for 15 min. The tube was removed from the bath to a metal test tube rack, 0.050 ml of 30% H_2O_2 was added, and the tube was again heated at 190°C-200°C for 10 min. The tube was removed to the rack, and the sides of the walls were washed down with 0.500 ml of water. In succession, with mixing between additions,* 0.100 ml of buffer, 0.100 ml of 5% molybdate solution, and 0.300 ml of water were added. The tube was centrifuged briefly at low speed to collect all liquid at the bottom of the tube, was shaken again, and was allowed to stand 10 min prior to reading the optical absorbancy at 700 mμ and a 1-cm path length. In the presence of phosphorus, a blue color, stable with time, was formed.

V F3: Results

The following relation, based on 13 observations approximately equally spaced in 4 groups of 3 each, was closely followed throughout the 0-10 μgm range when anhydrous Na_2HPO_4 was used as a standard:

$$\mu gm \ P = 8.585 \ A^{1\ cm}_{700\ m\mu}$$

$$P.E. = 0.26 \ \mu gm \ .$$

The results of the assay are shown in Table XII, p. 124.

Clearly, the blocking group R does not contain phosphorus.

---

*The mixing was accomplished by flipping the test tube with the forefinger.

TABLE XII

Assay for Phosphorus

Compound Taken	Amount Compound Taken[a]	Amount P Found	
Lys[b]	1.00 μm	0.00 μm	(2)[c]
Glu-lys[b]	1.00	0.00	(2)
Glu[b]	1.00	0.00	(2)
Adenylic Acid	1.00	0.95 ± .03	(2)
R-val-his	1.00	0.03 ± .07	(2)

a) The actual amount of each compound taken was approximately 0.1 μm. For ease of comparison, all results have been expressed on the basis of 1.0 μm taken.

b) These compounds were obtained from the papain hydrolysate of $\beta^{A_{Ic}}$T-1 (cf. pp. 63-67) by taking aliquots from the Dowex-50 X8 chromatogram (Fig. 9, p. 68) from which R-val-his was isolated. In this manner any impurities from the isolation procedures which might have given a false positive reaction for phosphorus could be allowed for. As can be seen from the last column in this Table, whatever impurities were present did not contain phosphorus. This is interesting in view of the fact that prior to the papain hydrolysis, $\beta^{A_{Ic}}$T-1 had been purified on cellulose phosphate (cf. pp. 58-61).

c) The number in parentheses is the number of independent determinations which were made.

V G: ASSAY FOR CARBOHYDRATE

V G1: Introduction

Protein samples containing carbohydrate can be tested either directly or indirectly. By directly is meant that no attempt is made to isolate the carbohydrates from other chemical classes which may be present; by indirectly is meant that the carbohydrates are separated as a distinct chemical class before being tested. Both methods were used with success in the present case, although when quantitative data were desired, the direct method was found to be more accurate. In either method appropriate controls were carried through the entire procedure.

For the indirect tests carbohydrates were isolated as a class by column chromatography on Dowex-50, after they had been hydrolyzed with HCl (101).

Regardless of the method which was used to isolate or assay the sample, it was important to avoid contaminating the reagents and samples with exogenous carbohydrate such as lint or the glue on the backs of Scotch Tapes of various types. Such materials are in abundant usage in many laboratories, and their presence may go unsuspected until detected as an abnormally high blank. Even though precautions were taken to avoid such contamination, typical blanks contained approximately 0.1 μm glucose equivalents.* About 2/3 of this was due to the reagents, while the remaining 1/3 appeared to come from the Dowex-50

---

*That amount of any carbohydrate which gives the same color reaction as 1 μm (μgm) glucose in a specific assay is by definition 1 μm (μgm) glucose equivalents.

columns which were used in the indirect method. It might be thought that such a large blank would overwhelm endogenous carbohydrate from the sample and thereby invalidate the results. In practice this did not happen: the accuracies of the assays were sufficiently high that even when the carbohydrate in the blank was five times larger than that in the sample, a reliable quantitative assay of the sugar in the sample was obtained.

The two assays that were used, the "anthrone" and "phenol" tests, were specific for carbohydrates in general, and not for any carbohydrate in particular. Both were modified to make them applicable to submicromole quantities of sugar. The anthrone test (26, 95, 134), which gives positive reactions with hexoses, 6-deoxyhexoses, hexuronic acids, and aldopentoses,* was found to be both accurate and reproducible when used as described below. The phenol test (27), though slightly less accurate, was more rapidly and easily performed. Positive reactions are given by all classes of reducing sugars. The latter test had the advantage that to the eye the qualitative contrast between a positive and negative test was more readily apparent than in the anthrone test; however, if a spectrophotometer was used to quantitate the results both tests were equally sensitive. The phenol test suffered from the disadvantage that the orange color of a positive test could be artifically

---

*Positive reactions with anthrone for some sugars are critically dependent on the exact experimental procedures. Thus, by suitably modifying the conditions (cf. Refs. 95, p. 490; and Ref. 134), hexoses and 6-deoxyhexoses can be distinguished from pentoses and hexuronic acids, and these in turn from heptoses and 2-deoxypentoses.

augmented by yellow or orange colored noncarbohydrate contaminants. In the anthrone test a positive result was indicated by a blue-green color; because there are not a great many blue or green colored substances this test was less sensitive to contaminating materials, and in particular to yellow or orange ones. Together the two tests served as good checks for each other.

V G2: Experimental

Anthrone Test

The compound, containing 0-30 μgm glucose equivalents (cf. footnote, p. 125), was dried in a 13 x 100-mm Pyrex test tube. Exactly 2.00 ml of anthrone reagent (0.500 gm anthrone/250 ml fresh* conc. H_2SO_4) was added, and the tube was cooled 10 min in a 5°C water bath.** Exactly 1.00 ml of H_2O was carefully layered onto the cooled mixture, and the tube was kept at 5°C for 5 more min. At the end of this time the contents were mixed, while still in the bath, with a glass stirring rod. A separate rod was used for each tube and was left inside the tube throughout the following procedure. The tube was removed from the 5°C bath and was allowed to stand at 25 ± 3°C for 10 min, at which time it was plunged (t = 0 min) into boiling water at 100°C. The tempera-

---

*It was imperative that the H_2SO_4 come from a freshly opened bottle if reliable results were to be obtained: this was by far the most critical parameter in both the anthrone and phenol tests. If several tests were to be performed over a period of days, it was convenient to purchase the sulfuric acid in small bottles. The same bottle could be used on different days provided it was kept tightly capped between uses.

**When many tests were run simultaneously, it was convenient to carry out all the cooling and heating operations in a metal test tube rack which fitted into the various water baths.

ture of the bath was immediately lowered to 90°C by the addition of cooler water or of ice. The tube was kept at 90°C for exactly 16 min (t = 16 min) at which time it was plunged into an ice bath at 0°C. When the temperature within the tube* had dropped to 27°C the tube was removed from the bath and was allowed to stand at room temperature for 10 min. Directly thereafter (the blue color, though stable for periods of half an hour, was not stable over long periods of time such as two or three hours), the absorbancy at 625 mμ and 1-cm path length was determined.

Phenol Test

The compound, containing 0-20 μgm glucose equivalents, was dried in a 13 x 100-mm Pyrex test tube. Exactly 0.600 ml of 2.5% w/v phenol in H_2O and 1.500 ml of fresh conc. H_2SO_4 were added. The sulfuric acid was added from a calibrated fast flow pipet which had been made by cutting off the tip of a 10-ml graduated pipet and regrinding. It was important to drain the pipet in exactly the same manner as it was drained during the calibration. The tube was shaken, allowed to stand 10 min at room temperature, and again shaken, at which point it was placed in a water bath at 25°C-30°C for 20 min.** The absorbancy at 490 mμ and 1-cm path length was determined. The presence of carbohydrate was indicated by an orange color.

---

*A thermometer was placed directly inside an extra tube which had been carried through the entire procedure to this point.

**Medium sized beakers made perfectly acceptable water baths. The purpose of allowing the tubes to stand at room temperature prior to immersing them in the water bath was to dissipate some of the heat evolved from the reaction. Otherwise the temperature in a small bath would exceed 30°C when the tubes were placed in it.

V G3: Results

Anthrone Test

The following relation held between the concentration of glucose and the absorbancy:

$$\mu gm/ml^{*} = 18.49\ A^{1\ cm}_{625\ m\mu}$$
$$P.E. = 0.37\ \mu gm/ml.$$

For carbohydrates other than glucose the coefficient of $A^{1\ cm}_{625\ m\mu}$ differs from 18.49, and for these carbohydrates the relation given above is only approximate.

Phenol Test

The following relation held between the concentration of glucose and the absorbancy:

$$\mu gm/ml = 31.34\ A^{1\ cm}_{490\ m\mu}$$
$$P.E. = 0.78\ \mu gm/ml.$$

For sugars other than glucose the same considerations apply as for the anthrone test.

Results for R-val-his

Table XIII, p. 130, shows that the blocking group R does not contain carbohydrate. Glucose and adenylic acid, which contained known

---

*In converting concentrations to amounts (μgm) it was important to calculate the volume of the final solution from the known (handbook) density of sulfuric acid at different compositions. The reagent volumes are not additive and to ignore this fact can introduce errors up to 10%. This consideration also applies to the phenol test.

TABLE XIII

$$\frac{\mu m \text{ Glucose Equivalents Found}}{\mu m \text{ Compound}}$$

Compound	Indirect[a] Phenol	Indirect[a] Anthrone	Direct[a] Phenol	Direct[a] Anthrone
Glucose[b]	1.45 ± .35 (3)[c]	0.77 ± .17 (3)	1.30 ± .09 (4)	0.91 ± .03 (2)
Adenylic acid[b]	-	-	1.62 ± .10 (2)	-
R-val-his	-0.07 ± .11 (1)	-0.12 ± .03 (1)	0.24 ± .11 (1)	0.29 ± .03 (1)

a) Cf. p. 125 for definition of these terms.

b) Expected ratio is 1.00. Adenylic acid contains ribose instead of glucose.

c) Number in parentheses is the number of independent analyses which were made.

amounts of carbohydrate, were used as controls.

V II: ASSAY FOR STEROIDS

The microdetermination of steroids was made by reacting the compound to be tested with $SbCl_3$ (110, 113). Antimony trichloride reacts with steroids of all types to form highly colored compounds which fluoresce strongly in ultraviolet light (360 mμ). Compounds other than steroids do not give positive reactions. In many cases the color (viewed both in daylight and in ultraviolet light) that is formed in the reaction with $SbCl_3$ is characteristic of particular classes of steroids and can be used as a further identification aid.

V H1: Experimental

Exactly 0.15 μm of $\beta^{A_{Ic}}T$-1 which had been reduced with $NaBH_4$ (cf. p. 146) was spotted in 1 μl of water onto Whatman No. 1 chromatographic paper. Standards containing 0.07, 0.21, and 0.35 μm of testosterone in $CHCl_3$ and standards containing 0.10, 0.30, and 0.50 μm of cholic acid in a 1:1 mixture of acetic acid and acetone were also spotted. Testosterone and cholic acid are representatives of the steroids that are found in the keto-steroids and bile acids respectively. To guard against the possibility that the peptide moiety of reduced $\beta^{A_{Ic}}T$-1 might itself form, or prevent the formation of, a colored compound with $SbCl_3$, three controls were prepared. In the first 0.15 μm of βT-1, which does not contain the blocking group R, was spotted. In the second, a mixture of 0.15 μm of βT-1 and 0.07 μm of testosterone was spotted. In the third, a mixture of 0.15 μm of βT-1 and 0.10 μm of cholic acid was spotted.

After the samples had dried on the paper, the latter was sprayed* with a saturated solution** of $SbCl_3$ in chloroform. When the chloroform had evaporated from the paper, the latter was heated at 85°C for 5-15 min until maximum color had developed for the standards of testosterone and cholic acid. The paper was then examined with a long wavelength ultraviolet light to detect those samples which fluoresced.

In a modification*** of the above procedure, the paper was sprayed with a saturated solution of $SbCl_3$ in nitrobenzene.**** Otherwise the procedure was the same as in the preceding paragraph.

V H2: Results

When viewed in daylight, testosterone and cholic acid that had been treated with $SbCl_3$ in $CHCl_3$ appeared as violet and peach-colored spots respectively against the white background of the paper. In ultraviolet light these spots fluoresced bright pink and pinkish-yellow respectively against the weak purple background fluorescence from the

---

*In our hands, spraying was preferable to dipping: if the paper was dipped, many of the samples which had been spotted on the paper were washed off because of the high solubility of some of the steroid compounds in $CHCl_3$.

**Approximately 6.6 gm of $SbCl_3$ in 100 ml of $CHCl_3$.

***There are many simple modifications of the $SbCl_3$ test, and all give good results (113). Because the colors that are obtained in a positive test vary from modification to modification, a considerable amount of information about the steroid can be obtained by combining the color data from several modifications.

****Approximately 100 gm of $SbCl_3$ in 100 ml of $CHCl_3$. This modification usually results in a more intense fluorescence than the $CHCl_3$ modification.

Whatman No. 1 paper.*

The papers which had been treated with $SbCl_3$ in nitrobenzene showed gray and green spots in daylight for testosterone and cholic acid respectively. Both these spots fluoresced a bright yellow-white under ultraviolet light.

Behavior that was identical to that which has just been described was observed for the control samples that consisted of a mixture of testosterone or cholic acid with βT-1. The control sample of βT-1 to which no steroid had been added showed neither a color in daylight nor a fluorescence in ultraviolet light.** These observations demonstrate that the peptide moiety of $\beta^{A_{Ic}}$T-1 does not interfere with the $SbCl_3$ test.

$\beta^{A_{Ic}}$T-1, which contains the blocking group R, behaved exactly as βT-1.

V H3: Conclusions

Even had no more than 0.02 μm of a steroid been present in the 0.15 μm of reduced $\beta^{A_{Ic}}$T-1 which was examined, we could have detected it both by the color it would have given in daylight after treatment with $SbCl_3$ as well as by its ultraviolet fluorescence. The failure of

---

*The colors which are observed in daylight for a given steroid depend on a variety of factors. For a given steroid, the temperature at which, and the time during which, the paper is heated, the time which has elapsed between heating the paper and viewing it, and the concentration of the steroid can all affect the color. These considerations apply to a lesser extent for the colors which are observed in ultraviolet light.

**Compounds, such as βT-1, which did not contain steroids usually caused a slight quenching of the weak purple background fluorescence of the Whatman No. 1 paper.

$\beta^{A_{Ic}}{}_{T-1}$ to give either a color or a fluorescence with $SbCl_3$ demonstrates that the blocking group R in this compound is not itself, nor does it contain, a steroid.

V I: SUMMARY OF PRECEDING ASSAYS

The group R which blocks the N-terminus of the $\beta^{A_{Ic}}$ polypeptide chain of hemoglobin A_{Ic} is not a nonketoacyl-group of less than five carbon atoms, is not carboxy, carbamyl, pyridoxal or other aromatic aldehyde. The group R does not contain phosphorus or carbohydrate. R is not itself, nor does it contain, a steroid.

From biological considerations alone, a priori it seemed that R stood a good chance of being one of the above groups. Failure to find such groups was very disappointing. It was now clear that nothing short of a complete modern classical structure analysis on a highly purified compound containing R was likely to end in success. The following two sections of the thesis are devoted to such an analysis.

VI: THE MOLECULAR WEIGHT OF THE BLOCKING GROUP R

VI A: INTRODUCTION

The molecular weight of a compound is of primary importance in establishing its structure. Knowledge of this parameter severely limits the number of possible atomic arrangements. To determine the molecular weight accurately requires a highly purified sample of the compound. Because of the limited amounts of A_{1c}, $\beta^{A_{1c}}T-1$, and R-val-his available, it has not yet been possible to isolate sufficient material of the required purity to make feasible an exact determination of either its molecular weight or of its elemental composition. Nevertheless, an estimate of the molecular weight of R has been made which is in error by at most seven carbon atoms,* and probably by less than three (out of a total of about 22 carbon atoms).

VI B: EXPERIMENTAL

VI B1: Methodology

A weighed amount of R-val-his was hydrolyzed to its constituent amino acids with HCl. The amino acids were analyzed on the amino acid analyzer to determine the number of moles of R-val-his that had been hydrolyzed. The molecular weight was obtained by dividing the weight of R-val-his by the number of moles found by amino acid analysis.

In practice, the experimental procedure was not as simple as that

---

*More precisely, by $\pm$ 84 atomic weight units. We have expressed the error as so many carbon atoms because it is somewhat easier to visualize the latter than to visualize atomic weight units. By expressing the error in this manner we do not mean to preclude the possibility that R might contain atoms other than carbon.

just described. The R-val-his was not a highly purified sample, but was contaminated with impurities which had been picked up during the chromatographic procedures (cf. pp. 48-50). It was therefore necessary to determine that percentage of the weighed sample which was due to these contaminants. This was accomplished by weighing several chromatographic zones taken from the same chromatogram as that from which R-val-his had been isolated. These control zones were taken on either side of the zone which contained R-val-his so as to allow for the possibility that the amount of contaminant was a function of the effluent volume. Some of these control zones contained peptides whose molecular weights were about the same as that of val-his, and so were in every sense equivalent to the zone from which R-val-his was obtained.

VI B2: Detailed Procedures for the Molecular Weight Determination

A 13 x 100-mm Pyrex test tube was dried at 110°C for 1 hr, then allowed to cool over P_2O_5 in a desiccator prior to weighing the tube on a microbalance.* The desiccator was not evacuated, for it was found that more reproducible weighings resulted by cooling the samples at atmospheric pressure: this was probably due to the fact that the stream

---

*Before each tube was weighed, it was brushed lightly with a camel's hair brush to remove any lint or dust particles. Occasionally a tube would give very erratic scale readings which were probably caused by electrostatic charges on the tube surface. Such erratic readings could be avoided by wiping the tube with a barely damp chamois which had been prepared by wringing a thoroughly wet chamois as dry as possible, blotting it between two pieces of filter paper, and storing it in a covered Petri dish. The freshly wiped tube was brushed with the camel's hair brush, was placed on the weighing hooks, and the reading was recorded after exactly five minutes. The very thin film of water which was left on the tube by the chamois evaporated almost instantaneously.

of air which results upon opening an evacuated desiccator can induce electrostatic charges upon the dry objects within it. Each tube was dried and reweighed until the probable error of the average of three or more weighings was less than 8 μgm (cf. p. 140). After the tube had been dried to constant weight, the sample, containing approximately 600 μgm R-val-his, was added to the Pyrex tube and dried *in vacuo at room temperature* over P_2O_5. It was not possible to dry the sample by heating because of the lability of the blocking group R (cf. pp. 67-70). The drying procedure just described did not remove the blocking group R from R-val-his. After the tube with sample had been brought to constant weight, five 0.2-ml portions of doubly glass-distilled 6 F HCl were added to the tube; care was taken to rinse down the sides of the tube in the process. An additional 1 ml of 6 F HCl was added to the tube, and the latter was evacuated with a water pump and sealed. The sample was hydrolyzed for 22 hr at 110°C. The resulting amino acid mixture of valine and histidine was analyzed on the amino acid analyzer to determine the number of moles of R-val-his which had been present in the weighed sample.

Exactly the same procedure was followed with the controls.

VI C: RESULTS

Calculation of the molecular weight of the blocking group R is shown in Table XIV, p. 139.

TABLE XIV

Calculation of the Molecular Weight of R

Molecular weight L-ala_3 [a]	243 ± 2 [a]
Weight R-val-his + contaminants	774 ± 17 µgm
Weight column contaminants [b]	136 ± 28
Weight R-val-his	638 ± 33
µm R-val-his	1.232 ± .04
Molecular weight R-val-his	518 ± 27
Molecular weight val-his	254
Molecular weight R	265 ± 27 [c]

[a] The molecular weight of a synthetic sample of L-ala_3 was determined in triplicate by the method which is described in the text. The true molecular weight of L-ala_3, calculated from its atomic composition is 231.34. It thus appears that the method itself has an inherent accuracy of 5%. The fact that the experimentally determined molecular weight of L-ala_3 is higher than the true molecular weight is probably a reflection of the hygroscopic nature of the polypeptide. The small error (± 2 atomic weight units) between triplicate samples of L-ala_3 must be considered accidental, for the sum of the weighing and amino acid analysis errors was ± 7 awu.

[b] Determined by evaporating and weighing fractions 130-137, 302-315, and 367-376 of Fig. 9, p. 68. The weight of each zone was reduced by the weight of any amino acids or peptides which it contained (none, glu-lys, and lys, respectively), and was then adjusted proportionally for differences in the number of chromatographic fractions in that zone and the zone from which R-val-his (fractions 316-330, Fig. 9, p. 68) was taken. The average of these adjusted values for the three zones was 136 ± 28 µgm. It should be noted that the agreement between the three zones is quite good, for 17 µgm out of the 28 µgm was due to weighing and amino acid analysis errors. There was thus an error of only 11 µgm due to uncertainty in the amount of contaminants themselves.

[c] This error is divided as follows: weighing 9, amino acid analysis 9, and contaminants 9.

VI D: DISCUSSION

From footnote c) of Table XIV it appears unlikely that the estimate of the molecular weight of R can be improved upon by more than ± 1 carbon atom, for the only errors which can be significantly reduced are those due to contaminants (by using a purer sample of R-val-his) and the weighing error. In view of the fact that the accuracy of a microbalance is in the range of 1-2 μgm attempts were made to reduce the weighing error below the ± 8 μgm mentioned at the top of p. 138. These attempts were uniformly unsuccessful. Two factors probably contributed to the 8 μgm error: First, water vapor in the weighing room was rapidly adsorbed onto the Pyrex tube and/or the sample inside the tube; second, the temperature, humidity, and vibration in the weighing room was poorly regulated. The error of 8 μgm from the weighing was of the same order of magnitude as the accuracy of the amino acid analyzer so that it was not worthwhile to attempt to reduce it further unless this could be done easily.

The above considerations demonstrate that the molecular weight of R which was obtained is about the best that the method is inherently capable of giving. Further improvements in the knowledge of the molecular weight of R must come from more precise methods such as mass spectrometry.

VI E: MOLECULAR WEIGHT OF R BY MASS SPECTROMETRY

An attempt was made to obtain the molecular weight of R-val-his, and hence of R, by mass spectrometry. This attempt arose not only from

the considerations of the preceding page, but from the fact that not only the molecular weight, but the elemental composition, and frequently the entire structure of a compound can be deduced from a good quality mass spectrogram.

Discussions with Dr. Klaus Biemann,* who has had considerable experience in interpreting the mass spectra of small peptides as well as of other compounds, indicated that there was a reasonable chance of obtaining a good quality spectrogram from the sample of R-val-his shown in Table XIV even though the R-val-his was contaminated to the extent of one-fifth its weight with impurities from the chromatography which had been used to purify the blocked dipeptide.

Accordingly, we sent Dr. Biemann samples of highly purified L-val-L-his, N-acetyl-L-val-L-his (cf. Appendix I, p. 201), and the impure sample of R-val-his of Table XIV. The three control zones of footnote b, Table XIV were also sent. The complete structures of L-val-L-his and N-acetyl-L-val-L-his were deduced from their mass spectrograms by Dr. Biemann and a graduate student of his, Mr. John Hays. The spectrum of L-val-L-his was recognized as present in the mass spectrogram of R-val-his; however, the contaminants in the latter sample prevented any deductions to be made about the molecular weight or the structure of the blocking group R.

---

*Chemistry department, Massachusetts Institute of Technology, Cambridge, Mass.

VII: THE NATURE OF THE CHEMICAL BOND BETWEEN THE BLOCKING GROUP R AND THE N-TERMINAL AMINO ACID OF THE $\beta^{A_{Ic}}$ SUBUNIT

VII A: INTRODUCTION

In the preceding sections we have established that the structure of A_{Ic} is $\alpha_2\ \beta\ \beta^{T-1(val \rightarrow R\text{-}val)}$.* The group R, which blocks the N-terminus of one of the β polypeptide chains, is covalently bonded to the alpha amino group of the N-terminal amino acid valine. R has a molecular weight of approximately 265, and has been shown by specific chemical assays not to be acetyl, carboxyl, carbamyl, nor an aromatic aldehyde. R does not contain phosphorus, carbohydrates, or steroids. From starch gel electrophoresis (cf. pp. 93-96), it is unlikely that R is bonded to the N-terminal valine through an amide bond. The two covalent bond types which are most consistent with the electrophoretic behavior are the substituted amine type bond, and the Schiff base type bond (cf. p. 96).

The alkaline lability of the bond between R and valine favors the Schiff base type bond as does other evidence which has been obtained in preceding sections and which is recapitulated below. It is the purpose of this section to present the experiments which show conclusively that R is covalently bonded to N-terminal valine in a Schiff base linkage.

---

*$\beta^{T-1(val \rightarrow R\text{-}val)}$ is a more specific designation for $\beta^{A_{Ic}}$. The former designation follows Baglioni (cf. footnote p. 19) and Schroeder and Jones (Ref. 131, p. 97), and means that $\beta^{A_{Ic}}$ is identical with the β polypeptide chain of A_{II} except that the amino acid valine in the N-terminal tryptic peptide has been replaced by R-valine.

VII B: RECAPITULATION OF DATA

During the course of the many experiments that have been described in preceding sections of this thesis we observed the following facts which support, but do not prove, the hypothesis that R is bonded to the N-terminal valine of the $\beta^{A_{Ic}}$ chain in a Schiff base bond:

1. $\beta^{A_{Ic}}$T-1 or R-val-his which had been spotted on chromatographic paper and allowed to dry reacted only very weakly with 0.5% w/v ninhydrin/acetone. However, when the ninhydrin treated paper was heated to 80°C a deep purple color, comparable in intensity to that given by an equivalent amount of βT-1 or val-his, developed (cf. pp. 47-48, and 67).
2. When $\beta^{A_{Ic}}$T-1 was concentrated from large volumes of buffer by evaporation, sometimes only βT-1 was found in the residue; i.e., the blocking group R had been cleaved from $\beta^{A_{Ic}}$T-1 (cf. pp. 58-60, and 69).
3. Each time that purified $\beta^{A_{Ic}}$T-1 or R-val-his was dissolved in water and then reevaporated, the amount of $\beta^{A_{Ic}}$T-1 or R-val-his decreased, and a corresponding amount of βT-1 or val-his was formed. The amount of the latter two compounds which was formed was greater when the evaporation was carried out at 38°C than when it was done at 25°C (cf. p. 69).
4. If R-val-his was heated at 38°C for 16 hr with either 1 ml of water or 1 ml of 10^{-4} F NaOH, it was quantitatively converted to val-his (cf. p. 69).

5. $\beta^{A_{Ic}}$T-1 which had been chromatographed under basic conditions on the quaternary amine anion exchange resin Dowex-1 was quantitatively converted to βT-1 (cf. pp. 50-51, and 69).
6. The stability of the bond between R and valine at acidic pH's (all the successful isolations were carried out below pH 5) and its instability at neutral and alkaline pH's have analogues in the Schiff bases between some proteins and pyridoxal (cf. p. 111).
7. Finally, the electrophoretic behavior of $\beta^{A_{Ic}}$ during starch gel electrophoresis is consistent with a Schiff base type bond between R and valine (cf. pp. 93-96).

Points 1 through 6 emphasize the ease with which the blocking group R is removed from the alpha amino nitrogen under very mild hydrolytic conditions. This lability is typical of many Schiff bases (23).

If indeed, the bond between R and valine is of the Schiff base type as all our evidence suggests, then it should be possible to reduce this bond to a secondary amine by the following reaction:

$$R = N- + 2[H] \longrightarrow HR-\overset{H}{N}- .$$

The success of such a reduction would be of crucial importance. Not only would it prove the Schiff base linkage, but it would also provide a means of overcoming the lability of the bond between R and the N-terminal valine; for in the secondary amine which is formed by the reduction, R is permanently bonded to the nitrogen atom. This lability has been the major cause of our inability to isolate sufficient amounts

of R containing compounds, such as R-val-his, on which to carry out a modern classical structure analysis of R. In particular, misfortunes such as that described on pp. 59 and 60 could not occur.

VII C: REDUCTION OF $\beta\ \beta^{A_{Ic}}$ WITH $NaBH_4$

Reductions with $NaBH_4$ are usually carried out at alkaline pH because of the instability of the hydride anion under acid conditions. However, Meyers and Libano (94) and Crestfield et al. (25) have shown that even at acidic pH values the H^- ion has a sufficient half-life to accomplish the same reductions which are normally done in alkaline media. Because of the lability of the blocking group R under neutral or alkaline conditions, we chose to carry out the reduction in acidic solution.

A 249-mg sample of $\beta\ \beta^{A_{Ic}}$-globin, which had been isolated as described on pp. 34-40, was dissolved in 20 ml of water and was titrated to pH 3.5 with approximately 0.02 ml of 6 F HCl. The resulting solution was stirred with a magnetic stirrer for 1 hr. At the end of this time the solution was slightly turbid and a very pale yellow in color. A 37-mg portion of $NaBH_4$ was dissolved in 1 ml of water to which 4 λ of 4 F NaOH had been added. This solution was added dropwise with a pipet to the solution of globin. It was necessary to add 2 F HCl manually during the addition of $NaBH_4$ in order to keep the pH constant at 3.5. There was much foaming, so that every now and then the form was stirred into the solution with a glass rod. The amount of $NaBH_4$ which was used was a 56-fold excess over that which was required to reduce a Schiff base double bond and to reduce all sulfhydryl groups in the globin to

H_2S.* Two more 37-mg portions of $NaBH_4$ were added at one hour intervals. Three hours after the first addition of $NaBH_4$, the reaction mixture was placed in a dialysis sac and was dialyzed for three 8-hr periods against water at 2°C. The dialysate was lyophilized to dryness to yield 186 mg of reduced globin. Amino acid analysis of the reduced $\beta\ \beta^{A_{Ic}}$-globin gave the same amino acid composition as the unreduced globin (cf. Table VI, p. 99a), so that no extensive degradation of the protein occurred during the reduction.

VII D: ISOLATION OF REDUCED $\beta^{A_{Ic}}$T-1

Reduced $\beta\ \beta^{A_{Ic}}$-globin from the above procedure was hydrolyzed with trypsin (cf. p. 40), and the tryptic peptides were isolated on Dowex-50 (cf. p. 41).** The Dowex-50 chromatogram was similar in detail to that of Fig. 4, p. 44 except that zone #5 had moved into the region of zone #6, although a shoulder was still visible on the left hand side of the latter zone. $\beta^{A_{Ic}}$T-1 is normally found in zone #5 of Fig. 4 (cf. pp. 46-48). If in fact $\beta^{A_{Ic}}$T-1 were a Schiff base, then reduction would convert it into a secondary amine (cf. p. 145); and the reduced peptide, being more basic, would move more slowly on Dowex-50 than $\beta^{A_{Ic}}$T-1. The observed movement of zone #5 into the region of zone #6 was consistent with the success of the reduction. Indeed, when the zones from the Dowex-50 chromatogram were analyzed by paper

---

*The reduction of the sulfhydryl groups to H_2S was a conceivable but not very likely side reaction.

**Because of the smaller amount of material, a proportionally smaller Dowex-50 column was used to separate the peptides.

electrophoresis and paper chromatography (cf. p. 45), no peptide having the electrophoretic mobility (cf. Fig. 5, p. 46) of $\beta^{A_{Ic}}$T-1 was found. Instead, a histidine-positive peptide having the electrophoretic and paper chromatographic mobility of βT-1 was found in zone #6. This observation was again consistent with the success of the reduction: one would expect the paper chromatographic and electrophoretic mobilities of reduced $\beta^{A_{Ic}}$T-1 and βT-1 to be very similar: the difference between the basicities of a secondary amine (reduced $\beta^{A_{Ic}}$T-1) and a primary amine (βT-1) is slight. Because zone #6 is adjacent to zone #7 in which βT-1 is normally found (cf. pp. 46-48), it was possible that the βT-1-like peptide in zone #6 was not reduced $\beta^{A_{Ic}}$T-1, but rather was simply βT-1 from zone #7 which had overlapped into zone #6. This possibility was ruled out by rechromatographing the mixture of peptides in zone #6 on Dowex-1 by the method of Schroeder and Robberson (132).* Two peptides were isolated from the Dowex-1 chromatography: βT-9 (cf. Table IX, p. 107), which was expected, because it is usually found in zone #6 from the Dowex-50 chromatography; and a peptide which was contaminated with βT-9, and which from its amino acid analysis could have been either reduced βT-1** or βT-1. The latter peptide was

---

*This method is a more recent modification of the procedures described in Refs. 6 and 127. The major difference is that a different sequence of buffers is used to develop the Dowex-1 columns. The new buffers give a more gradual decrease in the pH gradient. This reduction in gradient results in better resolution of neutral peptides.

**An amino acid analysis of βT-1 would give the amino acid composition (val, his, leu, thr, pro, glu_2, lys) (cf. Table VIII, p.102a), whereas an analysis of reduced $\beta^{A_{Ic}}$T-1 would give the composition (his, leu, thr, pro, glu_2, lys). That is, the amino acid valine would be

again chromatographed on Dowex-50, but this time buffers of lower ionic strength were used in an attempt (successful) to free the peptide from the contaminating βT-9. The chromatographic system which was used has not been described before, and is therefore given in detail in the following paragraph.

A 0.6 x 60-cm Dowex-50 X2 column was prepared in the usual manner (127) and was equilibrated at 38°C with 65 ml of buffer B of Table I, p. 52.* The sample, containing up to 10 μm peptide, was dissolved in 1 ml of buffer A and driven into the column with an air line. The column was developed with a two-vessel gradient device (15) in which the ratio of the areas of the mixer and reservoir were 2:1 respectively. The mixer and reservoir were filled with 66 ml of buffer B and 33 ml of buffer C respectively. One-ml fractions were collected at a flow rate of 15 ml/hr. When both mixer and reservoir had emptied, they were refilled with 100 ml of buffer C and 50 ml of buffer F respectively, and the development of the chromatogram was continued until the gradient device had again emptied. At this point the gradient was discontinued and development was continued with buffer F for 40 fractions, at which point the chromatogram was finished.

The results of the Dowex-50 chromatogram which has just been described are given in Fig. 12, p. 150. The amino acid analysis of the

---

missing, because the alpha amino group of this amino acid would be bonded to the blocking group R as a secondary amine and would be incapable of reacting with ninhydrin.

*All buffers which are referred to in this paragraph are those in Table I, p. 52.

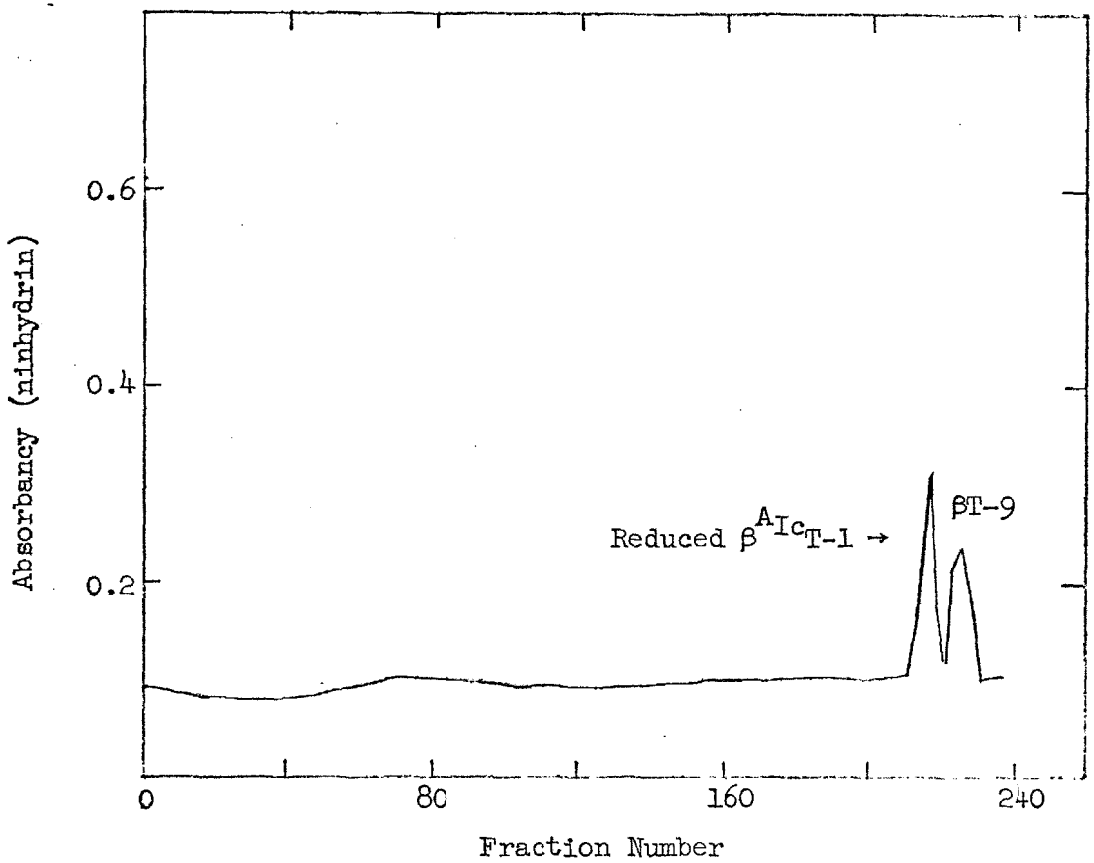


Fig. 12: Purification of reduced $\beta^{A_{Ic}}{}_{T-1}$ by chromatography on Dowex-50 X2.

peptide from the zone labeled "Reduced $\beta^{A_{Ic}}$T-1" in Fig. 12 is given in Table XV. For comparison, the analysis of BT-1, from zone #7, (cf. p.

TABLE XV

Amino Acid Composition* of Reduced $\beta^{A_{Ic}}$T-1 and βT-1

Amino Acid	Reduced $\beta^{A_{Ic}}$T-1	βT-1
lys	0.95	0.99
his	0.87	0.86
thr	0.95	0.95
glu	2.00	1.97
pro	1.01	1.01
val	0.04	0.99
leu	0.82	1.00

*The numbers in the table are the number of amino residues found in the peptide heading the column.

148) of the Dowex-50 chromatogram of the tryptic peptides from reduced β $\beta^{A_{Ic}}$, is also given in the Table; this peptide was purified from zone #7 by rechromatography on Dowex-1 in the manner described on p. 148. The important point in Table XV is the absence of the amino acid valine in the new peptide which otherwise has an amino acid composition identical to βT-1 (cf. second footnote, p. 148). The possibility that valine is missing because of a conceivable hydrolysis of the amide bond between valine and histidine (the latter is the amino acid penultimate to the N-terminus) during the reduction with $NaBH_4$ is unlikely from the fact that the normal yield of βT-1 was found in zone #7, as well as from the fact that $NaBH_4$ does not normally attack amide bonds. Positive proof that hydrolysis of the amide bond between valine and

histidine did not occur was obtained by hydrolyzing reduced $\beta^{A_{Ic}}T-1$ with papain (cf. p. 63). The resulting fragments were electrophoresed as described on p. 45 . The electrophorogram is shown in Fig. 13, p. 153. Had the amide bond between valine and histidine been split during the reduction with $NaBH_4$, the α-amino group of histidine would be free to react with ninhydrin. Therefore the papain hydrolysate of the peptide that we have labeled reduced $\beta^{A_{Ic}}T-1$ in Table XV would contain at least one peptide fragment which would contain histidine and which also would react positively with ninhydrin. As can be seen from Fig. 13 the papain hydrolysate contained only a single peptide which contained histidine; in Fig. 13 this peptide is indicated by the arrow. However, this peptide gave no reaction whatsoever with ninhydrin at 80°C. Thus, the α-amino group of the histidine residue in this fragment is not free, but as Table XV shows, can be made free by hydrolysis with HCl. From the known amino acid sequence of $\beta^{A_{Ic}}T-1$ (cf. Table IX , p. 107) the only group which can be blocking the α-amino group of the histidine containing peptide in Fig. 13 is the reduced R-valyl group.*

The great similarity in Fig. 13 of the electropherograms of the papain hydrolysates of βT-1, $\beta^{A_{Ic}}T-1$, and reduced $\beta^{A_{Ic}}T-1$ makes it highly probable that the peptide which is indicated by the arrow in Fig. 13 is in fact reduced R-val-his; this assignment is further

---

*Reduced $\beta^{A_{Ic}}T-1$ itself gives a weakly positive ninhydrin reaction because of the presence of the ε-amino group of lysine. The main purpose of the papain hydrolysis was to obtain a peptide that contained histidine and which was free from lysine so that a positive or negative reaction of this peptide with ninhydrin could be interpreted unambiguously.

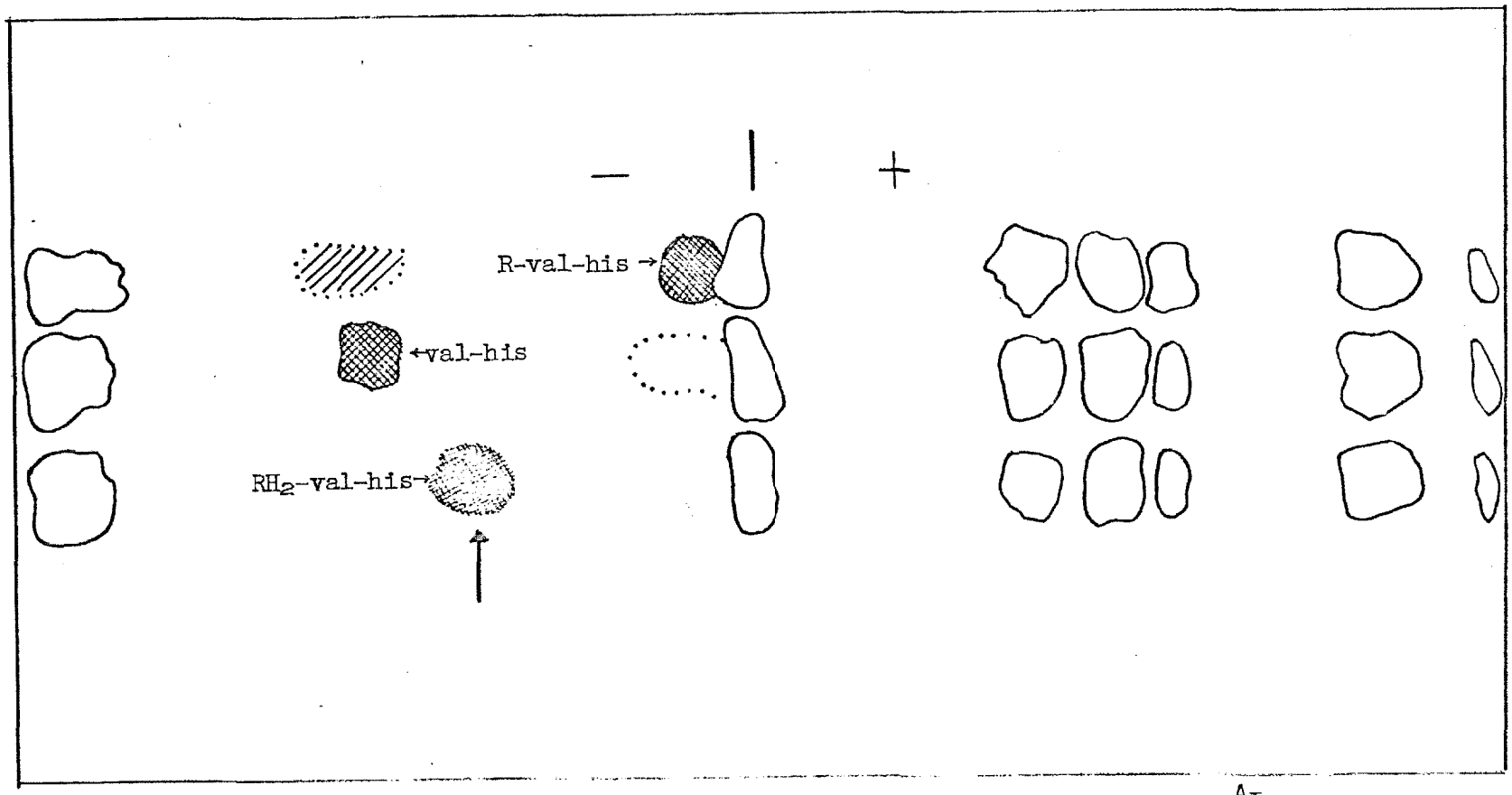
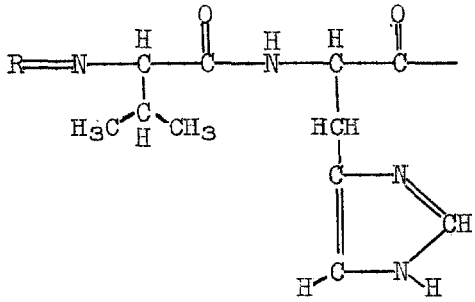


Fig. 13: Paper electrophoretic patterns at pH 6.4 of papain hydrolysates of β^{AIc}T-1 (top), βT-1 (middle) and β^{AIc}T-1 that has been reduced with $NaBH_4$ (bottom). The vertical arrow points to reduced R-val-his (RH_2-val-his). Solid and dotted lines outline spots that gave a strong and very faint color, respectively, with ninhydrin. The crosshatched spots reacted very strongly for histidine and the singly hatched spot very weakly for histidine. The vertical bar shows the line of application.

strengthened by the total absence of either val-his or R-val-his in the papain hydrolysate of reduced $\beta^{A_{Ic}}$T-1.

VII E: CONCLUSIONS

The successful isolation of reduced $\beta^{A_{Ic}}$T-1 proves the hypothesis of a Schiff base type bond between the blocking group R and the N-terminus of the $\beta^{A_{Ic}}$ polypeptide chain:



Further, the above structure implies that R itself derives from either an aldehyde or a ketone:

R = O .

The molecular weight of this aldehyde or ketone is 281 ± 27 (P.E.).*

VII F: DISCUSSION

It has been shown by Cordes and Jencks (23) that the pK_a for the proton of the nitrogen atom of tertiary butyl amine which is bonded in a Schiff base type linkage with benzaldehyde is 6.70 at 25°C. When electron withdrawing groups are substituted on the benzene ring, the pK_a is less than this figure, and when electron donating groups are present

---

*This number was obtained by adding the atomic weight of oxygen (16) to that of R (265).

in the benzene ring the pK_a is larger. From Fig. 13, the pK_a for the proton on the Schiff base nitrogen of R-val-his can be calculated to be 6.64 ± 0.14* at about 10°C. These authors have also shown that at pH 2 the half-life time at 25°C for the hydrolysis of aromatic Schiff bases increases from 6 sec for p-nitrobenzene to 2.5 hr for p-methoxybenzene. These differences gradually become less as the pH increases until at pH 8 the half-life time is of the order of 1 min for all monosubstituted benzene ring systems.

One would expect Schiff bases formed from nonaromatic aldehydes to be less stable than those formed from aromatic aldehydes because the former have no resonance stabilization. The blocking group R is known to be nonaromatic (cf. p. 122). Yet hemoglobin A_{Ic} is stable over periods of at least a month. This relative stability of A_{Ic} compared to the Schiff bases that were discussed in the preceding paragraph indicates a rather large stabilizing force is present in the protein. From the viewpoint of theoretical chemistry it would be very interesting to know what this force is. Probably it is related to the tertiary and perhaps even the quaternary structure of hemoglobin A_{Ic}. An X-ray investigation of hemoglobin A_{Ic} in the vicinity of the N-terminus of the $\beta^{A_{Ic}}$ chain might be quite rewarding.

---

*The method of calculation is given in Appendix III, p. 215.

Part III

THE IN VITRO BIOSYNTHESIS OF A_{Ic}

VIII: THE IN VITRO BIOSYNTHESIS OF A_{Ic}

VIII A: PURPOSE

It is the purpose of the experiments which are described below to determine, insofar as it is possible, answers to the following questions: What is the origin of hemoglobin A_{Ic}? Is it synthesized by the cell independently of the major hemoglobin A_{II}, or is it a degradation or metabolic product of A_{II}? Conceivably, all of the above processes may occur simultaneously; if so, what is the rate at which each occurs? Finally, what is the relationship between the amount of A_{Ic} in an erythocyte and the physiological age of that erythrocyte?

VIII B: INTRODUCTION

Historically, many insights into the biology and biochemistry of living, *in vivo*, systems have been gained by *in vitro* experimentation. A major contribution of the latter type experiment to science is to suggest important qualitative principles operative in the living system. It is rare that a *principle* found true *in vitro* is false *in vivo*. Much more care must be exercised, however, in extrapolating the *quantitative* results of an *in vitro* system to the living system. Although in principle it should be possible, with sufficient care and knowledge, to design an *in vitro* experiment to resemble the living system as closely as is desired, in practice this is difficult. The test of course is the number of experimental phenomena which accurately mimic the real situation. Usually the experimenter is only interested in one facet of the system under study, and he therefore looks for only that facet. As a result it can sometimes happen that he has selected the very one

which fails to correspond to reality. In the field of hemoglobin biosynthesis this has actually occurred (121a). Fortunately in this case the experimentalist was able to perform the experiment both *in vivo* and *in vitro* and correct the error. This study has been undertaken with the preceding points in mind.

In vitro protein synthesis in reticulocytes, liver, and HeLa cells has been shown to take place with the aid of intracellular particles called polysomes. Polysomes consist of aggregates of smaller particles which are called ribosomes. Each ribosome is approximately 180 Å in diameter and is composed of approximately equal amounts of RNA and protein. It is probable that in the healthy reticulocyte the pentasome (*i.e.*, a polysome consisting of an aggregate of five ribosomes which may be sequentially arrayed in the form of a closed loop) is the major site of protein synthesis. The rate at which protein synthesis proceeds is linearly proportional to the number of and the activity of these pentasomes. This rate decreases with cell age or mishandling. It is not the purpose of this introduction to survey the experiments upon which the above statements are based; however, for the interested reader, two recent and pertinent papers are those by Philipps (114) and Glowacki and Millette (36). Adequate references to earlier work may be found therein.

The experiments which were referred to in the preceding paragraph were done primarily with rabbits. Both technical and ethical considerations have made corresponding evidence for the synthesis of proteins in humans difficult to obtain. What evidence there is has been obtained

from in vitro studies of the hemoglobins of genetically abnormal (52, 115, 138, 139) and of anemic or otherwise hemoglobinopathic (18, 115) persons. In vivo studies have been limited to the hemoglobins of terminal cancer patients (118). Whether or not such studies can be extrapolated to normal individuals is open to question; however, at present there is no strong reason to believe that the general course of protein synthesis in human red cells differs from that in the rabbit.

The purpose of the preceding two paragraphs has been to lay an experimental foundation for the theoretical treatment which is developed below, and in particular for the hypothesized constancy of the rates at which A_{Ic} and A_{II} are synthesized under conditions of constant ribosomal activity. Because the interconversion of these two hemoglobins was the system to be investigated, no experimental basis existed for the first order kinetics which were assumed for this interconversion. These kinetics were chosen both because they were the simplest possible, and because they seemed reasonable in light of the chemical structures of A_{Ic} and A_{II}. Anticipating much of what is to follow, it can be stated here that the theory successfully predicted the behavior which was found experimentally and only that behavior.

VIII C: THEORETICAL BASIS FOR EXPERIMENTS

VIII C1: Incubation Experiments: Theory and Definition

If A_{Ic} and A_{II} are made independently at constants rates A and B respectively, and if each is made from the other at a rate proportional to the amount of the other present, then

$$\frac{dA_{Ic}}{dt} = A + kA_{II} - k_{-}A_{Ic}$$

$$\frac{dA_{II}}{dt} = B + k_{-}\,A_{Ic} - kA_{II}$$

where k is the proportionality constant for the formation of A_{Ic} from A_{II}, and k_{-} that for the formation of A_{II} from A_{Ic}. If at time zero no A_{Ic} or A_{II} is present, the solution to these differential equations is

$$A_{Ic} = \frac{(Ak_{-} - Bk)}{(k + k_{-})^2}\left[\,1 - e^{-(k + k_{-})t}\,\right] + \frac{(A + B)k}{k + k_{-}}t$$

$$A_{II} = \frac{(Bk - Ak_{-})}{(k + k_{-})^2}\left[\,1 - e^{-(k + k_{-})t}\,\right] + \frac{(A + B)k_{-}}{k + k_{-}}t.$$

For short times these equations can be expanded in a power series:

$$A_{Ic} = At - \tfrac{1}{2}(Ak_{-} - Bk)t^2 + \tfrac{1}{6}(Ak_{-} - Bk)(k + k_{-})t^3 - \cdots$$

$$A_{II} = Bt + \tfrac{1}{2}(Ak_{-} - Bk)t^2 - \tfrac{1}{6}(Ak_{-} - Bk)(k + k_{-})t^3 + \cdots .$$

These equations would apply for example to the <u>in vitro</u> synthesis of radioactive A_{Ic} and A_{II} if at time t = 0 radioactive L-valine were added to a suspension of reticulocytes in an appropriate nutritive medium. The equations would allow one to calculate the amounts of radioactive A_{Ic} and A_{II} present at time t provided, of course, that the assumptions on which the equations were derived held for the <u>in vitro</u> system.

An experiment of the type just described is called an <u>incubation</u> experiment. With two exceptions, such an experiment is in principle capable of giving a complete answer to the questions initially asked.

If the amount of radioactive A_{Ic} formed is plotted against time, and the amount of radioactive A_{II} formed is similarly plotted, the initial slope of these plots gives A and B respectively. The coefficients α and β of t^2 and t^3 respectively in the power series expansion of A_{Ic} can be determined by the usual methods of curve fitting from the observed deviations from linearity in the plot. From these coefficients the rate constants k and k_- can be calculated:

$$k_- = \frac{-(3B\beta + 2\alpha^2)}{(A + B)\alpha}$$

$$k = \frac{-(3\beta + \alpha k_-)}{\alpha}$$

One exception which was referred to above occurs if in the power expansion of A_{Ic} and A_{II} the coefficients of all terms nonlinear in t vanish. This can happen in two ways:

1) $k/k_- = A/B$

2) $k = k_- = 0.$

In this case A and B are still uniquely determined as before but k and k_- are ill-defined to the extent of the above two possibilities. The other exception occurs when k or k_- or both become very large:

3)
$$A_{Ic} = \frac{A + B}{(1 + k_-/k)}t$$
$$A_{II} - \frac{A + B}{(1 + k/k_-)}t .$$

The case where both k and k_- become large corresponds to a rapid chemical equilibrium between A_{Ic} and A_{II}. The ratio k/k_- is the equilibrium

constant and is equal to A_{Ic}/A_{II} at <u>any</u> time. If k/k_- is known, (A + B), but not A and B individually, can be found from the slope of a plot of the amount of radioactive A_{Ic} or A_{II} vs. time. The case where either k or k_- but not both become large can be recognized by the fact that then the amount of radioactive A_{II} or A_{Ic} respectively which is present at any time will be zero. Now the only information which can be extracted is the sum (A + B). The case of rapid chemical equilibrium cannot be distinguished from possibility 1) of the first type of exception (cf. bottom p. 162). Similarly, the case where either k or k_- but not both become large cannot be distinguished from the possibility 2) of the first type of exception where B or A respectively is zero.

VIII C2: <u>Chase Experiments: Theory and Definition</u>

Although in principle an incubation experiment is capable of completely answering the questions we have posed, in practice the accuracy of the data, its incompleteness, or the fact that the experiment has not or cannot be continued for a long enough time may cause the deviations from linearity to be unobservable. If this is the case one cannot extract the maximum of information by an analysis of the above type. Sometimes these difficulties can be reduced by conducting a <u>chase</u> experiment: if at some time during the course of an incubation experiment a large excess, say one thousandfold, of <u>non</u>radioactive L-valine is added to the incubating mixture, all but a negligible fraction of the hemoglobin which is synthesized after that time will be nonradioactive. Thus, one can watch the changes in the amounts of radioactive A_{Ic} and A_{II} that are already present under the assumption A = B = 0. The same

differential equations apply as before. If the amounts of radioactive A_{Ic} and A_{II} that are present at the time the excess of nonradioactive L-valine is added are A^i_{Ic} and A^i_{II} respectively, then at any later time t, measured from the addition of the nonradioactive valine, the amounts of each that are present are

$$A_{Ic} = \frac{A^i_{Ic} + A^i_{II}}{k + k_-} k + \frac{A^i_{Ic}k_- - A^i_{II}k}{k + k_-} e^{-(k + k_-)t}$$

$$A_{II} = \frac{A^i_{Ic} + A^i_{II}}{k + k_-} k_- - \frac{A^i_{Ic}k_- - A^i_{II}k}{k + k_-} e^{-(k + k_-)t} .$$

For short times these become

$$A_{Ic} = A^i_{Ic} - (A^i_{Ic}k_- - A^i_{II}k)t + \tfrac{1}{2}(k + k_-)(A^i_{Ic}k_- - A^i_{II}k)t^2 - \cdots$$

$$A_{II} = A^i_{II} + (A^i_{Ic}k_- - A^i_{II}k)t - \tfrac{1}{2}(k + k_-)(A^i_{Ic}k_- - A^i_{II}k)t^2 + \cdots$$

The usefulness of these equations lies in the fact that the coefficient of t^{n-1} will give the same amount of information as the coefficient of t^n in an incubation experiment. The sensitivity of the experiment is thus increased by an order of magnitude.

VIII C3: <u>Correction for Nonconstant Ribosomal Activity</u>

The above equations were derived on the assumption of a constant rate of hemoglobin synthesis. As we have seen in the introductory discussion this assumption has been verified experimentally when the ribosomal activity has remained constant. A complication which is normally present to some degree or other in an actual incubation is the gradual decrease in ribosomal activity over the period during which the experi-

ment is conducted. This is of course reflected in a decreased rate of protein synthesis. If there is a way to determine the ribosomal activity independently of the amount of protein synthesized, then the amount of protein which is actually made during a given time can be accurately corrected to that which would have been made had the ribosomal activity remained constant. It is this corrected value which should be compared to the theoretical value that is calculated by the equations given above. The equations that are given below for making this correction are valid <u>only</u> for all the <u>exceptional</u> cases which were discussed at the bottom of p. 162 and on p. 163, and even for these cases <u>only</u> when the ribosomal activity decreases linearly with time.*

The correction is made in the following manner. Let the ribosomal activity be given as a function of time by the equation

$$\underline{a} = a - b_1 t \ .$$

The constants a and b_1 can be determined by fitting the experimental points to a least squares curve or by some other appropriate method. Let A^o_{Ic} and A^o_{II} represent the amounts of radioactive A_{Ic} and A_{II} respectively which are actually present at time t, and as before (cf. p. 161) let A_{Ic} and A_{II} represent the amount of A_{Ic} and A_{II} which would

---

*A completely general treatment together with its particularization to the equations which are given in the text for the exceptional cases can be found in Appendix IV, p. 223. In the general treatment all cases, not only the exceptional ones, are considered; and the decrease in ribosomal activity is allowed to assume an arbitrary form. The corrections in the text are particularly simple and proved to be sufficient to interpret the actual experiment.

have been present at time t if the ribosomal activity had remained constant at the initial value. A^{o}_{Ic} and A^{o}_{II} will, of course, be smaller than A_{Ic} and A_{II}. Then for all the exceptional cases* (cf. p. 162) the following relations hold:

$$A_{Ic} = \frac{A^{o}_{Ic}}{(1 - \frac{b_1}{2a} t)}$$

$$A_{II} = \frac{A^{o}_{II}}{(1 - \frac{b_1}{2a} t)}$$

In these equations A^{o}_{Ic} and A^{o}_{II} are the experimentally observed amounts of radioactive hemoglobin. These observed amounts are divided by the quantity $(1 - \frac{b_1}{2a}t)$ to obtain the corrected amounts. It is the latter which should be compared with the theoretical values calculated from the equations on pp. 161-162.

VIII C4: Illustrations of the Behavior Predicted by the Theory for Incubation and Chase Experiments

Fig. 14 and Fig. 15, pp. 166a, 167, show the predicted behavior for two hypothetical incubation and chase experiments. In both experiments the rate of synthesis of A_{II} by the ribosomes has been assumed to be twenty times the rate of synthesis of A_{Ic}.** In Fig. 14 the first

---

*In practice exceptional cases can be recognized from the fact that in these cases either the experimentally observed specific activity of A_{Ic} is equal to that for A_{II} for all times t, or the specific activity of either A_{Ic} or A_{II} is zero.

**The experimental ratio of the cellular concentrations of A_{II} to A_{Ic} is approximately 20:1 (cf. Fig. 16, p. 183).

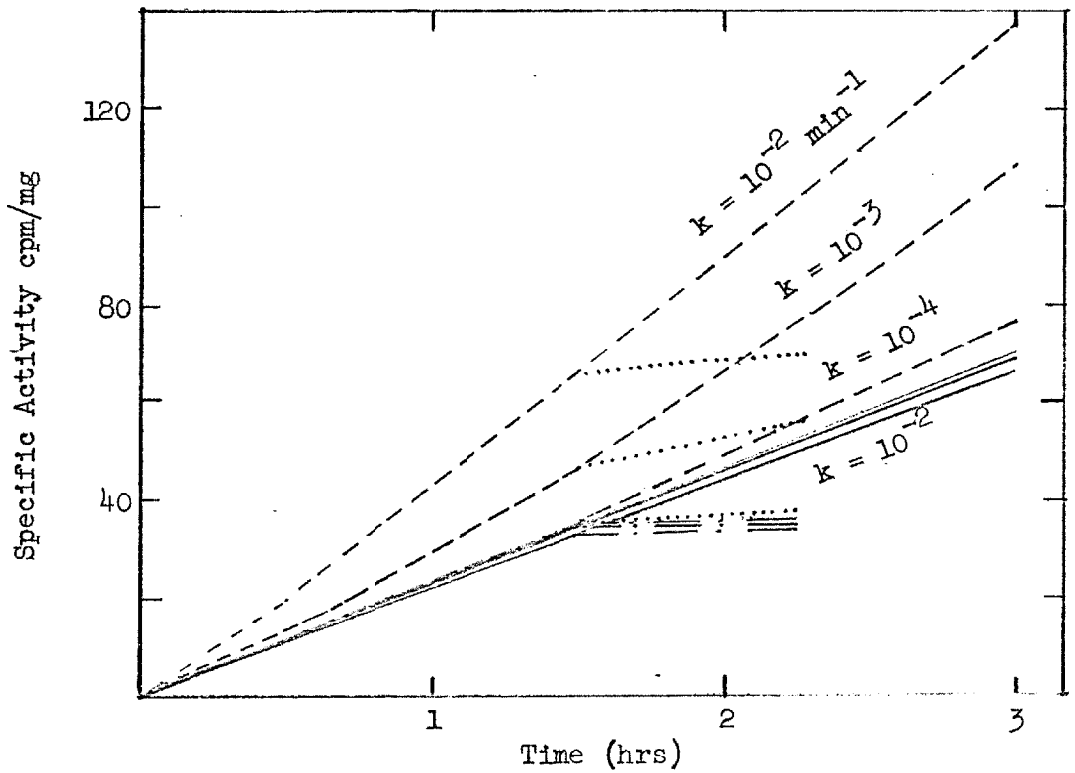


Fig. 14: Predicted behavior for incubation and chase experiments. For all the curves shown B = 20A and k/k_{-} = 0.1. The units of k are min^{-1}. Incubation experiments: A_{Ic} – – –; A_{II} ———. Chase experiments: A_{Ic} ·····; A_{II} –·–·–. See text for further explanation.

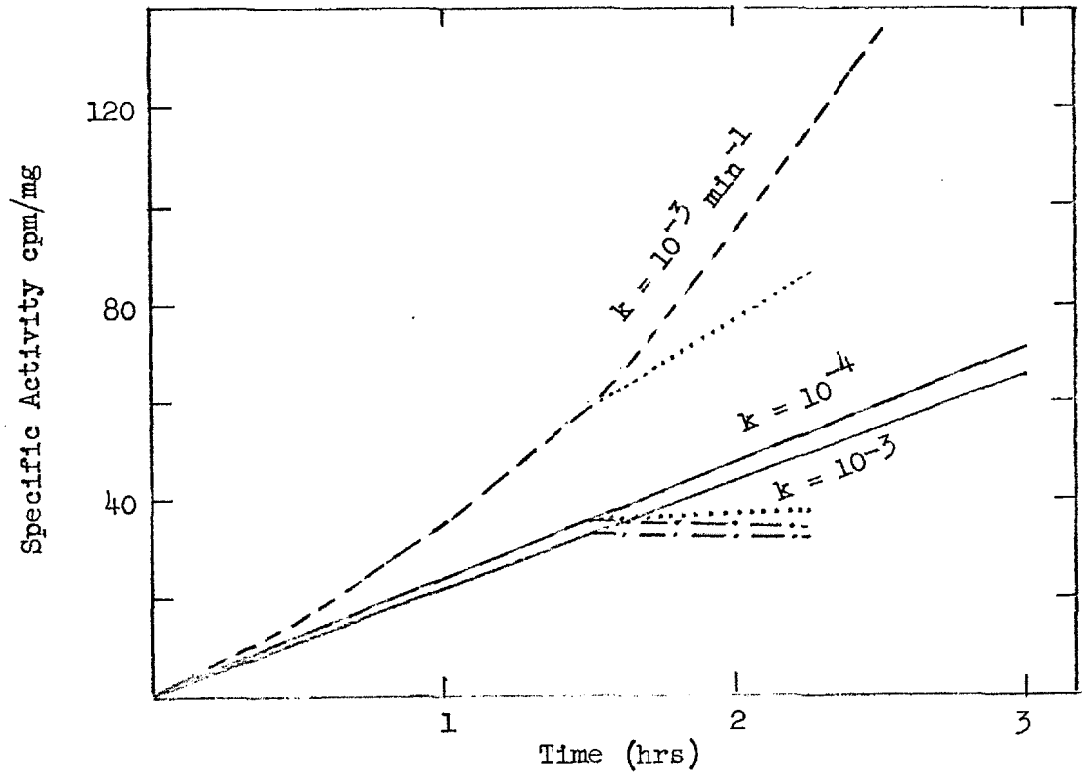


Fig. 15: Predicted behavior for incubation and chase experiments. For all the curves shown B = 20A and $k/k_{-} = 10$. The units of k are min^{-1}. Incubation experiments: A_{Ic} – – –; A_{II} ———. Chase experiments: A_{Ic} ·····; A_{II} –·–·–. See text for further explanation.

order rate constants have been chosen so that there is a slight net synthesis of hemoglobin A_{Ic} from A_{II}.* In Fig. 15 the first order rate constants were chosen so that there is moderate net synthesis of A_{Ic} from A_{II}.* The specific activity units of the ordinate in these Figures are arbitrary, though for a comparison with the experimental results to be given later they have been given the numerical values shown. In both these Figures the important consideration is the shape of the specific activity vs. time curve for A_{Ic} with respect to the corresponding curve for A_{II}: if there is appreciable synthesis of A_{Ic} from A_{II}, the curves for A_{Ic} and A_{II} in the incubation experiments diverge markedly, and the chase curve for A_{Ic} has a definite positive slope; if there is no or only slight net synthesis of A_{Ic} from A_{II} the curves for A_{Ic} and A_{II} in the incubation experiments are more nearly coincident, and the chase curve for A_{Ic} is more nearly horizontal. A careful comparison of these hypothetical curves with the experimental curves to be presented later will permit quantitative conclusions to be drawn about the magnitude of the rates of ribosomal synthesis and interconversion of hemoglobin A_{Ic} and A_{II}.

Figures 14 and 15 also show that if there is no more than a moderate interconversion of A_{Ic} and A_{II}, an incubation experiment of less than one hour duration is not likely to detect the interconversion simply because for times less than this the curves are not very divergent. For

---

*Similar figures result if there is a net synthesis of A_{II} from A_{Ic}, except that the curves for A_{Ic} then fall below those for A_{II} rather than above them as in Figs. 14 and 15.

this reason the incubations which are described below were conducted for periods varying from 15 min to 3 hr. Incubations for periods longer than 3 hr were not done, because for such long term incubations the biological integrity of the *in vitro* system is uncertain,* whereas Philipps (114) has given evidence that for periods up to 2 hr the biological integrity of the reticulocyte is unimpaired, at least with respect to hemoglobin synthesis.

VIII D: EXPERIMENTAL

VIII D1: Introduction

It is not easy to perform a radioactive incubation experiment in the case of human reticulocytes. The reason for this is twofold. First, to obtain a high degree of incorporation of the radioactive compound into the protein being studied, a subject must be found whose blood contains a high percentage of immature young red cells which have a high ribosomal activity; such individuals are almost always anemic, or otherwise ill, and the amount of blood which can be taken from them is severely limited.** Second, the length of time that these patients are hospitalized or willing to donate blood is unpredictable; as a result

---

*In a closed *in vitro* system there is no way for the cell to get rid of waste products. Therefore after a certain period of time, an impairment of the biological function of the cell is to be expected.

**In order to find such individuals, we contacted various doctors who agreed to call us when they had a suitable patient. Such calls came suddenly and at the most unexpected times. Consequently, it was necessary to prepare a great many of the reagents several days in advance so that when the call did come the incubation could be done expeditiously before the cells lost their activity.

it is seldom that an experiment can be repeated a second time with blood from the same patient. It is therefore not possible to vary conditions in such a manner as to optimize them. In the experiments which are described below, the quantitative aspects of the experimental results would be strengthened by additional data. The results are quite sufficient however to establish certain qualitative principles, and it is these principles that we wish to emphasize.

The conditions which were used in preparing the donor's cells for incubation, in the incubation and chase experiments themselves, and in the subsequent isolation of the radioactive ribosomes and soluble proteins were essentially those that Philipps (114) found to be suitable for *in vitro* incubation of rabbit reticulocytes with C^{14}-amino acids. A summary of these conditions is given in Table XVI, p. 171, which may be conveniently consulted in conjunction with the descriptions to follow.

VIII D2: Preparation for the Incubation and Chase Experiments

Several days before it was expected that the doctor would call saying that he had a suitable patient, the "reagent mixture,"* human serum, and solution of L-valine-C^{14} were prepared and stored frozen. Chromatographic developer #4,** STM, TM, NKM, 0.3% BSA, 5% TCA, 10%

---

*Many solutions and compounds are abbreviated throughout the experimental description to conserve space. Their compositions may be found in Tables XVI-XX on pp. 171-174. Tables XVI-XIX define the composition of the incubation mixture; Table XX gives the composition of miscellaneous reagents which are needed in the work-up of the cells prior to and after the incubation.

**The composition of developer #4 is given in the footnote to p. 90.

TABLE XVI

Incubation and Chase Conditions (a)

Reaction Mixture	Incubation	Chase
Packed red cells	1.5	-
Reagent mixture (b)	4.9	2.5(f)
Unfiltered human serum (c)	5.0	2.5
2% Ferrous ammonium sulfate	0.025	0.0125
5% L-glutamine solution	0.060	0.030
Vitamins solution (d)	0.050	0.025
L-Valine-C^{14} (e)	0.250	-
Chase	-	5.0(g)
Total volume	11.8	10.1
[Val-C^{14}] in reaction mixture	0.25 mM	0.125 mM
[Val-C^{12}] in reaction mixture	-	96.5 mM
[Total valine]	0.25 mM	96.6 mM
Specific activity of total valine	6.5 mc/mm	0.008 mc/mm

(a) All entries in Table are in ml unless otherwise indicated.

(b) See Table XVII.

(c) Hyland Laboratories, Los Angeles, California.

(d) Vitamins Solution Liquid, Basal (Eagle) 100 X, Hyland Laboratories, Los Angeles. For composition see Table XIX.

(e) L-valine-C^{14} with a specific activity of 6.5 mc/mm was prepared by diluting 0.100 mc of uniformly labeled L-valine-C^{14} with a specific activity of 185 mc/mm (obtained from New England Nuclear Corporation, Boston, Mass.) in a total volume of 1.00 ml of 0.01 F HCl with 0.207 ml of 0.0717 F L-valine-C^{12} in 0.01 F HCl. The concentration of total valine was then 12.73 mM.

(f) 0.386 F in L-valine-C^{12}: Prepare by dissolving 0.5 gm of L-valine-C^{12} in 10 ml of reagent mixture. Centrifuge at low speed to remove any undissolved valine and take 2.5 ml of the supernatant. Save an aliquot of the supernatant so that the exact concentration of L-valine (in this case it was 0.386 F) can be determined later by amino acid analysis.

(g) Taken from the incubation experiment.

TABLE XVII

Composition of Reagent Mixture

Earle Solution, Spinner 10 X (a)	50	ml
Amino acid mixture (b)	100	ml
$MgSO_4$ solution, 0.195 F (c)	1	ml
Pencillin G (d)	0.25	ml
Streptomycin sulfate (e)	0.25	ml
Sodium bicarbonate, 7.5% (f)	22	ml

Add distilled water to the above to give a total volume of 500 ml. Store frozen in five 100-ml plastic bottles until use. The pH of this mixture at 37°C under 95% O_2-5% CO_2 is 7.5.

(a) Hyland Laboratories, Los Angeles. The composition of this solution is NaCl, 6.8 gm/l; KCl, 0.40 gm/l; $MgSO_4 \cdot 7H_2O$, 0.20 gm/l; glucose, 1.00 gm/l; $NaHCO_3$, 0.20 gm/l; phenol red, 0.01 gm/l.

(b) See Table XVIII.

(c) 0.195 F; Dissolve 23.47 gm of $MgSO_4$ in 1 liter NKM.

(d) 40,000 units/ml H_2O.

(e) 200 mg/ml H_2O.

(f) Dilute 7.5 gm of $NaHCO_3$ to 100 ml with distilled H_2O.

TABLE XVIII

Composition of Amino Acid Mixture

Amino Acid (a)	mg
L-Lys·HCl (b)	914
L-His (b)	776
L-Arg·HCl	412
L-AspNH_2·H_2O	150
L-Thr	238
L-Ser (b)	210
L-Pro	230
L-Gly	150
L-Ala	89
L-Val (b)	938
L-Met	149
L-Leu (b)	821
L-Ileu	262
L-Phe (b)	677
L-Tyr	453
L-Try (b)	204
L-Cys/2·HCl	176

Dissolve in 900 ml of sterile H_2O, adjust to pH 7.4 with NaOH, and dilute to 1 liter. Store frozen.

(a) That amino acid which is to be added later as a radioactive isotope should be omitted in preparing the above mixture. For example, for our experiments no L-valine was added.

(b) Found limiting in a 4-hr incubation: Borsook, et al., Jour. Biol. Chem., 229, 1059, 1957.

TABLE XIX

Composition of the Vitamins Solution, Basal

Compound	mg/l
Choline	1.0
Folic Acid	1.0
Inositol	2.0
Nicotinamide	1.0
Pantothenate	1.0
Pyridoxal	1.0
Riboflavin	0.1
Thiamine	1.0

TABLE XX

Composition of Miscellaneous Reagents[a]

STM:	0.5 F sucrose, 0.04 F KCl, 0.003 F $MgCl_2$ in pH 7.0 Tris-HCl.[b]
TM:	Identical with STM with the sucrose omitted.
NKM:	0.153 F NaCl, 0.005 F KCl, 0.005 F $MgCl_2$; 9.00 gm/l NaCl, 0.37 gm/l KCl, 1.016 gm/l $MgCl_2 \cdot 6H_2O$.
BSA, 0.3%:	Dissolve 30 mg of bovine serum albumin in 10 ml H_2O.
TCA, 5% and 10%:	Dissolve 5 gm and 10 gm of trichloroacetic acid in 95 ml and 90 ml of H_2O respectively.
Mark's solution:	0.0015 F $MgCl_2$ in pH 7.2 0.001 F Tris-HCl.[b]

[a]In order of appearance in text.

[b]Tris = trishydroxymethylaminomethane.

TCA, 0.85% NaCl, 1.5 F NaCl, and "Mark's solution" were prepared and stored at 2°C. STM tended to develop molds and was checked periodically to see if new reagent needed to be made. One-half gm of L-glutamine and 0.20 gm of ferrous ammonium sulfate hexahydrate were weighed into dry and separate 10-ml glass-stoppered graduated mixing cylinders.

Sufficient Erlenmeyer flasks were sterilized and stored at room temperature.

The incubator was checked to see that the temperature control was working and that the mechanical shaker was in good operating condition.

The tank containing the 95% O_2-5% CO_2 gas mixture to be used in maintaining the appropriate atmosphere during the incubation was checked to make sure it contained sufficient gas.

The centrifuges were inspected to insure that they and their refrigerators were operating properly.

Immediately after the doctor had called, the reagent mixture, the human serum, and L-valine-C^{14} were set in beakers of water at room temperature to thaw. The refrigerators to the centrifuges were turned on and the centrifuge heads were placed in a refrigerator to cool. The incubator was set for 37°C and gentle, slow shaking. The container to be used to transport the blood was filled with ice. In those cases, such as cold agglutinin disease (cf. p. 176), where the blood was given special treatment, necessary solutions and equipment were taken to the doctor's office. We had arranged with the doctor beforehand that he would not draw the blood until the experimenter arrived at his office. The doctor was also advised as to the type of anticoagulant to be used.

Just before we left for the doctor's office the Erlenmeyer flasks which were to contain the incubation mixture were set on a bed of ice to chill.

VIII D3: Collection and Work-up of Red Blood Cells

The *in vitro* biosynthesis of hemoglobins A_{II} and A_{Ic} was studied with the cells from two individuals. Pertinent data from their medical histories are presented in the Tables which summarize the experimental results. One of these patients had a mechanical heart valve anemia, and the other had idiopathic cold agglutinin disease. The latter presented special experimental difficulties which are elaborated in the following paragraph.

The donor's blood was drawn into a heparinized syringe, was transferred to a chilled test tube, and was transported at 0°C to the laboratory. Anticoagulants other than heparin can be used provided they do not contain sugars: excess glucose has been found (114) to have a strongly inhibitory effect on protein synthesis. The blood from the patient with cold agglutinin disease was known to clot spontaneously below 37°C because of serum factors. Her blood was collected in a heparinized syringe which had been prewarmed to 37°C and centrifuged. The serum was discarded. Her cells were washed twice at 37°C with NKM, and were *then* transferred to a test tube at 0°C and transported to the laboratory.

When the blood sample had arrived in the laboratory, the serum, if still present, was centrifuged off at 0°C and 1,100 g for 5 min, and the cells were washed twice with 3 times their volume of NKM by centrifuga-

tion.* During the centrifugations, the preweighed ferrous ammonium sulfate hexahydrate was diluted to 10 ml with NKM, and the L-glutamine to 10 ml with 0.85% NaCl. Warming the solution of glutamine slightly (a steam line was convenient) facilitated the dissolving of the glutamine. The chilled Erlenmeyer flasks were filled with the appropriate amounts (see Table XVI) of the "reagent mixture," human serum, 5% glutamine solution, 2% ferrous ammonium sulfate solution, and vitamins solution. (When the total volume of the incubating mixture was 12 ml, a 125-ml Erlenmeyer gave the mixture a depth which was just deep enough to provide good mixing with gentle shaking by the incubator and which was thin enough to allow oxygen and carbon dioxide to diffuse into it.) The packed cells were added to the mixture with a smoothed-tip large bore graduated pipet which was made by cutting off and fire-polishing the end of a 5-ml pipet. The flasks were covered with Sanitab** rubber caps and were placed in the incubator. Two gas lines, an inlet and outlet, which terminated in hypodermic needles were inserted into each cap. After preincubating the flasks for 10 min to equilibrate the mixture with the 95% O_2-5% CO_2 atmosphere, the solution of radioactive valine was added. This was t = 0 min. The radioactive valine was conveniently added with a long tip, $\frac{1}{2}$-ml pipet which was calibrated to the tip's end. To eliminate the possibility of ingesting any radioactive material, we used a Clay-Adams suction pipet control** to draw up and discharge the radioactive L-valine solution. Flasks to be used in the chase experiments

---

*Until the incubation was begun all operations were carried out at 0°C.

**Los Angeles Chemical Co., South Gate, California.

were kept chilled until 10 min before the chase was to begin. At that time they were preincubated as described above.

After the incubation or chase had proceeded for the desired amount of time the reaction was stopped by adding one-half times the total incubating volume of NKM at 0°C and by placing the flask in ice water.

VIII D4: Isolation of Ribosomes and Hemoglobin

To isolate the radioactive ribosomes and soluble proteins, the suspension of cells from the incubation or chase was centrifuged at 0°C* and 1,100 g for 5 min and the supernatant was removed with a suction line.** The cells were washed twice with three times their volume of NKM and then were lysed by stirring in exactly 3 times the cell volume of Mark's solution for 30 sec, then adding 0.34 times the cell volume (before the addition of Mark's solution) of 1.5 F NaCl to restore isotonicity. The lysate was centrifuged at 1,100 g for 5 min and the pellet, consisting of unlysed cells and other heavy cellular fractions was discarded. The supernatant was centrifuged 10 min at 15,000 g and the pellet was again discarded. The supernatant from this operation contained the soluble proteins and the ribosomes. It was layered onto 3 ml of STM in a plastic centrifuge tube (empty space was filled with TM) and centrifuged for 3 hr at 105,000 g. The supernatant containing the soluble proteins was decanted and saved (it was worked up, as de-

---

*All operations are carried out at 0°C up to the point of determining the optical absorbancy.

**This line should be connected to a suction flask, not directly to the drain or pump: the usual safety precautions involved in the use of radioactive materials should be observed.

scribed below, concurrently with the isolation of the ribosomes). The pellet from the 3-hr centrifugation contained the ribosomes. This pellet was drained, rinsed gently (so as not to disrupt the pellet) 3 times with TM, and was then homogenized with 7 ml of TM in a glass tube which fitted snugly around the ground glass homogenizing pestle. The homogenate was layered onto 5 ml of STM in a centrifuge tube and was centrifuged for 5 hr at 105,000 g. The purified ribosomal pellet from this final centrifugation was homogenized in 4 ml of TM at 25°C and was diluted 100-400-fold to remove any turbidity. The optical absorbancy was determined at 25°C at 260, 280, and 415 mμ. TM was used as a blank.

The decanted supernatant from the first centrifugation through STM contained hemoglobins and other proteins. Carbon monoxide was bubbled gently through this solution to convert the oxyhemoglobin to the more stable carbonmonoxyhemoglobin. The solution was dialyzed against 3 changes of developer #4 at 2°C, was concentrated at 2°C through collodion bags,* and its optical absorbancy was determined at 260, 280, 415, and 520.5 mμ. In order to obtain purified fractions of hemoglobin A_{Ic} and A_{II}, 1 ml of this solution, which contained 40-100 mg** hemoglobin was chromatographed at 6.8°C on a 1 x 30-cm IRC-50 cation exchange column with developer #4 (1, 22, 76).

---

*(Cf. footnote, p. 32.)

**$\frac{\text{mg hemoglobin}}{\text{ml}} = 1.813 \times A_{520.5}$ for a mixture of carbonmonoxyhemoglobin and ferricyanohemoglobin (58).

VIII D5: Determining the Specific Activity of the Hemoglobin

The specific activity of the radioactive soluble proteins and of the radioactive purified hemoglobins was determined by diluting an aliquot containing 0.3 mg or less* of protein to 1 ml with water in a 1 x 10-cm test tube. One milliliter of 10% TCA was added and the contents were shaken on a vortex type mixer. The tube was heated at 90°C for 1 hr to precipitate the protein and to hydrolyze (if present) any RNA. The hot solution was shaken with a vortex mixer and was transferred quantitatively with 5% TCA onto a 25-mm, 0.45 μ pore size, Millipore** filter. Before being used, these filters were soaked in 5% TCA for at least 1 hr. The filters containing the radioactive protein were dried by allowing them to stand in air at room temperature on a clean glass plate for 20 min. When dry, they were glued onto counting planchets with rubber cement and were dried 45 min under a heat lamp. They were counted for radioactivity on a Nuclear Chicago low-background gas-flow counter. No correction was made for counter efficiency.

VIII D6: Specific Activity of Ribosomes

The specific activity of the ribosomal homogenate was determined in an identical manner except that 1 ml of homogenate with an optical absorbancy of 1.00 at 250 mμ (after correction for any contaminating hemoglobin) was taken as an aliquot, and the precipitation with 10% TCA was carried out at 10°C instead of 90°C. It was necessary to correct the specific activity for any radioactivity from contaminating hemoglobin.

---

*If less, sufficient 0.3% BSA solution was added to make up the deficit.

**Millipore Filter Corp., Bedford, Massachusetts.

VIII E: RESULTS

VIII E1: Ribosomal Specific Activity

The ribosomal pellets which were isolated as described in the preceding section were a bright red in color. This indicated heavy contamination with hemoglobin (and hence possibly with other proteins). Depending on the pellet which was examined, between 2-95% of the optical absorbancy at 260 mμ was due to hemoglobin. This contamination could mean that heme was added to the polypeptide chain while it was still on the ribosome (40, 143), possibly contributing to its release therefrom, or simply that the conditions were not optimal for isolating human ribosomes from reticulocytes. Rabbit ribosomes which have been isolated as described are largely free from any contaminating hemoglobin (114). For the total of 12 samples which were examined, the ratio (after correction for contaminating hemoglobin) of the optical absorbancy at 280 mμ to that at 260 mμ was 0.80 ± .08 for the ribosomes. In view of the large variability of the amount of contaminating hemoglobin the constancy of this ratio is remarkable and suggests that contamination by other proteins is small.

The ribosomal specific activities (after correction for contaminating hemoglobin) from one set of incubations are plotted as a function of time in Fig. 17, p. 187. The gradual decrease of ribosomal activity with time, as shown in Fig. 17, was observed for the ribosomes from both the donor with a mechanical heart valve anemia and the donor with idiopathic cold agglutinin disease. Such a decrease has been observed by others (2, 39, 50, 142). Experiments by Philipps (114) suggest that

the decrease is probably due to nonoptimal conditions which may exist during the incubation.

VIII E2: Specific Activities of Hemoglobins A_{Ic} and A_{II}

The isolation of radioactive hemoglobin A_{Ic} and A_{II} is shown in the chromatogram of Fig. 16, p. 183. The purity of hemoglobin A_{Ic} and A_{II} is attested to by the constancy of the ratio of optical absorbancies at 280 mμ and 415 mμ and by the similar constancy of the specific activity across a chromatographic zone.

The specific activities of the purified hemoglobins from the two incubations are summarized in Tables XXI and XXII, pp. 184-185. The three values of specific activity which are given for each time represent, from top to bottom, samples taken from the front, middle, and tail of the chromatographic peak. Three facts should be noted in these Tables: first, the fact that both hemoglobin A_{Ic} and A_{II} are radioactive; second, the fact that the ratio of the specific activities of these two hemoglobins is approximately unity; and third, the fact that during the chase experiments there was no appreciable change in the specific activity of either A_{Ic} or A_{II}.*

---

*The 30-min chase of the 15-min incubation of the reticulocytes from the patient with mechanical heart valve anemia appears to contradict this statement. However, the observed high values of the specific activities of A_{Ic} and A_{II} for this chase cannot be correct because total counts are not conserved. The most probable explanation for the observed high values is a decrease in the optical absorbancy for these samples. The cause of this decrease is not known, but it would result in assigning a false low value to the concentration of the hemoglobin solutions, and this in turn would give false high values, as observed, of the specific activity.

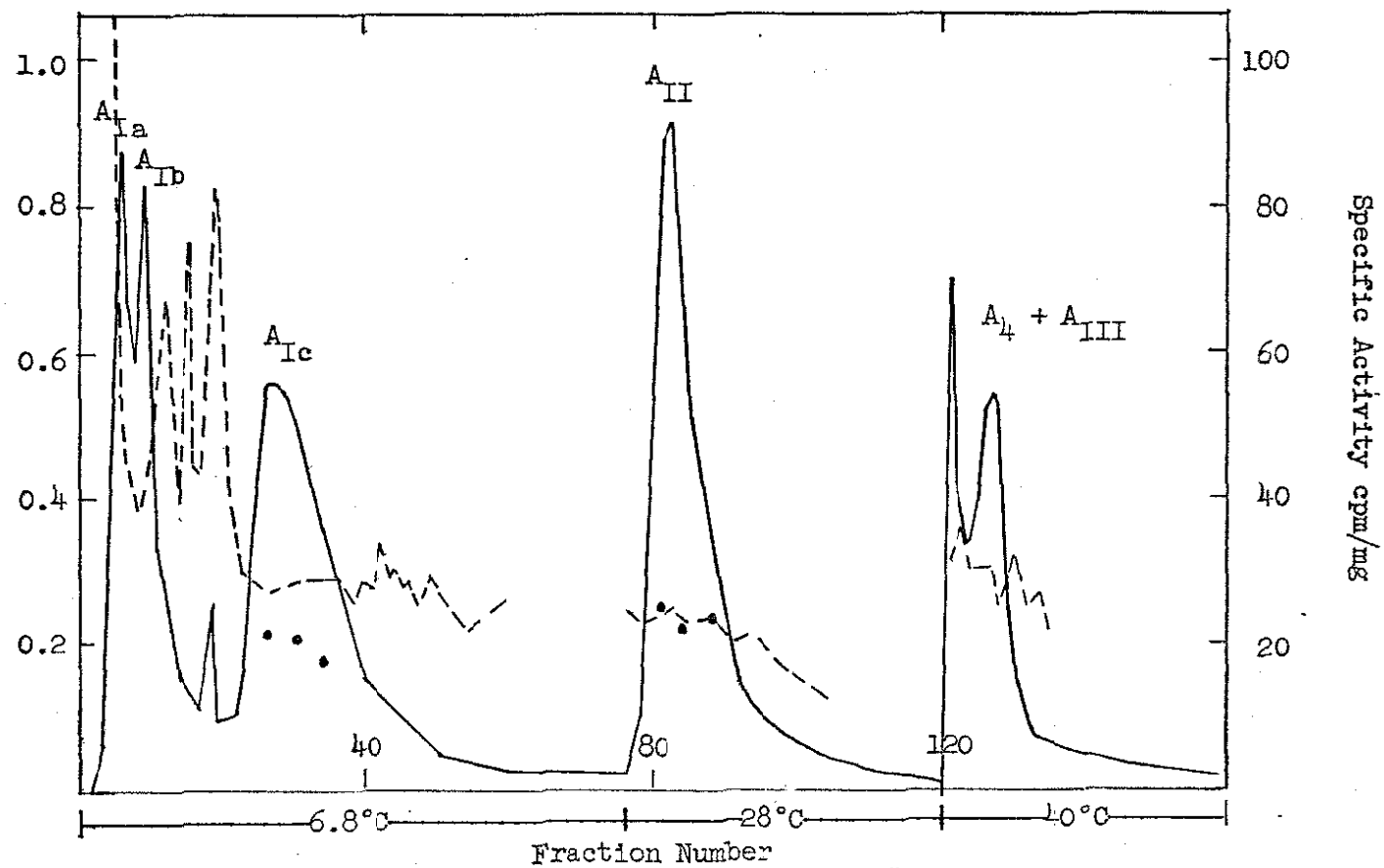


Fig. 16: Isolation of A_{Ic} and A_{II} by chromatography with developer #4 on a 1 x 30-cm Bio-Rex 70 column. For fractions 1-20 and 21-160 2-ml and 4-ml fractions, respectively were collected. The flow rate was 6 ml/hr. Right hand ordinate: ——— Absorbancy 415 mμ; – – – $A_{280\ m\mu}/A_{415\ m\mu}$. Left hand ordinate: • • • Specific Activity. The absorbancy of the A_{II} peak (only) has been reduced by a factor of 21.7.

TABLE XXI

Specific Activities[a] of Hemoglobin Biosynthesized in Vitro[b,c]

Time of Incubation	A_{Ic}	A_{II}	A_{Ic}/A_{II}
45 min	23.32	22.54	
	18.01	13.33	
	20.03	10.88	
Average:	20.45 ± 1.80	15.58 ± 4.15	1.32 ± .37
90	21.94	25.64	
	21.05	22.66	
	17.42	23.36	
Average:	20.14 ± 2.28	23.89 ± 1.05	0.84 ± .10
180	59.70	42.83	
	45.37	52.37	
Average:	52.53 ± 6.84	47.60 ± 1.10	1.10 ± .18
Chase of 90-min Incubation			
Duration of Chase			
45	22.13	26.91	
	19.68	18.12	
	22.78	23.60	
Average:	21.53 ± 1.10	22.88 ± 3.00	0.94 ± .13
Average value of A_{Ic}/A_{II} for all incubations and chases:			1.05 ± .14

a) In cpm/mg Hb; not corrected for counter efficiency.

b) The blood donor was a 70 year old white female who had cold aglutinin disease. Her hematocrit and reticulocyte count were 19% and 25% respectively.

c) Filtered human serum (Hyland Laboratories, Los Angeles, Calif.), rather than unfiltered, was used in this incubation. The vitamins solution (cf. p. 171) was also omitted.

TABLE XXII

Specific Activities[a] of Hemoglobins Biosynthesized in Vitro[b]

Time of Incubation	A_{Ic}	A_{II}	A_{Ic}/A_{II}
15.5 min	44.85 44.67 70.70	70.50 59.02 63.75	
Average:	53.40 ± 10.11	64.42 ± 3.89	0.83 ± .16
30.0	330.00 204.50 284.42	352.36 384.90 420.16	
Average:	272.97 ± 42.85	385.81 ± 22.89	0.71 ± .12
60.0	207.88 184.03 273.37	514.91 504.21 549.98	
Average:	221.76 ± 32	523.00 ± 16.16	0.42 ± .06
Chase of 15-min Incubation			
Duration of Chase			
15.0	31.94 39.95 42.32	68.38 68.28 76.05	
Average:	38.07 ± 3.67	70.90 ± 3.01	0.54 ± .06
30.0	68.75 102.46 182.20	96.58 117.63 122.47	
Average:	117.80 ± 39.33	112.23 ± 9.28	1.05 ± .36
Average value of A_{Ic}/A_{II} for all incubations and chases:			0.71 ± .17

a) In cpm/mg Hb; not corrected for counter efficiency.

b) The blood donor was a 27 year old Negro female. She had a mechanical Starr-Edwards heart valve which had been inserted as treatment for an incompetent aortic valve. Subsequent to the insertion of the Starr-Edwards valve she developed a brisk hemolytic anemia. Her hematocrit and reticulocyte count were each 25%.

The data from Table XXI are plotted in Fig. 17,* p. 187. For reasons which are described on p. 190, the points have been fitted by a least squares line which has been constrained to go through the origin. The deviations of some points from this line are larger than can be explained by the errors which are given in Table XXI. We believe this to be due to the manner in which the experiment was conducted: The incubations for different times were done in separate flasks because of the small amount of blood which was available. If for some reason an incubation in a single flask failed (6 out of 8 incubations were successful), the other flasks would remain unaffected. The occasional failure reflected in part the sensitive nature of the in vitro system to handling. The deviations from linearity would probably be less for a time study which was performed by taking aliquots from a single flask. In any case, the qualitative conclusions which will be drawn in the next section are not dependent on the detailed fit of the experimental points to the least squares line.

VIII F: CONCLUSIONS

The fact that both A_{Ic} and A_{II} become radioactive during the incubation shows that A_{Ic} is not a product which results from the degradation of hemoglobin A_{II} and which is found only in older cells.

---

*The data in Tables XXI and XXII are primary data without corrections of any sort. The plot in Fig. 17 has been corrected (cf. pp. 164 and 166) for the decrease in ribosomal activity during the incubation. The latter correction was made so that Fig. 17 could be compared directly with the theoretical plots of Figs. 14 and 15, pp. 166a-167.

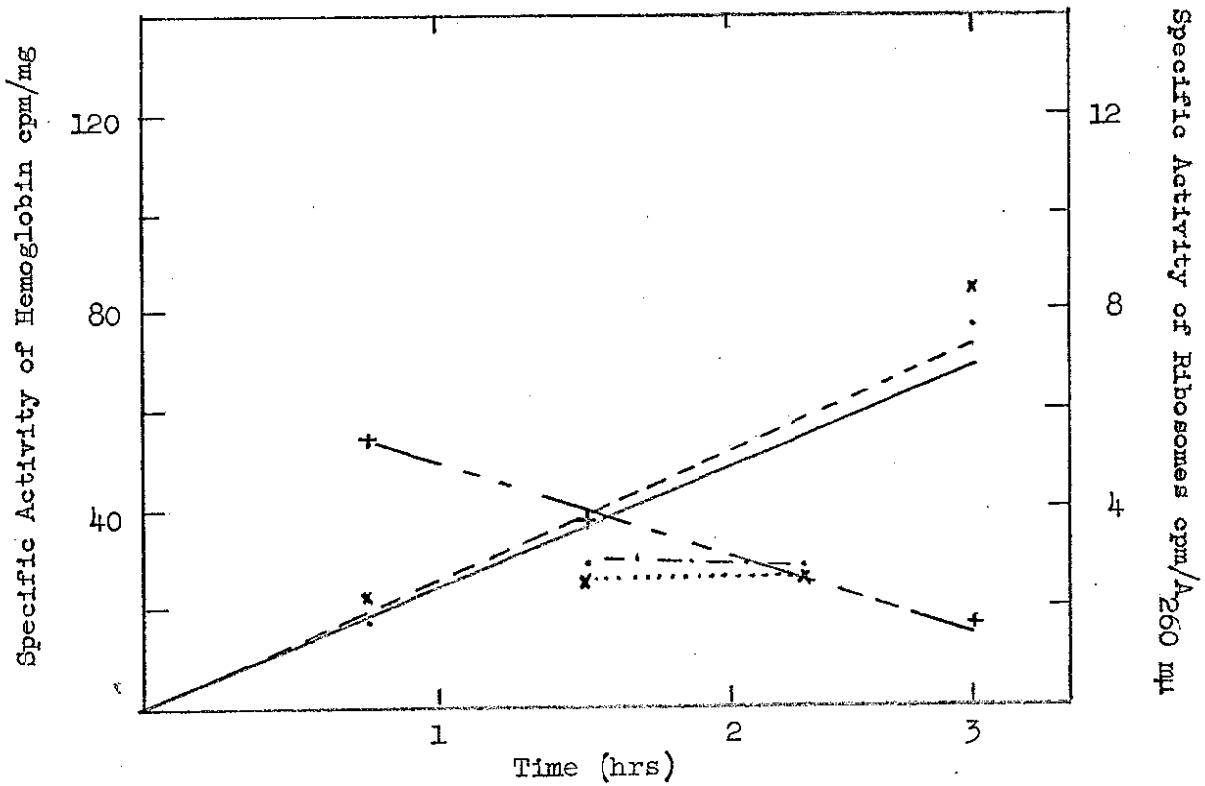


Fig. 17: Experimentally observed behavior for incubation and chase experiments. Incubation experiment: A_{Ic} x – – x; A_{II} ———. Chase experiment: A_{Ic} x.....x; A_{II} –.–.–. Ribosomes during the incubation: + – – – – – – +. See text for further explanation.

Rather, A_{Ic} appears to be a normal biosynthetic product of young red cells.

The fact that the synthesis of A_{Ic} is concurrent with that of A_{II} implies that either the ketone or aldehyde R=O* is hooked onto the β chain while the latter is still on the ribosome, or that R is attached to the β chain relatively soon after its release from the ribosome. In either case the de novo synthesis of hemoglobin A_{Ic} and A_{II} can be said to occur independently of each other.

The facts that within the experimental error the specific activities of A_{Ic} and A_{II} are equal at any given time,** and that during the chase experiments no significant change in the specific activities of A_{Ic} or A_{II} occurs imply that

1) A_{Ic} and A_{II} are not rapidly interconvertible over the time period which was studied, or
2) though interconvertible at moderate rates, any net gain or loss is just offset (by "numerical accident") by the rate at which the component is synthesized on the ribosome, or

---

*Chemical characterization has established that A_{Ic} differs from A_{II} only in that the N-terminal amino group of the former is blocked in a Schiff base type bond by a group R whose molecular weight is about 265 (cf. pp. 135-155).

**This statement is true for seven out of the total of nine incubation and chase experiments which were run. The two exceptions are the 60-min incubation and 15-min chase in Table XXII, p. 185. In view of the experimental considerations which have already been discussed (p. 186), we do not believe these two exceptions invalidate the general conclusion.

3) that A_{Ic} and A_{II} are in rapid equilibrium.

Conclusions 1) and 2) imply that the rates at which A_{Ic} and A_{II} are synthesized by the ribosomes are proportional to the average cellular concentration of A_{Ic} and A_{II} respectively. It should be noted that examples are known where the rate of hemoglobin synthesis both is (52), and is not (18) proportional to the cellular concentration. Conclusion 3) implies only that the total rate (i.e., $A_{Ic} + A_{II}$) of hemoglobin synthesis is proportional to the total ($A_{Ic} + A_{II}$) average cellular concentration. A priori there is no reason to throw out the "numerical accident" hypothesis which is implied by conclusion 2) simply because it seems improbable. Many improbable things occur in cells, and there may exist good reason for such a numerical coincidence to occur. With respect to conclusions 1) and 3), the "dynamic equilibrium" hypothesis of Guidotti et al. (cf. p. 92, and Ref. 48) together with the fact that A_{Ic} and A_{II} are isolable by column chromatography would favor the non-interconvertibility of A_{Ic} and A_{II} rather than a rapid equilibrium between them. However, it should be recognized that in the cell, conditions might favor a rapid equilibrium: there might be an enzyme "A_{Ic}-A_{II} equilibrase," for example.

VIII G: DISCUSSION

What follows is a discussion of the theoretical behavior to be expected of the specific activities when the only sources of radioactive A_{Ic} and A_{II} are synthesis by the ribosomes at a constant rate and interconversion of A_{Ic} to A_{II} or vice versa by first order kinetics.

We state as a fact (the detailed calculations were given earlier in the theoretical section) that under the above two assumptions the time independent equality of the specific activities of A_{Ic} and A_{II} rigorously implies all the experimental conclusions reached above and only these. The equality of specific activities further implies that the specific activity should increase linearly with time. This is the reason for fitting the data to a least squares straight line in Fig. 17, p. 187. This predicted linearity is at least approximately borne out by the experimental points. Both theory and experiment lend support to each other.

If in fact, A_{Ic} and A_{II} are not rapidly interconvertible (cf. pp. 188-189), the theory can tell us more. Figures 14 and 15 (pp. 166a-167) show the theoretical behavior to be expected of A_{Ic} and A_{II} when there is a very slight net conversion of A_{II} to A_{Ic}, and a moderate rate of conversion respectively. By comparing these curves with Fig. 17 we can set upper limits on k_- and k, viz:

$$k_- \leq 0.001 \ \text{min}^{-1}$$

$$k \leq 0.0001 \ \text{min}^{-1}$$

for the *in vitro* systems which were studied.

Finally, to end on a practical note, it appears that the use of unfiltered human serum in the incubation mixture and the inclusion of the vitamins solution increases the rate of hemoglobin synthesis. This

can be seen by comparing the specific activities in Tables XXI and XXII, pp. 184 and 185.*

---

*It is possible that the increased specific activities in Table XXII are due to a greater biological activity of the reticulocytes of the donor with a mechanical heart valve anemia relative to the activity of the reticulocytes of the donor with cold agglutinin disease. At any rate, the use of unfiltered serum and the vitamins solution does not adversely affect the system, and until it has been demonstrated to the contrary, their addition to the reagent mixture seems desirable.

Part IV

IX: SUMMARY OF CURRENT KNOWLEDGE ABOUT HEMOGLOBIN A_{Ic}

IX A: Chemical
IX B: Biological

IX A: CHEMICAL SUMMARY

The following summary has been limited entirely to listing those facts which delineate the actual chemical structure of hemoglobin A_{Ic}. No attempt has been made to outline the experiments on which these facts are based. Similarly, little or no mention has been made of the interaction of hemoglobin A_{Ic} with other systems: for example, the chromatographic behavior of A_{Ic} on various types of chromatographic columns is not mentioned at all. This is not to say that these aspects of the chemistry of hemoglobin A_{Ic} are not important. On the contrary, from a practical viewpoint they are the most important parts of the thesis. But after all has been said and done, the really interesting question is "What is hemoglobin A_{Ic}?", and it is to answering that question that this summary has been directed.

IX A1: Hemoglobin A_{Ic}

The structure of hemoglobin A_{Ic} is $\alpha_2\ \beta\ \beta^{A_{Ic}}$. To the extent that they have been investigated, no differences have been found between the primary amino acid sequences of the α and β polypeptide chains of hemoglobin A_{Ic} and the corresponding chains of hemoglobin A_{II}. The latter hemoglobin is the major component in normal adults and comprises 80-90% of the total hemoglobin. It is highly probable that the corresponding chains between these two hemoglobins are identical. The extent to which this identity has been experimentally verified is summarized in Figs. 18 and 19, pp. 194 and 195. The $\beta^{A_{Ic}}$ polypeptide chain differs from the β polypeptide chain only in that the N-terminus of the former is blocked by a group which we designate as R.

Fig. 18: Primary Amino Acid Sequence of α Chains from Hemoglobin A_{Ic}

The symbols $\underline{|\alpha T\text{-}1\rightarrow}$ etc., represent the tryptic peptides as designated by Baglioni (cf. footnote p. 19). The arrow points in the direction from the N-terminus towards the C-terminus of the polypeptide chain. The order of the tryptic peptides has been <u>assumed</u> to be the same for A_{Ic} as the known order for A_{II}. The listed order of the amino acids in this Figure is the same as the known order (15, 16, 61, 80, 81, 126, 131) in Hemoglobin A_{II}. The number above each amino acid in each tryptic peptide is the <u>experimentally</u> determined number of residues of that amino acid in A_{Ic}. These numbers have been taken directly from Table VII, p. 101a. Within a given peptide, the order of those sequences joined by hyphens is certain; the order of sequences separated by commas is not known. The small number below every 10th amino acid represents the position of the amino acid in the polypeptide chain of A_{II} and is given for convenient reference.

αT-1 αT-2 αT-3

1.12 0.83 0.73 1.01 1.01 1.06 0.97 0.92 1.00 1.00 1.00 0.97 0.97 0.14 1.06
(val, leu, ser, pro, ala, asp) - lys (thr, asg, val) - lys (ala, ala, try, gly) -
10

αT-4

1.00 1.16 1.03 1.08 0.98 1.08 1.03 0.99 0.93 1.03 1.08 0.99 1.08 1.19 0.99 0.98
lys (val, gly, ala, his, ala, gly, glu, tyr, gly, ala, glu, ala, leu, glu) - arg
20 30

αT-5 αT-6

0.46 0.98 1.20 0.97 0.98 1.02 0.94 0.94 1.06 0.88 1.01 0.92 1.22 0.89 0.92 1.01
(met, phe, leu, ser, phe, pro, thr, thr) - lys (thr, tyr, phe, pro, his, phe, asp,
40

αT-7 αT-8 αT-9

1.01 0.95 0.89 1.18 0.95 1.06 1.07 1.09 0.95 0.97 0.93 0.97 1.10 1.00 0.98
leu, ser, his, gly, ser, ala, glm, val) - lys (gly, his, gly) - lys lys (val,
50 60

0.99 0.97 0.99 0.99 0.99 0.97 0.99 0.98 0.99 0.82 0.98 0.97 0.97 0.39 1.04 0.97 0.99
ala, asp, ala, leu, thr, asg, ala, val, ala, his, val, asp, asp, met, pro, asg, ala,
70

αT-10 αT-11

0.99 0.96 0.99 0.99 0.96 0.97 0.99 0.82 0.99 0.82 1.07 1.00 1.00 1.01 1.02 0.94
leu, ser, ala, leu, ser, asp, leu, his, ala, his) - lys leu - arg (val, asp, pro,
80 90

αT-12 + αT-13

1.01 1.02 0.96 1.01 1 1 1 1 1 1 1 1 1 1 1 1
val, asg, phe) - lys (leu , leu, ser, his, cys, leu, leu, val, thr, leu, ala, ala,
100 110

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
his, leu, pro, ala, glu, phe, thr, pro, ala, val, his, ala, ser, leu, asp, lys,
120

αT-14

1 1 1 1 1 1 1 1 1 1 1 1 0.94 1.04
phe, leu, ala, ser, val, ser, thr, val, leu, thr, ser) - lys tyr - arg
130 140

Fig. 19: Primary Amino Acid Sequence of the β Chain from Hemoglobin A_{Ic}

The symbols $\lfloor\beta T\text{-}1\rangle$ etc., represent the tryptic peptides as designated by Baglioni (cf. footnote p. 19). The arrow points in the direction from the N-terminus towards the C-terminus of the polypeptide chain. The order of the tryptic peptides has been <u>assumed</u> to be the same for A_{Ic} as the known order for A_{II}. The listed order of the amino acids in this Figure is the same as the known order (15, 16, 61, 80, 81, 126, 131) in Hemoglobin A_{II}. The number above each amino acid in each tryptic peptide is the <u>experimentally</u> determined number of residues of that amino acid in A_{Ic}. These numbers have been taken directly from Table VIII, p. 102a. Within a given peptide, the order of those sequences joined by hyphens is certain; the order of sequences separated by commas is not known. The small number below every 10th amino acid represents the position of the amino acid in the polypeptide chain of A_{II} and is given for convenient reference.

The $\beta^{A_{Ic}}$ chain differs from the β chain only in the presence of the blocking group R at the N-terminus (residue 1) of the former.

βT-1 βT-2

0.91 0.93 1.02 0.96 1.00 1.03 1.03 1.04 0.80 1.02 1.01 0.88 1.02 1.00 0.56 1.06
val - his - leu - thr - (pro, glu) - glu - lys ser - ala - val - thr - ala - leu - (try, gly) -
10

βT-3 βT-4

0.94 0.96 1.02 0.96 1.02 1.04 0.96 1.00 1.00 1.04 1.02 1.04 1.00 1.00 0.98 0.98
lys val - asg - val - (asp, glu, val, gly, gly, glu, ala, leu, gly) - arg leu - leu -
20 30

βT-5

0.79 0.79 0.90 1.04 0.19 0.96 1.04 0.98 0.87 0.87 1.01 0.80 0.87 1.03 0.92 1.02
val - val - tyr - pro - (try, thr, glm) - arg phe - phe - glu - (ser) phe - gly - (asp, leu,
40

βT-6 βT-7

0.80 0.93 0.91 0.92 0.97 1.05 0.47 1.03 0.92 0.91 0.94 0.97 1.02 0.97 0.83 1.15 1.06
ser, thr, pro, asp, ala, val, met, gly, asg, pro) - lys val - lys (ala, his, gly) - lys
50 60

βT-8 βT-9

1.00 1.01 1.12 1.00 1.00 1.00 0.89 0.98 1.00 1.12 1.00 1.02 1.12 0.98 0.98 1.12 1.04
lys val - leu - gly - ala - phe - ser - (asp, gly, leu, ala, his, leu, asp, asg, leu) - lys
70 80

βT-10,11

1.35 0.88 0.92 1.33 0.88 1.05 1.14 0.95 1.05 0.86 0.83 0.97 1.13 1.05 0.86 1.25 0.97
(gly, thr, phe, ala, thr, leu, ser, glu, leu, his, cys, asp, lys, leu, his, val, asp,
90

βT-12

1.16 0.95 0.97 0.92 0.84 0.84 0.84 1.02 1.05 0.89 0.84 0.89 1.06 0.89 0.84 1.23
pro, glu, asg, phe, arg) (leu, leu, gly, asg, val, leu, val, cys, val, leu, ala,
100 110

βT-13

0.83 0.83 1.06 1.02 1.03 1.01 1.00 0.92 1.02 1.02 1.04 1.01 0.97 0.97 0.89 1.01 1.02
his, his, phe, gly) - lys glu - (phe, thr, pro, pro, val, glu, ala, ala, tyr, glm) - lys
120 130

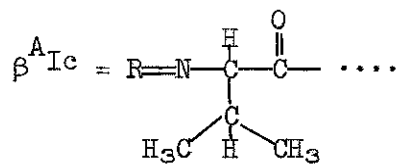
βT-14 βT-15

1.22 1.22 0.97 1.01 1.22 0.97 0.98 0.97 1.00 0.97 1.01 1.03 1.02 0.98
val - (val) ala - gly - val - ala - (asg, ala, leu, ala, his) - lys tyr - his
140

The group R is not a nonketo acyl group of less than five carbon atoms, is not a carboxyl group, and is not a carbamyl group. Rather, R is derived from a ketone or aldehyde with the structure R=O. Hemoglobin A_{Ic} is the Schiff base that is formed between R=O and hemoglobin A_{II} when the former condenses with the N-terminal amino group of one (and only one) of the two β chains of the latter with the elimination of water:

$$R = O + \alpha_2\ \beta_2 \longrightarrow \alpha_2\ \beta\ \beta^{A_{Ic}} + H_2O\ ,$$

where the structure of $\beta^{A_{Ic}}$ in the vicinity of the N-terminus is



At approximately 10°C the pK_a for the Schiff base nitrogen atom is 6.6 ± 0.2. The double bond of the Schiff base is readily reduced to the secondary amine with sodium borohydride.

Between pH 2 and pH 11, the double bond of the Schiff base linkage in hemoglobin A_{Ic} is stable towards aqueous hydrolysis at 2°C for periods of at least 24 hr. At pH 6.8 and 10°C hemoglobin A_{Ic} is stable toward hydrolysis for at least a month. A similar stability is found at pH 8 and 25°C for periods of at least 24 hr. Likewise β $\beta^{A_{Ic}}$-globin is stable for periods of at least 24 hr at 25°C and for periods of at least a week at pH 1.8 and 25°C. The time periods of stability that are given above were taken from actual experiments and are minimum values. The actual periods of stability may be much longer.

The stability that is found for the imine bond of the Schiff base linkage in hemoglobin A_{Ic} is atypical of Schiff bases in general. More typically the half-life time for the hydrolysis of this bond is of the order of a few seconds to a few hours. The anomalous stability of A_{Ic} is probably due to interaction of the N-terminal region of the $\beta^{A_{Ic}}$ chain with other parts of the molecule. This interaction apparently protects the imine bond of the Schiff base from hydrolytic attack. This interpretation is supported by the fact that the Schiff base imine bond in low molecular weight degradation products of hemoglobin A_{Ic} is readily hydrolyzed. For example, this bond can be hydrolyzed simply by dissolving the N-terminal octapeptide R=val-his-leu-thr-pro-glu-glu-lys or the N-terminal dipeptide R=val-his in water at room temperature and letting the solution stand a few minutes. The hydrolysis is relatively rapid at neutral and alkaline pH and relatively slow below pH 5. Higher temperatures favor the hydrolysis. Presumably in these smaller compounds there is nothing with which the Schiff base region of the molecule can interact to protect itself from hydrolytic attack.

IX A2: The Chemical Structure of R=O

The following facts about R=O are known with certainty:

1) The molecular weight of R=O is 281 $\pm$ 27 (P.E.)
2) R=O is not an aromatic aldehyde.
3) R=O does not contain phosphorus.
4) R=O does not contain carbohydrate.
5) R=O is not and does not contain a steroid.
6) R=O does not contain any amino acids.

The probability that the following statements are true is high; however, their truth is not certain:

7) R=O does not contain any nitrogen atoms the basicity of which corresponds to $\underline{p}K_a > 2.8$.

8) R=O does not contain any acidic groups (e.g., carboxylic, or phenolic) the acidity of which corresponds to a $\underline{p}K_a < 7.1$.

Statements 7) and 8) make it likely that R=O contains only carbon, oxygen, and hydrogen.

IX B: BIOLOGICAL SUMMARY

Hemoglobin A_{Ic} is a naturally occurring constituent of the blood of normal adult humans. It comprises 5.3% of the total hemoglobin in the body, though this percentage can vary between 4-7% for individuals with certain pathological disorders.

During <u>in vitro</u> biosynthesis of hemoglobin, hemoglobin A_{Ic} is synthesized concurrently with the major hemoglobin component A_{II}. The latter hemoglobin constitutes 80-90% of the total hemoglobin in normal adults. The concurrent synthesis of hemoglobins A_{Ic} and A_{II} demonstrates that hemoglobin A_{Ic} is not a biological product that is associated only with older red cells. On the contrary, it is a normal constituent of all red cells, irrespective of their age.

During the <u>in vitro</u> biosynthesis of hemoglobin A_{Ic}, R=O (cf. Section IX A, p. 193) is condensed with the N-terminal amino group of the β chain either while the latter is still on the polysome, or relatively soon after its release therefrom.

Finally, detailed analysis of the course of the *in vitro* biosynthesis of hemoglobin A_{Ic} and A_{II} demonstrated that either these two hemoglobins are noninterconvertible or else they are in a rapid chemical equilibrium.

Part V

APPENDICES

APPENDIX I

SYNTHESIS OF L-VALYL-L-HISTIDINE AND N-ACETYL-L-VALYL-L-HISTIDINE

A I: SYNTHESIS OF L-VALYL-L-HISTIDINE

A IA: Introduction

The method of synthesis followed Schwyzer *et al.* (133). In this method 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide is used to effect the formation of the amide bond at room temperature. The above reagent, though more difficult to prepare, has an advantage over the more commonly used dicyclohexylcarbodiimide in that the 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]urea which results from the former during the course of the synthesis is soluble in the polar organic solvents in which the formation of the peptide is effected. The dipeptide, which is insoluble in these solvents, can therefore be purified by simply washing it with the polar organic solvent. On the other hand dicyclohexylurea, which is formed when dicyclohexylcarbodiimide is used to effect the formation of the amide bond, is like the dipeptide, insoluble in polar organic solvents, and consequently purification of the dipeptide is more difficult.

A IB: Synthesis of 1-Cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide

Preparation of N-Cyclohexyldithiocarbamic Acid (43)

A mixture of 34.7 ml of cyclohexylamine in 500 ml of anhydrous diethyl ether was placed in a 1-liter round-bottom flask, and 9.1 ml of CS_2 was slowly added (considerable heat was evolved). After the CS_2 had been added, a white crystalline solid began to precipitate. The mixture was allowed to stand for 30 min with occasional swirling, at which time solvent was removed by evaporation *in vacuo*. The product

was a white solid which was left in the 1-liter flask for the next step. The yield was 100%.

Preparation of Cyclohexylisothiocyanate (43)

To the flask from the preceding step were added 40.5 gm of $HgCl_2$ which had been dissolved in 600 ml of water. The mixture was refluxed for 1 hr and was then distilled at atmospheric pressure. The residue in the flask became black during the distillation. The product appeared as a slightly dirty oil which collected in globules in the aqueous distillate. This distillate was extracted five times with 100 ml of diethyl ether. The ether was removed from the extract by evaporation at room temperature with a stream of air. The wet oily residue was transferred to a 25-ml glass-stoppered cylinder and was dried by shaking the wet oil with anhydrous Na_2SO_4. The dry oil was transferred to a centrifuge tube, and the last traces of sodium sulfate were removed by low speed centrifugation. The product, which had a penetrating odor and was slightly lachrymatory, was obtained in 50% yield.

Preparation of 1-Cyclohexyl-3-[2-morpholinyl-(4)-ethyl]thiourea (43)

A mixture of 500 ml of anhydrous diethyl ether, 15.9 gm of cyclohexylisothiocyanate, and 14.8 gm of N-(2-aminoethyl)morpholine* was refluxed in a 1-liter round-bottom flask until (about 10 min) the white crystalline product began to separate out. After the flask had cooled, the ether was removed <u>in vacuo</u>, and those crystals which had not adhered

---

*Aldrich Chemical Co., Inc.; 2369 No. 29th St., Milwaukee, Wisconsin.

to the flask wall were collected. The yield was 95%.

Preparation of 1-Cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide (43)

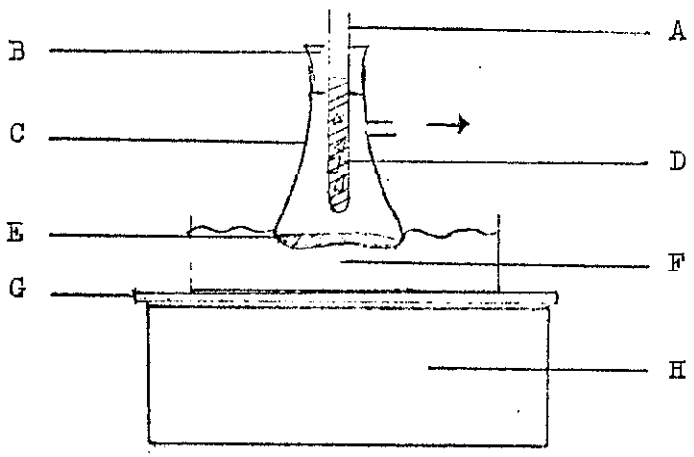
A mixture of 4 gm of 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]thiourea, 6 gm of yellow HgO, and 50 ml of acetone was refluxed for 6 hr in a 200-ml round-bottom flask, at which time the mixture was filtered through a sintered crucible* into a second 200-ml flask containing 6 gm of fresh HgO. The contents of the second flask were refluxed another 6 hr, filtered through a sintered crucible, and the filtrate was evaporated *in vacuo*. The oily residue from the filtrate was washed twice by evaporating it with benzene to remove traces of acetone. The product appeared as a colorless to pale yellow oil of a not unpleasant odor.** The crude oil was purified by distilling it at 145°C and at less than 200 μ Hg pressure*** onto a cold finger. The apparatus shown in Fig. 20, p. 205, was used for this evaporation. The layer of crude product was not more than 5 mm in depth, and the distance between this

---

*After use, the crucible was most easily cleaned with hot concentrated nitric acid.

**Sometimes white crystals, which we did not characterize, formed during the evaporation. These were filtered off, washed several times with benzene, and the washings were added to the original oil. The benzene was again evaporated off. This procedure was repeated until the residue from the benzene evaporation was free of crystals. It is likely that the crystals were 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl] urea formed from the product and small traces of water in the benzene. Drying the benzene over sodium prior to using it would probably obviate this minor difficulty.

***Any good forepump will suffice. A water aspirator should not be used because the product will react with water vapor to form the urea mentioned in the preceding footnote.



A. 18 x 150-mm Pyrex test tube
B. Rubber stopper
C. 250-ml Erlenmeyer heavy wall filter flask
D. Dry ice-methyl cellosolve cooling mixture
E. Crude 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide
F. Oil bath set at 145°C
G. Asbestos sheet
H. Combination heater and magnetic stirrer

Fig. 20: Apparatus for evaporative distillation of crude 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide.

layer and the bottom of the cold finger was approximately 2.5 cm during the distillation. Before the heater was turned on, or the cold finger was put into place, the crude product was outgassed at room temperature until it had ceased to bump. As the distillation proceeded the purified product appeared as a frozen mass on the coldfinger. When the distillation was complete, the heating bath was removed, and the apparatus was allowed to cool in vacuo to room temperature. The pump was turned off, the coldfinger was removed from the flask, and the cooling mixture was poured out. During the latter step care was taken not to contaminate the product with stray droplets of the cooling mixture. The coldfinger was placed in an upright position over a collection tube, and the purified product was thawed by pouring water at room temperature into the coldfinger. The final product was a perfectly clear colorless oil that was obtained in 50% yield. The net yield, based on the starting amount of cyclohexylamine, was about 25%.

A IC: Synthesis of L-Valyl-L-histidine

Preparation of L-Histidine Methyl Ester (9)*

A solution of 2% NH_3 in $CHCl_3$ was prepared by adding 2 ml of liquid NH_3 to 100 ml of $CHCl_3$. Approximately 10 ml of this solution was added to 0.50 gm of L-histidine methyl ester dihydrochloride** in a 10-ml glass

*The procedure which is given was by far the simplest of several procedures (44) which were tried. The slight extra effort which was required to transport the liquid NH_3 was repaid by 100% yield of the free ester.

**Nutritional Biochemicals Corp., Cleveland, Ohio.

stoppered cylinder. The cylinder was inverted several times until the precipitate of NH_4Cl appeared to be distributed uniformly throughout the solution. The solution was then allowed to stand for 30 min with occasional shaking; it was then filtered through a sintered glass disc into a weighed (± 0.01 gm) 100-ml round-bottom flask. The NH_4Cl which remained in the cylinder was washed several times with 2% NH_3 in $CHCl_3$, and the washings were filtered into the 100-ml flask. The contents of the flask were evaporated in vacuo to exactly 0.35 gm. The free ester was a viscous syrup of a very faint fruity odor.

Preparation of Carbobenzoxy-L-valyl-L-histidine Methyl Ester (133)

The syrup from the preceding step was transferred to a 15-ml graduated centrifuge tube with 0.2-ml portions of acetonitrile and was diluted with that solvent to a total volume of 2.1 ml. Exactly 0.53 gm of carbobenzoxy-L-valine* was stirred into the acetonitrile solution until all the valine had dissolved. The solution was cooled to 10°C, and exactly 0.55 gm of 1-cyclohexyl-3-[2-morpholinyl-(4)-ethyl]carbodiimide was added. The mixture was allowed to stand at room temperature for 16 hr.** Within 1 hr the mixture had solidified to a gel. Additional acetonitrile was added to a total volume of 12 ml, or until the mixture was again fluid. At the end of 16 hr the white gelatinous pro-

---

*Nutritional Biochemicals Corp., Cleveland, Ohio.

**When the carbodiimide was added, a fine black precipitate began to form in minute amounts. This precipitate was not characterized. It did not interfere with subsequent steps, and no attempt was made to remove it. Conceivably the precipitate was a compound formed from traces of Hg which somehow carried through to this step.

duct was filtered onto a Buchner funnel and was rinsed several times with acetonitrile. The acetonitrile in the filtrate was removed by evaporation in vacuo to yield an immobile highly viscous syrup which when reëvaporated with several portions of acetonitrile yielded additional product. This additional product was added to that already on the Buchner funnel. The product was washed with acetonitrile several more times* and was dried on the funnel by suction to 0.54 gm* of white to tan crumbly crystals. The crude crystals were dissolved in a minimal amount of methanol at 50°C and were filtered by gravity through a coarse frit. The pale yellow filtrate was evaporated at 25°C under a stream of air, and the residue was placed in the refrigerator at 10°C for 30 min. The gel-like crystals were transferred to a small funnel and were compressed and broken up into a powder with a spatula. The powder was dried in a vacuum desiccator. The melting point of this powder was 152°C-154°C, which compared favorably with that found by Schwyzer et al. (133), viz, 152°C-156°C.

Preparation of Carbobenzoxy-L-valyl-L-histidine Monohydrochloride (133)

A mixture of 0.52 gm of carbobenzoxy-L-valyl-L-histidine methyl ester in 2.6 ml of methanol was warmed to 38°C in a weighed 15-ml centrifuge tube. To the warm solution 1.56 ml of 1 F NaOH was added, and the mixture was allowed to stand at 38°C for 90 min** with constant

---

*When the filtrate was evaporated, the syrupy residue weighed 0.75 gm. How much of this was product is not known. The 0.54 gm of product which was recovered on the funnel was only 38% of the starting material.

**When the sodium hydroxide was added, the solution turned yellow and a white flocculent precipitate formed.

stirring. At the end of this time, the tube was centrifuged, and the supernatant was decanted into an unweighed 15-ml centrifuge tube. This supernatant was evaporated at 38°C to a volume of 1.6 ml. As the evaporation proceeded the color of the solution deepened to a yellow-orange, and white crystals began to separate out. The mixture of crystals and solution was extracted twice with 1.6 ml of ethyl acetate. The colorless organic phase was discarded, and the clear yellow aqueous phase was acidified at 0°C with 0.52 ml of concentrated HCl. On acidifying the solution, the yellow color disappeared and a dense, almost solid, mass of white crystals formed. The acidified solution was allowed to stand at 0°C for 1 hr, at which time it was filtered, and the white crystalline product was washed with distilled ice-water and dried *in vacuo*. The crystals melted at 115°C-120°C. Schwyzer *et al.* (133) reported a melting point of 121°C-123°C. The yield was 53%.

Preparation of L-valyl-L-histidine Dihydrobromide (45, 133)

Anhydrous glacial acetic acid was prepared by freezing 25 ml of commercial reagent grade material in a 100-ml beaker which was placed in water at 5°C. The freezing was made to proceed from the outside edge inward.* Shortly before the entire mass had solidified, the liquid that had not yet frozen was decanted, and the frozen acid was placed in a vacuum desiccator and pumped on at room temperature for 2 hr.** At

---

*Because glacial acetic acid supercools readily, the liquid must not be moved or agitated during the freezing process: otherwise the entire solution will freeze at once. If the latter occurs, the acid must be thawed and refrozen.

**An oil forepump must be used. A water pump will simply reintroduce water. The purpose of the pumping was to remove water vapor from

the end of this time air was readmitted through a $CaCl_2$ drying tube, the desiccator was sealed, and the acid was allowed to thaw at room temperature.

The anhydrous acetic acid was then saturated with anhydrous hydrogen bromide.

A mixture of 1.7 ml of anhydrous HOAc saturated with HBr and 0.29 gm of carbobenzoxy-L-valyl-L-histidine monohydrochloride in a 5-cm test tube* was allowed to stand at room temperature for 1 hr. At the end of this time 1.7 ml of anhydrous diethyl ether was added to precipitate the dihydrobromide, and the mixture was cooled for 3 hr at 2°C. The crystals were filtered onto a Buchner funnel, washed with anhydrous ether, and vacuum dried. A 68% yield was obtained.

Preparation of L-valyl-L-histidine (133)

Approximately 0.2 gm of L-valyl-L-histidine dihydrobromide was dissolved in 1 ml of water and passed through a 1 x 14-cm Amberlite 4B anion exchange column which had been equilibrated with water. The column was developed with water, and the first 30 ml of effluent was collected in a 100-ml round-bottom flask. This effluent was evaporated *in vacuo*, and the residue was reevaporated three times with methanol. The residue was transferred with warm methanol to a small test tube, the methanol was evaporated with a stream of air, and the crystals were vacuum dried in a desiccator to a pale yellow, almost white,** powder

---

the desiccator; the 2 hr pumping period could probably be reduced considerably.

*The tube was protected from moisture by a $CaCl_2$ drying tube.

**The trace of yellow was probably from the Amberlite 4B resin.

which decomposed at 194°C. Schwyzer et al. (133) reported a decomposition point of 205-206°C.

The last traces of yellow color could be removed by chromatographing the dipeptide on Dowex-1 as described by Schroeder et al. (127, 132), or more quickly, by boiling the crude powder with a small volume of anhydrous methanol, allowing the solution to cool, and discarding the pale yellow supernatant. The purified dipeptide was a pure white powder. The overall yield, based on the starting amount of carbobenzoxy-L-valine, was 15%. The structure was confirmed by paper electrophoretic comparison with an authentic sample of L-valyl-L-histidine obtained from a papain digest of βT-1 (cf. p. 63), by amino acid analysis, by Edman degradation (cf. p. 105) to establish the amino acid sequence of the dipeptide, and finally by mass spectrometry.*

A ID: Synthesis of N-acetyl-L-valyl-L-histidine (46, 47)

A mixture of 4.8 mg purified L-valyl-L-histidine and 0.050 ml of a 16.5% v/v solution of acetic anhydride in anhydrous acetic acid (cf. p. 209) was sealed into a 0.5 x 6.0-cm Pyrex tube. The sealed tube was immersed in a boiling water bath for 15 min. During this time the tube was frequently inverted so as to thoroughly mix the contents.** At the end of this time, the tube was cooled and opened, and the acetic acid was evaporated at room temperature in a desiccator under vacuum. The

---

*The mass spectroscopy of both L-val-L-his and N-acetyl-L-val-L-his was done at the Massachusetts Institute of Technology through the courtesy of Dr. Klaus Biemann and his associates (cf. p. 141).

**The tube was suspended by sewing thread at each end. The tube was inverted by shortening or lengthening one or the other of the threads.

residue, a yellow glass-like syrup, was dissolved in a small quantity of hot water and the evaporation in vacuo was repeated to remove traces of HOAc; this was done twice. The yellow glass-like crystals which resulted from the aqueous evaporations were evaporated once with hot benzene, then twice with absolute methanol to bring the product into a more crystalline form. The whitish powder from the methanol evaporations was recrystallized by dissolving it in a minimal volume of hot methanol, then chilling the solution at 2°C for 18 hr. The final product was a pure white powder. The structure of this product was confirmed by paper electrophoresis, its failure to give a positive ninhydrin reaction, by amino acid analysis, and finally by mass spectrometry.

APPENDIX II

FORMULAS FOR THE CALCULATION OF THE ERRORS IN THE QUANTITIES f, R_y^{calc}, K_e, K_e', AND K_e'' IN A SELF-HYBRIDIZATION EXPERIMENT

A II: ERRORS IN f, R_y^{calc}, K_e, K_e', AND K_e''

The following measures of the relative errors in f, R_y^{calc}, K_e, K_e', and K_e'' were obtained by differentiation of the expressions for these quantities which are given on pp. 80-81 of the thesis:

$$\frac{\Delta f}{f} = \left[1 - \frac{4}{(1-P)^2 f(1+2R_x)}\right]\left[\frac{P}{(1-P)}\right]\frac{\Delta P}{P} + \left[\frac{R_x}{(1-P)}\right]\left[1 + \frac{4}{(1+2R_x)f}\right]\frac{\Delta R_x}{R_x}$$

$$\frac{\Delta R_y}{R_y} = \left(1 - \frac{1}{R_y}\right)\left(\frac{1+P}{1-P}\right)\frac{\Delta P}{P} + \left(\frac{1}{R_y} - 1\right)\frac{\Delta f}{f}$$

$$\frac{\Delta K_e}{K_e} = \left(\frac{4}{2-f}\right)\frac{\Delta f}{f}$$

$$\frac{\Delta K_e'}{K_e'} = \left(\frac{3-2f}{(1-f)}\right)\frac{\Delta f}{f} \quad ,$$

$$\frac{\Delta K_e''}{K_e''} = \left(\frac{2-f}{1-f}\right)\frac{\Delta f}{f} \quad .$$

The latter three equations point up a desirable property possessed by K_e but not by K_e' or by K_e'': the relative error of the former is always of the order of three times the relative error in f irrespective of the value of f; on the other hand the relative errors in K_e' and K_e'' become infinite as the fraction of self-hybridization approaches 100%.

APPENDIX III

CALCULATION OF ACID-BASE DISSOCIATION CONSTANTS FROM PAPER ELECTROPHORETIC DATA

A III: CALCULATION OF ACID-BASE DISSOCIATION CONSTANTS FROM PAPER ELECTROPHORETIC DATA

Let us assume that the velocity $\underline{v}$ with which a molecule moves during paper electrophoresis depends only

a) on its ionic charge $\underline{q}$:

$$v_q = kq$$

where k is a constant of proportionality which is a function of the electric field E in which the molecule is moving, of the shape of the molecule, of the interaction of the molecule with the electrophoretic buffer and paper, and of the temperature at which the electrophoresis is conducted;

b) and on the average rate of flow $\underline{v}_b$ of buffer through the paper during the electrophoresis:

$$v_b = v_b \text{ (a constant)} .$$

Assumption a) is equivalent to assuming that the electrical force Eq on the molecule is just balanced by a "frictional" force $k'v_q$ that is proportional to the velocity v_q of the molecule. Thus $k = E/k'$. We shall only be concerned with molecules that have similar shapes, and that interact (at least approximately) equally with the buffer-paper system. Within this approximation k is a constant that is equal for all such similar molecules.

The quantity $\underline{v}_b$ in assumption b) might arise from a slight inequality in the level of the two buffer reservoirs: buffer could thus siphon through the paper from the reservoir having the higher level of

buffer to the other reservoir. Physically, this siphoning action is *analogous to the effect of a river current on several different swimmers:* each swimmer is carried forward with a velocity equal to that of the current, irrespective of any additional velocity he might have from his own efforts at swimming. Thus v_b is a constant that is equal for all similar molecules.

Consider the four specific molecules lys, leu, glu, and X which we shall assume bear net ionic charges of +1, 0, −1, and Q respectively.* After electrophoresis for a time t, these four molecules will have moved the following distances d:

$$d_{lys} = (1k + v_b)t$$

$$d_{leu} = (0k + v_b)t$$

$$d_{glu} = (-k + v_b)t \ .$$

$$d_X = (Qk + v_b)t \ .$$

If the distances each of the above molecules has moved is measured relative to leucine and this distance is designed D, then:

$$D_{lys} = kt$$

$$D_{glu} = -kt$$

$$D_X = Qkt \ .$$

---

*The exact charges upon these molecules will depend of course upon the pH of the electrophoretic buffer, as well as upon the ionization constants of the ionizable groups of these molecules in the buffer-paper system. For clarity of exposition we shall assume that the pH of the electrophoretic system is 6.50. In this case lys, leu, and glu have net ionic charges equal to those given in the text.

Our simple theory thus predicts that at pH 6.50 lysine and glutamic acid will move equal distances but in opposite directions from leucine. This prediction has been verified experimentally: measurements which were taken from five independent electropherograms of a mixture of lysine, leucine, and glutamic acid demonstrated that

$$D_{lys} = -(1.00 \pm 0.03)D_{glu} \ .$$

These electropherograms were obtained as described on p. 45 . It thus appears that our theory is a reasonable one for molecules which differ from each other no more than lysine, leucine, and glutamic acid.

If lys, leu, glu, and X are electrophoresed on the same sheet of paper, then the charge Q on the molecule X can be calculated from

$$Q = D_X/kt$$

where k is given by

$$k = (D_{lys} - D_{glu})/2 \ .$$

We now proceed to relate Q to the ionization constant of the molecule X. For simplicity we shall assume that the molecule X has only a single ionization constant the value of which is unknown.* It is this constant that we wish to calculate. For concreteness we take the molecule X to be the base A or its conjugate acid AH^+. Consider the

---

*Several ionizable groups may be present in X, but the ionization constant of all but one of these must be known if the unknown constant is to be obtained by the method to be described.

equilibrium

$$AH^+ \rightleftharpoons A + H^+ .$$

The ionization constant for this equilibrium is given by the expression

$$K_a = (A)(H^+)/(AH^+) \quad .$$

The average* charge Q (due to the above equilibrium) that is on the molecule X is then

$$Q = \frac{(AH^+)}{(AH^+) + A}$$

If the ionization constant K_a is now expressed in terms of (H^+) and Q we obtain

$$\underline{p}K_a = \underline{p}H + \log \left(\frac{Q}{1 - Q}\right) .$$

Similar equations hold if the molecule X is taken to be the acid AH or its conjugate base A^- with the equilibrium

$$AH \rightleftharpoons A^- + H^+ .$$

Then

$$K_a = (A^-)(H^+)/(AH) ,$$

$$-Q = \frac{(A^-)}{(AH) + (A^-)} ,$$

---

*At any given instant the charge on a given molecule AH^+ is of course +1, and that on a molecule A is zero. The average charge Q that is calculated in the text can be looked upon as the fraction of the time that the molecule X spends in the acid form AH^+.

and $$\underline{p}K_a = \underline{p}H - \log \left(\frac{-Q}{1 + Q}\right) .$$

Summarizing, from a knowledge of the $\underline{p}H$ of the electrophoretic buffer and from measurements of the distances which the compounds, lys, leu, glu, and a compound X move on electrophoresis in this buffer, the ionization constant of a single unknown ionizable group in the compound X can be calculated provided the ionization constants of all the other ionizable groups in X are known.

To check the validity of the above theory the $\underline{p}K_a$ for the ionization of the proton on the imidazole nitrogen of histidine was calculated as described above. The calculated value, based on six independent determinations was $\underline{p}K_a = 6.35 \pm 0.06$. The literature value (42) for free histidine at 25°C in dilute aqueous media is 6.00.* Our value of 6.35 indicates that the imidazolium ion of histidine is more difficultly ionizable under our conditions of electrophoresis than it is in dilute aqueous media at 25°C. This may be due to the fact that the electrophoresis was carried out at a temperature of approximately 10°C, not at 25°C.

Calculations from the electropherogram that is shown in Fig. 13, p. 153, give the following estimates at 10°C for the $\underline{p}K_a$ of the imidazolium ion of val-his and reduced R-val-his:

---

*Literature (24, 42) values range at 25°C from 6.00 for free histidine to 6.83 for compounds such as carnosine in which the α-amino group of histidine is bound in an amide linkage.

$$\underline{p}K_a^{\text{val-his}} = 6.62$$

$$\underline{p}K_a^{\text{reduced val-his}} = 6.34 \ .$$

These values are in agreement with the literature values which are found for other histidine-containing peptides. We conclude from this agreement that molecules of the type val-his are sufficiently similar to lys, leu, glu, and his to fulfill our initial assumptions (cf. p. 216). Therefore we shall now calculate the $\underline{p}K_a$ of the conjugate acid of the Schiff base R-val-his. If we take the average of the above two values, viz. 6.48 ± 0.14, as the $\underline{p}K_a$ for the imidazolium group of R-val-his, then the ionization constant at 10°C for the proton on the Schiff base nitrogen is found from Fig. 13 to be

$$\underline{p}K_a^{\text{R-val-his}} = 6.64 \pm 0.14 \ .$$

Cordes and Jencks (23) found values of $\underline{p}K_a$ ranging from 5.40 to 7.70 at 25°C for the Schiff bases between tertiary butyl amine and various mono-substituted benzaldehydes (also see pp. 111 and 154).

Although the above agreements speak well for the calculation of $\underline{p}K$ values from electrophoretic data, the reader should be cautioned that there are cases in which the method fails completely. If the $\underline{p}K_a$ for the proton on the imidazole nitrogen in the compounds N-acetyl-val-his or the octapeptide val-his-leu-thr-pro-glu-glu-lys is calculated from their paper electrophoretic behavior at $\underline{p}H$ 6.50, absurd values are obtained. This is not surprising. The N-terminal amide bond in

N-acetyl-val-his would not be expected to bear much resemblance to the N-terminal α-amino group of say val-his or leucine. And certainly, to expect a theory which is based on the electrophoretic behavior of free amino acids to hold for an octapeptide is to expect too much. These considerations should not discourage the experimentalist from using paper electrophoretic data to calculate pK values of unknown compounds, but they should caution him to ascertain the validity of the underlying assumptions of the theory before such a calculation is made.

APPENDIX IV

CORRECTION FOR NONCONSTANT RIBOSOMAL ACTIVITY

A IV: CORRECTION FOR NONCONSTANT RIBOSOMAL ACTIVITY

Let the ribosomal activity $\underline{a}$ be given as a function of time by the equation

$$\underline{a} = a - \sum_{1}^{n} b_i t^i .$$

The constants a and b_i can be determined by fitting the experimental points to a least squares curve or by some other appropriate method.

The rates at which the ribosomes are synthesizing hemoglobin A_{Ic} and A_{II} respectively are then

$$\underline{A} = c\underline{a} = ca - c \sum_{1}^{n} b_i t^i = A - c \sum_{1}^{n} b_i t^i$$

$$\underline{B} = d\underline{a} = da - d \sum_{1}^{n} b_i t^i = B - d \sum_{1}^{n} b_i t^i ,$$

where c and d are constants of proportionality, and A = ca, and B = da. It should be noted that A and B (but not $\underline{A}$ and $\underline{B}$) are identical to the A and B on pp. 160 and 161: They are the constant rates at which A_{Ic} and A_{II} respectively would be synthesized by the ribosomes if all the b_i were zero, i.e., if the ribosomal activity remained constant throughout the incubation.

Let A_{Ic} and A_{II} represent the amount of radioactive A_{Ic} and A_{II} which would be present at time t if the ribosomal activity remained constant during the incubation, and let A^o_{Ic} and A^o_{II} represent the amount of radioactive A_{Ic} and A_{II} which would be present at time t if the ribosomal activity decreased with time during the incubation. In the latter case the net rates at which A_{Ic} and A_{II} are being synthesized

at any time t are

$$\frac{dA^{o}_{Ic}}{dt} = \underline{A} + kA^{o}_{II} - k_{-}A^{o}_{Ic}$$

$$\frac{dA^{o}_{II}}{dt} = \underline{B} + k_{-}A^{o}_{Ic} - kA^{o}_{II} ,$$

where k and k_{-} are identical to the k and k_{-} of p. 161. These equations reduce to those at the top of p. 161 when all the b_i are zero. When the b_i are not all zero, the equations can in all cases be solved without approximation by the "Method of Undetermined Coefficients" (116). We shall give the solution only for the case where the ribosomal activity decreases linearly* with time, i.e., $b_1 \neq 0$, and $b_i = 0$ for $i > 1$, so that

$$\underline{a} = a - b_1 t .$$

With this assumption, the solution to the above differential equations is

$$A^{o}_{Ic} = A_{Ic} - \Delta_1 - k\,\Delta_2$$

$$A^{o}_{II} = A_{II} + \Delta_1 - k_{-}\,\Delta_2 ,$$

---

*The equations are not significantly harder to solve if the ribosomal activity decreases nonlinearly in time; however, the form of the final equations becomes very complex. Further, the accuracy with which the decrease in ribosomal activity can be measured frequently does not justify anything more than a linear approximation; and even if it does, from a calculational viewpoint it is simpler, and quantitatively just as satisfactory, to decompose the nonlinear function into several linear functions, each of which closely approximates the actual function for a particular period of time.

where A_{Ic} and A_{II} are identical to the A_{Ic} and A_{II} of p. 161, and where Δ_1 and Δ_2 are given by

$$\Delta_1 = \frac{b_1(Ak_- - Bk)}{a(k + k_-)^3}\left\{(k + k_-)t - \left[1 - e^{-(k + k_-)t}\right]\right\}$$

$$\Delta_2 = \frac{b_1(A + B)}{2a(k + k_-)}\, t^2 \ .$$

In practice, the simplest way to use the above equations is to determine A and B experimentally as described* on p. 162, and a and b_1 as described on p. 224. Values of k and k_- are guessed, and A^o_{Ic} and A^o_{II} are calculated by the two equations at the bottom of the preceding page. If the correct values of k and k_- have been chosen (i.e., the values which correspond to physical reality) the calculated values of A^o_{Ic} and A^o_{II} should agree to within the experimental error with the observed values for all times, provided that the basic assumptions which were made on pp. 160 and 161 are valid.

For all the exceptional cases which were discussed on pp. 162 and 163 the correction given by the equations at the bottom of p. 225 for a linear decrease in ribosomal activity reduces to the particularly simple forms

---

*At t = 0 the correction terms Δ_1 and Δ_2 as well as their time derivatives are zero, so that the method which is described on p. 162 for determining A and B still suffices.

$$A^{o}_{Ic} = A_{Ic} \left(1 - \frac{b_1}{2a} t\right)$$

and

$$A^{o}_{II} = A_{II} \left(1 - \frac{b_1}{2a} t\right) .$$

In practice exceptional cases can be recognized from the fact that either the experimentally observed specific activity of A_{Ic} is equal to that of A_{II} for <u>all</u> times t, or the specific activity of either A_{Ic} or A_{II} is zero.

Part VI

PROPOSITIONS

P I: On the Complete Structure of the Blocking Group R of Hemoglobin A_{Ic}

P II: Considerations of Structure-function Relationships in the Reduction of Human Ferrihemoglobin A in the Presence of Cysteine

P III: A Proposed Reinvestigation of the Kinetics of Reduction of Horse Hemoglobin by Cysteine

P IV: A Quantitative Approach to the Problem of Multiple Hits in Paleogenetics

P V: On Nonspecific Cleavage in the Edman Degradation

PROPOSITION I: SUGGESTIONS FOR COMPLETING THE CHARACTERIZATION OF THE BLOCKING GROUP R OF HEMOGLOBIN A_{Ic}

Proposition: Two methods are proposed by means of which two highly purified low molecular weight compounds which contain the blocking group R of hemoglobin A_{Ic} can be prepared. It is likely that the mass spectra of these compounds would suffice to determine the structure of R completely.

P IA: Introduction

At the beginning of this thesis (cf. p. 15) it was stated that "characterization of the blocking group R is sufficiently advanced so that the directions which additional efforts should take are now well marked out." The facts upon which this statement was based are listed in the thesis summary (p. 192) and form the background for this proposition. From a practical viewpoint the most important of these facts is that the lability of the blocking group R can be obviated by reducing the imine double bond of the Schiff base with $NaBH_4$. We shall therefore assume that reduced $\beta^{A_{Ic}}$T-1 (cf. p. 147) and reduced R-val-his (cf. p. 152 and p. 64) are available as starting materials for further characterization. We shall further assume, in accordance with past experience, that the purity of $(\beta^{A_{Ic}}H_2)$T-1 and RH_2-val-his* is no more than can be obtained by the column chromatographic procedures that are described in the thesis proper. As we have seen (p. 140), compounds that have

---

*Henceforth we shall use the abbreviations $(\beta^{A_{Ic}}H_2)$T-1 and RH_2-val-his respectively to designate $\beta^{A_{Ic}}$T-1 and R-val-his, each of which has been reduced with $NaBH_4$ as described on pp. 146-154.

been isolated by column chromatography are not usually pure enough for the purpose of obtaining a mass spectrogram from them. The primary goal of each of the methods to be described below is to obtain compounds which are pure enough for mass spectrometric analysis, for at present we believe that it is by such an analysis that the structure of R can be most readily obtained.

P IB: Method I--Purification of RH_2-val-his by Thin Layer Silica Gel Chromatography

Experiments with model compounds have demonstrated that the following chromatographic system is likely to effect the purification of RH_2-val-his to the degree necessary for mass spectrometric analysis.*

A 20 x 20 x 0.1-cm chromatographic plate of Silica Gel HR** is prepared from a slurry of 11.76 gm of Silica Gel HR in 35.3 ml of H_2O by any standard (1) method. The freshly prepared plate is allowed to dry overnight in air at room temperature, is then heated for 15 min at 110°C, and is air cooled for 45 min. It is then prechromatographed in 9:1::MeOH:HCl v/v to free it from iron.*** The plate is activated by

---

*The R_f values for his, val-his, N-acetyl-val-his, and val in the chromatographic system that is described are 0.10, 0.13, 0.39, and 0.51 respectively. On a 20 x 20 x 0.1-cm plate of Silica Gel HR, a mixture consisting of up to 50 µm of each of several of these model compounds could readily be resolved. The purified resolved compounds were eluted from the gel in 86% yield.

**E. Merck AG, Darmstadt, Germany; Distributor in the United States is Brinkman Instruments, Inc., Westbury, New York. Silica Gel HR and not Silica Gel G should be used. The latter contains $CaSO_4$ as a binder, and this compound is water soluble so that when the peptide is eluted from Silica Gel G it will be contaminated with $CaSO_4$. Silica Gel HR does not contain any binder and thus is free from this difficulty.

***The iron appears as a rapidly moving pale yellow line which moves parallel to the solvent front. When the front reaches the top of the

heating it at 110°C for 15 min and cooling it in air for 45 min. The impure sample of RH_2-val-his is applied to the gel in 50 λ of 1 F HCl along a line 15 cm in length and 1 cm from the bottom edge of the plate. After the sample has been applied, the plate is allowed to air dry for 1 hr at room temperature and is then heated for 5 min at 110°C and air cooled for 45 min. The chromatogram is developed with n-butanol:acetic acid: water::65:17:18 v/v/v in a covered battery jar. The jar should be lined with filter paper that has been saturated with the solvent. After development, the plate is removed from the chromatographic jar and is air dried for 5 min and is then heated at 110°C for 10 min. The peptides and amino acids on the plate can be located by covering all but a narrow strip of the dry gel with a metal plate. The exposed strip is sprayed with ninhydrin* to locate those peptides with free α-amino groups and then with diazotized sulfanilic acid** to locate those peptides which

---

plate, the latter is removed from the MeOH-HCl mixture and allowed to dry. That portion of the gel which is near the top of the plate and which contains the iron is then scraped off with the metal edge of a ruler and is discarded. (Cf. Ref. 1, p. 30.)

*The ninhydrin reagent (cf. Ref. 1, p. 94) is prepared by mixing equal volumes of reagents I and II. Reagent I contains 50 ml of anhydrous EtOH, 10 ml of HOAc, 2 ml of 2, 4, 6-collidine and 0.10 gm ninhydrin. Reagent II contains 1% $Cu(NO_3)_2 \cdot 3H_2O$ in absolute ethanol v/v. The sprayed plate is heated at 80°C or on an asbestos-covered hot plate until the location of the peptides are obvious.

**This spray may be applied after the peptides have been located with ninhydrin as in the preceding footnote. Diazotized sulfanilic acid is prepared by mixing equal volumes of reagents I and II. Reagent I contains 0.7% $NaNO_2$ in absolute methanol w/v. Reagent II contains 0.7 gms sulfanilic acid in a mixture of 36.8 ml methanol, 8.0 ml conc. HCl, and 55.2 ml H_2O. Immediately after the diazotized sulfanilic acid has been applied, the plate is sprayed with reagent III which contains 9 gms Na_2CO_3 in a mixture of 31.8 ml methanol and 59.2 ml H_2O. Peptides which contain histidine appear as yellow to orange spots after a few minutes.

contain histidine. After RH_2-val-his has been located by the above sprays, that (dry) portion of the gel which contains this compound is scraped into a 10 ml weighed centrifuge tube. The gel is stirred with three times its weight of water, centrifuged, and the supernatant (which contains the peptide) is decanted. This elution procedure is repeated three times. The combined supernatants are then evaporated in order to obtain the purified RH_2-val-his.

P IC: Method II--Preparation and Purification of RH_2-val-O-$\overset{O}{\overset{\|}{C}}$-$\varphi(NO_2)_2$

The principle of the method to be described is to hydrolyze $(\beta^{A_{Ic}}H_2)$T-1 into its constituent amino acids, to convert the free amino and carboxyl groups of these amino acids to hydroxyl groups by diazotization and reduction respectively, and finally to convert these alcohols to their 3,5-dinitrobenzoates. Because of charge differences among these 3,5-dinitrobenzoates, the 3,5-dinitrobenzoates resulting from pro and RH_2-val can be separated from the other 3,5-dinitrobenzoates by electrophoresis; and finally, the 3,5-dinitrobenzoates from pro and RH_2-val can probably be separated from each other by thin layer chromatography. All reactions preceding the electrophoresis of the 3,5-dinitrobenzoates are carried out in a single test tube without transfer so that there is virtually no loss of compound.* After the 3,5-dinitrobenzoates have been formed, their separation from each other by electrophoresis and chromatography can be followed by the natural yellow color of these compounds.

---

*This quantitative aspect of Method II might favor it over Method I, because the isolation of RH_2-val-his from the papain hydrolysate of $(\beta^{A_{Ic}}H_2)$T-1 proceeds in only 28% yield (cf. p. 70).

The sequence of reactions that are proposed are as follows:

$$RH_2\text{-val-his-leu-thr-pro-glu-glu-lys} \xrightarrow{1)^{*}\ {}^{+}N_2\text{-}\varphi\text{-}SO_3^-}$$

$$HO\text{-}\varphi\text{-}SO_3^- + RH_2\text{-val-(his-}\varphi\text{-}SO_3^-\text{)-leu-thr-pro-glu-glu-lys} \xrightarrow[110°C,\ 24\ hr]{2)^{**}\ 6F\ HCl}$$

$$HO\text{-}\varphi\text{-}SO_3^- + RH_2\text{-val} + \text{his-}\varphi\text{-}SO_3^- + \text{leu} + \text{thr} + \text{pro} + 2\text{glu} + \text{lys} \xrightarrow{3)^{***}HNO_2}$$

$$HO\text{-}\varphi\text{-}SO_3^- + RH_2\text{-val} + HO\text{-his-}\varphi\text{-}SO_3^- + HO\text{-leu} + HO\text{-thr} + \text{pro} + 2HO\text{-glu}$$
$$+ (HO\text{-})_2\text{lys} \xrightarrow{4)^{****}\ LiAlD_4}$$

$$HO\text{-}\varphi\text{-}SO_3^- + RH_2\text{-val-OH} + (HO\text{-})_2\text{his-}\varphi\text{-}SO_3^- + (HO\text{-})_2\text{leu} + (HO\text{-})_2\text{thr} + \text{pro-OH}$$
$$+ 2(HO\text{-})_3\text{glu} + (HO\text{-})_3\text{lys} \xrightarrow{5)\ (NO_2)_2\ \varphi\text{-}\overset{O}{\overset{\|}{C}}\text{-}Cl}$$

---

*The coupling of the imidazole nucleus with diazotized sulfanilic acid is probably not a necessary step. This step, however, costs nothing and the extra negative charge introduced by the $^-SO_3^-$ grouping will later be helpful during the electrophoretic separations.

**Each of the amino acids which result from this acid hydrolysis has a free α-amino group and a free carboxyl group with the exception of the secondary amines RH_2-val and pro which have only a free carboxyl group each.

***This step converts all free primary amino groups to the alcohols. We shall designate these alcohols as HO-leu, etc.

****This step converts all free carboxyl groups to the alcohols. The deuteride anion is used to provide a marker for mass spectrometry.

$$(DNB)\text{-}\varphi\text{-}SO_3^- + his\text{-}\varphi\text{-}SO_3^-(\text{-}DNB)_2 + \begin{matrix} leu(\text{-}DNB)_2^* \\ thr(\text{-}DNB)_2 \\ glu(\text{-}DNB)_3 \\ lys(\text{-}DNB)_3 \end{matrix} + \begin{matrix} pro\text{-}DNB \\ RH_2\text{-}val\text{-}DNB \end{matrix}$$

The dinitrobenzoate of p-hydroxysulfonic acid has a charge of approximately -1 at pH 6.0, and that of histidine sulfonate a charge of $-\frac{1}{2}$. The di- and tri- dinitrobenzoates that are derived from leu and thr, and glu and lys respectively have zero charge. The dinitrobenzoates which are derived from pro and RH_2-val have a charge of +1. Thus the latter two compounds can readily be separated from the remaining six dinitrobenzoates** by electrophoresis at pH 6.0.***

Finally the dinitrobenzoates that are derived from pro and RH_2-val can probably be separated from each other by thin layer chromatography on Silica Gel with either benzene:petroleum ether (b.p. 60°-80°C)::1:1 v/v or toluene:ethyl acetate::1:1 (cf. Ref. 1, p. 215).

An additional point, that is perhaps not unimportant, is the ease

---

*DNB is an abbreviation for $(NO_2)\varphi\text{-}\overset{O}{\overset{\|}{C}}\text{-}$.

**It would probably be preferable to carry out the electrophoresis on a Silica Gel HR support for it has been our experience that the yields of compounds which have been eluted from paper are low.

***Because of the relative ease with which ester bonds are hydrolyzed, the electrophoretic buffer should be chosen carefully.

with which RH_2-val-DNB can be converted to other derivatives (such as RH_2-val-O-$\overset{\overset{\displaystyle O}{\|}}{C}$-$CH_3$) that might be used for characterization.

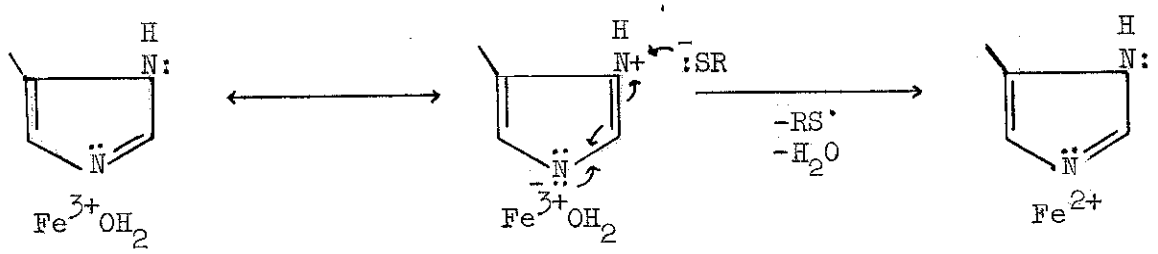
BIBLIOGRAPHY for PROPOSITION I

1. Kurt Randerath, Thin Layer Chromatography, Academic Press, New York, 1963.

PROPOSITION II: ON THE REDUCTION OF HUMAN FERRIHEMOGLOBIN A BY CYSTEINE

P IIA: Introduction

Holmquist and Vinograd (3) have proposed a mechanism that is reasonably successful in explaining the rate of reduction of human ferrihemoglobin A by cysteine as a function of hemoglobin and cysteine concentrations and as a function of pH. In this mechanism, the following three assumptions (among others) were made. First, each of the four iron atoms in the ferrihemoglobin molecule was assumed to be reduced independently of the others; i.e., interactions among the polypeptide chains were neglected. Second, the rate of reduction of each of the four iron atoms was assumed to be the same, even though the environment of the iron atom in the α chain differs from that in the β chain. Third, the reduction was assumed to occur via the heme-linked histidine in the fifth coordination* position of the iron atom:



*In hemoglobin A each iron atom has octahedral ligancy. Four of the ligand positions are occupied by nitrogen atoms of the planar porphyrin ring, the fifth is occupied by a histidine residue of the polypeptide chain, and the sixth by molecular oxygen or some other ligand such as H_2O, OH^-, CN^-, etc. In aqueous solutions of ferrihemoglobin A the ligand in the sixth coordination position is H_2O at acidic or neutral pH values and is OH^- at basic pH values.

Although these assumptions led to a mathematical model that was successful in explaining several features of the reduction, their validity is not obvious. It is the purpose of this proposition to suggest some experiments by which the correctness of these three assumptions can be judged directly.

P IIB: Estimation of the Interactions between Polypeptide Chains

Hemoglobin A (cf. p. 3 , this thesis), hemoglobin H (5, 7) and hemoglobin α_2 (1, 2) can be represented by the gross structural formulas $\alpha_2\ \beta_2$, β_4, and α_2 respectively. In these formulas each Greek letter designates a single polypeptide chain. The three dimensional structures of these molecules can be represented schematically by the formulas

$$\begin{matrix} {}^{1}\alpha\ {}^{4}\beta \\ {}^{3}\beta\ {}^{2}\alpha \end{matrix}\ , \qquad \begin{matrix} {}^{1}\beta\ {}^{4}\beta \\ {}^{3}\beta\ {}^{2}\beta \end{matrix}\ , \qquad \text{and} \qquad \begin{matrix} {}^{1}\alpha \\ {}^{2}\alpha \end{matrix}\ .$$

Hb-A Hb-H Hb-α_2

The numerical superscripts to the left of each Greek letter label the position in space of each polypeptide chain.

For a given set of experimental conditions let us represent the rates of reduction of hemoglobin A, H, and α_2 by r_A, r_H and r_{α_2} respectively. These rates can be written as the sum of the rates of reduction of the individual polypeptide chains plus the sum of terms which represent interactions between the various polypeptide chains:

$$r_A = (2r_\alpha + 2r_\beta) + (r_{1\alpha 2\alpha} + r_{3\beta 4\beta} + 2r_{1\alpha 4\beta} + 2r_{1\alpha 3\beta})$$

$$r_H = (4r_\beta) + (r_{1\beta 2\beta} + r_{3\beta 4\beta} + 2r_{1\beta 4\beta} + 2r_{1\beta 3\beta}) \quad \text{(Eq. 1)}$$

$$r_{\alpha_2} = (2r_\alpha) + (r_{1\alpha 2\alpha})$$

In these equations r_α and r_β are the rates at which the α and β polypeptide chains respectively are reduced in the absence of interaction, and $r_{1\alpha 4\beta}$, for example, represents the contribution to the rate of reduction of hemoglobin A from the interaction of the α chain in the spatial position labeled 1 on the preceding page with the β chain in position 4. Although r_α and r_β must always be positive or zero, each interaction term may be negative, zero, or positive.

We now introduce an interaction parameter I. This parameter is defined by the solution to the three equations

$$\begin{aligned} r_A &= 2r_\alpha + 2r_\beta + mI \\ r_H &= 4r_\beta + nI \\ r_{\alpha_2} &= 2r_\alpha + pI\,. \end{aligned} \quad \text{(Eq. 2)}$$

These three equations are equivalent to Eq. 1 if the sum of the interaction terms for r_A, r_H, and r_{α_2} respectively is set equal to mI, nI, and pI. Explicitly I is given by

$$I = \frac{2r_A - r_H - 2r_{\alpha_2}}{2(m - p) - n} \quad \text{(Eq. 3)}$$

The numbers m, n, and p are numbers which are in the last analysis to be

determined by experiment.* A simpler approach is to assign each of the interaction terms the same value I. In this approximation m, n, and p are integers that have the values 6, 6, and 1 respectively, and I is given by the expression**

$$I = \frac{1}{4}(2r_A - r_H - 2r_{\alpha_2}) \ . \qquad \text{(Eq. 4)}$$

The interaction parameter I has one very simple property that is independent of any approximations that are made about m, n, and p: if I is zero, no interactions that have a net effect on the reduction rates occur;*** if it is not zero, interactions do occur. The only way that I can be zero is for the numerator of I to vanish:

$$2r_A - r_H - 2r_{\alpha_2} = 0 \ . \qquad \text{(Eq. 5)}$$

---

*The experiment(s) would be very difficult to design and analyze because a knowledge of m, n, and p implies a knowledge of some of the individual interaction terms. We shall make no attempt to find these individual terms but shall content ourselves with answering the question: "Do interactions that have a net effect on the rate of reduction occur among the polypeptide chains or not?" We shall not try to answer the much more difficult question: "Between which chains do the interactions occur?"

**As formulated in Eqs. 2, 3, and 4, I is a function of both the hemoglobin and cysteine concentrations. From a strictly theoretical viewpoint it would be desirable to eliminate this concentration dependence because we are considering interactions within a single molecule. This can be accomplished by a simple mathematical transformation. For our present purposes the form of I that is given in the text is satisfactory.

***I = 0 does not imply that each interaction term (e.g., $r_{1\alpha 4\beta}$) is zero: For example, if two nonzero interaction terms had equal but opposite effects on the rate of reduction no net affect would be observed.

Each term in Eq. 5 is the experimentally measured rate at which hemoglobin A, H, and α_2 respectively is reduced by cysteine for a fixed set of conditions. These rates can be determined as described in Ref. 3.

P IIC: Test of the Assumption that $r_\alpha = r_\beta$

If Eqs. 2 are solved for r_α and r_β, it is seen that these two quantities are equal if and only if

$$r_A = r_H + (m - n)I \ . \qquad \text{(Eq. 6)}$$

If there is no interaction, or if the approximation is made that all the interaction terms are equal, Eq. 6 becomes

$$r_A = r_H \ . \qquad \text{(Eq. 7)}$$

Equation 7 is a simple experimental criterion for determining whether the iron atom in the α chain and the iron atom in the β chain are reduced at equal rates.

P IID: Test for the Involvement of the Heme-Linked Histidine in the Reduction Mechanism

It is not our purpose here to go into the original reasons that implicated the heme-linked histidine in the reduction mechanism. Very briefly, mechanisms that omitted a consideration of this histidine could not be brought into quantitative agreement with the experimental data (3).

However, it may be in order in the following two paragraphs to summarize a few theoretical concepts which suggest that a reduction path which involves an attack by the thiol anion RS^- on the heme-linked histidine in the fifth coordination position of the iron atom would be

favored over a path that involved a direct frontal attack by RS^- on the iron atom by way of the sixth coordination position (cf. footnote p. 236).

A reduction path through the sixth coordination position would be inefficient because of the competitive reaction

$$Fe^{3+}OH_2 + RS^- \rightleftharpoons Fe^{3+}OH^- + RSH \ .$$

Here only the intermolecular exchange of a proton occurs; no reduction of the iron atom takes place. Moreover, if the reduction did occur, it would of necessity be by way of the transition state intermediate ($Fe^{3+}OH_2{}^-SR$). The final transfer of one of the sulfur 3p electrons to the iron atom could occur in one of two ways: 1) The electron could travel through the σ bonds between the oxygen and hydrogen atoms of the water molecule, or through the 2s orbital of the oxygen atom to the d_{z^2}* orbital of the iron atom; 2) The RS^- anion could attack the iron atom directly; in this mode of attack the sulfur 3p electron would be transferred directly to the d_z orbital of the iron atom. Such an attack would require the displacement of the water molecule and would imply a transition state in which seven ligands are grouped around the iron atom. From an energetic viewpoint, the transfer of a sulfur electron by process 1) is not favorable because of the high electronegativity of the oxygen atom, and process 2) is not favorable because of steric crowding and electrostatic repulsion.

---

*The z direction is taken as perpendicular to the plane of the porphyrin ring.

On the other hand, if RS were to attack *via* the heme-linked histidine as outlined on p. 236, the easily polarized 3p electrons of the ionized sulfur atom could be transferred by way of the π-electron system of the imidazole ring to the d_{z^2} orbital of the iron atom. Moreover, such an electron transfer would not involve a transition state in which the iron atom would have to accommodate seven ligands. Both these considerations would suggest a transition state of lower energy than the corresponding transition state for an attack by way of the sixth coordination position.

The above considerations can be investigated experimentally as follows: Hemoglobin M_{Iwate} (6, 8) is a naturally occurring hemoglobin which is identical to hemoglobin A except that in the α chains of hemoglobin M_{Iwate} the heme-linked histidine is replaced by tyrosine. Thus, the reduction of ferrihemoglobin M_{Iwate} by cysteine should occur more slowly than the reduction of ferrihemoglobin A if the considerations of the preceding three paragraphs are correct. *In vivo*, this is indeed the case (4, 8). A quantitative determination of the reduction rate of hemoglobin M_{Iwate} *in vitro* would allow an estimate to be made of the energies of the two transition states that result from an attack of the thiol anion by way of the histidine in the fifth coordination position and by way of the sixth coordination position respectively.

BIBLIOGRAPHY for PROPOSITION II

1. A. I. Chernoff, Xth Congress Int. Soc. Haematology, Abstract L:3, Stockholm, 1964.

2. P. Fessas and D. Loukopoulos, Science, 143, 590, 1964.

3. W. R. Holmquist and Jerome Vinograd, Biochim. et Biophys. Acta, 69, 337, 1963.

4. E. R. Huehns and E. M. Shooter, J. Med. Gen., 2, no. 1, 48, 1965.

5. R. T. Jones and W. A. Schroeder, Biochemistry, 2, 1357, 1963.

6. R. T. Jones, et al., Federat. Proc. (Amer. Soc. Exp. Biol.), 23, Abstract 433, 1964.

7. R. T. Jones, et al., J. Am. Chem. Soc., 81, 3161, 1959.

8. T. Miyagi, Acta Haem. Jap., 26, 538, 1963.

PROPOSITION III

It is proposed to reinvestigate the reduction of horse ferrihemoglobin in the presence of cysteine in the belief that

a) earlier work may be incorrect

b) such a reinvestigation would throw light upon the relation between protein structure and function, and in particular might show that it is not valid to extrapolate structural results obtained from the X-ray study of one species to another species if a detailed understanding of the function of the protein is wanted.

P IIIA: Introduction

Cremer (3) showed that cysteine could reduce hemin. Morrison and Williams (6) demonstrated that cysteine reduced human ferrihemoglobin in the presence of air, H_2, CO, and *in vacuo*. Kiese (5) studied the reduction of horse ferrihemoglobin quantitatively using a variety of reductants, including cysteine. Holmquist and Vinograd (4) made quantitative measurements of the rate of reduction of human ferrihemoglobin A in the presence of cysteine and as a function of hemoglobin concentration, cysteine concentration, temperature, pH, ionic strength, stereospecificity, atmospheric composition, and the method of isolation of the hemoglobin; they also proposed a mechanism for the reduction which was partially successful in quantitatively describing the observed rates of reduction.

This proposition will be concerned solely with certain interesting differences between the work of Kiese (5) on horse hemoglobin and that of Holmquist and Vinograd (4) on human hemoglobin A.

It has been suggested (4) that Kiese's study of the reduction of horse hemoglobin by cysteine might profitably be reëvaluated in view

of the facts that he determined the reduction rates manometrically by determining the CO uptake of the reduced hemoglobin and that CO was later (4) found to be a strong inhibitor of the reduction of human hemoglobin by cysteine. Further (4, 5), for horse hemoglobin, the pH optimum for the reduction occurs at pH 8.1; for human hemoglobin the optimum is at pH 7.68. Between pH 5.5 and pH 7.68 there is a 5-fold increase in the rate of reduction of horse hemoglobin, while for human hemoglobin the rate increase is 34-fold. Finally, the shape of the pH vs. rate of reduction curve is significantly different for the two hemoglobins.

We suspect these facts to be inconsistent with what is currently known about the detailed primary structures of horse and human hemoglobins.

P IIIB: Basis for Proposition

Forty-three known differences exist in the primary sequences (7) of the polypeptide chains of horse* and human hemoglobins. Eighteen of these occur in the α chain and twenty-five in the β chain. In the α chain twelve of the differences involve the substitutions of neutral amino acids for each other (val $\leftrightarrow$ ala, for example); three involve the substitution of asparagine for a neutral amino acid (e.g., leu $\leftrightarrow$ asg); and one involves the replacement of aspartic acid by glutamic acid.

---

*Horses have two types of hemoglobins, an electrophoretically fast and slow moving component, respectively, which are present in a mole ratio of 3:2 (1, 2). We speak here of the slow component. Because horse hemoglobin heterogeneity was unknown at the time of his studies, Kiese may have been using a mixture of the two types. This ambiguity alone warrants reinvestigation of the reduction.

None of these three types of substitutions involves a change in the ionic charge of the polypeptide chain. The two substitutions involving a change of one unit of charge occur at residue 82 (asp ↔ ala) and residue 85 (asg ↔ asp). The total ionic charge on the α chain is the same for horse and human hemoglobin between pH 5 and pH 8.* It is plausible to assume that the val ↔ ala type of substitution results in minimum change in the molecule. Excepting these and the substitution at residue 85, which is located near the extreme edge of the porphyrin ring, all the other substitutions occur at positions which are distant in space from this ring. Further, none of the substitutions involve histidine, an amino acid which may be importantly involved in the reduction mechanism (4; see also Proposition II, p. 236, this thesis). Although small changes in hydrogen bonding may occur as a result of the above substitutions, it seems reasonable to suggest that the differences in the rates of reduction of horse and human ferrihemoglobin are not due to differences between the α polypeptide chains.

The above discussion would place the onus on the β chains to account for the different reduction rates between horse and human hemoglobin. An analysis analogous to that carried out above for the α chain shows that for the β chain there are fifteen substitutions of the val ↔ ala type, none of the type leu ↔ asg, and three of the type asp ↔ glu. None of these involves a change in charge. There are three substitutions of the type glu ↔ val, three of the type his ↔ val, and

---

*At pH 5, 6.6 and 8.0 this charge is +12, +7, and +2 respectively.

one of the type his ↔ arg. Between pH 5 and 8, the horse β chain bears a charge of -2 relative to the human β chain.* Again limiting our considerations to those substitutions which result in a change of charge and which are near the iron-porphyrin moiety, only the substitution at residue 69 (his ↔ gly) can reasonably be expected to exert an influence on the reduction. However, this influence should be small for residue 69 is at the edge of the porphyrin ring.

P IIIC: Conclusions

The conclusion to be drawn from the above considerations is that the large observed differences in the reduction rates of horse and human hemoglobins cannot reasonably be explained by differences in their currently known structures. Therefore we propose that the reduction of horse hemoglobin be studied under conditions identical to those of Ref. 4. If the experimental differences discussed above persist under these conditions, then it is clear that some very interesting structure-function relationships exist in the hemoglobin molecule, and that further investigations along the line of Proposition II might help delineate these relationships. The persistence of these differences would also suggest that the extrapolation of X-ray data from one species to another can at best be considered a poor approximation where structure-function relationships are involved.

---

*For horse hemoglobin the net ionic charges at pH 5, 6.6, and 8.0 respectively are +6, +1½, and -3.

BIBLIOGRAPHY for PROPOSITION III

1. A. D. Bangham and H. Lehmann, Nature, 181, 267, 1958.

2. R. Cabannes et Ch. Serain, C. R. Séances Soc. Biol., 149, 1193, 1955.

3. W. Cremer, Biochem. Z., 192, 426, 1928.

4. W. R. Holmquist and Jerome R. Vinograd, Biochim. et Biophys. Acta, 69, 337, 1963.

5. M. Kiese, Biochem. Z., 316, 264, 1943.

6. D. B. Morrison and E. F. Williams, Science, 87, 15, 1938.

7. Walter A. Schroeder and Richard T. Jones, Fort. der Chemie Org. Naturstoffe, XXIII, 85, 1965. This is a well organized review in which the sequences discussed may be conveniently compared side by side. Primary references with respect to the sequence work can also be found here.

PROPOSITION IV: ON A PROBLEM IN PALEOGENETICS

Proposition: Consider a polynucleotide which contains L individual nucleotides. Let exactly X mutagenic events* occur randomly along the length of this polynucleotide. After the X mutagenic events have occurred, in general a number x, which is less than L, nucleotide sites will have been hit;** for example, all X mutagenic events could conceivably have occurred at the same nucleotide site. Let N(x) designate the average number of nucleotide sites which have been hit. A method is proposed by which N(x) can be obtained.

P IVA: Introduction

In light of increasing knowledge about the primary amino acid sequences of homologous proteins,*** both within and among various phylogenetic species, it has seemed plausible to attempt to correlate the amino acid differences between two such proteins with a time of origin, measured from the present, of a "common ancestor" protein which is homologous to both. In favorable cases the primary amino acid sequence of portions of the common ancestor protein can be deduced. Proceeding in this way one can build up a biochemical "tree of life" which can be compared to those evolutionary and phylogenetic relationships that are already known from classical biology. The name paleogenetics or paleobio-

---

*A mutagenic event is defined as one step change of one nucleotide to a different nucleotide: i.e., C → T = (C → T) is one mutagenic event; C → T = (C → T → C → T) is three mutagenic events; C → T (at one nucleotide site) and A → G (at another nucleotide site) are together two mutagenic events; etc. C, T, A, G are abbreviations for cytidine, thymidine, adenosine, and guanosine respectively.

**A nucleotide site is said to have been hit each time that a mutagenic event has occurred at that site.

***Homologous proteins are proteins that can formally be considered to be derived from each other by point mutations in the deoxyribonucleic acids coding for those proteins.

chemistry has been suggested (4, 5, 6) for that branch of science which concerns itself with molecular restoration (4) studies of the above type. Among the proteins for which such paleogenetical studies have been made one may mention the hemoglobins of a great many species (4, 5, 6), the cytochromes of various organisms (3), and the A and B fibrinopeptides of the artiodactyls, of rabbits, and of humans (1). Infrequently the relationships so revealed differ radically (1) from those that have been deduced from a great body of classical biological evidence. Until many more homologous proteins have had their primary structure established than are at present known, such anomalies must remain unresolved.

Amino acid difference among proteins are at present thought to arise from mutational changes occurring in the nucleic acid segments which code for these proteins. The number of amino acids which have mutated in the proteins may differ from the number of mutations which have affected the nucleic acid segment for several reasons: 1) Viability--if a particular nucleotide mutation leads to a nonviable organism one will not be able to observe this mutation as an amino acid substitution; 2) Multiple hits at the same site--i.e., several mutations occur at a single nucleotide position in the segment instead of each mutation altering a different nucleotide along the segment; 3) Back mutation--i.e., a multiply hit single nucleotide site may end up as the same nucleotide as it was originally; 4) If each amino acid in the protein is coded by a triplet of nucleotides, then several of the mutations may fall within the same triplet--this would give rise to only a single amino acid substitution; 5) Degeneracy--some amino acids may be coded for by more than one nucleotide triplet. Further, even though the rate of mutation

may be accurately known over a certain region of space, what one must frequently examine is a particular subregion of this space; for example, if the mutation rate along a chromosome were known, the mutations themselves would show up as changes in the amino acid compositions of many nonhomologous proteins, only one, or even only a part of one of which is at hand for study.

From the considerations of the preceding paragraph it is clear that no simple relationship exists between mutagenic events and observed protein mutations. As a consequence, a detailed understanding of paleogenetical studies will require the quantitative evaluation of each of the above factors so that their relative importance can be assessed. It is the purpose of this proposition to suggest one method by which the problem of multiple hits at the same nucleotide site can be handled in a quantitative manner. This problem has been explicitly formulated at the top of p. 249.

P IVB: Calculation of N(x)

Let us make the following definitions:

> An x-part partition of X is a decomposition of X into a set of x (nonzero) positive integer summands $\{a_1,\cdots,a_x\}$, where $\sum_i a_i = X$. Partitions having the same a_i are considered to be identical even though the order of the a_i may differ in two such partitions. Let a particular x-part partition of X be denoted by $(x,X)_j$, and let $n_{a_{ij}}(x)$ be the number of integers in this partition having the value a_i. Note that $\sum_{a_i \neq a_k} n_{a_{ij}}(x) = x$.*

---

*To make these abstract definitions more concrete, consider the following example. A particular 3-part partition of 6 is the set of

> Define N_{jx} as the number of ways of realizing $(x,X)_j$ along a polynucleotide which contains L individual nucleotides.

The definition of N_{jx} requires that we associate the partitions $(x,X)_j$ in some way with the polynucleotide of length L. We do this as follows: The partition (x,X) <u>means</u> that x nucleotide sites have been hit a total of X times; and the particular x-part partition of X $(x,X)_j = \{a_1, a_2, \cdots, a_x\}$ <u>means</u> that the first* nucleotide site has been hit a total of a_1 times, the second site a_2 times, and the xth site a_x times. Now the first nucleotide site can be hit a_1 times in a number of ways: *for example if* $X = 30$ *and* $a_1 = 3$, *the first, second, and third* mutagenic events** could occur at the first site, or the second, fifth, and twenty-seventh mutagenic events could occur there. Similar considerations hold for the other nucleotide sites. The total number of ways in which the first site can be hit a_1 times, the second site a_2 times, the xth site a_x times, and the $x+1^{th}$, $x+2^{th}$, ..., and L^{th} sites zero times is <u>by definition</u> N_{jx}.

---

integers $\{4,1,1\}$. Here, $a_1 = 4$, $a_1 = 1$, and $a_3 = 1$. We denote this particular partition by $(3,6)_1$, where the subscript $j = 1$ is to remind us this partition refers specifically to the set of integers $\{4,1,1\}$. For $(3,6)_1 = \{4,1,1\}$, $n_{a_{11}} = 1$, $n_{a_{21}} = 2$, and $n_{a_{31}} = 2$. As stated in the definition, $a_1 + a_2 + a_3 = 4 + 1 + 1 = 6$; and $n_{a_{11}} + n_{a_{21}} = 3$. A different 3-part partition of 6 would be the set of numbers $\{3,2,1\}$ and this partition could be labeled $(3,6)_2$, for example.

*The nucleotides in the polynucleotide can be numbered in any convenient manner. For example, the 3'-terminal nucleotide could be taken as the first nucleotide and the 5'-terminal nucleotide as the Lth nucleotide.

**The mutagenic events are numbered in any convenient manner.

Now the average number N(x) of polynucleotide sites that have been hit is by definition

$$N(x) = \sum_{x} xP(x) \qquad \text{(Eq. 1)}$$

where P(x) is the probability that <u>exactly</u> x sites have been hit. But P(x) is by definition

$$P(x) = \frac{\sum_{j} N_{jx}}{\sum_{x \leq L} \sum_{j} N_{jx}} \equiv \frac{\sum_{j} N_{jx}}{L^{X}} . \qquad \text{(Eq. 2)}$$

The denominators in Eq. 2 are the total number of ways X mutagenic events can hit L nucleotide sites. Thus if we can find an expression for N_{jx}, N(x) will be given explicitly by Eq. 1. This expression for N_{jx} is derived in the following paragraph.

If x nucleotide sites have been hit in a polynucleotide of L nucleotides, then L - x have not been hit. This can happen in

$$W_1 = \frac{L!}{(L - x)!x!} \qquad \text{(Eq. 3)}$$

ways. Now let us limit our consideration to those sites which have been hit at least once. In particular let us assume these sites have been hit in the precise manner defined by the physical meaning attached to $(x,X)_j$. These x sites can be hit in a total of

$$W_2 = X!x! \qquad \text{(Eq. 4)}$$

ways, because X! is the number of ways X mutagenic events can occur along the polynucleotide, and x! is the number of ways the x a_{ij} can be permuted among themselves.* Not all of these W_2 ways represent distinct physical situations, for the a_{ij} hits at the ith site can occur in a_{ij}! ways and each of these ways leads to the same physical result. Similarly if in the partition there are $n_{a_{ij}}$ integers having the value a_{ij}, these integers can be permuted among themselves in $n_{a_{ij}}$! ways without altering the physical result. The total number of ways x sites can be hit by X mutagenic events is thus

$$W_3 = \frac{W_2}{\prod\limits_{a_{ij} \neq a_{kj}} n_{a_{ij}}! \prod\limits_{a_{ij}} a_{ij}!} \qquad \text{(Eq. 5)}$$

Therefore $$N_{jx} = W_1 W_3 = \left[\frac{X!}{\prod\limits_{a_{ij} \neq a_{kj}} n_{a_{ij}}! \prod\limits_{a_{ij}} a_{ij}!} \cdot \frac{L!}{(L-x)!} \right]_j . \qquad \text{(Eq. 6)}^{**}$$

This completes the solution to our problem.

The method that has been given above for calculating N(x) in terms of partitions illuminates the physical details of the mutation process.

---

*The factor x! arises from the fact that in the <u>definition</u> (cf. p. 251) of $(x,X)_j$ the order of the a_{ij} in the partition was irrelevant while the <u>physical meaning</u> attached to $(x,X)_j$ (cf. p. 252) was such that identical partitions in which the a_{ij} occur in different orders refer to different physical situations.

**Ref. 8 lists 0! - 9999! explicitly, so that it is not necessary to calculate these factorials.

However, writing out the partitions that are needed in this method is frequently tedious,* and it clearly would be desirable to have a formula for $N(x)$ that does not require the calculational labor of Eq. 6. Such a formula can be obtained as follows: Define $m(X,x)$ to be the number of ways X mutagenic events can hit x nucleotide sites where each site is hit <u>at least</u> once.** Thus $\sum_j N_{jx}$ is given by

$$\sum_j N_{jx} - m(X,x)W_1 \qquad \text{(Eq. 7)}$$

where W_1 is given by Eq. 3. We now calculate $m(X,x)$. The total number of ways X mutagenic events can hit x nucleotide sites is x^X. $m(X,x)$ is therefore given by x^X less those number of ways in which X mutagenic events can hit x nucleotide sites when k sites are not hit at all,

---

*This is if anything, an understatement. The basic difficulty is that no general formula exists for $n_{a_{ij}}$. In some applications, the calculation of the probability for back mutation, for example, the individual a_{ij} and $n_{a_{ij}}$ of each partition must be known, and there is no way to get them except to write down the partitions one by one in some systematic manner that insures against leaving any partition out. In this respect Ref. 7 is very helpful, for it lists the total number $P(x,X)$ of each possible (x,X) as well as the total number $p(X,X) = \sum_{x=1}^{X} P(x,X)$ of all possible partitions for a given X. Fortunately, $N(x)$ can be calculated by means of the dodge that is described in the text without the necessity of having to write out any of the partitions.

**Mathematically, the number we have designated by $m(X,x)$ is the number of mappings of X <u>onto</u> x.

where k takes on successively the values 0, 1, 2, $\cdots$, x - 1.* But as in Eq. 3 *the number of ways in which k sites can be hit, and x - k sites* not hit is

$$W_4 = \frac{x!}{k!(x-k)!}$$

Those k sites which have been hit can be hit in a total of m(X,k) ways by definition. The total number of ways in which the X mutagenic events can hit x sites when some of the sites are not hit at all is thus

$$W_5 = \sum_{k=0}^{x-1} \frac{x!}{k!(x-k)!} m(X,k) \qquad \text{(Eq. 8)}$$

Therefore $$m(X,x) = x^X - W_5 = x^X - \sum_{k=0}^{x-1} \frac{x!}{k!(x-k)!} m(X,k) \qquad \text{(Eq. 9)}$$

N(x) can now be calculated from equations 9, 7, 2, and 1. It should be noted that because of Eq. 9 no knowledge whatsoever about partitions is required in calculating N(x).

Finally we notice that Eq. 9 is a recursion formula for m(X,x); i.e., starting from m(X,0) = 0, all the other m(X,k) and finally m(X,x)** can be calculated from Eq. 9 alone. This important fact

---

*This follows from the fact that we defined m(X,x) to include only those situations where every one of the x sites is hit at least once. Those situations in which some site or sites are not hit at all must be subtracted out.

**m(X,x) is simply a particular m(X,k) with k = x.

reduces the calculation of N to a simple iterative procedure which can readily be carried out by a computer.*

P IVC: Discussion

In conclusion we note that though the formulas that we have derived for N(x) are exact under the assumptions that we have made, it is not obvious that these assumptions are true although it is likely that they are sufficiently near the truth to be useful. In particular we have assumed that any given time t, the number of mutagenic events X that occurred along the polynucleotide is a fixed number that is either known exactly or that can be estimated. Among other possible assumptions one could consider the case where the mutagenic events occur randomly in time so that at any given time t the number of mutagenic events that have occurred along the polynucleotide could vary from zero to some finite (or infinite) number. Such an assumption would possibly lead to somewhat simpler formulas for N(x) because less information about X would be required. In principle these formulas could be derived from Eq. 1 by averaging N(x) over X provided that the probability distribution of X was known or assumed. However, because of the fairly complicated form of Eq. 9, it would probably be simpler to derive possible alternative formulas for N(x) from scratch.

---

*The calculation can also of course be made by hand, and in point of fact such a manual calculation is not prohibitively difficult if the factorial tables of Ref. 8 are used.

BIBLIOGRAPHY for PROPOSITION IV

1. R. F. Doolittle and R. Blombaeck, Nature, 202, #4928, 147, 1964.

2. Richard V. Eck, Science, 140, 447, 1963.

3. E. Margoliash, Can J. Biochem., 42, #5, 745, 1964.

4. L. Pauling, and Emile Zuckerkandl, Acta Chem. Scan., 17, 89, 1963.

5. E. Zuckerkandl and L. Pauling, Horizons in Biochemistry. M. Kasha and B. Pullman, eds.; Academic Press, New York, 1962, p. 189.

6. E. Zuckerkandl, Scientific American, May, 1965, p. 110.

7. Tables of Partitions, Royal Society Mathematical Tables, Volume 4, University Press, Cambridge, 1958 (see especially Eq. 1.1, p. ix, and Eq. 2.2a, p. xi).

8. Tables of Factorials 0! - 9999!, National Academy of Sciences, National Research Council Publication 1039, Washington, D. C., 1962.

PROPOSITION V: ON NONSPECIFIC CLEAVAGE DURING THE EDMAN DEGRADATION

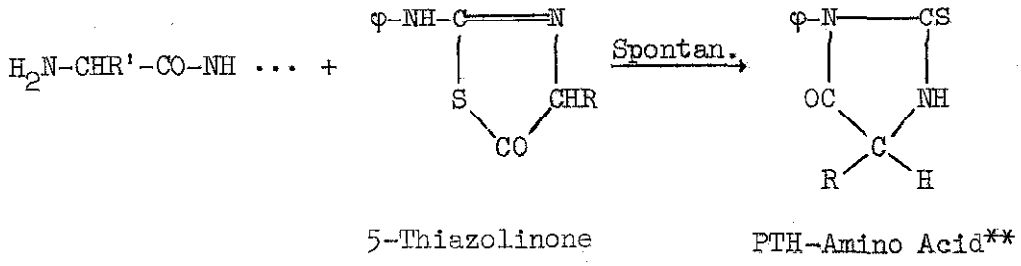
Proposition: In the degradation of peptides by the Edman method with direct identification of the N-terminal PTH-amino acid, if "histidine is (or becomes) N-terminal, the histidyl residue is removed during reaction with phenyl isothiocyanate and the succeeding residue then reacts also so that after cyclization, both residues are detected on the chromatograms" (5). The identification of N-terminal histidine cannot therefore be made with certainty because of the interference by the amino acid which follows histidine in the peptide chain. An explanation is proposed for this behavior, and methods are suggested by which the ambiguity in the identification of N-terminal histidine can be removed.

P VA: Introduction*

The primary amino acid sequence of a peptide or protein is frequently determined by the Edman degradation (3) for the sequential removal of amino acid residues at the N-terminus of the peptide or protein by the consecutive reactions

$$\varphi NCS + H_2NCHR\text{-}CO\text{-}NH\text{-}CHR'\text{-}CO\text{-}NH \cdots \xrightarrow{OH^-} \qquad 1)$$

$$\varphi NH\text{-}CS\text{-}NH\text{-}CHR\text{-}CO\text{-}NH\text{-}CHR'\text{-}CO\text{-}NH \cdots \xrightarrow{H^+} \qquad 2)$$



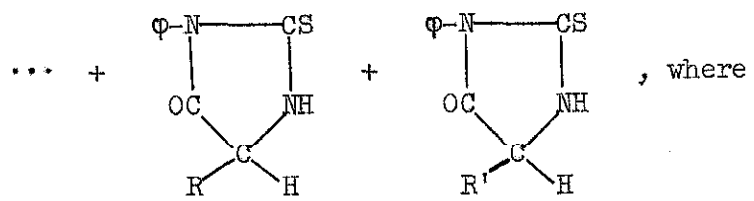
---

*Some of the introductory material has been taken more or less directly from Schroeder (5). The detailed experimental procedures for the Edman degradation can also be found in that reference.

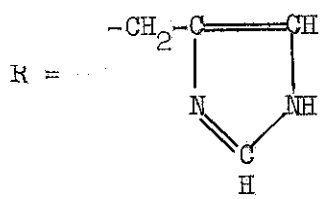
**PTH=phenylthiohydantoin.

The PTH-amino acid which results from this sequence of reactions is then identified by paper chromatography. These reactions proceed smoothly in all cases excepting those in which the N-terminal amino acid is histidine. In the latter case the following reactions occur:

$$2\varphi NCS + H_2NCHR\text{-}CO\text{-}NH\text{-}CHR'\text{-}CO\text{-}NH \cdots \xrightarrow[2)\ H^+]{1)\ OH^-} \qquad 3)$$



, where



,

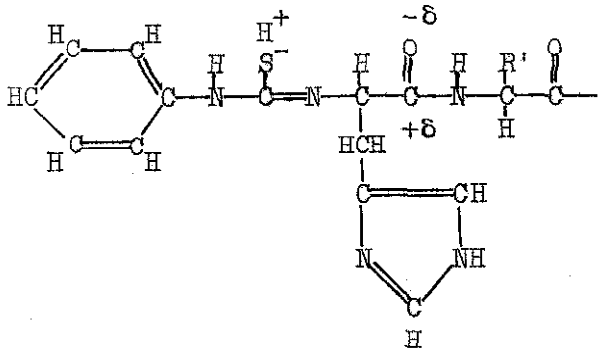
and R' may be the side chain of any amino acid. In the absence of other information it is impossible to tell which of the two PTH-amino acids from Reactions 3) arose from the N-terminus.

Schroeder (5) has suggested that the abnormal behavior that has been summarized in Reactions 3) is caused by some unexplained deterioration of the phenyl isothiocyanate over a period of time. We believe, however, that the cause of this abnormal behavior is inherent in the peptide structure, and that though such deterioration may contribute to this abnormal behavior, this contribution is more in the nature of aggravating an already existing situation rather than in the nature of a fundamental cause.

P VB: Proposed Explanation for Reactions 3)

First it is clear that the underlying cause of Reactions 3) must lie in Reaction 1) of p. 259, because Reaction 2) is carried out in the absence of phenyl isothiocyanate and the latter reagent is necessary if the phenylthiohydantoin of the amino acid penultimate to the N-terminus is to be formed. The formation of the PTH-amino acid penultimate to the N-terminus further implies that the N-terminal histidyl residue is cleaved from the peptide at Reaction 1), because in order for phenyl isothiocyanate to react with the amino acid penultimate to the N-terminus, the latter amino acid must have a free alpha amino group.* The question we must answer then is: "When histidine is the N-terminal amino acid of a peptide, why is the peptide bond between this histidine and the adjacent amino acid so readily cleaved?". Having answered this question we can then consider ways to correct the situation.

Let us consider in detail the nature of the adduct formed between phenyl isothiocyanate and the peptide in Reaction 1). The structure of this adduct may be written:



*The fact that the N-terminal histidine is cleaved from the peptide at Reaction 1) does not necessarily imply that cyclization occurs at this step.

Formation of the 5-thiazolinone (cf. p. 259), then occurs by nucleophilic attack of the sulfur atom of this adduct on the +δ carbon atom of the amide bond.* We note the following points: 1) Preferential cleavage of the amide bond between the N-terminal amino acid and the amino acid which is adjacent to it is made possible only by increasing the +δ charge on the amide carbon. 2) In the Edman procedure this increased +δ charge is achieved by adding the electron deficient phenyl isothiocyanate group to the α amino group of the N-terminal amino acid.** 3) The imidazole ring of histidine is also an electron deficient grouping; moreover, the imidazole ring has two electronegative nitrogen atoms both of which can be protonated. Each of these peculiarities of the imidazole ring increases the +δ charge on the amide carbon atom above and beyond the increase caused by the phenyl isothiocyanate grouping.

We postulate that it is the increased +δ charge that is imposed on the amide carbon atom by the peculiarities of the imidazole ring of histidine which is responsible for the abnormal behavior that is summarized in Reactions 3, p. 260.

This postulate explains why only N-terminal histidine results in the observed abnormal behavior. First why is histidine the only amino acid which causes such behavior? Because with the exception of tyrosine

---

*For amino acids other than N-terminal histidine the cleavage of the N-terminal residue (to form the 5-thiazolinone) from the remainder of the peptide does not occur until acidification as in Reaction 2): under acidic conditions the amide nitrogen atom is protonated and this fact makes it a good leaving group.

**There are other ways of achieving the same result: e.g., in Sanger's end-group method (cf. p. 73 this thesis), the electron deficient 2,4-dinitrophenyl group is added to the α amino group.

and phenylalanine, histidine is the only electron deficient (in the sense of being aromatic) amino acid. The fact that histidine does result in abnormal behavior while phenylalanine and tyrosine do not is explained by the presence in the imidazole ring of histidine of two electronegative nitrogen atoms. Such electronegative atoms are not present in the aromatic ring of phenylalanine or tyrosine. Further, the protonation of the imidazole ring would further increase the +δ charge on the amide carbon. Such protonation of the phenyl ring in phenylalanine and tyrosine is not possible. Second, why must the histidine be present at the N-terminus to cause the abnormal reactions? Because in the N-terminal position the imidazole ring of histidine is only three bonds removed from the amide carbon; in any other position the imidazole ring is at least four bonds removed.

P VC: Suggestions for Reducing the Influence of the Imidazole Ring

From the considerations of the preceding section it follows that there are basically only two ways by which to reduce the lability of amide bond between N-terminal histidine and the adjacent amino acid: first, to destabilize the +δ-δ dipole of the amide carbon atom; and second, to destroy the aromatic electronegative character of the imidazole ring. In the following two paragraphs each of these possibilities is considered.

The effect of the +δ-δ dipole of the amide carbon can be diminished in two ways: by reducing the polarity of the environment in Reaction 1), and by reducing the pH at which Reaction 1) is carried out. As described by Schroeder (5), Reaction 1) is carried out as follows. The peptide is applied as an aqueous solution to a paper strip. The strip

is air dried and then is saturated with phenyl isothiocyanate in dioxane. The basic pH that Reaction 1) requires is provided by suspending the paper strip above a mixture of pyridine-dioxane-water. The polarity of the reaction environment could be decreased in several ways: e.g., by drying the paper strips over P_2O_5 after the sample has been applied, but before the phenyl isothiocyanate is applied to the strip; or by applying the phenyl isothiocyanate to the strip in benzene rather than in dioxane; and by replacing the pyridine-dioxane-water system with a pyridine-benzene system. Because of the known sensitivity of the N-terminal amide bond to its environment* it might be further advisable to dry** the benzene and pyridine prior to use and perhaps it would help to use saturated hydrocarbon solvents rather than benzene, because unsaturated compounds are inherently polar in nature. It might also be preferable to replace pyridine with a saturated amine, because the latter would be less polar. Conceivably all of the above changes might be necessary to reduce the effect of the $+\delta-\delta$ dipole. Finally reducing the pH of the reaction environment would reduce the concentration of nucleophilic anions (such as the thiol anion of the adduct that has been described on p. 261) that might attack the $+\delta$ carbon atom of the amide bond and cause cleavage.*** The reduction in pH might be achieved by

---

*This sensitivity is demonstrated by the fact that when histidine is N-terminal Reactions 3) sometimes do, and sometimes do not occur even when the experimental conditions that are described in Ref. 5 have not obviously been altered. (W. A. Schroeder private communication.)

**To remove the highly polar water molecule.

***The pH must be reduced at the paper strip, not just in the liquid mixture above which the strip is suspended. Also the pH should probably not be taken below pH 7.5 for then the protonation of the imidazole ring might offset the advantages of the reduced nucleophile concentration.

reducing the pyridine concentration, or by replacing pyridine by a less basic amine, or by the addition to the liquid above which the paper strip is suspended of an anion such as cyanide.* Also, in addition to the above precautions, it is clear that the temperature at which Reaction 1) is carried out should be as low as is possible consistent with the success of the procedure.

The second possibility to reduce the influence of the imidazole ring is simply to destroy the aromatic character of that ring prior to reacting the peptide with phenyl isothiocyanate. One way to accomplish this destruction might be by a suitable modification of the Bamberger fission (1, 2, 4); however, there are several aspects of this fission which might require rather thorough study before the method could be applied to proteins and peptides with success.

In conclusion, of the two possibilities described above, the first, i.e., *a destabilization of the +δ-δ dipole of the amide carbon atom,* appears to require the least experimental modifications with respect to the normal (5) procedure. Moreover these modifications are relatively simple, requiring in the main no more than maintaining anhydrous conditions and replacing one solvent by another.

---

*Below pH 9 any cyanide ion that has been added would be present essentially as HCN. The latter would be present in the vapor above the liquid--i.e., at the surface of the paper strip.

BIBLIOGRAPHY for PROPOSITION V

1. J. N. Ashley and C. R. Harrington, J. Chem. Soc. 2586, 1930.

2. E. Bamberger, Ann., 273, 267, 1893.

3. P. Edman, Acta Chem. Scand., 4, 283, 1950.

4. H. Heath et al., J. Chem. Soc. 2215, 1951.

5. W. A. Schroeder, "Degradation of Peptides by the Edman Method with Direct Identification of the PTH-Amino Acid," Enzyme Structures, Vol. VII of Methods in Enzymology, Academic Press, in press, 1965.

6. W. A. Schroeder (private communication).

Part VII

BIBLIOGRAPHY

BIBLIOGRAPHY

1. D. W. Allen, et al., J. Am. Chem. Soc., 80, 1628, 1958.

2. E. H. Allen and R. S. Schweet, J. Biol. Chem., 237, 760, 1962.

3. Anonymous, Blood, 17, 125, 1961.

4. M. L. Anson and A. E. Mirsky, J. Gen. Physiol., 13, 469, 1930.

5. Gilbert Ashwell, Ann. Rev. Biochem., 23, 101, 1964.

6. D. R. Babin, et al., Biochim. et Biophys. Acta, 86, 136, 1964.

7. C. Baglioni, Biochim. et Biophys. Acta, 48, 392, 1961.

8. Ibid., 59, 437, 1962.

9. J. C. Bailey, J. Chem. Soc. London, 3461, 1950.

10. H. Banaschak and F. Jung, Biochem. Z., 327, 515, 1956.

11. W. Bartley, Biochem. J., 53, 307, 1953.

12. G. H. Beaven, et al., J. Haemat., 6, 201, 1960.

13. E. A. Beet, Ann. Eugen. London, 14, 279, 1949.

14. Biochemistry of the Amino Acids, 2nd edition. Alton Meister, ed.; Academic Press, New York, 1965, 1, 375-437.

15. R. M. Bock and N. S. Ling, Anal. Chem., 26, 1543, 1954.

16. G. Braunitzer, et al., Z. Physiol. Chem., 325, 283, 1961.

17. Ibid., 331, 1, 1963.

18. E. R. Burka and P. A. Marks, Nature, 204, 659, 1964.

19. R. E. Canfield and C. Anfinsen, J. Biol. Chem., 238, 2684, 1963.

20. Amoz I. Chernoff, Abstracts 10th Congress of the International Society of Haematology, L: 3, Stockholm, 1964.

21. A. I. Chernoff and N. M. Pettit, Jr., Blood, 24, 750, 1964.

22. M. D. Clegg and W. A. Schroeder, J. Am. Chem. Soc., 81, 6065, 1959.

23. E. H. Cordes and W. P. Jencks, J. Am. Chem. Soc., 85, 2843, 1963.

24. C. D. Coryell and L. Pauling, J. Biol. Chem., 132, 1769, 1940.

25. A. M. Crestfield, et al., J. Biol. Chem., 238, 622, 1963.

26. R. Dreywood, Ind. Eng. Chem., Anal. Ed., 18, 499, 1946.

27. M. Dubois, et al., Anal. Chem., 28, 350, 1962.

28. P. Edman, Acta Chem. Scand., 4, 283, 1950.

29. P. Edman and J. Sjöquist, Ibid., 10, 1507, 1956.

30. Carl Faurholt, J. Chim. Phys., 22, 1, 1925.

31. J. K. W. Ferguson and F. J. W. Roughton, J. Physiol., 83, 68, 1934.

32. P. Fessas and D. Loukopoulos, Science, 143, 590, 1964.

33. G. Giustina, et al., Giorn. Biochim., 1, 357, 1952.

34. Ibid., 1, 475, 1952.

35. G. Giustina and R. Margaria, Ibid., 2, 153, 1953.

36. E. R. Glowacki and R. L. Millette, J. Mol. Biol. 11, 116, 1965.

37. J. Goldstein, et al., J. Biol. Chem., 236, PC 77, 1961.

38. Ibid., 238, 2016, 1963.

39. H. M. Goodman and A. Rich, Nature, 199, 318, 1963.

40. S. Granick and R. D. Levere, Prog. Hemat., 4, 1, 1964.

41. F. C. Green and L. M. Kay, Anal. Chem., 24, 726, 1952.

42. G. P. Greenstein and M. Winitz, Chemistry of the Amino Acids. John Wiley and Sons, New York, London, 1961, p. 498.

43. Ibid., Proc. 10-85, 1020, 1961.

44. Ibid., Proc. 10-110, 1061, 1961.

45. Ibid., Proc. 10-167, 1245, 1961.

46. Ibid., Proc. 21-5, 1831, 1961.

47. Ibid., Proc. 27-7, 1989, 1961.

48. G. Guidotti, et al., Proc. Nat. Acad. Sci. U.S.A., 50, 774, 1963.

49. P. B. Hamilton, Anal. Chem., 30, 914, 1958.

50. G. Hardesty, et al., Proc. Nat. Acad. Sci., U.S.A., 50, 1924, 1963.

51. Dorothea Heyl, et al., J. Am. Chem. Soc., 70, 3429, 1948.

52. J. D. Heywood, et al., J. Clin. Invest., 43, 2368, 1964.

53. R. J. Hill, et al., J. Biol. Chem., 237, 1549, 1962.

54. R. L. Hill, et al., J. Biol. Chem., 235, 3182, 1960.

55. C. H. W. Hirs, J. Biol. Chem., 219, 611, 1956.

56. W. R. Holmquist, unpublished data, 1962.

57. W. R. Holmquist and W. A. Schroeder, Biochim. et Biophys. Acta, 82, 639, 1964.

58. W. R. Holmquist and J. R. Vinograd, Biochim. et Biophys. Acta, 69, 337, 1963.

59. B. F. Horton and T. H. J. Huisman, "Studies on the Heterogeneity of Haemoglobin, VII: Minor Haemoglobin Components in Haematological Diseases," British J. Haemat., in press, 1965.

60. B. F. Horton and T. H. J. Huisman, Amer. J. Human Gen., 15, 394, 1963.

61. E. R. Huehns and E. M. Shooter, J. Med. Gen., 2, No. 1, 48, 1965.

62. T. H. J. Huisman, et al., J. Lab. Clin. Med., 52, 312, 1958.

63. T. H. J. Huisman, Nature, 188, 589, 1960.

64. T. H. J. Huisman and C. A. Meyering, Clin. Chim. Acta, 5, 103, 1960.

65. T. H. J. Huisman and A. M. Dozy, J. Chromatography, 7, 180, 1962.

66. T. H. J. Huisman, Adv. Clin. Chem., 6, 231, 1963.

67. T. H. J. Huisman, "Studies on the Heterogeneity of Hemoglobin, VIII: Chromatographic and Electrophoretic Investigations of Various Minor Hemoglobin Fractions Present in Normal and In Vitro Modified Red Blood Cell Hemolysates," J. Chromatography, in press, 1965.

68. J. A. Hunt and V. M. Ingram, Nature, 181, 1062, 1958.

69. J. A. Hunt and V. M. Ingram, Biochim. et Biophys. Acta, 42, 409, 1960.

70. Ibid., 49, 520, 1961.

71. W. D. Hutchinson, unpublished results.

72. V. Ingram, Biochim. et Biophys. Acta, 28, 539, 1958.

73. V. Ingram, Hemoglobin and Its Abnormalities. Charles C. Thomas, publisher, Springfield, Illinois, 1961.

74. J. H. Jandl, et al., J. Clin. Invest., 39, 1818, 1960.

75. R. T. Jones, unpublished results: the molecular weights were determined with an ultracentrifuge.

76. R. T. Jones, "Chromatographic and Chemical Studies of Some Abnormal Human Hemoglobins and Some Minor Hemoglobin Components," Ph.D. Thesis, Calif. Inst. of Tech., Pasadena, 1961.

77. R. T. Jones and W. A. Schroeder, Biochemistry, 2, 1357, 1963.

78. R. T. Jones and W. A. Schroeder, J. Chromatography, 10, 421, 1963.

79. J. H. Jonxis and J. F. Delafresnaye, eds., Abnormal Hemoglobins, a Symposium. Charles C. Thomas, publisher, Springfield, Illinois, 1959.

80. W. Konigsberg and R. J. Hill, J. Biol. Chem., 237, 2547, 1962.

81. W. Konigsberg, et al., ibid., 238, 2028, 1963.

82. E. Korber, Dissert. Dorpat, 1866, cited by H. Bischoff, Z. Ges. Exp. Med., 48, 472, 1926.

83. H. G. Kunkel and A. G. Bearn, Fed. Proc., 16, 760, 1957.

84. H. Kunkel and G. Wallenius, Science, 122, 288, 1955.

85. Dominique Labie, "L'Hemoglobine A_4," Doctoral Thesis, Paris, 1964.

86. G. Laurent, et al., Bull. Soc. Chim. Biol., 44, 419, 1962.

87. A. L. Levy and C. H. Li, J. Biol. Chem., 213, 487, 1955.

88. R. C. Leif, The use of pyrogallol was suggested by Dr. R. C. Leif.

89. J. B. Lingrel and H. Borsook, Biochemistry, 2, 309, 1963.

90. Julio Ludoweig and A. Dorfman, Biochim. et Biophys. Acta, 38, 212, 1960.

91. R. Margari, Schweiz. Med. Woch., 82, 990, 1952.

92. G. Matsuda, et al., Blood, 16, 984, 1960.

93. C. A. Meyering, et al., Clin. Chim. Acta, 5, 208, 1960.

94. A. I. Meyers and Willette Y. Libano, J. Org. Chem., 26, 4399, 1961.

95. Methods in Carbohydrate Chemistry, Roy L. Whistler and M. L. Wolfrom, eds. Academic Press, New York and London, 1952, 1, pp. 390, 479.

96. D. E. Metzler and E. E. Snell, J. Am. Chem. Soc., 74, 979, 1952.

97. D. E. Metzler, J. Am. Chem. Soc., 79, 485, 1957.

98. E. Milla, et al., Giorn. Biochim. Z., 2, 80, 1953.

99. S. Moore and W. H. Stein, J. Biol. Chem., 192, 663, 1951.

100. S. Moore, et al., Anal. Chem., 30, 1185, 1958.

101. S. Moore and W. H. Stein, Biochemistry, 3, 180, 1964.

102. M. Morrison and J. L. Cook, Science, 122, 920, 1955.

103. A. G. Motulsky, Amer. J. Trop. Med. Hyg., 13, 147, 1964.

104. C. J. Muller, "A Comparative Study on the Structure of Mammalian and Avian Hemoglobins," Doctoral Thesis, Univ. Groningen, 1961, Van Gorcumand Co. N.V. Assen Netherlands.

105. M. Murayama, Nature, 194, 933, 1962.

106. Ibid., 202, 258, 1962.

107. J. L. Naiman and P. S. Gerald, J. Lab. Clin. Med., 61, 508, 1963.

108. George Nakamura, Anal. Chem., 24, 1372, 1952.

109. C. M. Nechtman and T. H. J. Huisman, Clin. Chim. Acta, 10, 165, 1964.

110. R. Neher, Steroid Chromatography, second edition. Elsevier Publishing Co., Inc., New York, 1964, pp. 129, 135.

111. Organic Chemistry, Henry Gilman, ed. John Wiley and Sons, New York, 1943, Vol. II, pp. 1096-1097.

112. Linus Pauling, et al., Science, 110, 543, 1949.

113. Paper Chromatography, I. M. Hais and K. Macek, eds. Academic Press, New York, 1963, pp. 368, 799.

114. Georg R. Philipps, Nature, 205, 567, 1965.

115. Cornelius Pik, "Studies Over De Synthese Van α, β, en γ Ketens In Reticulocytenrijk Bloed Met Behulp Van Gemerkte Aminozuren," Doctoral Dissertation, Rijkuniversiteet Te Gronigen, 1965.

116. Earl R. Rainville, A Short Course in Differential Equations. The Macmillan Co., New York, 1949, Chapter 7.

117. H. M. Ranney, J. Clin. Invest., 33, 1634, 1954.

118. H. M. Ranney and P. Kono, J. Clin. Invest., 38, 508, 1959.

119. H. S. Rhinesmith, et al., J. Am. Chem. Soc., 79, 609, 1957.

120. Ibid., 79, 4682, 1957.

121. Ibid., 80, 3358, 1958.

121a. Ronald Rieder and David J. Weatherall, J. Clin. Invest., 44, No. 1, 42, 1965.

122. J. Rosa and D. Labie, Nature, 196, 901, 1962.

123. D. L. Rucknagel and J. V. Neel, Prog. Med. Gen., 1, 158, 1961.

124. F. Sanger, Biochem. J., 39, 507, 1945.

125. A. G. Schnek and W. A. Schroeder, J. Am. Chem. Soc., 83, 1472, 1961.

126. W. A. Schroeder, et al., Proc. Nat. Acad. Sci. U.S.A., 47, 811, 1961.

127. W. A. Schroeder, et al., Anal. Chem., 34, 1570, 1962.

128. W. A. Schroeder, et al., Biochim. et Biophys. Acta, 63, 532, 1962.

129. W. A. Schroeder, et al., Biochemistry, 2, 992, 1963.

130. W. A. Schroeder, "Degradation of Peptides by the Edman Method with Direct Identification of the PTH-Amino Acid," Enzyme Structures, Vol. VII of Methods in Enzymology, Academic Press, in press, 1965.

131. W. A. Schroeder and R. T. Jones, a review: "Some Aspects of the Chemistry and Function of Human and Animal Hemoglobins," Fort. Chem. Org. Natur., XXIII, pp. 85-160, 1965.

132. W. A. Schroeder and Barbara Robberson, Anal. Chem., submitted, 1965.

133. R. Schwyzer, et al., Helv. Chim. Acta, 41, 1273, 1958.

134. T. A. Scott and E. H. Melvin, Anal. Chem., 25, 1650, 1953.

135. E. Silvestroni and I. Bianco, Blood, 22, 545, 1963.

136. Wm. C. Stadie and H. O'Brien, J. Biol. Chem., 117, 439, 1937.

137. George R. Stark and Derek G. Smyth, J. Biol. Chem., 238, 215, 1963.

138. S. Swarup, et al., Bull. Calcutta School Trop. Med., 12, No. 4, 155, 1964.

139. Ibid., 13, No. 1, 7, 1965.

140. Synthetic Methods of Organic Chemistry, W. Theilheimer, ed. S. Karges, Basel, publisher, 18, 841, 1964.

141. J. Vinograd, et al., J. Am. Chem. Soc., 81, 3168, 1959.

142. F. P. Wettstein, et al., Nature, 197, 430, 1963.

143. K. W. Winterhalter, Pathol. Microbiol., 27, 508, 1964.

144. H. G. Wittman and G. Braunitzer, Virology, 9, 726, 1959.

145. E. Zuckerkandl, Sci. Amer., May, p. 110, 1965.