SELECTIVE FISSION OF PROTEINS.

by John Alexander Black.

Cyanogen bromide is known to be a chemical agent which causes selective fission at methionyl bonds in proteins under mild acid conditions and at room temperature. Its action on three proteins, horse-heart cytochrome c, horse-heart myoglobin and bovine serum albumin, is studied by separation of the reaction products by means of gel filtration, followed by their amino acid analysis, and in the case of bovine serum albumin, tryptic digestion and fingerprinting.

Cytochrome c, which contains two methionine residues, yielded three peptides, two of which correspond to the fragments expected from the published sequence and are formed by fission of the peptide chain at both methionine residues. The other appears to result fr fission at one methionine locus, that nearest the C-terminus of the peptide chain. A fourth peptide should be present in the reaction products as a result of complete fission and although evidence is given for its existence, it has not been isolated.

Myoglobin, also containing two methionine residues, gave four peptides, three of which account for all the amino acid residues in myoglobin and are therefore the products of complete fission. The fourth peptide is accounted for by non-cleavage of the methionyl bond joining two of the previous three peptides in the intact protein. N-terminal analyses of the four peptides allow the three peptides obtained by complete fission to be placed in the order of their appearance in the intact protein.

From its methionine content of four, bovine serum albumin was expected to give five fragments. Initial separations on Sephadex G 75 indicated the presence of this number of components in the reaction mixture. However improved separations on coupled columns of Sephadex G 200 and G 75 suggest that fifteen to twenty components are present, indicating incomplete cleavage at each of the four methionine residues of albumin. From the latter separation, three fractions have been isolated which account for the amino acid composition and tryptic peptides of the complete protein. Two of the fractions appear to contain pure peptides of calculated molecular weights 20,000 and 15,000. From the amino acid compositi tryptic peptides and gel filtration behaviour of the third fraction, it appears to contain three peptides of approximate molecular weight 10,000. Cyanogen bromide treatment of partially reduced albumin gave a slightly simpler mixture of products from which a new peptide of calculated molecular weight, 12,000, was isolated. Fingerprinti results suggest that it is one of the peptides present in the previo fraction containing three components. Evidence has also been obtai from fingerprinting for the existence of two of these components joined together in a single polypeptide chain.

Selective fission of proteins.

by

John Alexander Black.

Thesis presented for the degree of Doctor of Philosophy, The University of Glasgow.

June 1964.

Acknowledgments.

I wish to thank Professor J.N. Davidson, F.R.S. for providing the facilities for this research and to express my gratitude to Dr. G. Leaf for his advice and guidance as a supervisor.

I also wish to acknowledge helpful discussions on fingerprinting with Dr. B. Cohen of the Department of Genetics, who kindly loaned an apparatus for most of the separations described herein. I am indebted to Mr. D. Duthie for expert technical assistance and to Mrs. J. Nash for typing.

I finally wish to acknowledge the receipt of a N.A.T.O. Research Studentship and an Honorary D.S.I.R. Research Studentship during the three years of this work.

Contents.

INTRODUCTION.

THE ROLE OF PROTEINS.

The important role of proteins in biochemical phenomena has been acknowledged since the initial investigations on the materials such as egg white, blood serum and milk curd at the beginning of the nineteenth century. The term protein itself was first used in 1838 by Mulder to describe a nitrogen containing substance, present in both plants and animals, which served for the regulation of chemical metabolism and without which it appeared that life would not be possible. Subsequent investigations by several workers on the analyses of albumen, fibrine etc., led Liebig in 1847 to reject the radical concept proposed by Mulder in which animal and vegetable albumens, fibrines and caseines were all considered to be composed of one basic unit. Since that time the use of the word protein has been extended to embrace the class of substances which have been shown to be responsible for the properties noted by the original workers.

The essential role of proteins in the structural and dynamic aspects of living matter which was a tentative hypothesis in 1838 has now become an established fact. Proteins are unique among the components of living matter in the diversity of their functions in the living cell and organism.

Some proteins such as keratin which is a major component

of skin, hair and horns appear to have a purely structural role. Other structural proteins form the cell membrane and are responsible for the internal organisation of the cell. Some of the hormones, the regulators of the delicate balance in the metabolic activities of the intact animal or plant, are protein in nature. Insulin which is one of the controlling factors in the balance of carbohydrate metabolism in some mammals is the best known example. Most of the chemical reactions which take place in living cells with a delicacy that the organic chemist is seldom able to emulate, are catalysed by enzymes which are either entirely protein in character or composed of a protein-non protein prosthetic group complex. Proteins act as oxygen carriers in the blood and electron carriers in the cell. Nucleic acids in the cell nucleus may be associated with proteins as nucleo-proteins. Most of the toxins produced by bacteria are protein in nature. Protein-protein and protein-polysaccharide interactions are the basis of the mechanisms of immunity and allergy, and protein-carotenoid interactions are evident in the visual process.

Not only is there a wide diversity of function, but within each function there is generally a high specificity of action. The enzymes for example are specific for one type of chemical reaction and in some cases the

specificity is increased to an absolute requirement for a particular substrate. The existence of each living cell is therefore the result of the combined functions of a multitude of proteins each unique in its contribution. The unique role of each protein must be a consequence of its unique composition and structure.

THE STRUCTURE OF PROTEINS.

The discovery of amino acids as products of protein hydrolysis beginning with the isolation of glycine in 1820, led up to the independent proposals of Emil Fischer and Franz Hofmeister in 1902 that proteins are formed from the constituent amino acids joined in a linear manner through peptide bonds in which the α carboxyl group of one amino acid is condensed with the α amino group of the adjoining amino acid. In the light of the supporting evidence which has since been accumulated the peptide hypothesis is now assumed to be generally valid. The multiplicity of protein structures arises from variations in the total number of amino acids, from differing proportions of the twenty amino acids normally found in proteins, and from variations in the linear arrangement of these amino acids. This linear arrangement has been termed the primary structure of a protein (Linderstrøm -Lang, 1952).

In order to define such complicated macromolecules in descriptive terms three additional structural levels has been assigned to proteins - secondary, tertiary (Linderstrøm-Lang, 1952) and quaternary (Bernal, 1958).

The theoretical flexibility of the peptide chain is found to be restricted in most native proteins. This is caused by regular folding of the chain to allow hydrogen

bonding between the carbonyl and imide groups of the backbone and so forming the secondary structure of the protein. The hydrogen bonding can be intermolecular (Pauling and Corey, 1951) in which case a pleated sheet structure arises or intramolecular (Pauling, Corey and Branson, 1951) in which case the resulting configuration is an α helix. The majority of scleroproteins of the β keratin type appear to exist in the pleated sheet form, whereas the α helix is assumed to be the predominant form in the more compact globular proteins.

Further folding of the peptide chains and segments of α helices in three dimensional space gives rise to the complex, rigid tertiary structure. The maintenance of the tertiary structure is due to several different types of chemical interaction between parts of the folded protein chain.

Quaternary structure was proposed to define the degree of polymerisation of tertiary units.

Primary structure.

Of the four levels of protein structure only primary yields completely to chemical methods of investigation. The present era of primary structural determination was opened by the elucidation of the structure of insulin by Hyle, Sanger, Smith and Kitai in 1955. The approach and

techniques used by Sanger and his co-workers have been the model for all subsequent work. Partial hydrolysis of the intact protein gives small peptides in which the amino acid sequence can be determined by end group determination and amino acid analysis. The peptides are arranged in order by the isolation of overlapping peptides obtained by a different method of hydrolysis.

Secondary structure.

At present direct information about secondary and tertiary structure can only be obtained by the application of X-ray diffraction. X-ray analysis of β keratin fibres led to the proposal of the pleated sheet structure (Pauling and Corey, 1951). The α helix was originally suggested to explain the more compact structure of α keratin (Pauling, Corey and Branson, 1951). Supporting evidence for the α helix was provided by the deuterium exchange method of Hvidt, Johansen, Linderstrøm-Lang and Vaslow (1954) and also by the optical rotation studies of Yang and Doty (1957). However definitive evidence for the existence of α helices in globular proteins was first provided by the three dimensional Fourier synthesis from the X-ray data of Kendrew, Dickerson, Standberg, Hart, Davies, Phillips and Shore (1960).

Tertiary structure.

Tertiary structural investigations were initiated by the X-ray crystallography of sperm whale myoglobin at a 6 Å level of resolution (Kendrew, Bodo, Dintzis, Parrish, Wyckoff, and Phillips, 1958). The results indicated the polypeptide chain conformation within the molecule. Continued refinement leading to the present work at a 1.5 Å resolution has made possible the location of most of the atoms in the protein (Kendrew, Watson, Strandberg, Dickerson, Phillips and Shore, 1961). Such intimate detail has allowed calculation of the various interatomic distances and so provided the first proof of the bonds involved in maintaining the tertiary structure. Comparison with the incomplete chemical studies on sperm whale myoglobin has allowed the first formulation of the primary, secondary and tertiary structures of a protein.

Quaternary structure.

The phenomena of association and dissociation among protein molecules has been known since the early work on the ultracentrifugation of proteins. The recent literature on the dissociation of a wide range of globular proteins has suggested an overall biological importance for the subunit structure.

The best documented evidence for quaternary structure

in a protein is provided by the haemoglobins in which the naturally occuring tetramer can be broken down into two pairs of identical chains. The tertiary structure of the α and β chains and the manner of their association to give the quaternary structure has been shown by the X-ray work of Perutz, Rossmann, Cullis, Muirhead and Will (1960). While the isolated chains do have some oxygen binding capacity full physiological activity is only shown by the intact tetramer. The chains of the haemoglobin tetramer can be separated by pH alteration or by exposure to urea. However the A and B chains of insulin require chemical reduction of the disulphide bridges before separation can be achieved (Sanger, 1949). The isolated chains show no hormonal activity. Similarly the activity of α chymotrypsin requires the association of three different disulphide linked chains (Meedom, 1956). The work on the γ globulins (Porter, 1962) has shown that they contain two pairs of similar chains joined in the quaternary structure by disulphide bonds. The separated chains show different, partial activities. Activation of procarboxypeptidase A by trypsin involves disruption of a complex formed from three different chains (Brown, Cox, Greenshields, Walsh, Yamasaki and Neurath, 1961). β mercaptopyruvate transulphurase has been observed to exist in two forms (Fanshier and Kun, 1962), a highly

active unstable monomer and a less active stable tetramer. In the majority of enzymes, activity seems to depend on the association of inactive subunits. The activity of E. Coli alkaline phosphatase requires the dimer form which may contain non-identical subunits (Schwartz and Lipmann 1961; Schlesinger, 1964). It has been demonstrated with the aldolases (Drazi, Cheng and Horecker, 1962) that three inactive monomers interact to give an active trimer which appears to contain only one active center per trimer.

These examples show that quaternary structure is observed in proteins covering a wide range of physiological activities though there are some differences in the types of tertiary subunits and in their mode of association to give the quaternary structure. Thus the trend of present evidence suggests that for the majority of globular proteins several inactive subunits which may or may not be identical, interact by any of the forces considered to be important in the preservation of tertiary structure including covalent disulphide bonds, to give the fully active quaternary molecule.

X-ray determination of protein structure.

Although X-ray diffraction can yield the required information on primary, secondary and tertiary structures on which the properties of proteins depend, it has certain

inherent difficulties and limitations.

The original X-ray work on proteins was carried out with keratin which contains a considerable degree of molecular order in its natural fibrous state. With the globular proteins however this molecular order which is necessary for X-ray analysis is only obtained with crystalline preparations. In addition to the requirement for a crystalline protein the isomorphous replacement technique for obtaining X-ray data demands the availability of three or more different crystals, identical except for small defined regions of differing electron density due to the presence or absence of a heavy atom. This exacting requirement for suitable physical states of the protein has in the past proved to be a real stumbling block. The use of crystalline proteins has given rise to the objection that the conformation within the crystal might be different from the conformation of the active protein in dilute solution. The force of this criticism has been reduced by the demonstration by Doscher and Richards (1963) that crystalline ribonuclease has enzymatic activity.

When the diffraction data has been collected considerable calculations are required for the Fourier synthesis. This problem becomes more intense as the resolution increases. It is at this point that collaboration between the crystallographer and the protein chemist is most fruitful. With the simpler molecules the information

obtained can generally be interpreted in only one way, however the large protein molecules involve so many unknown parameters that the determination of a structure whose electron density corresponds to the diffraction patterns obtained is a long and tedious process involving much trial and elimination of possible structures. If the protein chemist can suggest relationships within the molecule which depend on a particular spatial folding of the polypeptide chain this process can be greatly simplified.

X-ray investigations are currently in progress on chymotrypsinogen (Kraut, 1964), chymotrypsin (Blow, 1964), lysozyme (Steinrauf, Reddy, and Dickerson, 1962), ribonuclease (Avey, Carlisle and Shukla, 1962), and insulin (Einstein, McGavin and Low, 1963). Most of these investigations are at a 4-6 Å level of resolution. Unfortunately it is not until a resolution of 1.5 or 2 Å isachieved that the wealth of detail necessary to pinpoint significant interactions within the molecule is obtained. Even at this resolution the complete primary structure can not be determined by physical methods alone since it is impossible to distinguish between the side chains of such residues as alanine and lysine unless the free movement of the side chain is restricted by interaction with another part of the protein molecule. Thus complete structural determination requires the combination of chemical and physical methods.

IMPORTANCE OF PRIMARY STRUCTURE.

The difficulties encountered with X-ray analysis mean that chemical study remains one of the most important methods for obtaining knowledge of proteins. This importance has been enhanced by the growing evidence that secondary, tertiary and so quaternary structure are dependent on primary structure alone. Crick (1958) in his hypothesis that nucleotide sequences in nucleic acids determine specific amino acid sequences of polypeptide chains maintained that folding of the polypeptide chain was simply a function of the order of the amino acids. It has been proved recently, at least for a few relatively simple protein molecules, that the native structure is the most probable thermodynamically and that the stabilising interactions are entirely due to the chemical properties of the linear amino acid sequence.

The current theories of protein biosynthesis mediated by polyribosomes imply the initial formation of a linear amino acid sequence. Experiments on the non-uniform labelling of haemoglobin (Dintzis, 1961; Bishop, Leahy and Schweet, 1960), bacterial α amylase (Yoshida and Tobita, 1960), lysozyme (Canfield and Anfinsen, 1963) and E. coli protein (Goldstein A., Brown B.J., 1961) suggest that synthesis starts at the amino terminal end and proceeds linearly to the carboxyl end.

The phenomena of reversible denaturation involving

partial alteration of the tertiary structure has been of interest for some considerable time, however the first reversible inactivation by cleavage of covalent bonds was demonstrated with insulin (Dixon and Wardlaw, 1960). The activity regained, although fairly low was considerably higher than that expected from a completely random rearrangement of the reduced disulphide bridges on oxidation. This has been followed by the complete chemical synthesis of an active insulin molecule (Katsoyannis, Fukuda, Tometsko, Suzuki and Tilak, 1964).

Similar reversible inactivation has been obtained with ribonuclease (White, 1960). The reduced form was shown to be a completely random chain containing no secondary structure. The original suggestion that the increase in activity over the statistical expectancy for the native configuration was due to the presence of more than one configuration possessing enzymic activity has since been disproved (White 1961) by showing that the reactivated material was identical with the native protein. Identity of reactivated egg white lysozyme and the native protein has been shown by Epstein and Goldberger (1963). Trypsin (Epstein and Anfinsen 1962), pepsinogen (Frattali, Steiner, Millar and Edelhoch 1963) and the much larger taka-amylase A (Isemura, Takagi, Maeda and Imai 1961) are also reactivated on oxidation of the completely random chain.

In 1963, Goldberger, Epstein and Anfinsen obtained a

microsomal preparation from rat liver which on addition of a soluble non-protein fraction reduced the half time of the in vitro reactivation of reduced pancreatic ribonuclease from 20 to 4.5 minutes. A similar enzyme has been purified one hundred fold from pig pancreas*. Since the time of reactivation obtained in the in vitro experiments is compatible with estimates of the time required for the cellular synthesis of a complete protein molecule, these findings suggest a role during protein biosynthesis for these apparently non-specific enzymes in the formation of disulphide bridges in a manner which is solely directed by the primary structure of the protein.

The weight of the present evidence indicates the validity of the simplifying assumption inherent in the theory of unidimensional genetic coding in protein biosynthesis. With such an intrinsic mechanism for tertiary structure formation it is not difficult to visualise the formation of the more complicated quaternary structures on the basis of the information contained in the amino acid sequences of their tertiary subunits.

An active aldolase trimer has been reconstituted from monomers which optical rotation studies showed to have considerably less helical content than the intact quaternary structure (Beal and van Holde, 1962; Stellwagen and Schachman, 1962). Similarly catalase has been reconstituted from

* Venetianer, Krause and Straub, (1964).

its two subunits which contained only half of the helical content of the original molecule (Samejima and Yang, 1963). Success has been reported (Haavik and Hatefi, 1961; Green, 1962) in attempts to reconstruct the more complicated electron transport activity from the component molecules. The E. coli pyruvate dehydrogenation complex has been resolved into three components and then reconstituted into an active multienzyme complex (Koike and Reed, 1961). In 1959 three groups (Ansevin and Lauffer, 1959; Anderer, 1959 and Wittman, 1959) reported the reversible dissociation of tobacco mosaic virus protein into subunits. The demonstration that the subunits possessed little or no three dimensional order (Wittman, 1959) provides direct evidence of the ultimate dependence of this very large quaternary molecule on the amino acid sequence of the constituent protein.

Genetics, evolution and primary structure.

The first evidence for a genetic control of proteins came from the observation that the incidence of the disease phenylketonuria, due to the absence of the enzyme phenylalanine hydroxylase, followed the known genetic laws. The isolation of the abnormal haemoglobins provided the first definite evidence for the genetic control of protein structure (Neel, 1949).

The work of Ingram (1956) on haemoglobin S spotlighted the dramatic effects small changes in primary structure can have when he showed that the lethal consequences of the homozygous sickle cell anaemia, in which the deoxygenated erythrocytes first assume a sickle shape and then rupture, was due to replacement of a glutamic residue in the β chain of normal haemoglobin by a valine residue in haemoglobin S. The severe limitations of mammalian genetics however preclude an investigation of the exact relationship between gene and primary structure.

Fortunately the microbial gene-protein system is much more amenable to study. Yanofsky (1962) has isolated approximately two hundred artificially induced mutants of the tryptophan synthetase A protein from E. coli. With these he has been able to show a relation between regions of the linear sequence of nucleotide containing coding units of the gene and specific amino acid residues in the primary sequence. This has given the first evidence that the triplet code letters suggested by the in vitro incorporation studies of Nirenberg, Matthaei, Jones, Martin and Barondes (1963) and Ochoa (1963) apply in the intact cell. Similar experiments have been performed with the alkaline phosphatase of E. coli (Levinthal, Garen and Rothman, 1961). From such work it should eventually be possible to obtain knowledge of genetic nucleotide

sequences from the primary amino acid sequence of the cellular protein.

Comparison of the elucidated sequences of several species of haemoglobins (Hill, Buettner-Janusch and Buettner-Janusch, 1963), cytochromes (Margoliash, 1963) and fibrinopeptides (Doolittle, Blomback, 1964) has allowed examination of evolution at a molecular level. Such comparisons are possible since certain of the amino acids in the sequence are essential for function and so are strongly selected for during evolution and thus can be used as markers for the correct alignment of sequences. Some of the other amino acids, with little functional role are subject to rapid flux which serves as a trail for following the course of the evolutionary changes. It is of interest that sequential evolution of the fibroinopeptides from the five artiodactyla investigated did not appear to correspond to the classical phylogenetic tree.

Protein structure and activity.

Introduction of the concept of enzymic active centers stimulated a great deal of research into their chemical natures. Various reagents have been found which react preferentially with certain amino acid residues in some enzymes. It is thus possible to implicate the reactive

amino acids in the enzymic active center. Diisopropylfluorophosphate shows selectivity in reaction with the serine of certain hydrolytic enzymes (Schaffer, May and Summerson, 1954). In the same way iodoacetate has been used for the detection of reactive methionine and histidine residues (Gundlach, Stein and Moore, 1960). Selective photooxidation of methionine has also been observed and used in the examination of the reactive methionine in chymotrypsin (Koshland, Strumeyer and Ray, 1962; Schachter and Dixon, 1964). Comparison of the amino acid sequences around the "active" amino acids have given rise to speculation on the reasons for their reactivity. However it is obvious from the studies on the complete sequences of ribonuclease and chymotrypsin that enzymic activity is a property of the tertiary structure and so confident speculation can only be made when this has been determined by X-ray methods or when sufficient information is available to allow the tertiary structure to be conceived through known interactions of the primary sequence.

The relation between structure and activity has also been approached by investigation of fragments of native proteins which still retain biological activity. Thus Smith and Hill (1960) in their studies of papain

have shown that one hundred and four residues can be removed from the N-terminal end of the protein without affectings its proteolytic action. Such studies help to localise the part of the protein molecule responsible for enzymatic activity. Other workers have reported the isolation of fragments of pepsin (Tokuyasu and Funatsu, 1962) and trypsin (Mosolov and Loginova, 1962) possessing activity. Synthetic experiments with the polypeptide hormones have given some indication of the structural requirements for hormonal action.

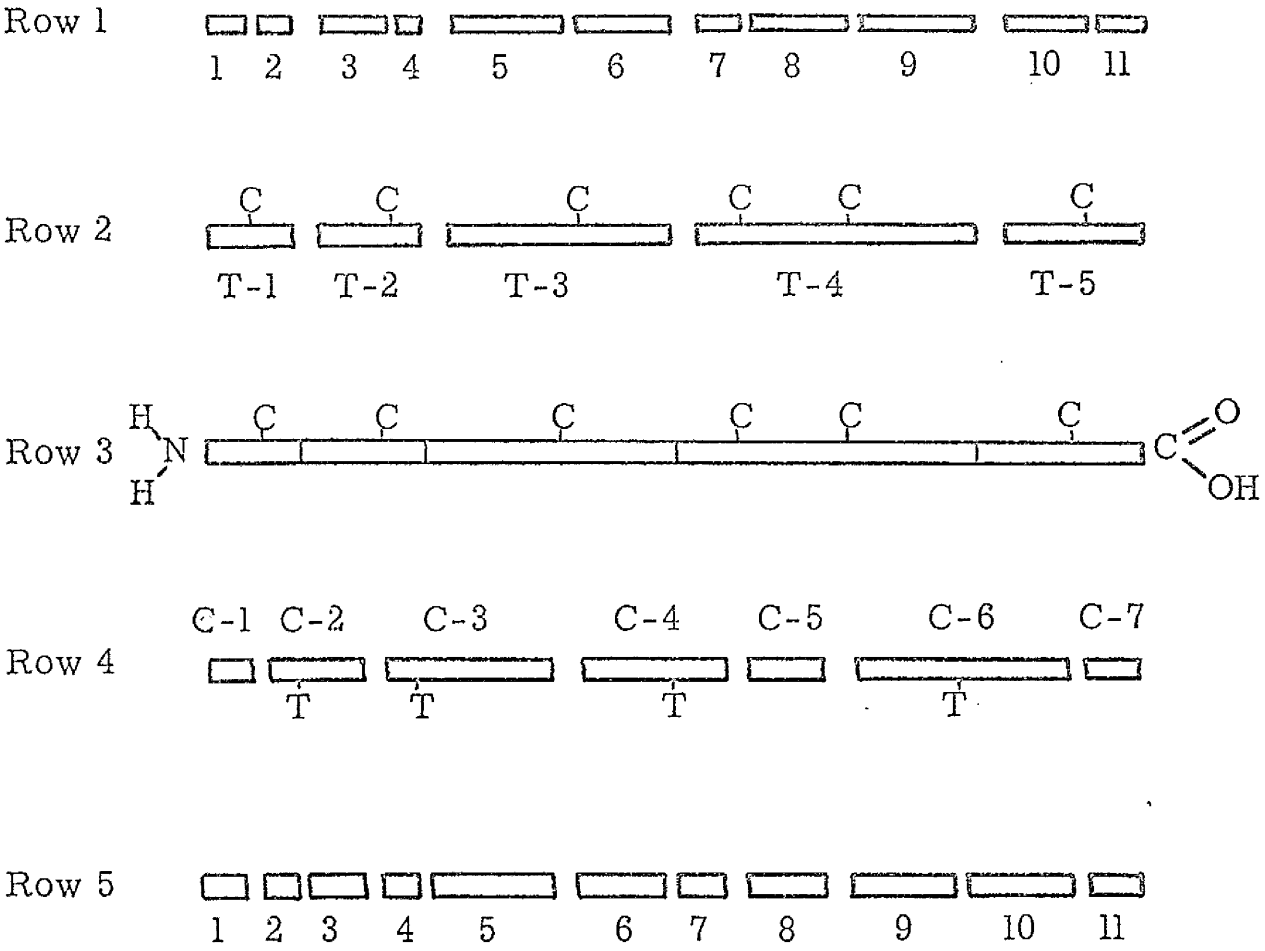


FIGURE 1.

A schematic illustration of the process used to order small peptides in their parent polypeptide chain. C and T are used as symbols for chymotrypsin or trypsin or their respective sites of cleavage.

PRIMARY STRUCTURE DETERMINATION.

The foregoing has illustrated some of the actual and potential uses of primary structural knowledge. Methods for primary sequence determination require a suitable source of the protein under investigation. After isolation, purification by the available techniques and determination of total amino acid composition, advantage is taken of any obvious simplifying procedures. Sanger's original work was greatly simplified by separation of the insulin molecule into A and B chains. Similarly disintegration of the large haemoglobin tetramer gave amenable single polypeptide chains containing approximately one hundred and forty amino acid residues. The most recent work on the structure of beef chymotrypsinogen A (Keil and Sorm, 1964, Hartley, 1964) was greatly facilitated by prior separation of the disulphide linked A, B and C chains. Although the principles involved in sequence determination are as formulated in the brilliant work of Sanger, the recent trend has been towards more specific methods of cleavage into small peptides.

Figure 1 (from Canfield and Anfinsen 1963[b]) is a schematic illustration of the current methodology. C and T represent points of cleavage of the most commonly employed cleavage agents, chymotrypsin and trypsin. Action of trypsin on the intact polypeptide chain shown in row 3 gives the products outlined in row 2. The tryptic peptides

are separated, purified and individually digested with chymotrypsin to give the peptides shown in row 1. These are separated and the N-terminal amino acid and amino acid composition determined. In the same manner, the same peptides are obtained by initial use of chymotrypsin followed by trypsin (Rows 4 and 5). Comparison of the N-terminal and amino acid composition of the peptides allows the row 1 and 5 peptides to be placed in order of appearance in the row 2 and 4 tryptic and chymotryptic peptides. Comparison of the tryptic and chymotryptic peptide compositions allows the original polypeptide chain to be represented as a linear arrangement of the peptides in Row 1. Where uncertainty still exists as to the order of amino acids within the small fragments the use of dilute acid hydrolysis or other enzymes such as papain, pepsin or subtilisin allows an unambiguous statement of the sequence.

Analysis of the amide content of the complete polypeptide chain and of the derived fragments can be obtained from the ammonia content of an acid hydrolysate. The position of the amide groups is determined either by amino acid analysis of the glutamic and aspartic containing peptides following digestion by leucine aminopeptidase or by examination of the electrophoretic behaviour of these peptides at pH 6.5.

Once the primary sequence of the individual polypeptide

chains has been worked out digestion of the intact protein with one of the proteolytic enzymes allows the separation of pairs of peptides joined by disulphide bonds of the original protein. Splitting of the disulphide bond followed by isolation of each of the peptides and comparison of its composition with the sequence in the individual polypeptide chains permits the correct location of the disulphide bridges between the chains to give the structure of the intact protein.

The peptides obtained in sequence work can be separated by ion exchange chromatography using either synthetic polyionic resins or modified celluloses, by gel filtration, or by paper chromatography or paper electrophoresis. N-terminal identifications of peptides are made either with the fluorodinitrobenzene method introduced by Sanger (1945), the Edman Reaction (Edman, 1950) or by use of the exopeptidase, leucine aminopeptidase. Repeated application of the last two methods in favourable circumstances can be used to give the sequence of small peptides directly. C-terminal determination can be achieved by hydrazinolysis (Akabori, Ohno and Narita, 1952) or by treatment with carboxypeptidase. Amino acid analysis is mainly by automated methods employing ion exchange resins (Spackman, Stein and Moore, 1958) or by two dimensional paper chromatography.

Proof of the feasibility of the methods in the successful amino acid sequence determination of insulin heralded an ever growing spate of work on other proteins. Among the proteins for which amino acid sequences have been proposed are bovine pancreatic ribonuclease (Spackman, Stein and Moore, 1960), horse heart cytochrome c (Margoliash, Smith, Kreil and Tuppy, 1961), human heart cytochrome c (Matsubara and Smith, 1962), pseudomonas cytochrome c-551 (Ambler, 1963), tobacco mosaic virus protein (Tsugita, Gish, Young, Fraenkel-Conrat, Knight and Stanley, 1960; Anderer, Uhlig, Weber and Schramm, 1960), the α, β, and γ chains of haemoglobin (Braunitzer, Gehring-Muller, Hilschmann, Hilse, Hobom, Rudloff and Wittman-Liebold, 1961; Konigsberg and Hill, 1962; Schroeder, Shelton, Shelton and Cormick, 1962), and most recently chymotrypsinogen (Keil and Sorm, 1964; Hartley, 1964) and bovine trypsinogen (Walsh, Kauffman, Sampath Kumar and Neurath, 1964).

The list of successes is impressive, However, due to the ever increasing molecular size of the proteins under investigation there is a very real need for more specific methods for cleaving the larger proteins into large fragments falling within the molecular weight limits of the present methods, which are capable of coping with a chain of one hundred and fifty amino acid residues. The most specific proteolytic enzyme in routine use is trypsin

which splits arginyl and lysyl bonds. The arginine and lysine content of the larger proteins means that initial use of that enzyme produces a mixture of peptides of sufficient complexity to overtax the present separation methods. Even with the smaller proteins such specific methods would alleviate the accountancy problem and would introduce a greater certainty that the determined sequence was correct. The possibility of error is shown by the ambiguity which existed in the original formulation for ribonuclease obtained by two independent groups of investigators (Spackman, Stein and Moore, 1960; Anfinsen, Sela and Fritch, 1956). There is considerable difference in the two proposals for the structure of chymotrypsinogen (Keil and Sorm, 1964; Hartley, 1964) mainly in the alignment of peptides. These can be vastly simplified

(Keil and Sorm) a - i j d b c - e f - k g h

(Hartley) a - i b d c e - j f - g k h

The letters denote sequences which are for the most part identical, - represents sequences where there is no apparent correlation. Obviously although both are in partial agreement over amino acid sequences within the peptides, there is considerable disagreement over peptide order. This example stresses the need for simplifying

methods or auditing techniques. Two approaches have been made, one enzymic and the other purely chemical. A combination of both methods was found necessary by Walsh et al. in their study on the two hundred and twenty nine residue containing bovine trypsinogen.

Selective enzymic fission.

The obvious solution is to use an enzyme which is specific for bonds involving only one rarely occuring amino acid. Examination of the range of known proteolytic enzymes shows however that this is not possible. Investigations on the formation of fibrin from fibrinogens by the action of thrombin (Doolittle and Blomback, 1964) have indicated that thrombin has a specificity for arginyl-glycine links. This specificity is too narrow to make thrombin a useful tool in amino acid sequence studies. The isolation of an enzyme, clostripain from Clostridium histolyticum with a specificity similar to trypsin has been reported (Gros and Labousse, 1960). The important difference from trypsin is that arginyl bonds are split three to five hundred times faster than lysyl bonds. This enzyme has potential applications although the lack of absolute specificity for arginyl bonds would be a disadvantage where long incubation times were necessary. Of the other proteolytic enzymes available trypsin with its specificity for arginyl and lysyl bonds is the most useful.

If the specificity could be restricted to one or other of these bonds the value of trypsin as a proteolytic tool would be greatly increased. In the light of the current knowledge of enzymes and their action, there are two potential solutions.

It should be possible to chemically modify the environment of the active centre in such a way that cleavage at one of the bonds is hindered by steric or other mechanisms. Present knowledge of the enzymic active site makes logical approach impossible although the work of Elmore, Baird, Roberts and Smyth (1964) with synthetic substrates suggests that the mechanism of action at the two amino acid residues is sufficiently similar to make successful increase in specificity by this approach extremely unlikely. Nevertheless the observation that acylation of trypsin reduces its activity towards proteins without affecting the activity towards synthetic substrates (Terminiello, Sri Ram, Bier and Nord, 1955) indicates that chemical modification of the enzyme can have some effect on its substrate specificity.

The more feasible approach is by chemical modification of the substrate protein. This can be by introduction of a substituent onto either the lysine or arginine side chain so that tryptic action is blocked or by alteration of the side chain of another amino acid in such a way that the enzyme cannot distinguish it from its natural point of action.

The ideal substrate blocking agent should give a soluble derivative, and after tryptic digestion and separation of the peptides should be readily removable so that trypsin can then attack at the unblocked susceptible residues.

Substrate blocking was first performed by Redfield and Anfinsen (1956) who used dinitrophenylated oxidised ribonuclease as a tryptic substrate in their studies on the amino acid sequence of ribonuclease. They found cleavage to be restricted to the arginyl bonds and simple location of arginine containing peptides from digests of other enzymes allowed the original peptides to be placed in order. The dinitrophenyl peptides however posed serious separation difficulties due to the strong absorption of the dinitrophenyl group to the paper or other separation media. Attempted removal of the dinitrophenyl group caused peptide bond hydrolysis. The technique was improved by introduction of benzyl-oxycarbonyl groups which can be removed after tryptic hydrolysis by treatment with anhydrous formic acid saturated with hydrogen bromide (Anfinsen, Sela and Tritch 1956). Unfortunately the carbobenzoxyl derivatives are sparingly soluble. N-acetyl and guanidinated derivatives were tried by Weil and Telka (1957) and shown to block tryptic action at lysyl bonds. The use of soluble succinyl derivatives by reaction of the protein with succinic anhydride has been proposed by Li and Bertsch (1960). However succinylation like

acetylation and guanidination is not readily reversible.

Reaction with carbon disulphide, the ethyl thiol ester of trifluoroacetic acid, or methyl acetimidate all yield masked ε - lysine groups susceptible to easy regeneration. Reaction with carbon disulphide under the experimental conditions is not quantitative, it is subject to ready reversibility and so gives a small amount of non specific proteolysis at incompletely masked lysyl bonds (Merigan, Dreyer and Berger, 1962). It also appears to remove N-terminal amino acids. The acetamido derivative, although more soluble than the trifluoroacetyl derivative requires exposure to ammonia at pH 11.3 for 8 hrs. for removal (Ludwig and Byrne, 1962). The substituent trifluoroacetyl groups can be removed by short treatment either with 1M piperidine (Goldberger and Anfinsen, 1962) or 0.5 M ammonia (McLaren, 1963).

It is possible to prevent tryptic attack at arginine by treatment of the protein with sodium in liquid ammonia (Berger, Kurtz and Noguchi, 1958). This converts arginine into tryptic resistant ornithine residues. The method has not yet been used in sequence work. Tryptic cleavage at arginyl bonds has also been blocked by chemical modification with benzil (Itano and Gottlieb, 1963). This gives slightly insoluble derivatives from which the substituent cannot readily be removed. So far the method has only been tested on the B chain of insulin and on salmine.

Lindley in 1956 reacted polycysteine with β bromoethylamine to give poly s-(β-aminoethyl-)cysteine which

was split to S-(β-aminoethyl-)cysteine by treatment with trypsin. Tieze, Gladne and Folk (1957) used the method on insulin and it was shown to be quantitative by Raftery and Cole (1963) who suggested its value in sequence work on the trypsin resistant core found in certain proteins. Ebata and Akabori (1960) and Ebata (1961) have shown that in model compounds conversion of the free γ glutamyl carboxy groups to γ hydrazides makes peptide bonds next to these residues susceptible to splitting by trypsin. The introduction of glycyl residues into proteins with the Leuchs anhydride of glycine makes peptide bonds next to serine susceptible to cleavage by trypsin (Shalitin, 1961).

Trifluoroacetylation has been used by McLaren (1963) in confirmatory work on horse heart cytochrome C sequence, by Carlton and Yanofsky (1963) in studies on the primary structure of tryptophan synthetase A protein, and by Dus and Kamen (1963) in comparative structural studies on some bacterial heme proteins. Walsh et al (1964) have used a combination of trifluoroacetylation with the trypsinophilic activation of cysteine residues in their work on bovine trypsinogen.

Selective chemical fission.

Selective chemical methods of hydrolysis depend on intramolecular assistance to promote reaction of one of the functional side chains with a component of the peptide bond, ultimately leading to cleavage of the peptide bond. Amino acid residues possessing such actual and potential functional side chains are cysteine, cystine, methionine, serine, threonine, aspartic acid and asparagine, glutamic acid and glutamine, lysine, arginine, tyrosine, tryptophan, histidine. Witkop (1961) differentiated between preferential cleavage which is competitive or hydrolytic, and selective cleavage which is non-competitive and non-hydrolytic. A close examination of the available methods of selective chemical cleavage however suggests that the difference between the two types of cleavage is more apparent than real. In all cases hydrolysis of the peptide bond is the essential feature, the difference, if any, lies in the rate of hydrolysis. In some cases the intramolecular assistance is such as to make the hydrolysis almost spontaneous. In less effective methods there may be considerable possibility of random hydrolysis at other peptide bonds. In the following discussion each method is described without reference to a specific type of mechanism.

The labilizing influence of the hydroxyl group on the

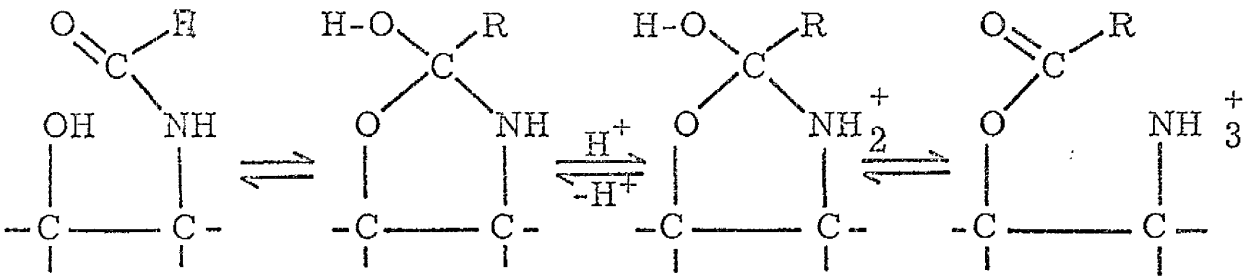


FIGURE 2.

The mechanism of nitrogen to oxygen acyl migration involved in the preferential acid hydrolysis of serine and threonine peptides.

acid hydrolysis of serine or threonine peptides has been known for a considerable time (Fischer and Abderhalden, 1907). The hydrolysis proceeds by nitrogen to oxygen acyl migration in the acidic media followed by preferential hydrolysis of the ester so formed (Fig.2.)

Habeeb (1961) has studied the method using concentrated sulphuric acid to stimulate the N→O shift. The O-peptide was then stabilised by acetylation and the ester split by mild alkali. However tryptophan, phenylalanine, tyrosine and cystine all suffered extensive modification under the strong acid conditions and the products obtained were difficult to separate. The use of anhydrous hydrogen fluoride has been reported to give more quantitative results on the dipeptides which were tested (Sakakibara, Shin and Hess, 1962). The studies have still to be extended to polypeptides and proteins.

Dilute acid causes preferential release of aspartic acid from proteins. The difficulty lies in finding conditions which favour the accelerating effect of the β carboxyl group on the cleavage of the aspartyl peptide bond over the preferential hydrolysis of peptides adjacent to serine and threonine. Grannis (1960) and Schultz, Allison and Grice (1962) have found 0.03 N HCl at 105° to be the most specific method. However Grannis reported that there was a significant hydrolysis at serine and

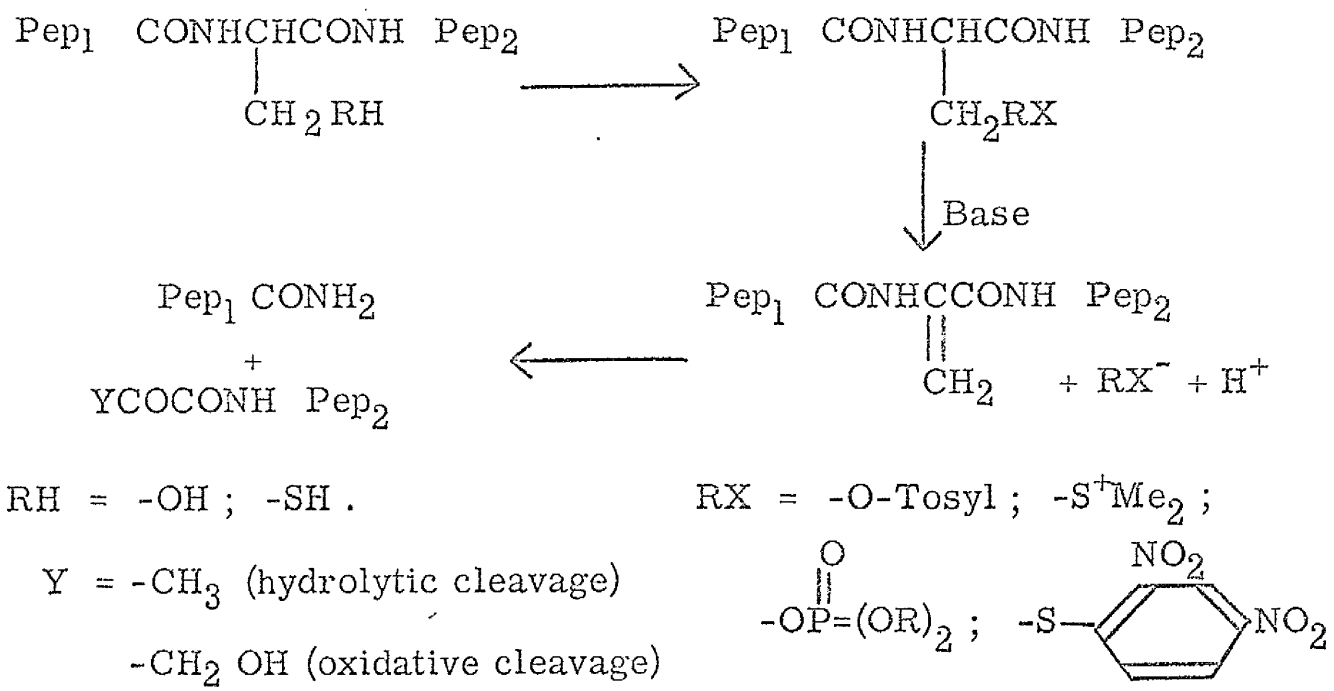
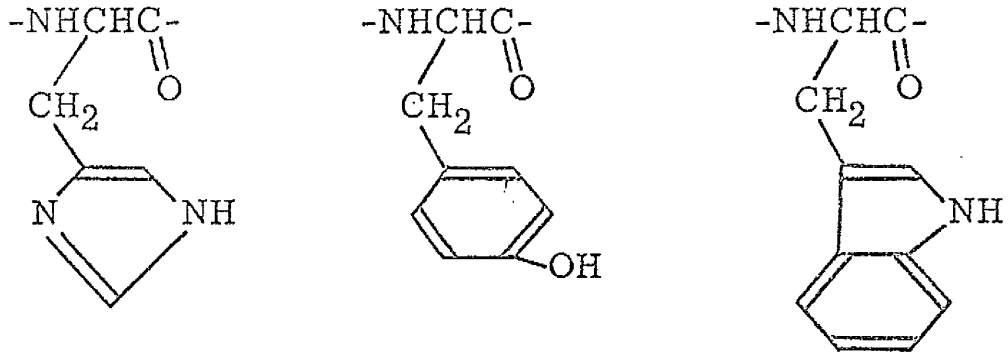


FIGURE 3.

Mechanism of cleavage of cysteine and serine peptides by β-elimination.



Histidine Tyrosine Tryptophan

FIGURE 4.

Comparison of the structures of histidine, tyrosine and tryptophan.

and threonine residues. The lack of specificity remains the major objection to all methods involving acid hydrolysis.

Patchornik and Sokolovsky (1964) have made an extensive study of a method giving cleavage at cysteine and serine residues. This involves conversion of the amino acid residue into a derivative which possesses a good "leaving group" on the β carbon atom, which is used to give a β-elimination reaction with formation of a dehydroalanine derivative. (Fig.3.). The dehydropeptide can be cleaved either hydrolytically of oxidatively. Oxidation with aqueous bromine at pH 5 followed by hydrogen peroxide at pH 7 is the method of choice. The present difficulty lies in the selective introduction of suitable leaving groups. This is particularly difficult with serine. The reaction with fluorodinitrobenzene is almost specific for the thiol group of cysteine at low pH values, and the thiodinitrophenolate ion eliminates in high yield making the method attractive for fragmentation of proteins at the cysteine residue (Sokolovsky, Sadeh, Patchornik, 1964).

N-bromosuccinimide has been found to cleave proteins at tryptophyl, histidyl and tyrosyl bonds. The similarities in the molecular architecture of the three amino acids is shown in Fig.4. The mechanism of the

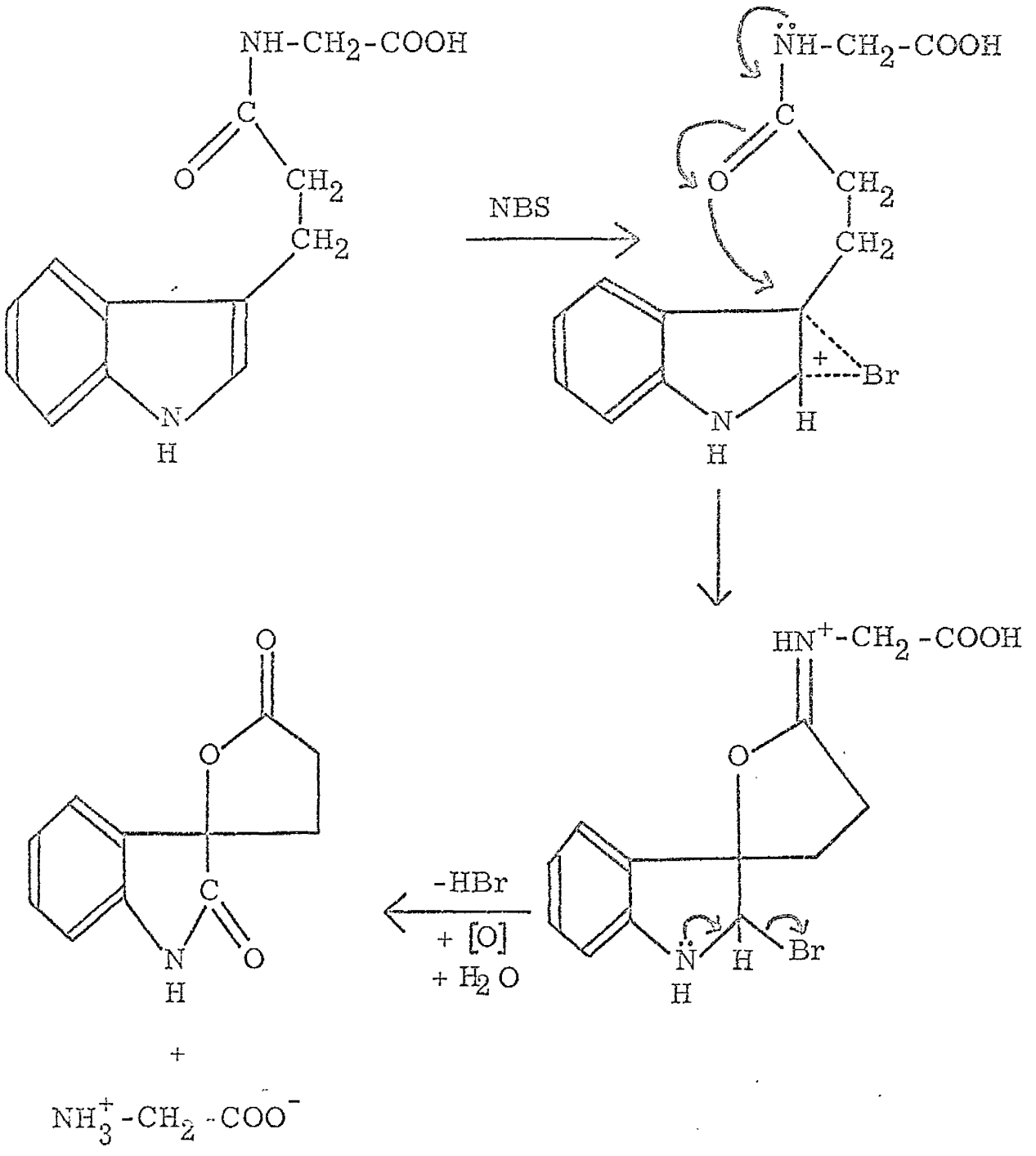


FIGURE 5.

Mechanism for the cleavage of indole-3-propionylglycine by N-bromosuccinimide (NBS).

reaction in each case is similar to that outlined for the cleavage of indole-3-propionylglycine (Patchornik, Lawson, Gross and Witkop, 1960) illustrated in Fig.5. The initial reaction yields a highly unstable intermediate which gives rise to a rapid, concerted 1,5-intramolecular displacement reaction. The double bond formed during the formation of the iminolactone is immediately hydrolysed under the conditions of the reaction to give a C-terminal lactone fragment and a new terminal amino group. Sasakawa (1963) has shown that a selective cleavage of C-tryptophyl bonds can be obtained by N-bromosuccinimide oxidation at pH 3.5. Raising the pH to 4,5 gives cleavage both at tryptophyl and tyrosyl residues. Shaltiel and Patchornik (1963) have proposed O-carbobenzyloxylation or O-acetylation of tyrosine in tyrosine containing proteins as a means for obtaining selective tryptophyl cleavage. N-bromosuccinimide has been used in the selective fission of tyrosyl bonds in ribonuclease which does not contain tryptophan (Wilson and Cohen, 1963). A selective cleavage of tyrosyl peptide bonds not involving N-bromosuccinimide has been obtained by electrolytic oxidation which had no effect on tryptophan residues (Iwasaki, Cohen and Witkop, 1963). Shaltiel and Patchornik (1963) have also obtained histidyl cleavage by heating the reaction mixture to 80° after reaction with N-bromosuccinimide. This reaction can only follow

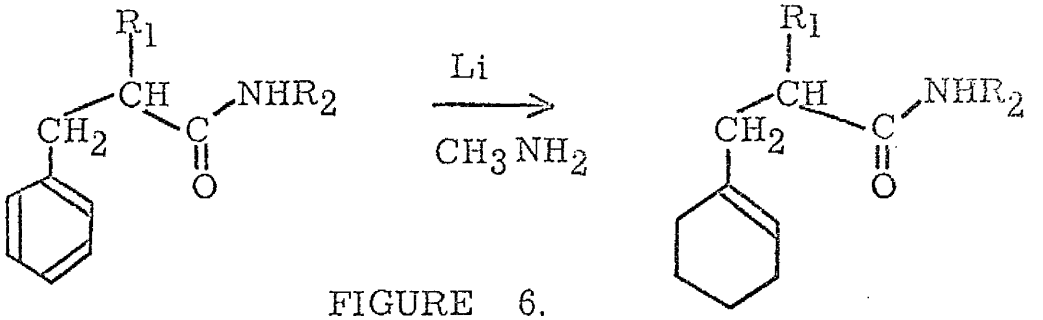


FIGURE 6.

Birch reduction of phenylalanine.

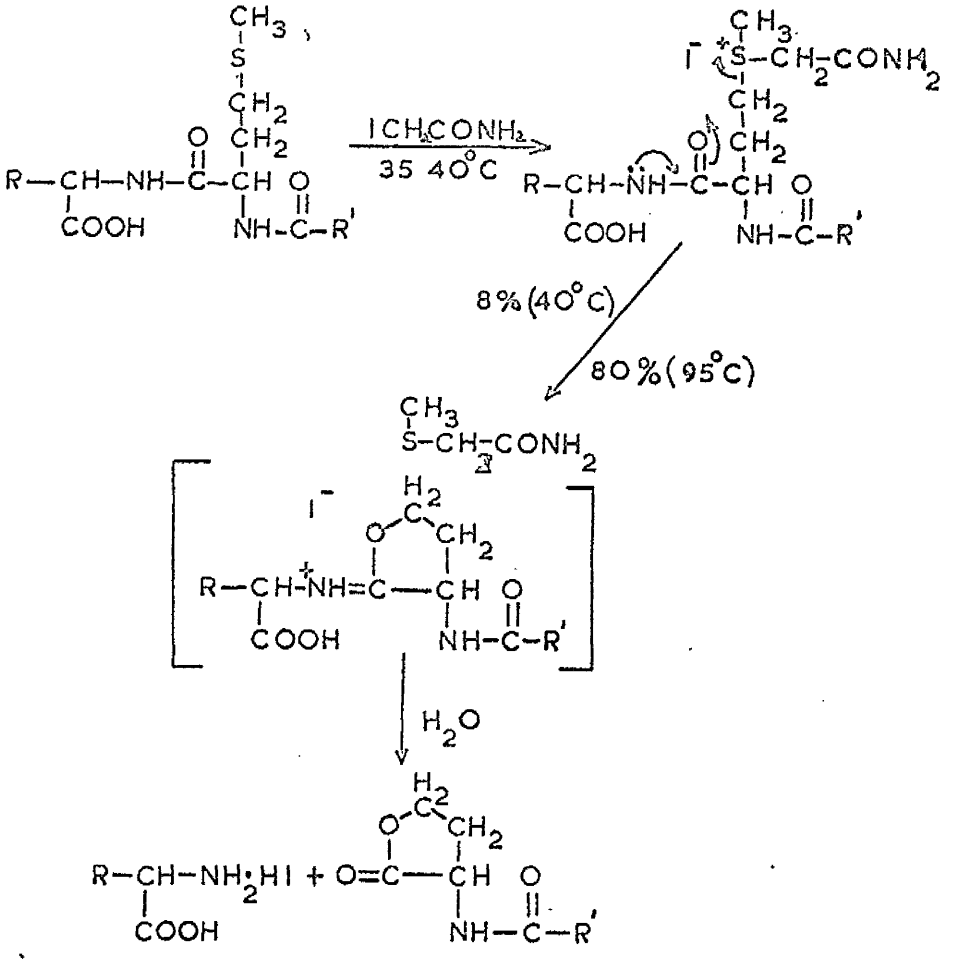


FIGURE 7.

Mechanism of selective cleavage of methionine peptides by iodoacetamide.

cleavage at tryptophan and tyrosyl residues. Wilchek and Patchornik (1962) have converted the phenylalanine side chain into a potential functional group by Birch Reduction of the benzene ring (Fig.6). Dipeptides of phenylalanine were found to be cleaved by N-bromosuccinimide in 65-70% yield. The method has still to be applied to proteins.

N-bromosuccinimide has some disadvantages as a selective cleavage reagent. In general the yields of peptides obtained are small (20-50%). Cysteine, cystine, methionine, tryptophan, tyrosine and histidine are all liable to attack and although peptide bond cleavage might not result, any chemical modification of side chains is a disadvantage in subsequent sequence work. N-bromosuccinimide also rapidly degrades free glycine and removes the ε-amino grpup of lysine. To date, N-bromosuccinimide cleavage has been of most value in auditing sequences next to tryptophan, tyrosine or histidine residues.

Selective cleavage of methionine peptides was first demonstrated by Lawson, Gross, Foltz and Witkop (1961) who tried several electrophilic agents. They obtained the best yield with the strongly electron withdrawing carbamylmethyl ($-CH_2-CO-NH_2$) group which acts as shown in Fig.7. Unfortunately with the alkyl halides tested heating was required before intramolecular lactonization was

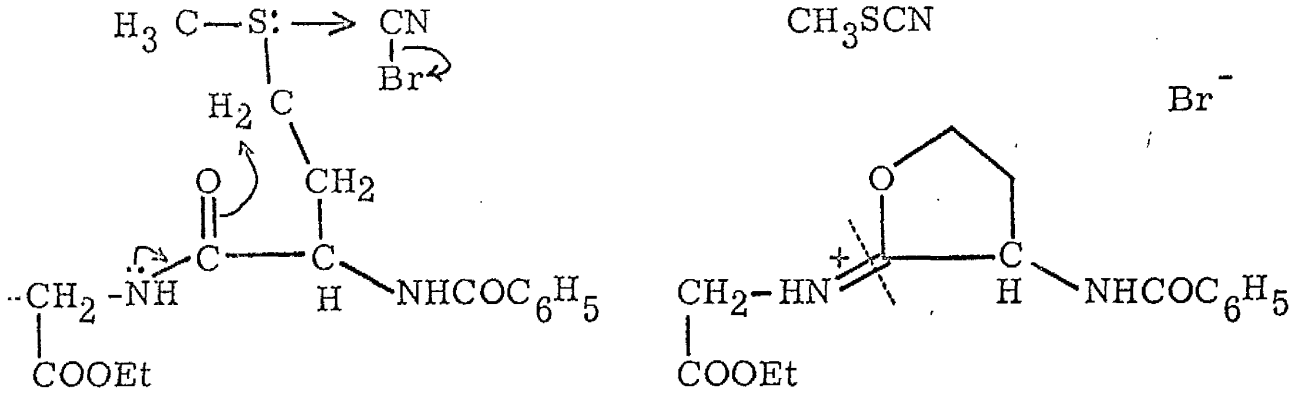


FIGURE 8.

Reaction of cyanogen bromide with ethyl N-benzoyl DL methionine glycinate.

obtained in any appreciable yield. A marked improvement in the cleavage of methionine peptides was achieved by the use of cyanogen bromide in 1961 by Gross and Witkop. Cyanogen bromide reacts with ethyl N-benzoyl DL methionine glycinate in aqueous alcohol at room temperature by intramolecular assistance in the displacement of methyl-thiocyanate, via the unstable cyanosulphonium salt (Fig.8). The products of the reaction are N-benzoyl DL homoserine lactone (70%), methylthiocyanate and ethyl glycinate (75-90%).

Reaction of bovine pancreatic ribonuclease with a 30 molar excess of cyanogen bromide in 0.1-0.3 N HCl yielded 25% of the chemical tail peptide, and 50% of free homoserine and homoserine lactone by cleavage of the methionyl-methionyl-lysine (29-30-31) sequence. Reaction of the "core" material with fluorodinitrobenzene revealed 85% cleavage of the methionyl-serine bond (Witkop, 1961). The method proved of great value in auditing the ribonuclease sequence of Spackman, Stein and Moore (1960).

Recently specific cleavage of C-methionyl bonds has been achieved by the action of hydrogen fluoride (Lenard, Schally and Hess 1964). The method has been applied to α melanocyte stimulating hormone with a resulting useful cleavage. Its exact potential is difficult to access until yields from reaction on a protein are available.

Cyanide has been found by Wood and Catsimpoolas (1963)

to cleave peptide bonds at the cystine amino group. One hundred molar excess at pH 7 with ribonuclease gave nine fragments in agreement with the theoretical expectancy. Yields and mechanism have still to be determined.

In the present work we set out to expand the knowledge of protein fission by cyanogen bromide. The initial work was on horse heart cytochome c, chosen as a small protein of known sequence which would therefore be expected to yield predetermined cleavage products. Subsequently cyanogen bromide was reacted with horse heart myoglobin, a slightly larger protein of unknown sequence. Having proved the value of the reagent on these smaller proteins, bovine serum albumin was chosen as a suitably large protein, to be representative of the sequence elucidation problems likely to be encountered with any protein containing more than 200 residues.

Since the initiation of this work Edmundson (1963) has used cyanogen bromide on sperm whole myoglobin where it has been particularly useful in fragmenting the insoluble tryptic core. It has been used by Hofmann (1964) on bovine trypsinogen and provided information which, when taken with the results of other methods, permitted the formulation of a working sequence for trypsinogen (Walsh et al., 1964). Craven, Steers and Anfinsen (1964) are applying the method to sequence work on β-galactosidase. It has been used

by Ambler (1963) to further degrade methionine containing tryptic peptides in his sequence work on pseudomonas cytochrome C-551. Koshland, Strumeyer and Nay (1962) have used cyanogen bromide to determine the reactive methionine residue in chymotrypsin following photo-oxidation.

MATERIALS.

MATERIALS.

Bovine serum albumin. Crystalline bovine serum albumin was obtained from Armour Pharmaceutical Co.Ltd., Eastbourne, England.

Cyanogen bromide. Cyanogen bromide was obtained from L.Light and Co., Ltd., Colbrook, Bucks., England.

Cytochrome c. Cytochrome c designated type III of 100% purity from horse heart was obtained from the Sigma Chemical Company.

Myoglobin. Myoglobin was prepared from horse heart, which had been removed from the animal immediately after slaughter. The heart was perfused with physiological saline as a preliminary step in the removal of contaminating haemoglobin (Theorell, 1932). As much as possible of the connective and adipose tissue was removed in the cold room and the strips of heart muscle passed through a meat mincer. Extraction and purification were then carried out at 4° by the method of Åkeson and Theorell (1960).

The minced muscle was suspended in half its weight of distilled water and centrifuged after standing for a short time. The extraction was repeated once with the same volume of ice cold water and after centrifugation the two supernatants were combined. The pH was adjusted to 7.0 and maintained at that value by addition of ammonia while solid ammonium sulphate was added to give 60% saturation. The precipitate obtained was centrifuged down and discarded. During the next two days

small portions of solid ammonium sulphate were added to the supernatant with adjustment of the pH to 7.0 when necessary until 90% saturation had been reached. Myoglobin precipitated as crystals which were collected in the centrifuge. The precipitate was suspended in a small amount of water and dialysed against 0.02 M phosphate buffer pH 6.9. When the precipitate had dissolved, a small amount of potassium ferricyanide was added to the buffer to convert all of the myoglobin or haemoglobin present to ferrimyoglobin or ferrihaemoglobin. The protein solution was then applied to an 80x 5 cm column of well packed carboxymethylcellulose which had been prepared and equilibrated with 0.02 M phosphate buffer pH 6.9. The column was developed with the same buffer at a flow rate of 4-5 ml/min. The colourless proteins were eluted first. The myoglobin, which moved slowly down the column, seperated into two broad bands. The haemoglobin and cytochrome C were retained at the top of the column. The two fractions containing myoglobin were combined and the myoglobin precipitated by ammonium sulphate saturation. A suspension of the protein was dialysed against distilled water until ammonium sulphate was no longer detectable in the dialysis water. The myoglobin was obtained in the solid state by lyophilisation. Estimates of the protein content of prepared solutions obtained from the spectra of the

pyridine haemochromogen and of the carbon monoxide derivative agreed well with the value obtained from the nitrogen content. Ultracentrifugation gave one peak corresponding to molecular weight of approximately 20,000.

Globin was prepared from myoglobin by dissolving 100 mg of the protein in 5 ml of water. Undissolved material was removed by centrifugation. Fifty millilitres of acid acetone (1% v/v conc. HCl) were added to the solution in a conical flask. The precipitate of globin was isolated by centrifugation and washed several times with acid acetone.

Sephadex. Sephadex for gel filtration was obtained from Pharmacia, Uppsala, Sweden. The grades used in the present investigation were G 50, G 75, G 100, G 200. The bead forms of G 50 and G 75 were used after they became available in November, 1963.

Trypsin. A crystalline preparation, free from chymotryptic impurity was obtained from the British Drug Houses Ltd., Poole, England.

Other reagents. All other reagents were A.R. grade where available.

EXPERIMENTAL

EXPERIMENTAL.

Preparation of reduced carboxymethylated bovine serum albumin.

It is routine practice when investigating the enzymic hydrolysis of proteins, to facilitate the hydrolysis by conversion of the native protein to the random polypeptide chain or chains. In the proteins containing disulphide bridges this involves oxidative or reductive splitting of the cystine residues. In the work on insulin (Sanger, 1949) and ribonuclease (Hirs, 1957) this was accomplished by performic acid oxidation. In addition to conversion of the cystine to cysteic acid residues, this procedure causes oxidative modification of any tryptophan residues present and oxidises methionine residues to methionine sulphone. In the present context modification of the methionine side chain was particularly to be avoided. Reduction of the disulphide bridges is most specific when carried out with one of the thiols - 2 mercaptoethanol, thioglycolic acid, or 2 mercaptoethylamine, of which mercaptoethanol is usually preferred, especially since White (1960) has shown that thioglycolic acid is frequently contaminated with highly reactive compounds which give thiolation of amino groups. Reduction is followed by alkylation with iodoacetic acid or iodoacetamide, forming carboxymethylcysteine residues which prevent random reformation of disulphide linkages. The experimental conditions involved in the early use of iodoacetate led to

considerable oxidation of methionine to methionine sulphoxides, however Sela, White and Anfinsen (1959) found that at pH 8 the reaction was confined to -SH groups with only a very slow reaction at histidine, tryptophan, lysine, tyrosine, and methionine. This observation has been used in the method of Crestfield, Moore and Stein (1963) for preparation of reduced alkylated proteins under conditions favouring retention of intact methionine residues. This method was used with slight modifications for the preparation of reduced alkylated bovine serum albumin.

One hundred milligrams of the protein were dissolved in 0.425 ml of EDTA solution (50 mg of disodium EDTA per ml) and 4.25 ml of tris buffer pH 8.6 (8.71 g of tris and 15 ml of 1.0 N HCl diluted to 50 ml with water). This was added to 5.12 g of urea in a 17 ml polythene stoppered vial. The level of liquid was made up to a 10.6 ml mark with water. The liquid was then stirred with a small magnetic stirrer until the urea had gone into solution. It was found that if the albumin was not dissolved prior to addition to the urea, it failed to dissolve completely. Nitrogen was then passed through the space above the solution for 2-3 min. The nitrogen stream was discontinued and 0.3 ml of mercaptoethanol added. The vial was then filled completely with a solution 8 M in urea and 0.2% (w/v) in EDTA, sealed with the polythene stopper and the reduction

allowed to proceed for 4 hr at room temperature with magnetic stirring. It was found that failure to stop the nitrogen stream before addition of the mercaptoethanol resulted in some of the highly volatile mercaptoethanol being removed with the nitrogen. After 4 hr the contents of the vial were transferred to a 25 ml beaker, filled with nitrogen and covered by a parafilm seal. A freshly prepared solution of 0.380 g of iodoacetic acid in 1.4 ml of 1.0 N NaOH was added to the reaction mixture and reaction continued for 15 min with stirring. The iodoacetate added is less on a molar basis than the amount of mercaptoethanol. The -SH groups of cysteine react most rapidly, and since the excess iodoacetate reacts faster with mercaptoethanol above pH 8 than it does with thioether sulphur, alkylation of methionine is kept to a minimum.

Removal of the urea by the method of Crestfield et al. (1963) involving chromatography on a 4x40 cm column of Sephadex G 75 was found to be time consuming. The presence of the high urea concentrations gave rise to very slow flow rates.

Isolation of the reduced alkylated protein was obtained by the method of Jirgensons and Ikenaka (1959). The reduced alkylated protein was precipitated at 4° with a mixture of acetone and 1 N HCl (39:1), washed thrice with the same solvent and twice with cold ether. Removal of the ether by evaporation left the reduced alkylated

protein in powder form.

Cleavage with cyanogen bromide. The conditions used were those of Gross and Witkop (1961) in their work on ribonuclease.

Ten micromoles of the protein were dissolved in 10 ml of 0.1 N HCl and reacted with a 30 molar excess of cyanogen bromide in a 50 ml round-bottomed flask on a magnetic stirrer for 24 hrs at room temperature. Excess cyanogen bromide and volatile side products of the reaction were removed by lyophilisation.

With the reduced alkylated bovine serum albumin it was necessary to use 10 ml of 0.1 N HCl which was 8 M in respect to urea in order to effect solution of the protein. Reaction was carried out with 1.5 μmoles of the protein instead of the 10 μmoles used with the smaller proteins. The same quantity of cyanogen bromide was used to give an even greater molar excess. After lyophilization, the products were dissolved in 10 ml of water, and the protein separated from urea by precipitation with acid acetone as described in the preparation of the reduced alkylated protein.

Cellulose acetate electrophoresis. Electrophoretic separation of the proteins or peptides depends on differences in electrical charge arising from differences in the content of amino acids with ionising groups on the side chain. Electrophoresis, with paper as a supporting

medium gave bad resolution due to tailing of the individual protein or peptide bands by interaction of the protein or peptide with the paper. The use of cellulose acetate gave sharp resolution with negligible tailing.

Cellulose acetate electrophoresis strips were obtained from the Oxoid division of Oxo Ltd., London. The electrophoretic separation was carried out on the Shandon Low Voltage Electrophoresis Apparatus. Buffers used were pyridine-acetic acid-water pH 3.4 (10 ml pyridine, 100 ml glacial acetic acid, 890 ml of water), pH 6.5 (100 ml pyridine, 10 ml glacial acetic acid, 890 ml of water) and barbitone-barbitone sodium pH 9.0 (163 ml 0.025 M barbitone, 100 ml 0.5 M barbitone sodium, 737 ml of water). Electrophoresis was continued for 1 hr with a current of 30 mA. After electrophoresis the strips were stained by immersing in a 0.1% (w/v) solution of Ponceau S in 0.2 N acetic acid and then washed in a 5% (v/v) solution of acetic acid. Protein present on the paper then appeared as a deep red band.

Sephadex filtration. The separation of substances on the basis of molecular size differences has been recognised for many years to be the principle involved in dialysis. Crystalloids are capable of passing through the pores of the dialysis membrane whereas colloids of much larger molecular size are retained. This basic principle of

molecular sieving has been extended and applied by Porath and Flodin (1959) in the introduction of particulate dextran gels as a means of separating solutes differing in size. The cross-linked dextran gels, obtained by polymerisingdextran with epichlorohydrin, are now commercially available under the trade name of "Sephadex". The dry particles of Sephadex, when placed in an aqueous solvent, swell up, and after a suitable equilibration time can be used for the preparation of a chromatography column. If a mixture of solutes of varying molecular size is applied to the column and the column developed in the conventional manner with any suitable aqueous solvent, it is found that the solutes have been fractionated and appear in the eluant in decreasing order of molecular size. The gel particle is conceived to be a porous network which contains the imbibed water. Small molecules have already access to the water contained in the gel network, whereas large molecules are completely excluded and are therefore restricted to the solvent between the particles of gel. Molecules of intermediate size can diffuse only into limited regions of the gel, the limit of these regions varying with the size of the molecule. Thus when the column is developed the largest molecules, which are restricted to the interparticle fluid, appear in the eluant after a volume which is equal to the interparticle

fluid space (void volume) of the column. The smallest molecules which have ready access to the intraparticle fluid only appear after a volume equal to the void volume plus the inner volume. Solutes of intermediate size appear between the two extremes at points depending on their molecular size.

It has been found possible to vary the diameter of the pores in the porous network of the gel by varying the amount of epichlorohydrin used to polymerise the dextran. By this means a range of Sephadex is now available whose exclusion limits very from small polypeptides to the largest proteins.

All of the present work using gel filtration has been carried out in 0.2 N acetic acid. It has the advantage as a solvent of repressing any tendency of the peptides to aggregate during separation. The required weight of the appropriate Sephadex was weighed into a beaker and equilibrated with 0.2 N acetic acid overnight. With the Sephadex G 50 and G 75 which were used before the bead forms became available it was found necessary to take extensive measures to remove the fines by decantation otherwise very slow flow rates resulted. Since the introduction of the bead forms this has become unnecessary. The beds were prepared by adding a slurry of Sephadex from the beaker to the column which was almost filled with 0.2 N acetic acid. As the Sephadex settled,

the clear supernatant was removed and fresh slurry added. When approximately six inches had settled to the bottom of the column, the screw clip at the bottom of the column was opened to give a flow rate similar to that required in actual operation (10-20 ml/hr). When the column had been poured to the required level an eluant reservoir was attached to the top and it was allowed to run for 24 hr. If at the end of that time the level of gel had fallen, fresh slurry was added until the level was at the original mark. The dimensions of the columns used throughout this work were G 50, 100 x 2 cm; G 75, 150 x 1.75 cm; G 200, 140 x 2 cm.

The peptide mixtures were taken up in a minimal volume of 0.2 N acetic acid. With the reduced alkylated bovine serum albumin and its cleavage products it was necessary to use 50% (v/v) acetic acid in order to dissolve the material. Two methods of sample application were used. The first required removal of the supernatant followed by careful addition of the sample with a pipette to the top of the bed which was protected by a glass paper disc. The bottom outlet was then opened and the sample allowed to enter the bed. At the moment it disappeared through the surface a small amount of eluant was added to wash the walls of the column and the surface. When this had disappeared the original volume of supernatant was added, the eluant reservoir connected and elution started. In the other

method, which was mainly used with high concentrations of albumin peptides, no protection of the bed surface was necessary. The sample was added directly to the top of the bed without removing the supernatant, using a bent tip Pasteur pipette. Care was necessary to avoid disturbing the bed surface and also to avoid disturbing the narrow zone of sample. The latter method was quicker and avoided the risk of air entering the top of the bed, but required a sample whose density differed considerably from that of the eluant.

Elution was carried out at room temperature with 0.2 N acetic acid as eluant. The pressure of elution was atmospheric at the top of the bed. Flow rate was adjusted by a screw clip on the column outlet to give 10-20 ml/hr. Two and a half millilitre fractions were collected by drop counting in an automatic fraction collector.

Protein content of the eluted fractions was determined by ultraviolet absorption measured at 280 mμ in a Unicam S.P.500. In the separations of the cytochrome c peptides, absorption of the porphyrin ring was measured at 415 mμ. Ninhydrin assays were carried out when necessary on 100 μl samples after alkaline hydrolysis.

The contents of the tubes containing the desired fractions were pooled and lyophilised before further purification or investigation.

Ninhydrin determination. Detection of the presence of

protein in a solution by determination of the absorption at 280 mμ is dependent on the tryptophan and tyrosine content of the protein which are its main absorbing components at that wavelength (Beaven and Holiday, 1952). A better indication is obtained by hydrolysis of the protein to its constituent amino acids followed by a colorimetric reaction for amino nitrogen. Hydrolysis was carried out with alkali by the method of Hirs, Moore and Stein (1956).

One millilitre of 2.5 N NaOH was added to 100 μl aliquots of the fractions contained in a set of numbered pyrex test tubes. The tubes were placed, unstoppered, in an open boiling water bath for 2.5 hr., during which time the contents evaporated to a volume between 0.1 and 0.2 ml. Half a millilitre of 60% (v/v) acetic acid and 0.4 ml of 0.5 M citrate buffer pH 5 was added to each of the cooled tubes. This was followed by 0.6 ml of the ninhydrin reagent of Yemm and Cocking (1955) and each tube mixed with a vortex mixer. The tubes were capped and heated for 15 min at 100°. After cooling each tube was diluted with 2.5 ml of 50% (v/v) ethanol/water and the absorption determined at 570 mμ using a Unicam S.P. 600.

Difficulty was frequently experienced in the appearance of a white turbidity during the alkaline digestion. For some reason this did not appear in every tube. Although

there is a considerable dilution factor, some of the determined optical density could arise from dispersion of the incident light by this material.

An additional possible source of error in ninhydrin estimations is contamination with extraneous ammonia. This is difficult to eliminate although all reasonable precautions were taken.

Acid hydrolysis. A weighed quantity of the protein or peptide was sealed in a small diameter test tube with 2 ml of 6 N HCl and hydrolysed for 24 hrs in an oven at 115°. Where an accurate measurement of the methionine or S-carboxymethylcysteine content was required the tube was sealed under vacuum. After hydrolysis the hydrochloric acid was removed in a vacuum desiccator at room temperature.

Amino acid analysis. The separation of free amino acids from protein hydrolysates by means of column chromatography was first achieved by Synge (1944) using starch columns. The technique was adapted to quantitative analysis by Moore and Stein (1948), through ninhydrin estimation of the amino acid content of the effluent fractions. The introduction of synthetic ion exchange resins allowed a more versatile application and an improved resolution (Moore and Stein 1951, 1954). Continued improvements in the precision of the analysis permitted its adaptation to an automatic recording apparatus capable of completing an analysisin 24 hrs (Spackman, Stein and Moore 1958).

The technique involves elution of the netural and acidic amino acids from 0.9 x 150 cm heated columns of Amberlite IR-120 by sodium citrate buffers. The basic amino acids plus tryptophan and ammonia are eluted from a 0.9 x 15 cm column with a more alkaline sodium citrate solution. The eluate from the ion exchange columns, passing through a continuous length of tubing, is heated for 15 min at 100°, the colour measured photometrically and recorded automatically. The photometer gives continuous readings at 570 mμ and 440 mμ. The 440 mμ reading is required for the determination of proline. The method is capable of estimating loads from 0.1 to 3.0 μmoles of each amino acid with an accuracy of ± 3%.

Development of the original apparatus has been directed towards more rapid analysis of decreasing quantities of amino acids. This is desirable due to the large numbers of peptides obtained during sequence work. Many of these peptides can be obtained only in limited quantities therefore the 1.5 - 4 mg of material required for a complete analysis placed an undesirable limitation on the sequence worker.

Eastoe (1960, 1961) obtained analysis on 0.3 mg of protein with only a slight loss in accuracy by reduction of the column diameters from 0.9 to 0.4 cm. This trend has been taken further by Kirsten and Kirsten (1962) who used 0.5 mm bore teflon tubing as columns. Although the time

for analysis was slightly increased it was possible to obtain a complete analysis on 5 µg of protein using a specially designed photometer to record the ninhydrin colours.

Technical developments in the resins available permitted Piez and Morris (1960) to obtain analysis of a protein hydrolysate on a single 0.9 x 133 cm column of Dowex 50 x 12. This reduced by half the quantity of material required by the Moore and Stein apparatus. Elution of the column is by a continuously increasing gradient of pH and salt concentration developed with the aid of a nine chambered "Varigrad" buffer reservoir. It is possible to vary the elution pattern of the amino acids by alteration of the gradient profile, which depends on the compositions of the buffers in the individual chambers of the Varigrad (Piez and Morris, 1960; Trincas and Rinetti, 1962).

Hamilton (1960) showed that resolution of amino acids during ion exchange chromatography could be increased by the use of the smallest resin particle size compatible with practical column operating pressures. While the use of faster flow rates caused a loss of resolution, combination with smaller resin particles gave a resolution comparable with that obtained in the original apparatus in a much shorter time. By use of a single 0.636 x 125 cm column packed with 17.5×10^{-4} cm spherical resin particles

and with a modified photometer and strip chart recorder he has been able to determine 10^{-8} mole of amino acid with an accuracy of $\pm$ 5% (Hamilton, 1963). Elution is achieved by stepwise increase of pH and salt concentration of the buffer supplied to the temperature programmed column. The sensitivity has since been further increased to determination of 10^{-9} mole at $\pm$ 3% accuracy by change in design of the photometer cuvette and by improvements in the signal amplification from the photocells (Hubbard and Kremen, 1964).

A sensitivity of this order requires that the time of elution be maintained at around 21 hrs. In cases where a slight loss of sensitivity is acceptable it has proved possible to reduce the time of analysis by using shorter columns with a faster flow rate. Spackman (1963), using the two column system of Spackman et al (1958), was able to reduce the time required for a complete analysis to 6½ hrs. by increasing the pumping rate from 30 to 40 mls per hr in conjunction with shortened columns. Attention to resin particle size and recorder operation at the same time gave a two fold increase in sensitivity over the original apparatus. Continued development by Benson and Patterson (1964) has made a complete analysis possible in 4½ hrs. and by combining the operations in the correct sequence it is possible to carry out three analyses in an eight hour working day. An effective decrease in analysis time has been obtained by Krampitz and Wieneke (1963), who have developed

photometers and recorders capable of simultaneous measurement of the effluent from six columns.

Simmonds (1958) also applied automation to the ion exchange columns of Moore and Stein (1951, 1954). His apparatus collects fractions of eluate, performs the reaction with ninhydrin, and records separately the colour intensity of each eluted fraction. The original apparatus required 48 hrs for a complete analysis. However, it was developed by Simmonds and Rowlands (1960) to give simultaneous determination from six columns and so increased the capacity of the method. Further improvement by Inglis (1964) allows simultaneous determination from nine columns.

Fraser, Inglis and Miller (1964) have developed an automatic computation system for the amino acid analyses from this apparatus which only requires selection of the correct programme. The use of this system, which can be applied to the results from other types of analysers, together with other forms of automatic integrators have further reduced the operator time required in the quantitative amino acid analyses of protein or peptide hydrolysates.

Automatic analysis of the column effluent by polarographic estimation of the copper complexes of the amino acids has been attempted (Blaedel and Todd, 1961; Corfield and Robson, 1962) but seems to afford few advantages over the ninhydrin method. The method is only suitable for α amino acids and

cannot be used for the detection of other ninhydrin positive substances.

The instability of the reduced ninhydrin reagent used in the apparatus of Spackman et al (1958) can be monitored by internal standardisation. This involves introduction of known concentrations of ninhydrin positive chemicals to the unknown protein or mixture of amino acids. A successful internal standard must be stable to acid hydrolysis, must react with ninhydrin and must not overlap any conventional amino acid in the chromatographic system employed. When recovery of the internal standard falls below a certain level the ninhydrin reagent is renewed. Norleucine, L-2-amino guanido propionic acid (Walsh and Brown, 1962) and β-2-thienyl-D L-alanine (Siegel and Roach, 1961) are among the chemicals proposed as internal standards.

The performance of the photometers and recording apparatus can be checked by external standardisation. This involves introduction of a known quantity of a known amino acid into the effluent line from the column, either manually or automatically.

The analyser available for the present work was designed and built by Dr. George Leaf. The design was based on the original model of Spackman et al. (1958) and was later modified to incorporate the improvements of Piez and Morris (1960). Incorporation of the most recent developments would

FIGURE 9.

Schematic diagram of the components of the automatic system for the analysis of amino acids (Piez and Morris, 1960).

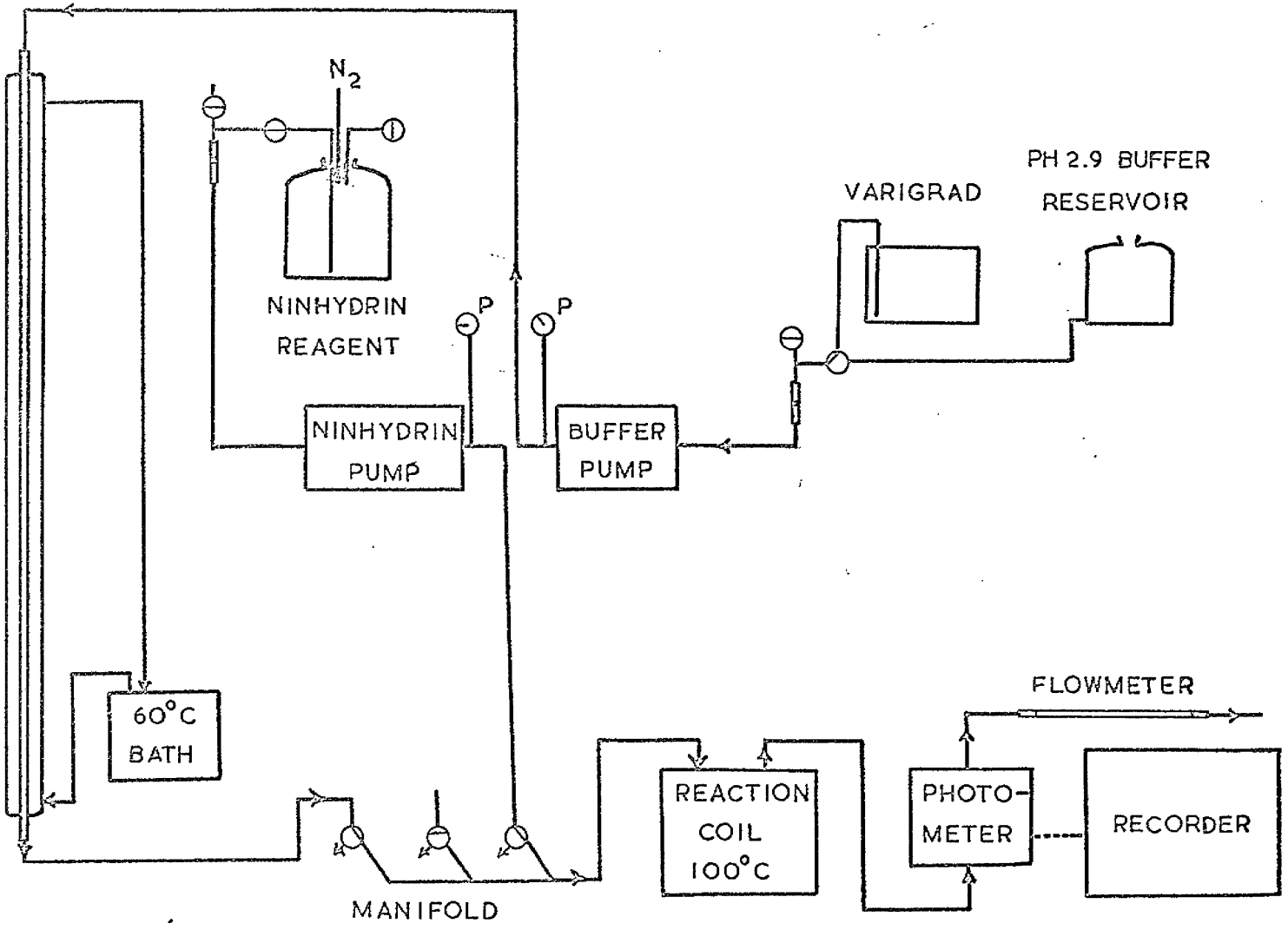

N2
NINHYDRIN
REAGENT
P
P
VARIGRAD
PH 2.9 BUFFER
RESERVOIR
NINHYDRIN
PUMP
BUFFER
PUMP
60°C
BATH
FLOWMETER
REACTION
COIL
100°C
PHOTO-
METER
RECORDER
MANIFOLD

FIGURE 9.

TABLE 1.

Buffer contents of the Varigrad compartments.

Chamber	pH 2.91 citrate (ml)	0.8 M sodium citrate (ml)	water (ml)
1	75	0	0
2	75	0	0
3	75	0	0
4	70	5	0
5	50	24	0
6	40	19	15
7	30	10	34
8	5	52	15
9	0	71	0

require major structural alterations. The design of Hamilton (1963), for instance, operated at pressures of approximately 450 p.s.i. whereas the maximum pressure obtained with the present apparatus is in the region of 45-50 p.s.i.. The buffer lines and column are not designed to sustain such pressures just as the photometers and recorders are insufficiently sensitive to evaluate the greatly reduced colour yields obtained with the smaller quantities of sample. A schematic diagram of the components of the analyser is shown in Fig.9.

During the analytical run the buffer pump is connected to the Varigrad. This was constructed to the original description of Peterson and Sober (1959). The composition of the solutions added to each of the nine intercommecting chambers is shown in Table 1. The supply to the pump is taken from chamber 1 and as the level in the chambers falls, so an increasing gradient of pH and sodium concentration is obtained. When the Varigrad is empty, the pump draws air into the buffer line, breaks an electrical circuit and so changes an automatic three way stopcock from the Varigrad to a supply of pH 2.91 citrate buffer. This buffer is pumped through the column and so prepares the resin for the next addition of sample.

The jacketed column is filled to a mark 133 cm from the sintered polythene retaining disc at the bottom, with Zeokarb

FIGURE 10.

Chromatographic analysis of a known mixture of amino acids plus cysteic acid and carboxymethyl cysteine. The blank values increase gradually during the run owing to the changing composition of the effluent. The completion of the gradient is indicated by the precipitate fall in the blank values following the arginine peak.

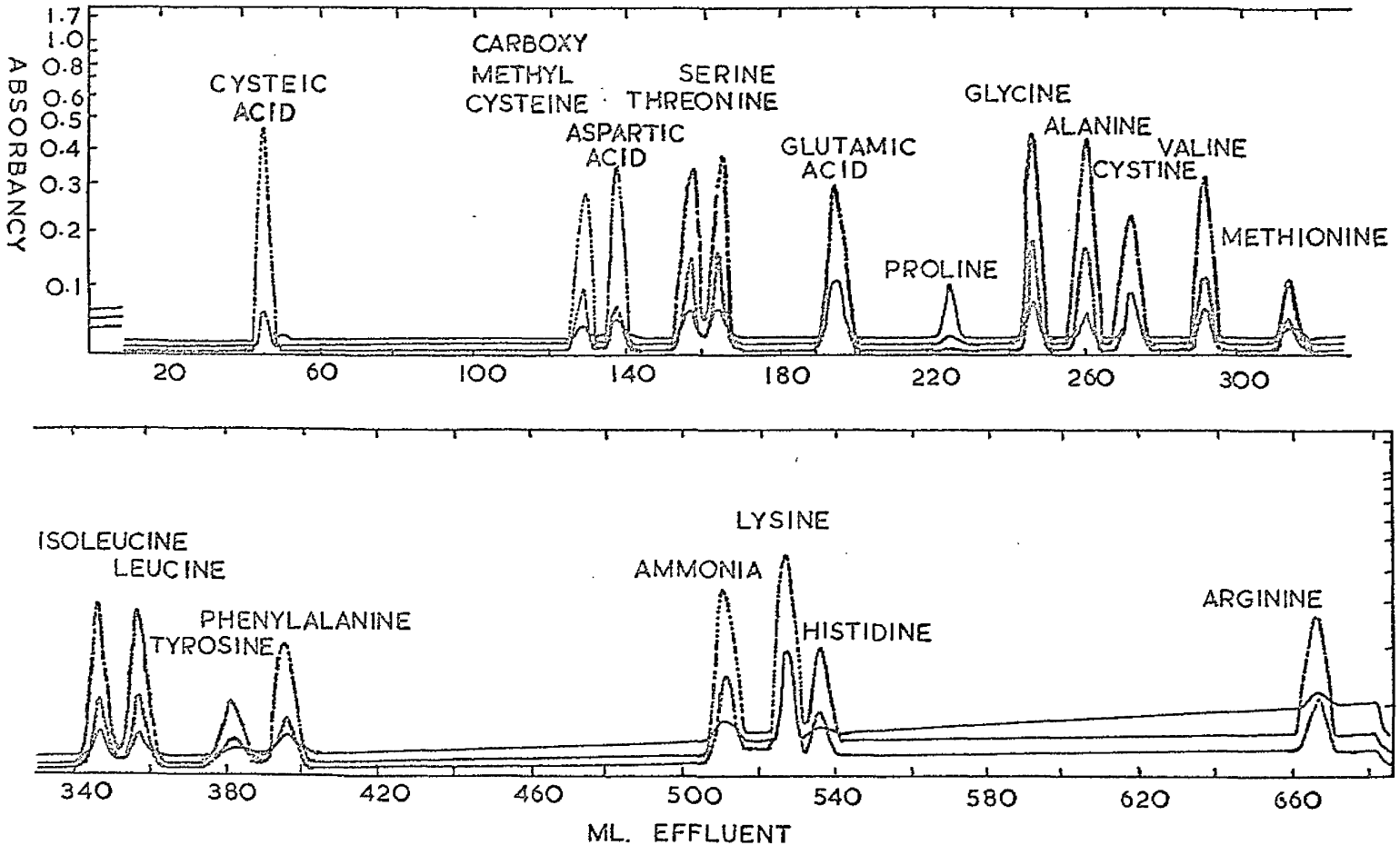

ABSORBANCY
1.7
1.0
0.8
0.6
0.5
0.4
0.3
0.2
0.1
CYSTEIC ACID
CARBOXY METHYL CYSTEINE
ASPARTIC ACID
SERINE
THREONINE
GLUTAMIC ACID
PROLINE
GLYCINE
ALANINE
CYSTINE
VALINE
METHIONINE
20
60
100
140
180
220
260
300
ISOLEUCINE
LEUCINE
PHENYLALANINE
TYROSINE
LYSINE
AMMONIA
HISTIDINE
ARGININE
340
380
420
460
500
540
580
620
660
ML. EFFLUENT

FIGURE 10.

225 x 12 obtained from the Permutit Co.Ltd., London. The resin was fractionated in the sodium form by the hydraulic procedure of Hamilton (1958) to give beads of the desired 25-35 μ diameter. The column is maintained at 60° by water circulated from a thermostatically controlled water bath.

The ninhydrin reagent is as described by Spackman et al. (1958) with the quantitative changes introduced by Piez and Morris (1960). From the addition of the stannous chloride all operations with the reagent are carried out under nitrogen.

The reaction coil, where the mixed ninhydrin reagent and column effluent are heated to 100°, is constructed of 0.7 mm bore teflon tubing of a length necessary to give a volume of 15 ml. Since both pumps are operating at 30 ml/hr this keeps the effluent and ninhydrin reagent at 100° for the required 15 min.

The use of three photometer cells placed in series in the effluent line allows measurement of the ninhydrin colour at 570 mμ, at 440 mμ for proline, and at 570 mμ with a light path a third of the first cell. The current produced by the three photocells is converted by the recorder into a three line trace. The presence of an amino acid in the effluent results in a peak on the recorder chart. A typical analysis record is shown in Fig.10.

After passing through a flowmeter the effluent is taken to waste. The flowmeter consists of a piece of constant bore glass tubing. It is possible to introduce an air bubble, half to three quarters of an inch long, into the effluent stream and by measuring the time required for the bubble to pass between two fixed points on the tubing to obtain an accurate measure of the performance of the pumps.

With a constant elution gradient, qualitative analysis of the amino acid mixture is provided by the relative position of each peak. Once the analyser has been standardisedón known quantities of amino acids, quantitative analysis is obtained by integration of the area under each peak. A sufficiently accurate measure of the area is provided by multiplying the height of the peak by the width at half the height. The use of the second cell at 570 mμ extends the range of the analyser since, if a greater than full scale deflection is obtained from the first photocell, a less than full scale deflection is usually obtained with the third photocell, which can then be integrated.

Several difficulties were experienced in the development and subsequent use of the apparatus.

When the analyser was first used it was found that arginine was not being eluted from the column before the appearance of the pH 2.91 buffer following the completion of the gradient. It was found that the remedy suggested

by Piez and Morris (1960) in shortening the column would have caused impaired separation of some of the other amino acids before arginine was completely eluted. The trouble was corrected by increasing the pH and sodium concentration in the latter part of the gradient through replacing the previous buffer used in chamber eight of the Varigrad with the same volume of 0.8 M sodium citrate as used in chamber nine. This also had the effect of giving earlier elution of all the amino acids after methionine but did not affect the resolution. It was later found that, due to excessive heat losses from the water line to the column and from the unlagged column itself, the column was not operating at the suggested optimal temperature. The results obtained when this was rectified indicated that a further increase in column temperature was desirable. A form of temperature programming is now employed. The column is operated at 60° until histidine emerges when the column temperature is manually increased to 68°. With this procedure it has been possible to revert to the original Varigrad composition while arginine is still eluted by a safe margin.

When the complete elution and resolution of the amino acids had been achieved the analyser was standardised using 1 ml samples of a 1 in 5 dilution of the Amino Acid Calibration Mixture Type I obtained from Spinco Division,

TABLE 2.

Values for the constant C obtained during the calibration of the amino acid analyser.

	1	2	3	4	5	6	7	C (mean)
Cys.A	7.27	5.07	8.26	7.75	4.74	6.78	7.3	7.47
Asp.	9.22	8.82	7.34	8.5	8.59	8.4	8.9	8.82
Thr.	10.12	9.06	7.02	9.65	8.7	9.03	8.53	9.39
Ser.	10.31	9.6	7.71	9.96	9.32	10.6	9.02	9.94
Glu.	9.17	9.32	9.46	8.83	8.5	9.11	8.63	9.14
Pro.	-	2.39	2.38	2.48	2.46	2.35	2.52	2.42
Gly.	10.27	10.27	9.92	9.6	9.35	9.46	9.87	9.66
Ala.	11.02	10.99	10.69	-	10.0	10.4	10.1	10.25
Cys.	6.46	6.44	6.25	-	6.06	6.00	5.85	6.06
Val.	11.08	11.02	11.0	10.42	10.3	10.9	10.0	10.56
Met.	10.82	10.53	10.42	10.12	9.82	9.05	6.32	9.87
Ileu.	12.59	12.16	11.28	11.42	10.6	11.92	10.75	11.32
Leu.	12.69	11.89	12.39	13.35	12.75	12.3	12.3	12.33
Norleu.	12.08	11.96	12.25	11.12	12.4	11.85	11.30	11.78
Tyr.	11.82	10.99	11.81	11.27	11.97	11.8	11.67	11.38
Phe.	12.38	12.88	12.26	11.78	11.73	12.0	11.85	11.93
NH_3	20.8	11.41	13.00	-	10.13	-	12.95	12.18
Lys.	13.73	12.84	12.91	-	13.30	12.4	14.8	12.84
His.	11.75	10.82	10.9	-	11.02	10.3	11.41	10.72
Arg.	11.21	11.21	10.92	9.72	11.17	-	-	11.18

Values for carboxymethyl cysteine (7.06) and homoserine (1.78) have also been obtained.

Beckman Instruments Inc., California, to which had been added cysteic acid and as an internal standard norleucine. Since at this dilution each ml contained 1 μmole of each amino acid plus ammonia, integration of the peaks obtained gave the ninhydrin colour yield of each amino acid. The figures obtained are shown in Table 2. Neglecting figures which were obviously in error, the mean was taken and used as the constant, c, in all subsequent calculations. The reproducibility as judged from the calibration runs and from duplicate runs of protein hydrolysates is close to the ±3% figure given by Spackman et al. (1958).

Rosen, Berard and Levenson (1962) gave details of a ninhydrin reagent which appeared to have several advantages over that of Spackman et al. (1958). Replacement of sodium acetate by sodium propionate, which has better solubility characteristics, eliminated the danger of clogging of the effluent lines which is always possible with sodium acetate. Replacement of stannous chloride by cyanide added to the column buffers instead of to the ninhydrin reagent, improved the stability of the ninhydrin reagent which no longer required to be kept under nitrogen.

The necessary alterations were made to the reagents and the analyser was recalibrated. It was found that the colour yields of the first amino acids were improved due to better buffering by the ninhydrin reagent. A

TABLE 3.

Comparison of the colour yields with the ninhydrin reagent of Spackman et al. (1958) and with that of Rosen et al (1962).

	Spackman et al.	Rosen et al.
Cys. A.	7.47	9.84
Asp.	8.82	11.62
Thr.	9.39	12.04
Ser.	9.94	12.42
Glu.	9.14	11.27
Pro.	2.42	2.69
Gly.	9.66	10.81
Ala.	10.25	11.86
Cys.	6.06	6.22
Val.	10.56	11.94
Met.	9.87	11.58
Ileu.	11.32	12.6
Leu.	12.33	12.86
Norleu.	11.78	11.18
Tyr.	11.38	11.02
Phe.	11.93	10.65
NH_3	12.18	7.60
Lys.	12.84	14.31
His.	10.72	8.45
Arg.	11.18	4.25

comparison of the yields obtained from the two ninhydrin reagents is shown in Table 3. Unfortunately the reagent had to be abandoned because of crystallisation in the effluent lines, the defect it had been introduced to remedy. Time did not permit suitable modifications of the reagent. However by taking care that the machine is never stopped for long periods of time with ninhydrin reagent in the heating coil, little difficulty has been experienced using the first ninhydrin reagent.

Just as the manual ninhydrin estimation is prejudiced by the presence of extraneous ammonia so error can also be introduced into the automatic analysis. Traces of ammonia are always present as impurities in the reagents used for making the citrate buffers. This results in an ammonia plateau which normally appears as a sharp rise in the blank just before the ammonia in the sample. If it is small, it does not affect the calculations even though the base line may not be horizontal. As a routine measure each 4 litre batch of pH 2.91 buffer was passed through a 45 x 4 cm column of Dowex 50 x 8. Even with this precaution in many cases there has been a large ammonia plateau which on occasion has made calculation of lysine impossible and calculation of histidine difficult due to the rapidly changing base line. This could be due to variable quantities of ammonia in the 0.8 M citrate or

to an undetected source of ammonia in the laboratory. The quantities of ammonia required for contamination are very small. (1 μmole of ammonia = 0.017 mg).

At one stage in the analyses a constant ninhydrin colour appeared in the effluent. This was traced to a contaminant of the methyl cellosolve which henceforth was passed through a 35 x 2 cm column of Amberlite IR-140.

Intermittent trouble has been experienced during this series of analyses through variable pump output. The pumps used are micro pumps obtained from the Distillers Company Ltd., Engineering Division, Epsom, Surrey. Pumping action is by positive piston displacement, direction controlled by ball valves. Pumping rate can be varied by vernier screw adjustment of the piston stroke. After the initial setting, pumping performance was determined by the flowmeter. The marks on the flowmeter tube were arranged so that at a pumping rate of 30 ml/hr, each bubble of air introduced took 90 sec. to pass between them. After checking the buffer pump on its own, the buffer pump and ninhydrin pump were then checked together. This was necessary since independent use of the ninhydrin pump would give crystallization in the effluent lines. Fortunately the ninhydrin pump was generally stable in operation. Serious trouble was usually indicated by an alteration of the readings on the pressure gauges.

However the buffer pump in particular showed a distressing ability to give a consistent, slightly abnormal output for variable lengths of time. It was found that frequent use of the flowmeter was the best method for detecting such variation.

Erratic operation of the buffer pump was frequently experienced after the pump had been dismantled for maintenance purposes. This was finally considered to be due to the presence of small, persistent bubbles of air in the pumping chamber, or on, or around the valve seat which were not dislodged by the low rate of liquid flow in the pump. Since the pumps are operating at four fifths of maximum stroke there was no possibility of increasing the liquid flow by reducing the pump motor gearing. It has been found that if a 10% (w/v) aqueous solution of Brig 35 is passed through the pump immediately after reassembly, the trouble is largely eliminated.

Inconsistent performance as a result of air bubbles was generally only experienced immediately after the pump had been reassembled. Erratic behaviour at other times was caused by the gland packing, which surrounds the piston and makes the piston chamber watertight, breaking down under the abrasive action of the piston movement and passing via the piston chamber to the area around the valve seats. The presence of this debris

interfered with the correct seating of the balls on the valve seats and so gave impaired pump performance. The only satisfactory solution was to dismantle and repack the pump at regular intervals. The consistently better performance of the ninhydrin pump is assumed to be due to the presence of the methyl cellosolve which by its lower surface tension will tend to remove air bubbles and the organic solvent will also give a better lubricating action between piston and packing.

In operation it was found that the pumping rate was liable to vary by±1% during any one run. This was considered to be an insignificant variation but in fact such a variation can have several effects on the accuracy of the results.

A change in the performance of either pump will have an effect on the ratio of buffer solution to ninhydrin reagent in the effluent lines. This could give rise, with the present ninhydrin reagent, to a solution with a slightly different pH from the normal for that point in the analysis. As the results with the propionate-buffered ninhydrin reagent show, the colour yield is dependant on the solution pH. Therefore a slowing of the buffer pump could give increased colour value, a slowing of the ninhydrin pump a decreased colour value. Increases in the pumping rates would have the opposite effects. Such

an effect would only be evident in the first part of the analysis where the column buffer pH is much lower than that of the buffer in the ninhydrin reagent.

A slowing of the buffer pump would also cause the resultant colour to be more dilute due to the reduced ratio of column effluent to ninhydrin reagent, which would also cause an increase in the blank value. The combination of these two effects would give a lower value for the height of the peaks. This would in part be compensated by the longer time required for any colour developed to pass through the photocells, which would result in a wider peak. Slowing of the ninhydrin pump would cause a more concentrated colour, a lowering of the base line and so an increase in the height of a peak. Due to the slower rate of flow the width of the peak would again be increased. Increase in either of the pumping rates would have the opposite effects.

Irregularities in the pumps sometimes became evident in fluctuations in the base line. This was most often found from the start of a run, up to the emergence of aspartic acid, during which time the pumps appeared gradually to settle down to a constant rate. However on occasion the baseline was uneven at other parts of the analysis. In these cases a straight line was drawn as a mean of the fluctuation and used to calculate the results.

If the baseline was straight the method of calculating the area of each peak gave consistent results. The main sources of error were in the estimation of the base line and to a greater extent the height of the curve particularly where it fell above 0.5 on the logarithmic scale, and in the estimation of the peak width by the dot counting method of Spackman et al. (1958).

The value of the constant, c, is of course subject to the possible errors mentioned above, however these were minimised by carrying out frequent calibration runs. The danger lies in assuming that the accuracy obtained in the calibration runs holds in all cases. Erratic pump operation has been observed and will give rise to inaccurate results. Consistent recovery of an internal standard is only an indication of analyser performance at the time when the standard is eluted from the column. Duplicate runs of unknown mixtures is time consuming and expensive in material but were carried out whenever possible. Since more accurate pumps are not presently available, every attempt was made by regular maintenance to restrict the limits of variation of the instruments in use.

Fingerprinting. Separation of mixtures of peptides obtained by enzymatic digestion of proteins by a combination of chromatography and electrophoresis on paper was first described by Ingram (1958) for the tryptic peptides of human haemoglobin. After electrophoresis at pH 6.4 at 19

V/cm for 2 hr., the paper was dried and an overnight ascending chromatogram run in butanol/acetic acid/water at right angles to the direction of electrophoresis. Location of the peptide spots by a ninhydrin dip gave a pattern which was characteristic for each protein. In this way Ingram was able to show that the differing properties of normal adult haemoglobin and sickle cell haemoglobin were due to a difference in one of the twenty eight tryptic peptides.

For the following work tryptic digestion of the proteins was carried out either in 0.05 M phosphate buffer pH 7.0 or 0.1 M ammonium carbonate buffer pH 8.0. Two to twenty milligrams of the protein or peptide were taken up in 1 ml of the buffer and digested with a fiftieth part by weight of trypsin for 2 hr. in a water bath at 37°. The digestion mixture was then lyophilised. The use of ammonium carbonate as buffer is preferable since this step removes most of the buffer salt. The lyophilisate was taken up in sufficient water to give a resulting concentration of approximately 1.5 mg/20 µl. Twenty microlitre samples were spotted on Whatman Chromatography paper. Two variations of the fingerprinting technique were tried.

The first was a modification of the method of Ingram. Whatman No.1 paper was used and electrophoresis was carried out in pH 6.4 pyridine/acetic acid/water (100:10:890 by

volume) for 1 hr on a Shandon High Voltage Electrophoresis Apparatus with a voltage drop of 60 V/cm. After drying, the second dimension was developed overnight in an ascending system of butanol/acetic acid/water (200 : 30 : 75 by volume). The results were suitable for separation of a simple mixture of peptides, however with the more complex mixture obtained from bovine serum albumin streaking of the peptide spots during electrophoresis gave a very poor resolution. In this case the use of a modification of the technique of Anfinsen, Åqvist, Cooke and Jönsson (1958) gave much better results.

Whatman No 3 paper was used and the electrophoresis was performed in a hanging strip apparatus after Durrum (1950) with pH 3.7 pyridine/acetic acid/water (1 : 10 : 289 by volume) with a voltage drop of 16 V/cm for 16 hr. Chromatography was carried out overnight in the descending direction with the previous butanol/acetic acid/water solvent. Very little difference was found in carrying out the chromatography first.

The peptides spots were located with a ninhydrin-collidine dip (600 ml absolute alcohol: 100 ml glacial acetic acid : 40ml collidine : 1 g ninhydrin). The dried papers were dipped, dried, and placed in an oven at 80-100° for 5 min. Peptide spots were outlined in pencil while the paper was held against a light source.

The Ehrlich Stain for indoles, Pauly for imidazole

and Sakaguchi for arginine were used, but it was found that sensitivity was lost if they were used sequentially after ninhydrin. Generally the small quantities of materials available prevented their separate application to several fingerprints. Comparison of different fingerprints of material from the same or similar sources was therefore based on similarities in the relative positions of the individual spots.

N-terminal determination. The fluorodinitrobenzene technique of Sanger (1945) was used as a qualitative method for determining the N-terminal amino acids of the peptides obtained.

The reaction with fluorodinitrobenzene was carried out either by the method of Koch and Weidel (1956) or by that of Levy (1954).

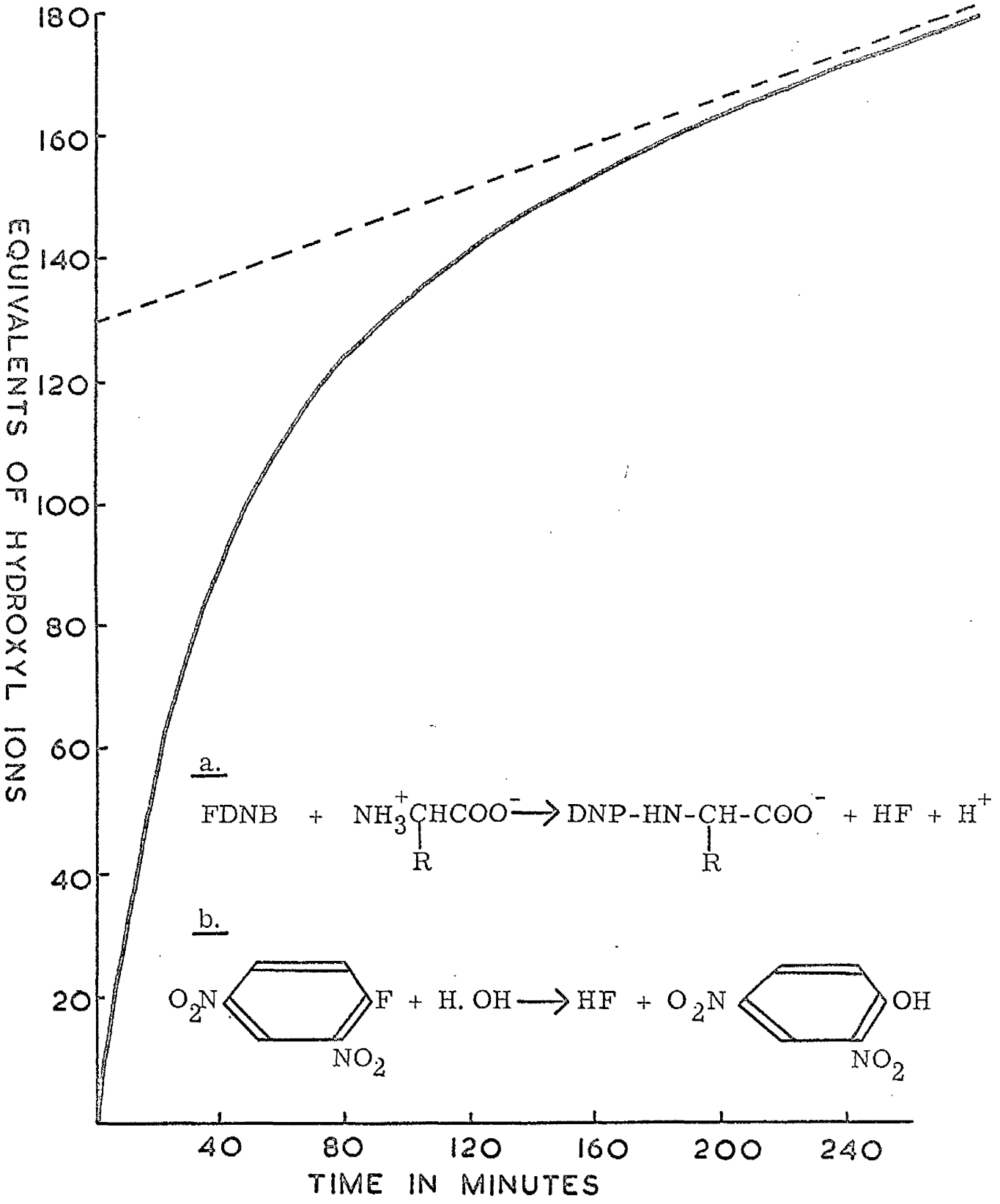
In the first method, which is the most convenient, 5-30 µmoles of the peptides to be reacted were conveyed in 2-3ml of distilled water to a 50 ml round-bottomed flask provided with a ground glass stopper. One millilitre of carbonate-hydrogen carbonate buffer pH 9 and 100 µl of fluorodinitrobenzene were added and after the stopper had been well secured by rubber bands the whole was shaken in the dark for 2 hr. at 40°.

With the method of Levy (1954), if the additions of alkali required to maintain the pH at a constant

FIGURE 11.

Graph obtained of alkali uptake during the reaction of FDNB with the products of cyanogen bromide treated cytochrome c when performed on a pH stat at pH 9. The initial rise in the curve is due to reaction of the amino groups with FDNB (a.). The final slow constant alkali uptake is due to decomposition of excess FDNB in the alkaline medium (b.). Extrapolation to time zero of this final slope as shown gives the alkali uptake due to reaction a alone.

FIGURE 11.


180
160
140
120
100
80
60
40
20
EQUIVALENTS OF HYDROXYL IONS
40
80
120
160
200
240
TIME IN MINUTES
a.
FDNB + NH3CHCOO → DNP-HN-CH-COO + HF + H+
R
R
b.
O2N
F + H. OH → HF + O2N
OH
NO2
NO2

value are carried out by an autotitrator, it is possible to follow the course of the reaction (Fig.11). Five to thirty micromoles of the peptides in 2-4 ml of water, 2 ml of 0.1M KCl and 100 µl of fluorodinitrobenzene were stirred vigorously in the reaction cell of a Radiometer autotitrator. The pH was maintained at 9 by autotitration of standard NaOH. A linear uptake of alkali (of the order of 0.044 µmoles of OH^- per ml per min) due to the decomposition of the fluorodinitrobenzene to dinitrophenol in the alkaline medium indicates that the reaction of the amino acids has reached completion. Extrapolation to zero time of the final linear slope due to dinitrophenol formation gives the number of µ moles of alkali taken up during the condensation reaction with the amino acids. This gives an indication of the extent to which the reaction has proceeded.

When the reaction by either method was over, the contents of the reaction vessel were transferred quantitatively to a hydrolysis tube and acidified with a few drops of 6 N HCl. The precipitate which appears on acidification was centrifuged down, washed with water, twice with ethanol, and finally twice with ether. The residue of ether was removed by evaporation, 1 ml of 6 N HCl added, the tube sealed, and hydrolysis carried out at 115° for 8 hr.

After hydrolysis the normality of the HCl was reduced by addition of 5 ml of water and the ether soluble DNP amino acids extracted with three 5 ml portions of ether. The extracts were combined in a sublimation tube and the solvent evaporated off by a current of air in a water bath at 70°. Dinitrophenol was then removed by sublimation under vacuum for 30 min. (Mills, 1952). The residue was taken up in 2-3 ml of ether, the solvent evaporated by a current of air and the residue again subjected to sublimation. Three such treatments were found to be sufficient to reduce the level of dinitrophenol such that it did not interfere with the chromatographic separation of the DNP amino acids. The residue was finally taken up in 5 ml of ether and a suitable amount used for paper chromatography.

The aqueous layer after ether extraction was diluted with water to 25 ml. A 5 ml sample was evaporated to dryness and the DNP amino acids taken up in a small volume of acid acetone (10 ml 0.1 N HCl made up to 100 ml with acetone) and applied to the paper for chromatography.

Two dimensional chromatography of suitable samples (50-500 µl of the ether soluble and 5 ml of the water soluble derivatives) on Whatman No 1 Paper was carried

FIGURE 12.

Two dimensional chromatogram of the ether soluble DNP amino acids performed in the solvent systems of Koch and Weidel (1956).

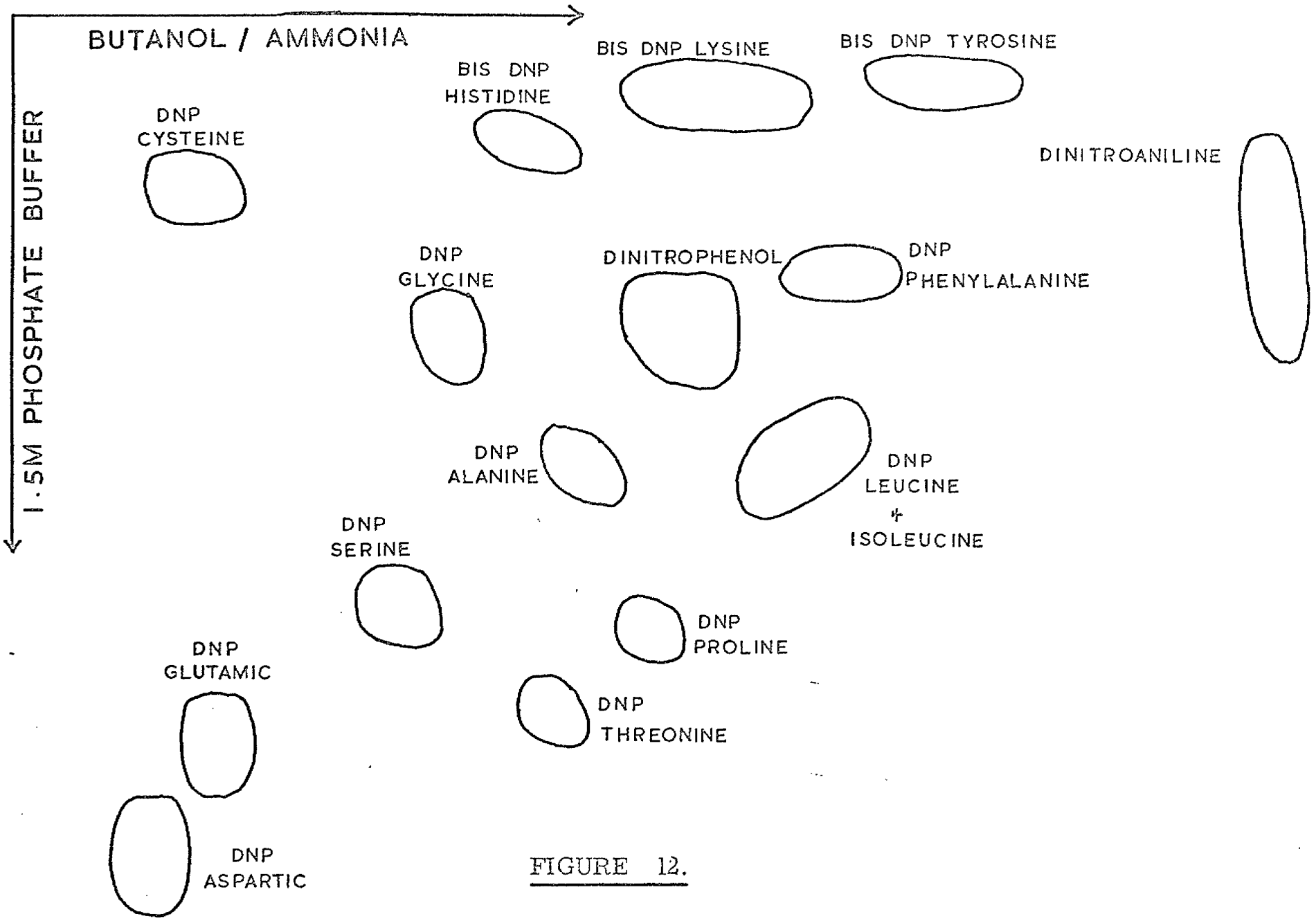

BUTANOL / AMMONIA
1.5M PHOSPHATE BUFFER
BIS DNP LYSINE
BIS DNP TYROSINE
BIS DNP HISTIDINE
DNP CYSTEINE
DINITROANILINE
DNP GLYCINE
DINITROPHENOL
DNP PHENYLALANINE
DNP ALANINE
DNP LEUCINE + ISOLEUCINE
DNP SERINE
DNP PROLINE
DNP GLUTAMIC
DNP THREONINE
DNP ASPARTIC

FIGURE 12.

FIGURE 13.

Two dimensional chromatogram of the water soluble DNP amino acids performed in the solvent systems of Hoch and Weidel (1956).

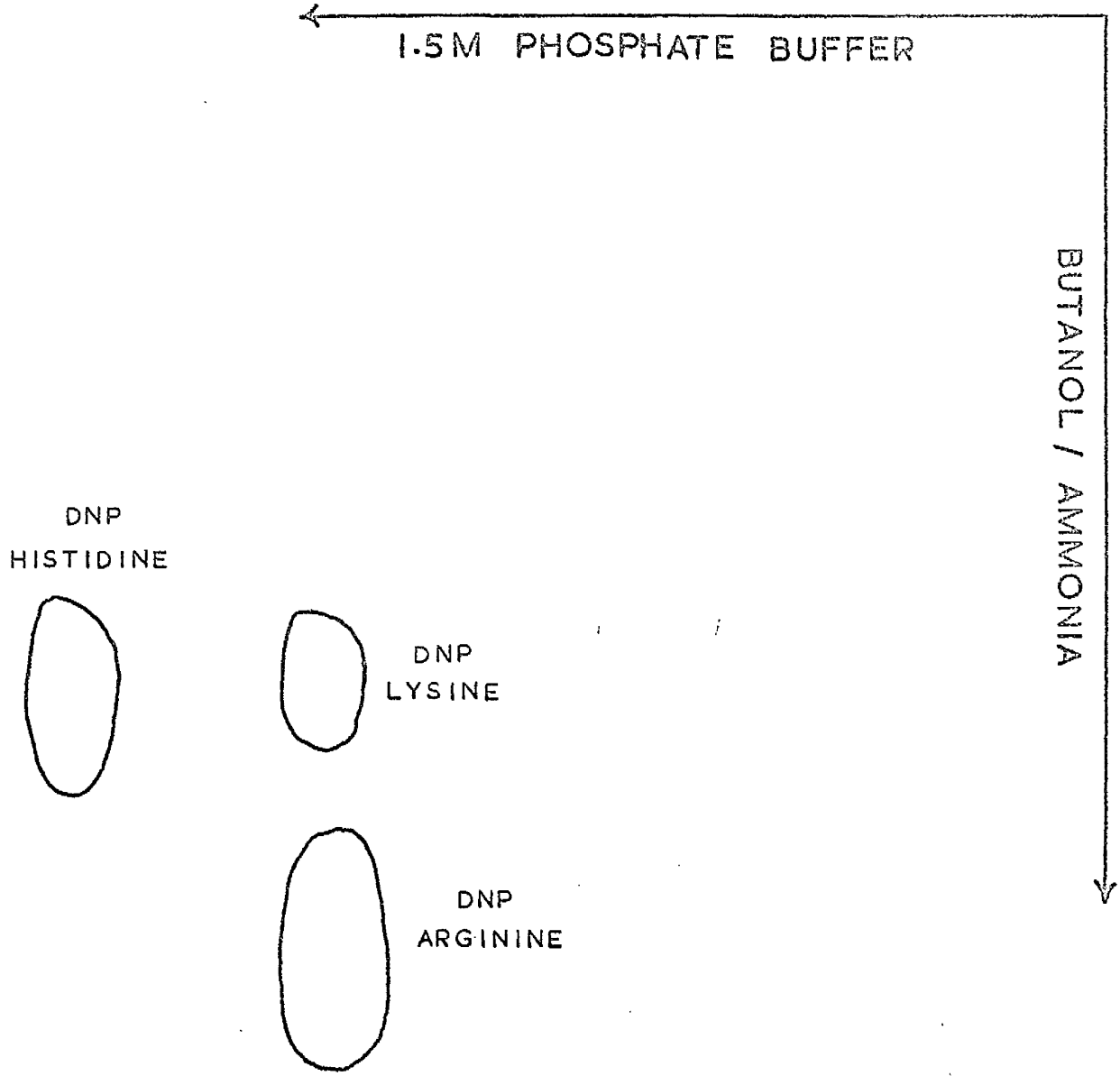

1·5 M PHOSPHATE BUFFER
BUTANOL / AMMONIA
DNP
HISTIDINE
DNP
LYSINE
DNP
ARGININE

FIGURE 13.

out by the system of Koch and Weidel (1956). Development of the first dimension in n-butanol saturated with 0.1% (w/v) ammonia was carried out descending for 24 hr. by which time the spot due to dinitroaniline had almost left the paper. The second dimension was run descending for 16 hr. in 1.5 M phosphate buffer pH 6. After drying the outlines of the spots were marked in pencil while viewed in ultraviolet light. In practice it was found that the spots due to bis DNP lysine and bis DNP tyrosine were never completely separated. The DNP leucines and DNP valine, with also possibly DNP methionine were always present as one rather diffuse spot (Figs.12 and 13). In some cases DNP phenylalanine was not completely separated from dinitrophenol.

The subtractive method of Hirst (1960) was used as a check on the determined N-terminal amino acid. A suitable aliquot of the water soluble derivatives was examined for free amino acids by automatic amino acid analysis. Comparison with prior analyses of the protein or peptide showed which amino acid had been removed as the ether soluble DNP derivative.

RESULTS

RESULTS.

Cytochrome c.

Native cytochrome c was reacted with cyanogen bromide under the conditions specified and after lyophilisation, the products were applied to the column of Sephadex G.50. Figure 14 shows the curve of optical densities at 280 and 415 mμ of the effluent fractions. The inflections in the curve suggested the presence of four components, the first of which will be due to unsplit cytochrome c. The fractions indicated in the figure were segregated, lyophilised and further purified.

In an attempt to improve the separation from cytochrome c, the porphyrin containing fraction 1A was run through the column of Sephadex G 75. Figure 15 shows the results obtained. The shoulder on the left hand side of the peak was proof of a partial separation. Fraction 1A was rerun through Sephadex G 50 to obtain purification from the contaminant indicated by the crossing of the 280 and 415 mμ curves on the right hand side of the peak. The shoulder on the right hand side of the peak in Figure 16 shows that a further separation had been achieved. Since the upper part of the peak was still skew, suggesting the continuing presence of cytochrome c,

fraction $1A_2$ was again applied to the Sephadex G 75 column. Comparison of the resulting Figure 17 with the previous pattern obtained from G 75 (Fig.15) shows that considerable purification had been achieved. Fraction $1A_3$ was used for amino acid analysis.

Considerable purification of fraction 2 was obtained by rechromatography on Sephadex G 50 (Fig.18). The readings at 415 mμ however showed that some of the porphyrin containing peptide 1A was still present. This was removed by reapplication of fraction 2_2 to the Sephadex G 50 (Fig.19). Fraction 2_3 was used for hydrolysis and amino acid analysis.

Fraction 3 was purified by a single run on Sephadex G 50 (Fig.20). Fraction 3_2 was used for amino acid analysis.

The amino acid analyses of the three fractions are shown in Table 4. Table 5 shows a comparison of these analyses with the amino acid composition to be expected from the amino acid sequence of Margoliash et al (1961). It can be seen that the amino acid composition of fraction 1A differs from the theoretical values. Such an amino acid composition could arise if fraction 1A were a mixture of the expected peptide and cytochrome c.

Tryptic digests of cytochrome c and fraction 1A

were fingerprinted with the results shown in Figure21. The obvious difference suggested that fraction 1A did not contain sufficient cytochrome c to seriously influence the amino acid analysis. Electrophoresis of fraction 1A on cellulose acetate at pH 3.4 and 9 gave a single band with a lower mobility than that obtained for cytochrome c. A mixture of fraction 1A and cytochrome c gave two bands on electrophoresis.

Further examination of the amino acid sequence of Margoliash et al. showed that fraction 1A could be accounted for by the cleavage of the methionine peptide bond nearest to the C-terminal end of the protein to give only two peptides (Table 6). Electrophoretic examination of the products of cytochrome c cleavage on cellulose acetate at pH 9 gave five bands of which three were coloured. The coloured bands are due to cytochrome c, fraction 1A isolated in these experiments, and the haem containing peptide expected from 100% cleavage of cytochrome c.

Qualitative N-terminal determination of split cytochrome c gave glutamic acid and isoleucine in accordance with the sequence of Margoliash et al (1961).

FIGURE 14.

Separation of the products of cyanogen bromide treated cytochrome c on Sephadex G 50. Conditions for chromatography are as described in the Experimental.

——— Optical density at 280 mμ.

----- Optical density at 415 mμ.

1←—1A—→1 Indicates the tubes from which the appropriate fraction (1A etc.) described in the text was obtained.

FIGURE 14.

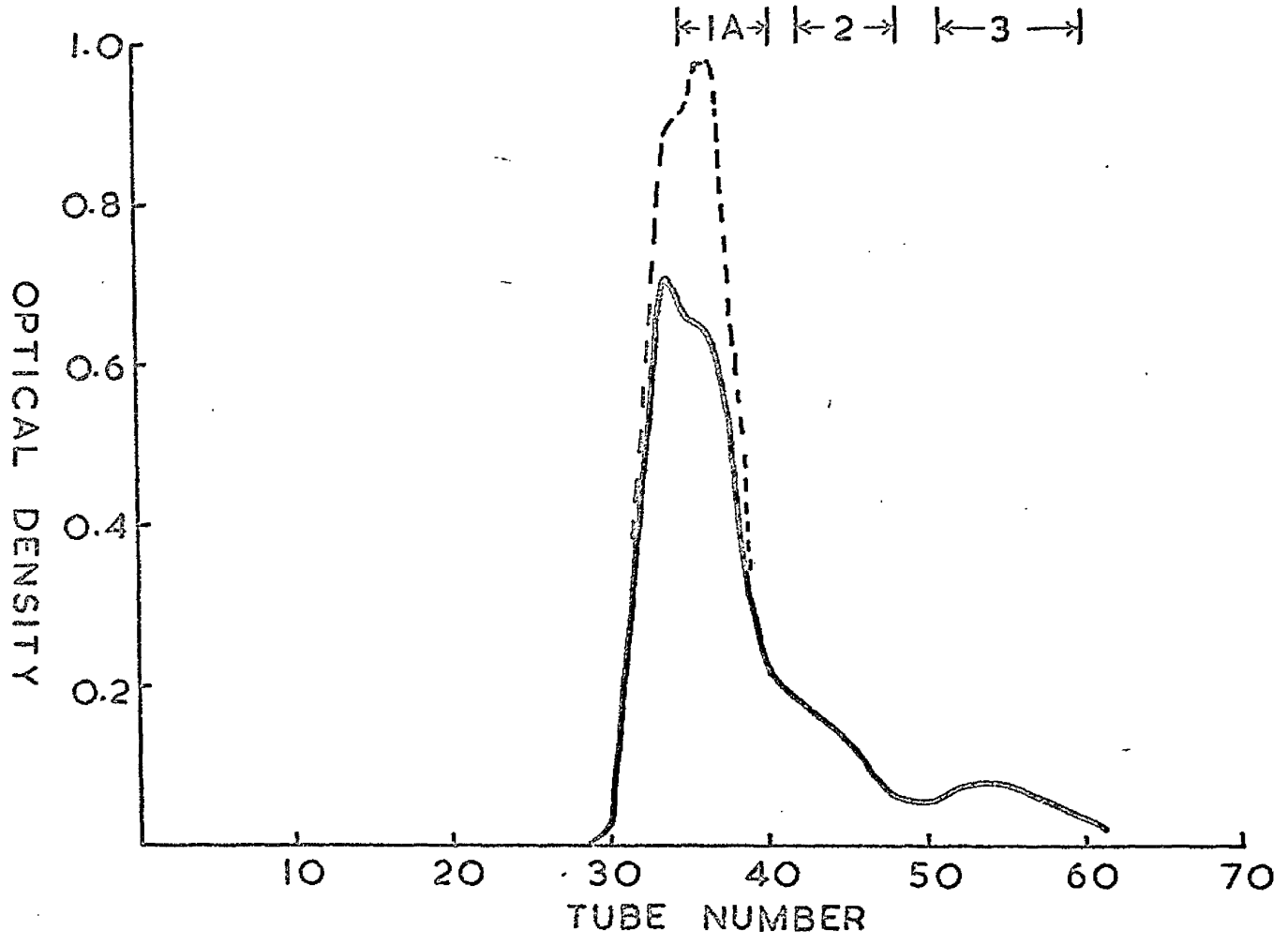

1A
2
3
1.0
0.8
0.6
0.4
0.2
OPTICAL DENSITY
10
20
30
40
50
60
70
TUBE NUMBER

FIGURE 15.

Separation obtained when fraction 1A (Fig. 14) was applied to the column of Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

----- Optical density at 415 mμ.

The tubes containing the fraction taken for further purification, $1A_1$, are indicated at the top of the figure.

FIGURE 15.

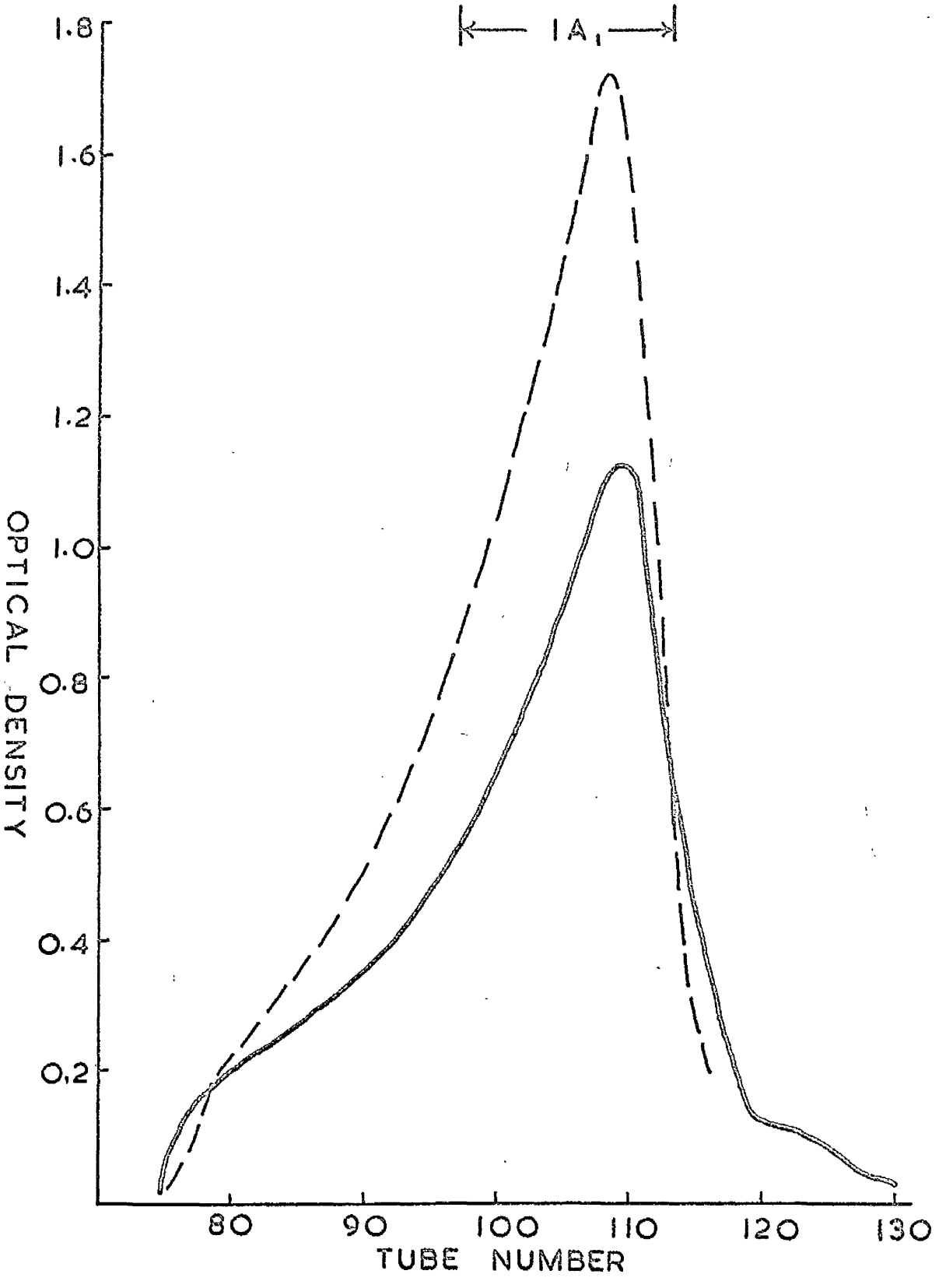

1A1
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
OPTICAL DENSITY
80
90
100
110
120
130
TUBE NUMBER

FIGURE 16.

Separation of $1A_1$ (Fig. 15) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

----- Optical density at 415 mμ.

Fraction $1A_2$ isolated from the area shown.

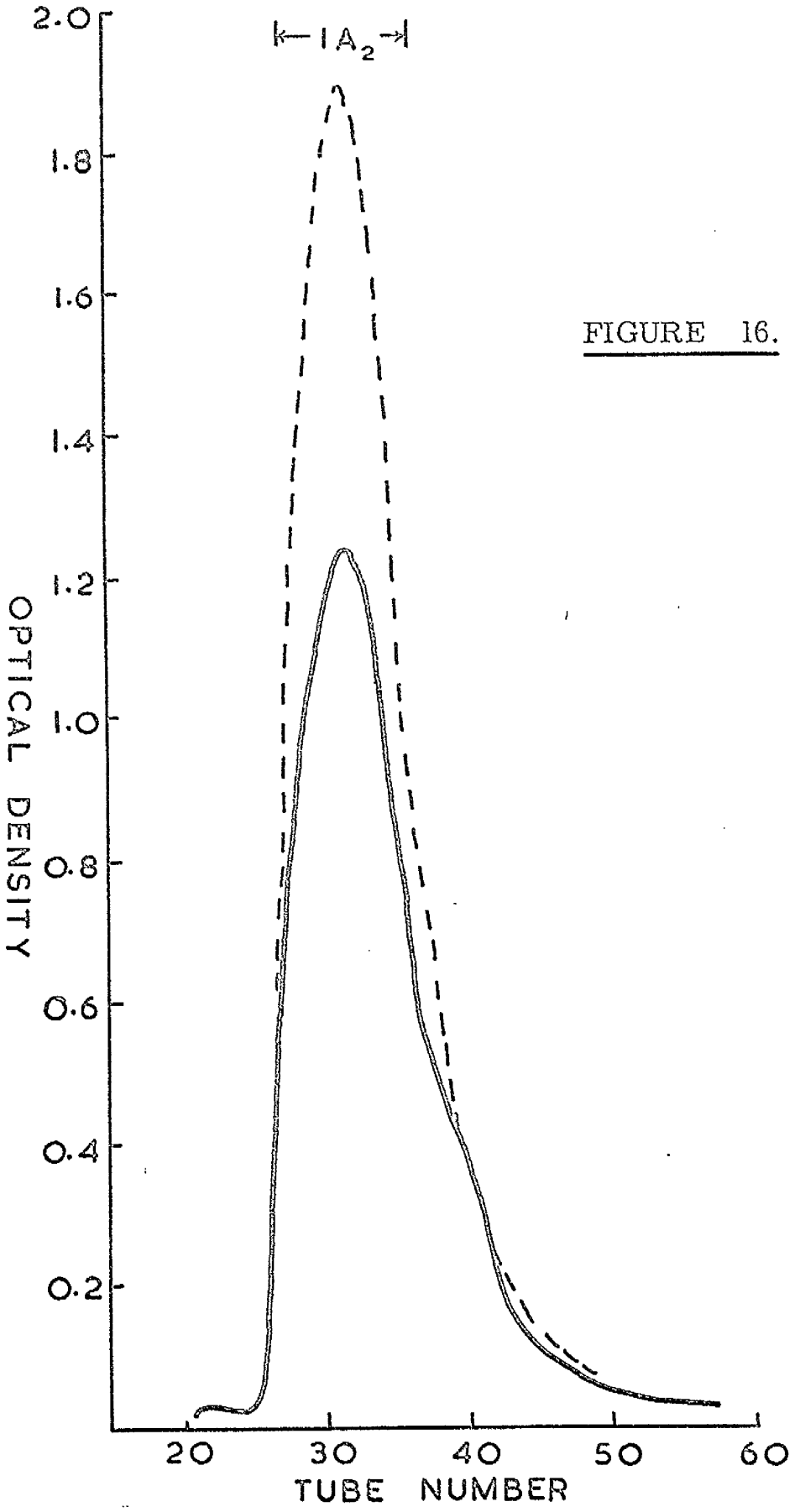

2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
OPTICAL DENSITY
A_2
20
30
40
50
60
TUBE NUMBER

FIGURE 16.

FIGURE 17.

Separation of $1A_2$ (Fig. 16) on Sephadex G 75.

Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

– – – – Optical density at 415 mμ.

Fraction $1A_2$ obtained from the area shown.

FIGURE 17.

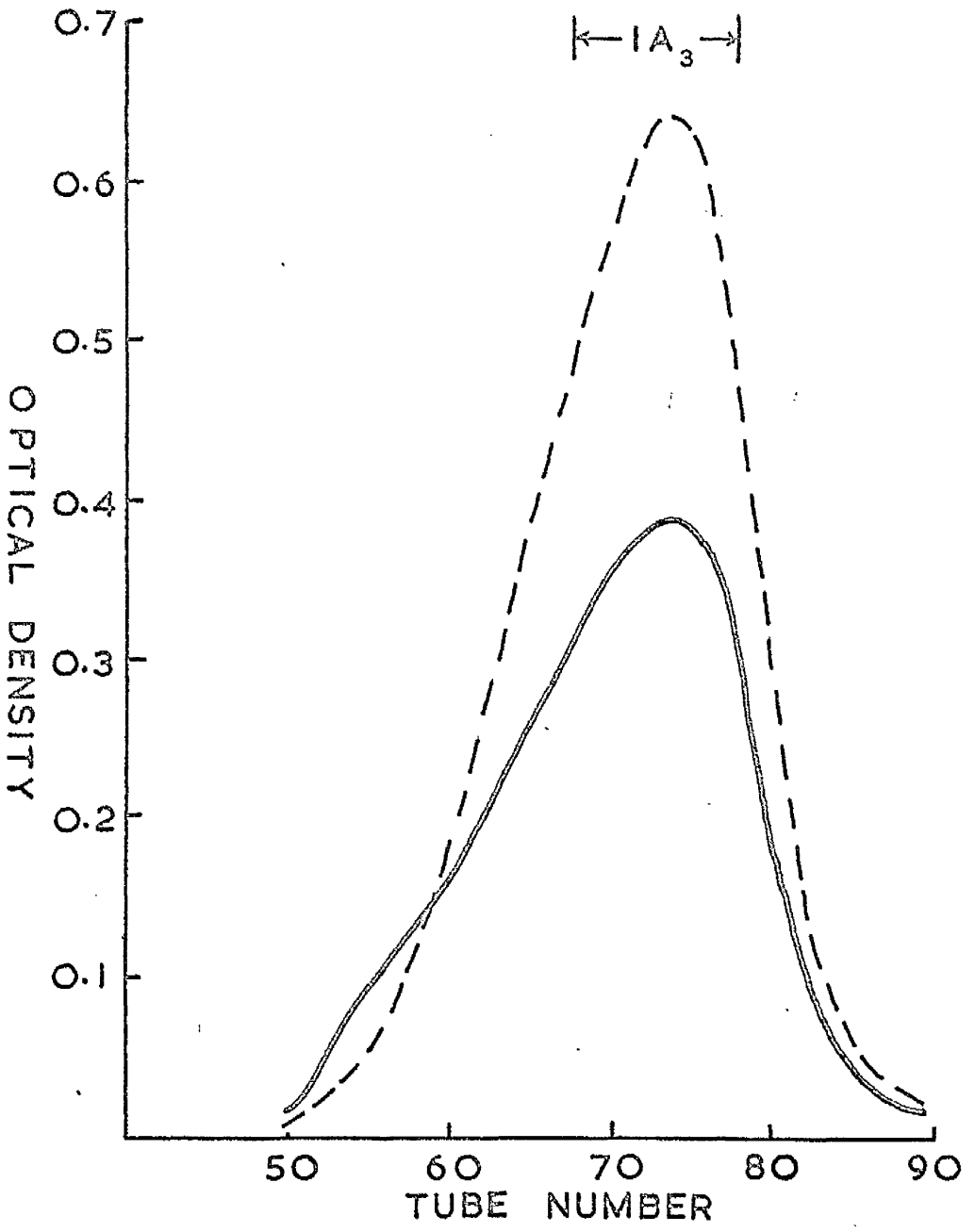

0.7
0.6
0.5
0.4
0.3
0.2
0.1
OPTICAL DENSITY
A_3
50
60
70
80
90
TUBE NUMBER

FIGURE 18.

Separation of fraction 2 (Fig. 14) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

----- Optical density at 415 mμ.

Fraction 2_2 isolated from the indicated area.

FIGURE 19.

Separation of fraction 2_2 (Fig. 18) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical Density at 280 mμ.

Fraction 2_3 obtained from the indicated area.

FIGURE 18.

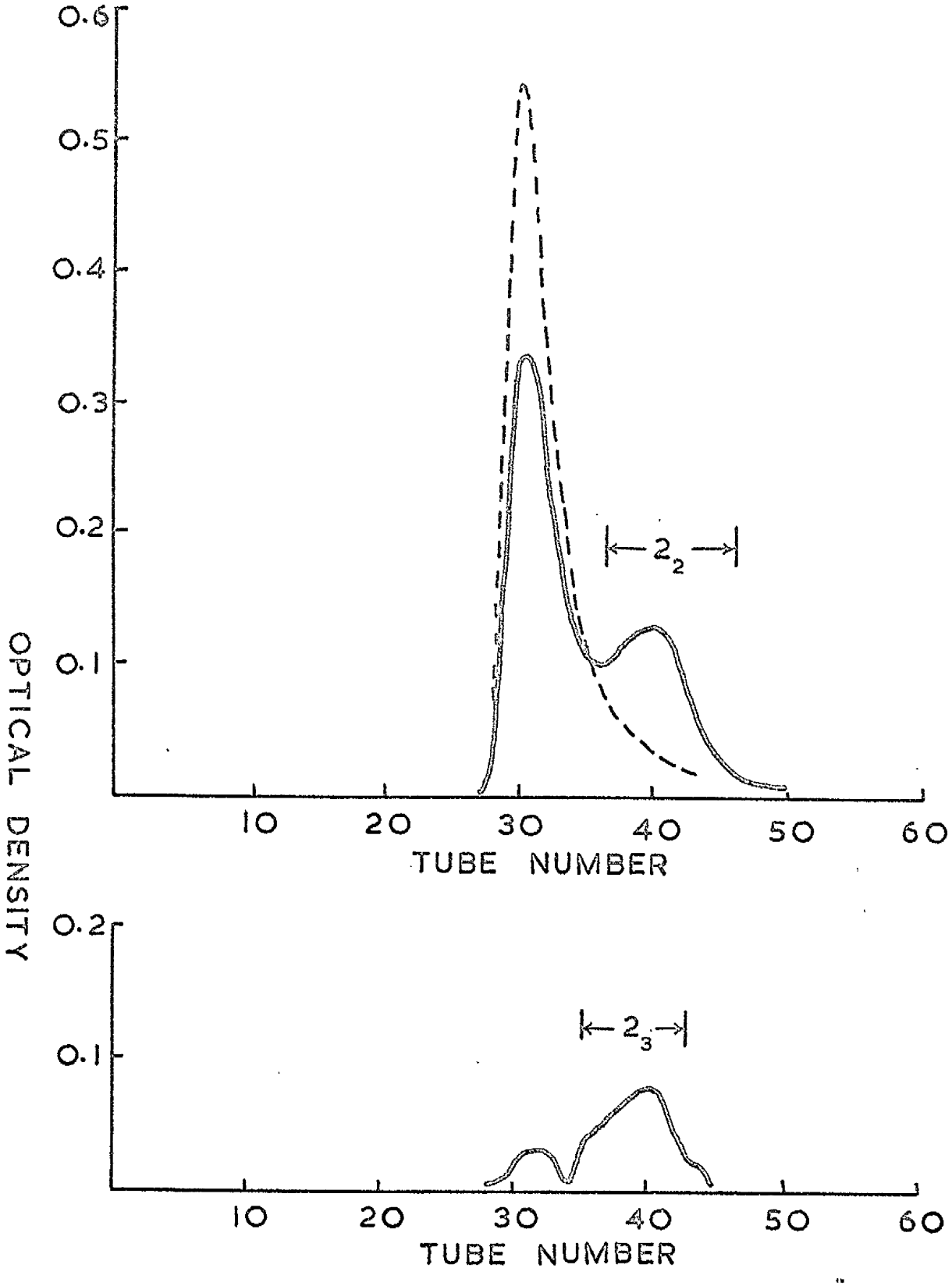


FIGURE 19.

FIGURE 20.

Purification of fraction 3 (Fig. 14) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction 3_2 obtained from the indicated tubes.

FIGURE 20.

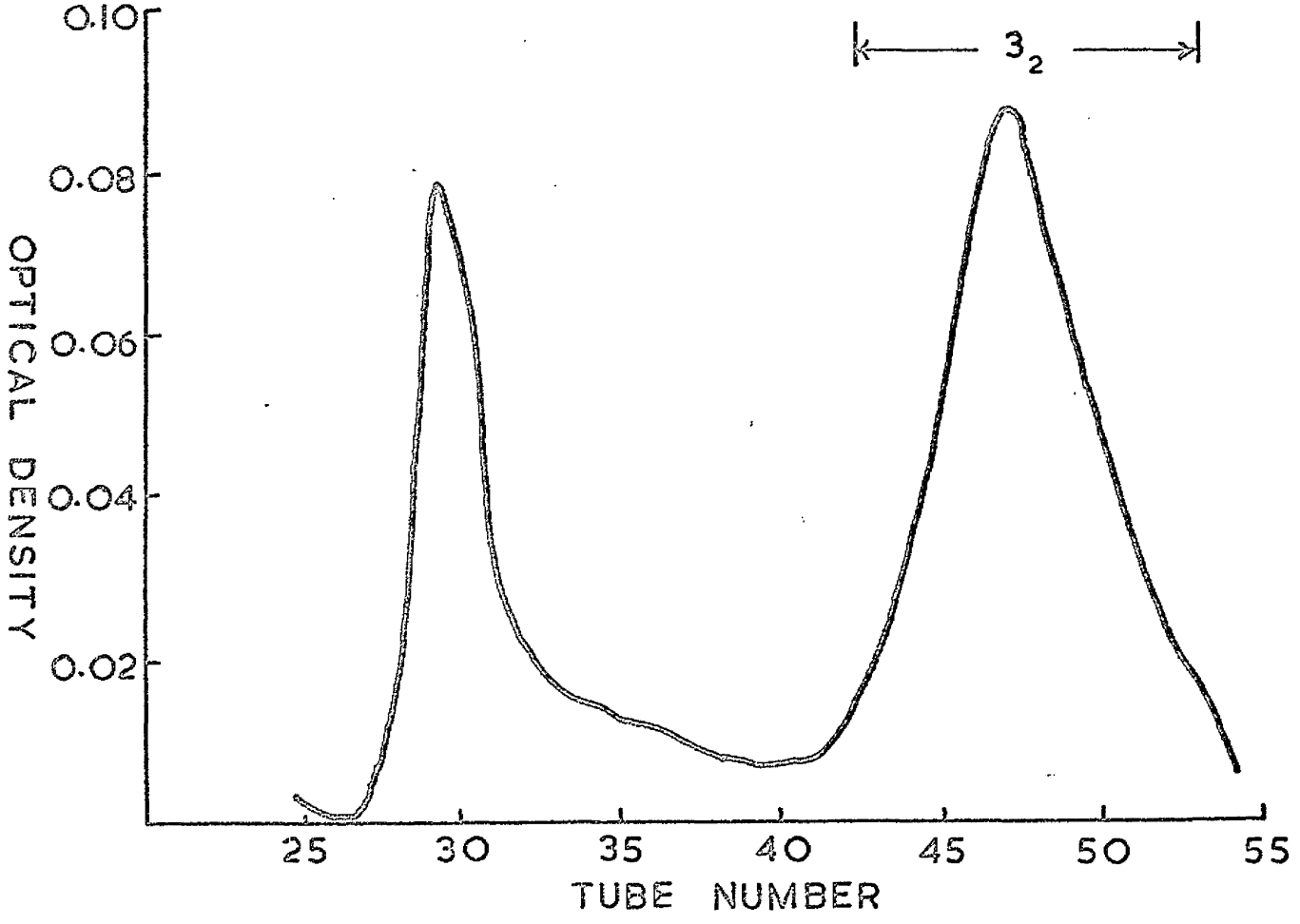
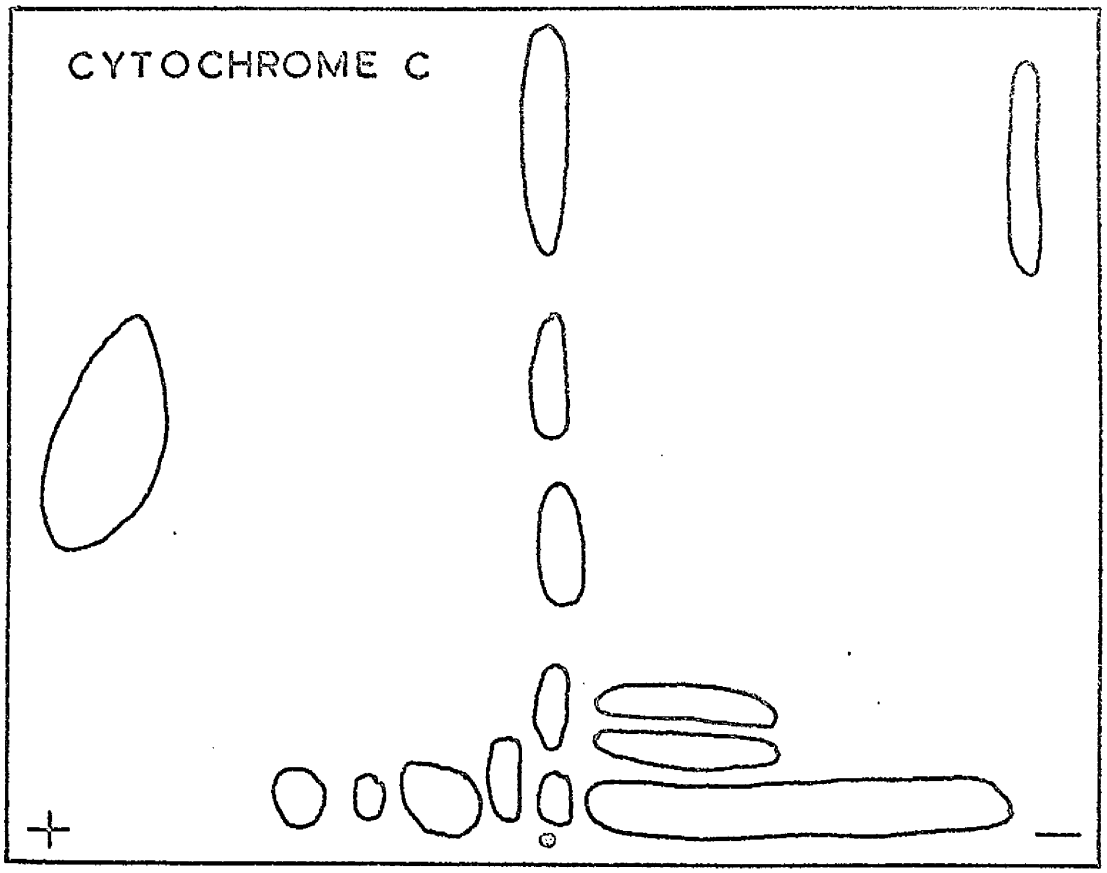

0.10
0.08
0.06
0.04
0.02
OPTICAL DENSITY
3_2
25
30
35
40
45
50
55
TUBE NUMBER

FIGURE 21.

Fingerprint comparisons of tryptic digests of cytochrome c and peptide 1A. Point of origin at bottom, center. Fingerprints obtained by the method of Ingram (1958) with slight modifications.

FIGURE 21.



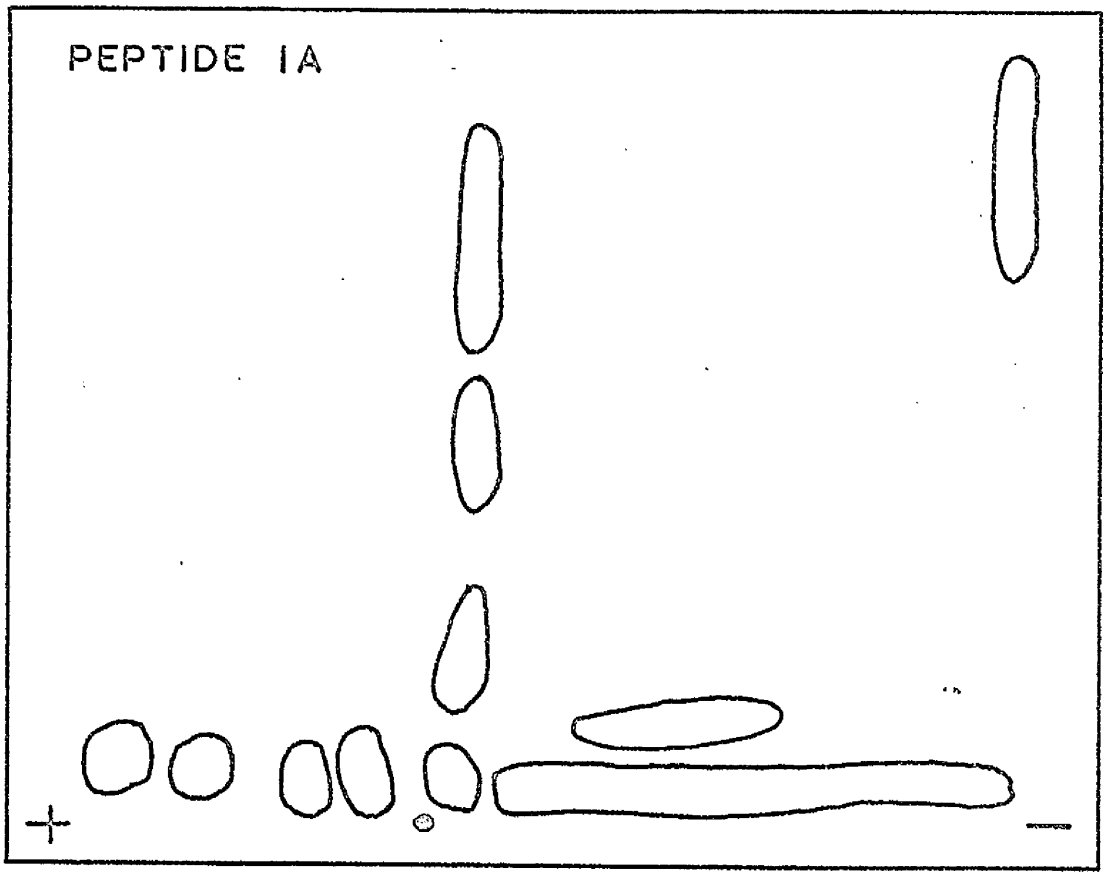


TABLE 4.

Amino acid analyses of peptides 1A, 2 and 3 from cytochrome c.

	1A	2	3
Asp	6.1	2.1	1.0
Thr.	3.6	1.8	0.8
Glu.	9.1	3.1	2.0
Pro.	2.9	0	2.0
Gly.	9.8	1.1	1.0
Ala.	4.0	2.8	0
Val.	2.9	0	0
Ileu.	3.5	2.9	0.9
Leu.	4.1	2.1	0.9
Tyr.	1.3	0.6	0.7
Phe.	3.1	0.9	0
Lys.	+	5.1	3.0
Arg.	1.2	1.0	0
Cys. A.	4.5	0	0

+ present but not determined for the reasons given in the text.

TABLE 5.

Comparison of peptides 1A, 2 and 3 from cytochrome c with the expected sequences.

	1A	1-65	2	80-104	3	65-80
Cys.AA	5	2	0	0	0	0
Asp.	6	5	2	2	1	1
Thr.	4	7	2	2	1	1
Glu.	9	7	3	3	2	2
Pro.	3	2	0	0	2	2
Gly.	10	10	1	1	1	1
Ala.	4	3	3	3	0	0
Val.	3	3	0	0	0	0
Ileu.	4	3	3	3	1	1
Leu.	4	3	2	2	1	1
Tyr.	1	1	1	1	1	2
Phe.	3	3	1	1	0	0
Lys.	+	11	5	5	3	3
His.	3	3	0	0	0	0
Arg.	1	1	1	1	0	0

+present but not determined for reasons given in the text.

TABLE 6.

Comparison of the amino acid composition of peptide 1A with that for the expected sequence 1-65 and the sequence which would be obtained by incomplete cleavage, 1-80.

	Found	1-65	1-80
$CySO_3H$	4.5	2	2
Asp.	6.1	5	6
Thr.	3.6	7	8
Glu.	9.1	7	9
Pro.	2.9	2	4
Gly.	9.8	10	11
Ala.	4.0	3	3
Val.	2.9	3	3
Met.	0	0	1
Ileu.	3.5	2	3
Leu.	4.1	3	4
Tyr.	1.3	1	3
Phe.	3.1	3	3
Lys.	-	11	14
His.	2.9	3	3
Arg.	1.2	1	1

Myoglobin.

A peptide map of the products of cleavage of myoglobin with cyanogen bromide (Fig. 22) indicated that there were probably five components. Assuming one of the components to be unchanged myoglobin, this is one more than demanded by theory from the methionine content of horse heart myoglobin.

The partial separation obtained by chrometography of the cyanogen bromide treated myoglobin on Sephadex G 75 is shown in Figure 23. In this case the inflections in the curves show resolution into unsplit myoglobin and possibly three other fractions. The purification of fractions 1A, 1B, and 2 on Sephadex G 75 is shown in figures 24, 25, and 26. Approximately 2 mg of each of the purified fractions was hydrolysed and used for amino acid analysis.

The shoulder on the right hand side of the main peak in Figure 26 suggested the presence of a peptide smaller than peptide 2. As shown in Figure 27 a better separation of this peptide, peptide 3, was obtained on Sephadex G 50. The further purification of peptide 3 on Sephadex G 50 is shown in in Figures 28 and 29. Table 7 shows the results of the amino acid analyses of the four peptides and of uncleaved horse heart myoglobin.

The N-terminal amino acid of each peptide was determined by the fluorodinitrobenzene method. The results were checked by amino acid analysis of an aliquot of the water-

-soluble DNP amino acids and free amino acids obtained in the N-terminal determination after ether extraction of the other soluble DNP amino acids. Comparison with the previous amino acid analyses showed which amino acid had been removed as the ether soluble DNP derivative. The results of the amino acid analyses after N-terminal reaction are shown in Table 8, together with the result of the N-terminal determination.

FIGURE 22.

Fingerprint of the products of cyanogen bromide action on myoglobin. Development by the method of Ingram with slight modifications.

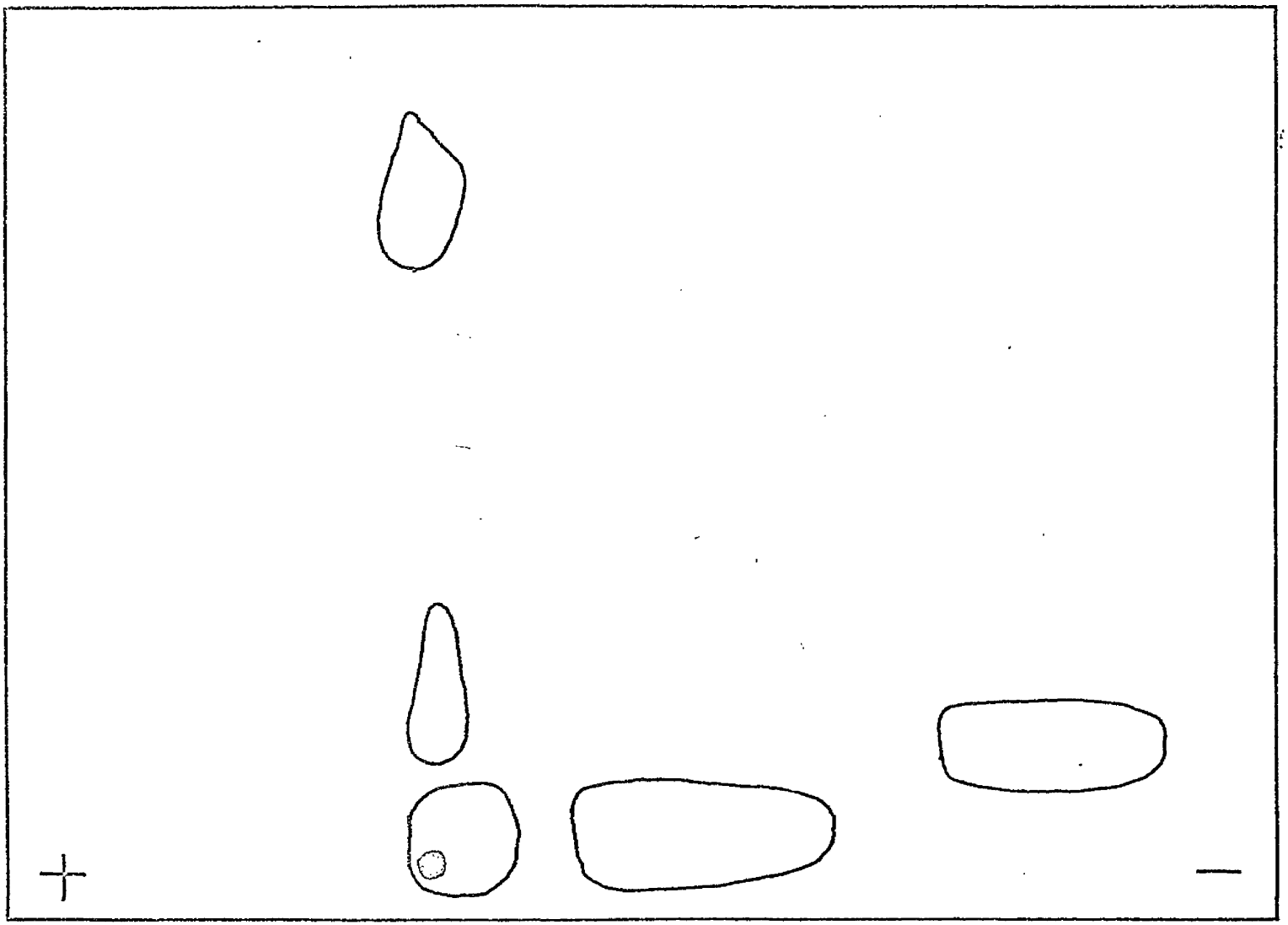


FIGURE 22.

FIGURE 23.

Separation of the products of cyanogen bromide action on myoglobin by chromatography on Sephadex G 75.

Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

1←1A→1 indicates the tubes from which the appropriate fraction (1A etc.) was obtained.

FIGURE 23.

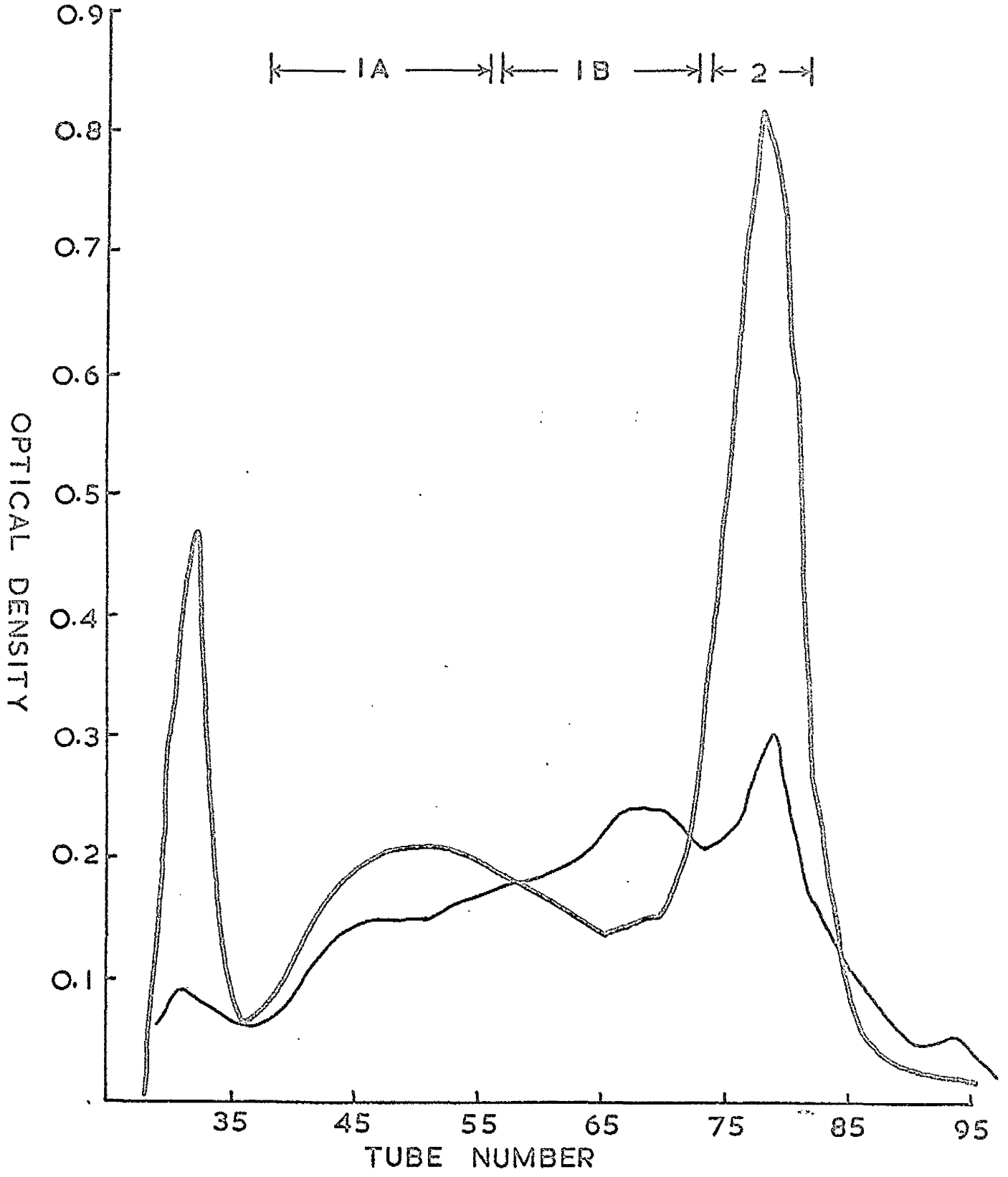

0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
OPTICAL DENSITY
1A
1B
2
35
45
55
65
75
85
95
TUBE NUMBER

FIGURE 24.

Purification of fraction 1A (Fig. 23) on Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

Fraction $1A_2$ obtained from the tubes indicated.

FIGURE 25.

Purification of fraction 1B (Fig. 23) on Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

Fraction $1B_2$ obtained from the tubes indicated.

FIGURE 24.

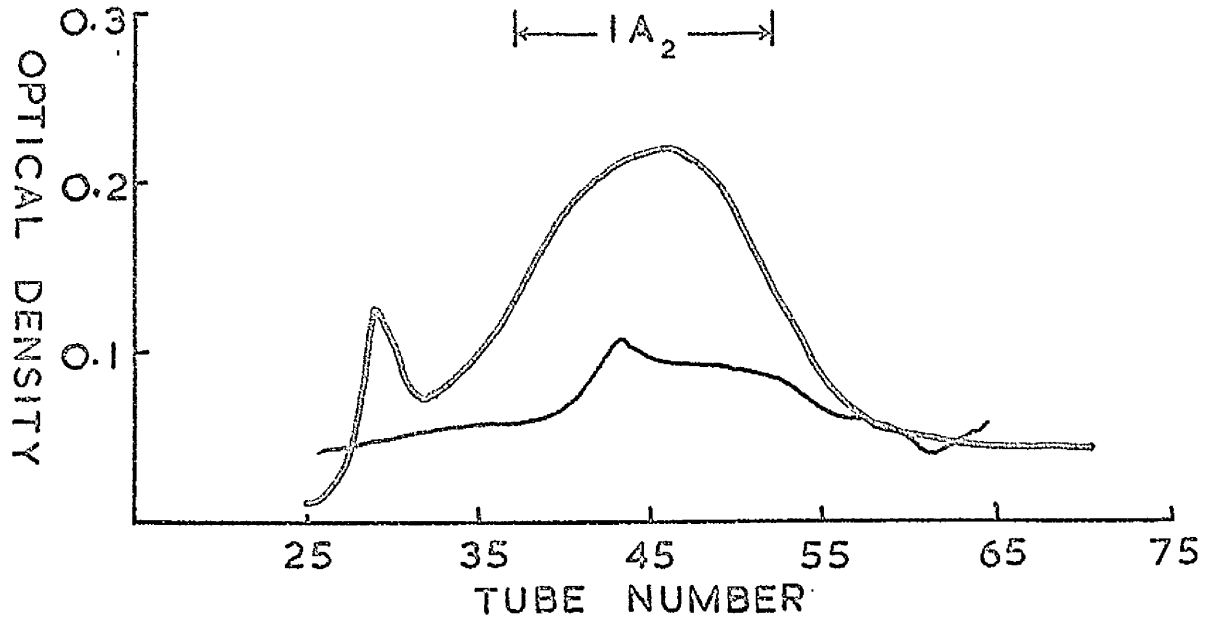


FIGURE 25

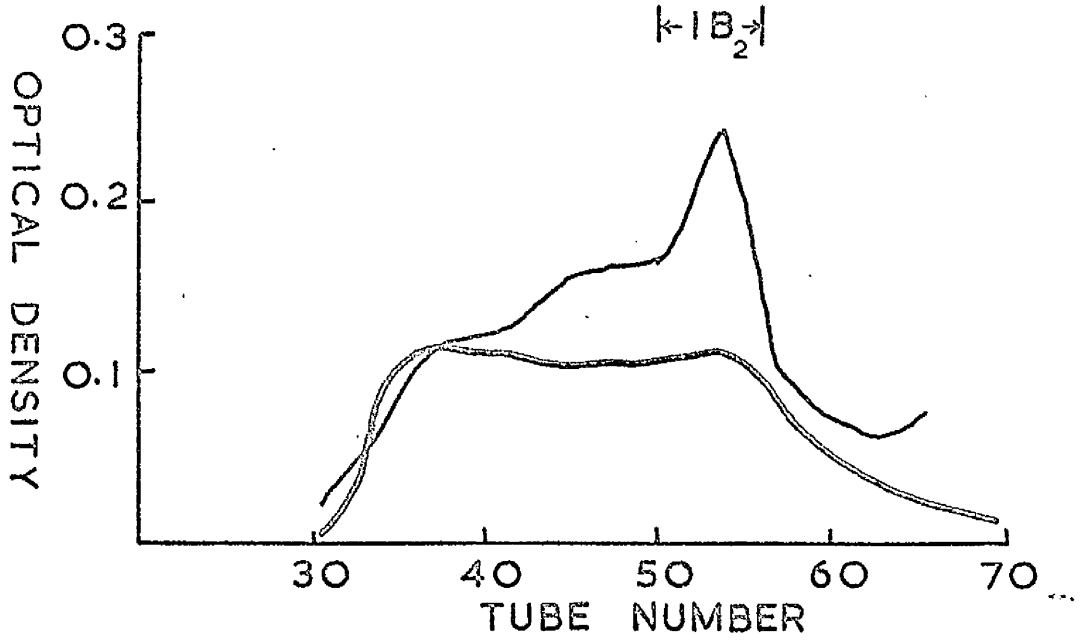


FIGURE 26.

Purification of fraction 2 (Fig. 23) on Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction 2_2 taken from the indicated area.

FIGURE 26.

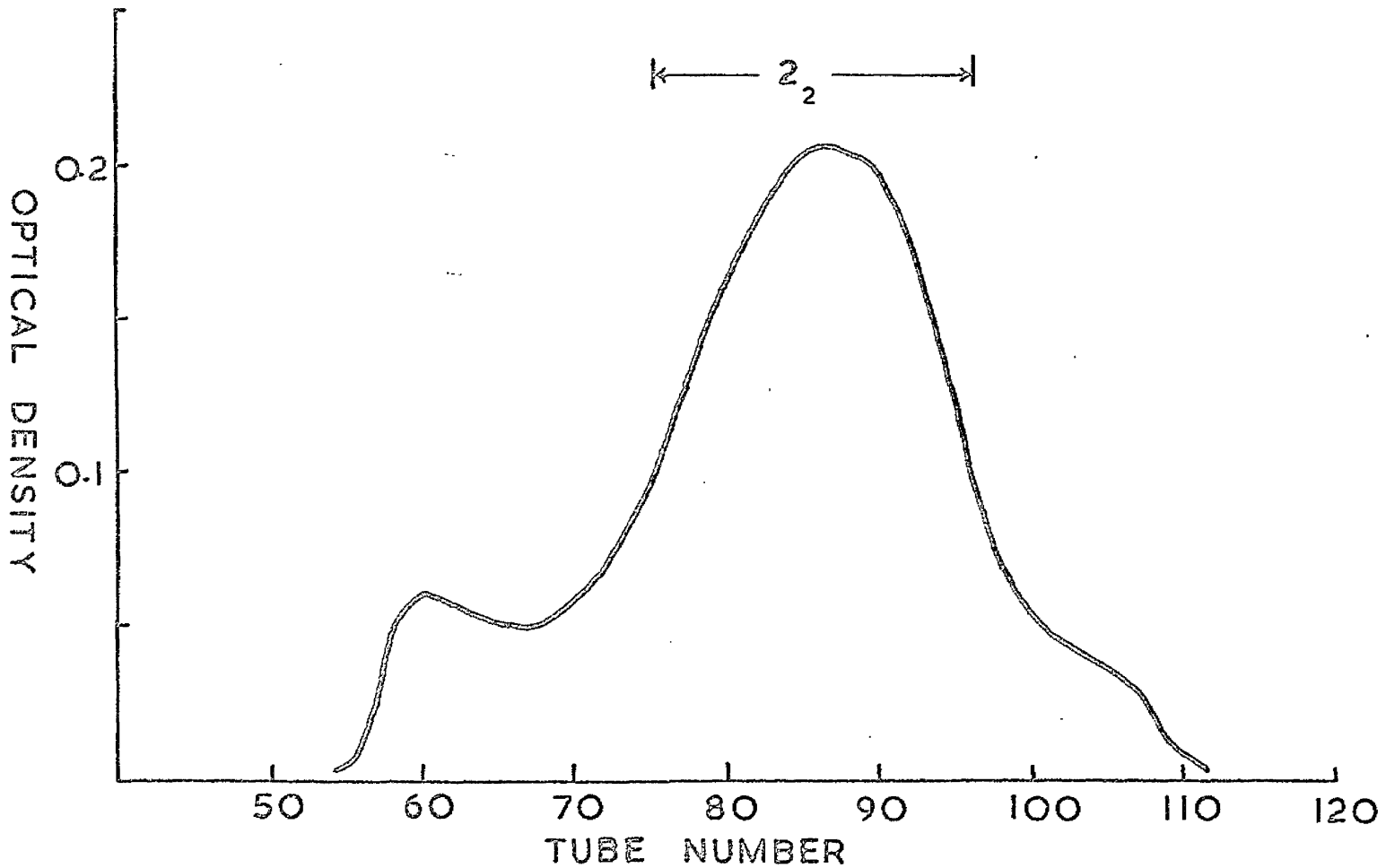


FIGURE 27.

Separation of products of cyanogen bromide action on myoglobin by chromatography on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction 3 obtained from the tubes marked.

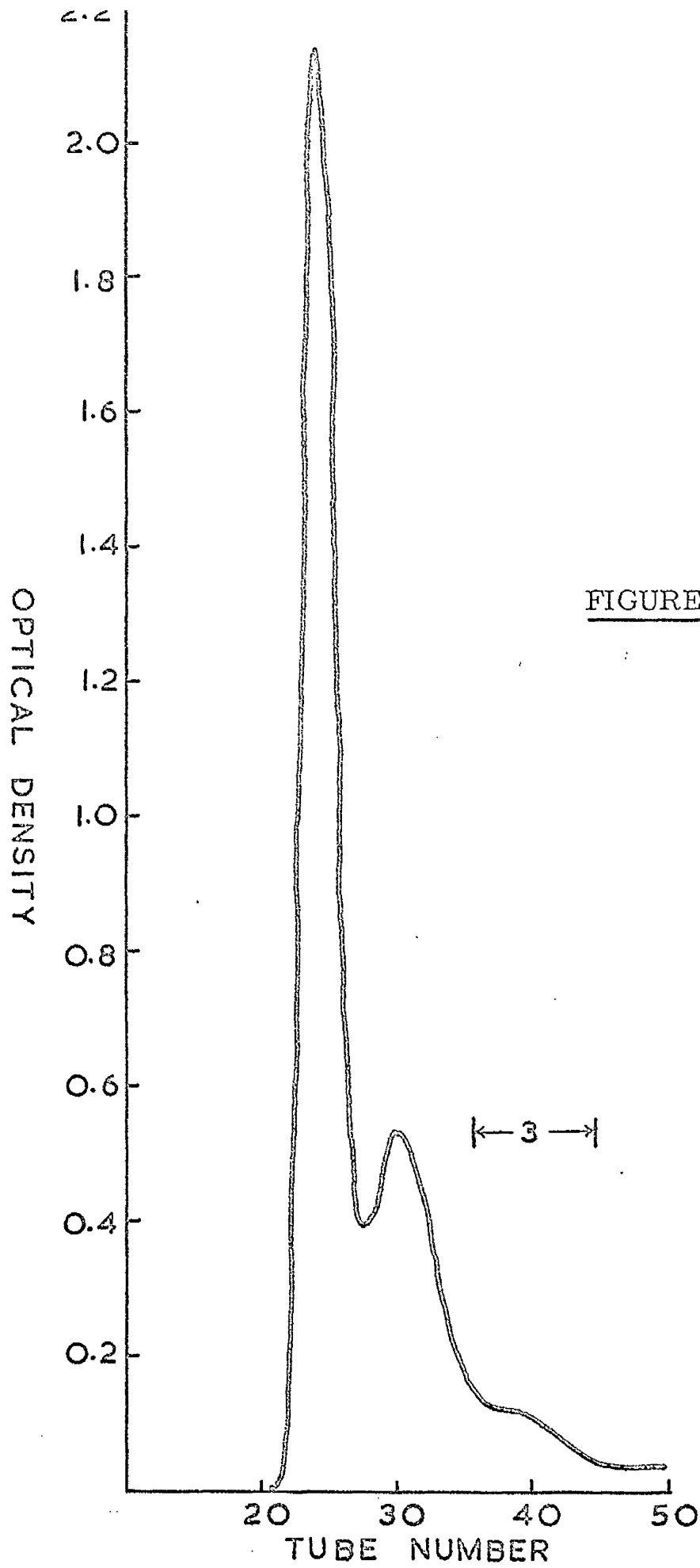

2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
OPTICAL DENSITY
3
20
30
40
50
TUBE NUMBER

FIGURE 27.

FIGURE 28.

Purification of fraction 3 (Fig. 27) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction 3_1 isolated from the tubes indicated.

FIGURE 29.

Separation of fraction 3_1 (Fig. 28) on Sephadex G 50. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction 3_2 isolated from the indicated area.

FIGURE 28.

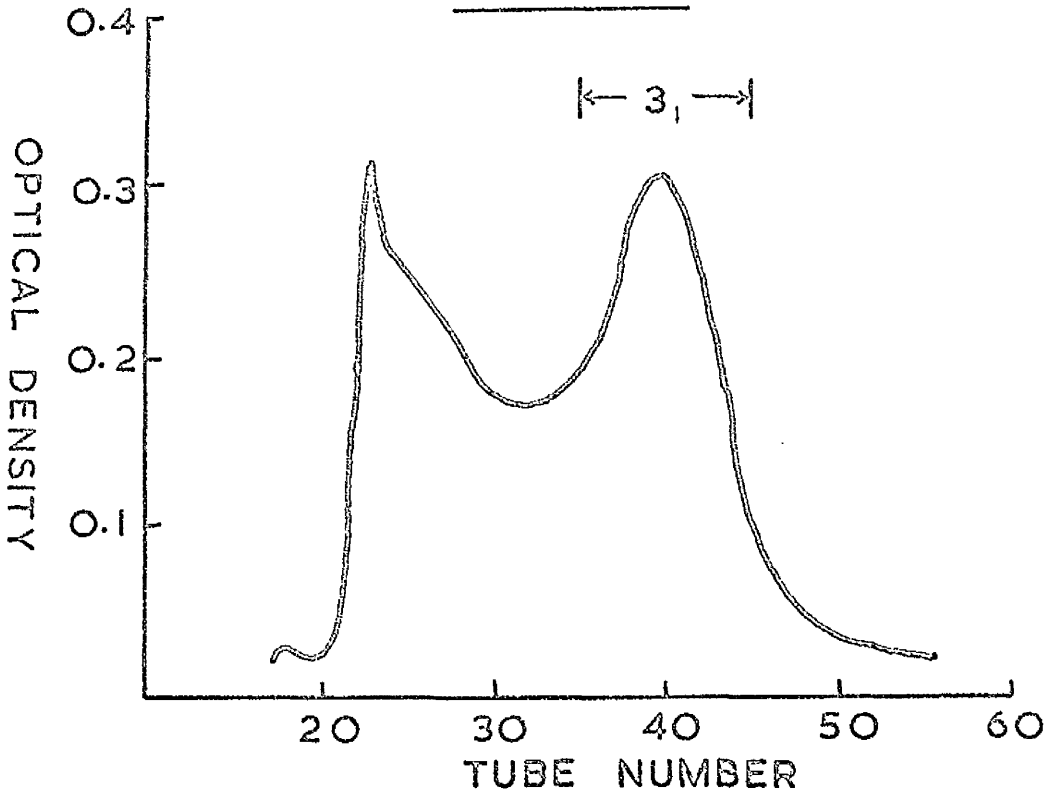


FIGURE 29.

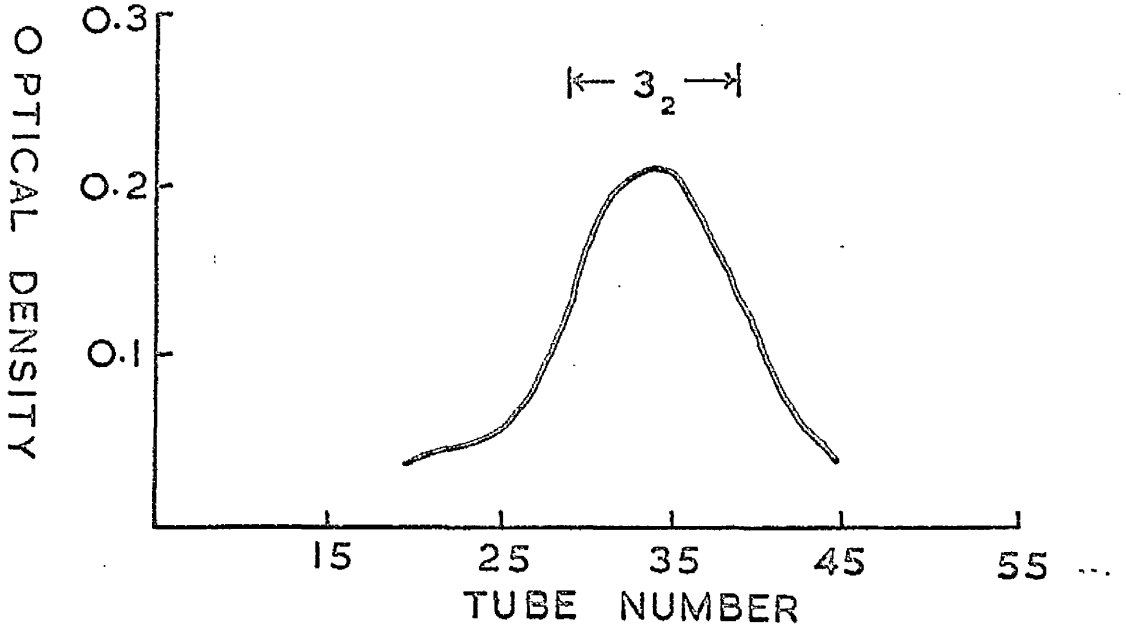


TABLE 7.

Amino acid analyses of myoglobin and the peptides 1A, 1B, 2 and 3.

	Myoglob.	1A	1B	2	3
Asp.	9.76	8.7	4.0	3.9	2.9
Ser.	5.41	5.7	2.9	0.9	1.3
Thr.	3.37	3.3	2.5	0.2	0.4
Glu.	20.1	16.6	6.9	8.9	4.2
Pro.	-	3.6	2.9	1.1	0.3
Gly.	14.8	12.3	6.9	5.7	3.0
Ala.	14.2	11.5	8.2	4.3	4.0
Val.	7.26	6.9	3.1	3.9	0.7
Ileu.	8.37	7.0	5.1	2.3	1.3
Leu.	17.0	13.4	8.0	5.9	3.9
Tyr.	1.0	0.7	0.7	0.2	0.6
Phe.	6.9	4.9	2.2	2.9	2.2
His.	11.1	9.8	7.2	3.0	0.9
Arg.	2.0	1.4	0.3	1.0	1.0

TABLE 8.

Amino acid analyses of the free amino acids obtained after N-terminal determination.

	Myoglob.	1A	1B	2	3
Asp.	10.3	8.7	4.6	3.9	3.1
Thr.	6.1	5.8	2.9	2.1	0.5
Ser.	3.7	3.3	2.0	0.7	0.3
Glu.	20.1	17.0	9.2	10.0	4.5
Pro.	4.0	4.0	2.6	+	0
Gly.	12.9	12.3	7.3	5.1	2.9
Ala.	14.0	12.4	7.8	4.0	4.1
Val.	7.1	8.1	4.0	3.3	0.6
Ileu.	8.0	7.8	5.5	2.3	1.4
Leu.	16.5	14.8	9.5	5.8	4.1
Phe.	6.7	5.0	3.0	3.0	2.3
Arg.	2.0	1.1	0.4	1.1	1.1
N-term.	Gly.	Gly.	Ser.	Gly.	Lys.

Bovine serum albumin.

The results of amino acid analyses of bovine serum albumin after 24 hrs hydrolysis, 72 hr hydrolysis, reduction and alkylation and full reduction and alkylation followed by cyanogen bromide treatment are shown in Table 9. The values are expressed as the nearest integer in Table 10. The results obtained with the reduced alkylated protein showed that, under the conditions used, reduction and alkylation were complete and so any methionyl bond cleavage taking place during cyanogen bromide treatment should cause fragmentation of the protein. The reduction in the methionine content of the cyanogen bromide treated protein showed that some form of reaction at methionine had taken place.

The first indication of a change in the nature of the material was the observation that after cyanogen bromide treatment the volume of acid acetone required to cause precipitation from the 8 M urea solution increased from 75 to 175 ml. An alteration in the average molecular weight of the material was shown by a change in the ultracentrifugation behaviour after cyanogen bromide treatment. Though no discrete separation was obtained by this technique, the peak obtained was much broader and had a lower rate of

sedimentation than that of the reduced alkylated protein. This type of behaviour would be expected for a mixture containing components of roughly similar molecular weight, such as could be produced by albumin fragmentation. Attempts were made to separate the cleavage products by cellulose acetate electrophoresis at pH 9, and at pH 3.4 in 8 M urea. In both cases only a single band was obtained.

Application of 20 mg of the reduced carboxymethylated, cyanogen bromide treated protein to the column of Sephadex G 75 gave the results shown in Fig.30. This gave an initial indication that cleavage into four or five fragments had been obtained. At this time Sephadex G 75 became available in bead form instead of the crushed form previously used. Fig.31 shows a repeat of the separation using 100 mg of split protein applied to a Sephadex G 75 bead column of the same dimensions and operated at the same flow rate. Comparison with Fig.30 shows that the bead form of Sephadex gives an improved separation even with an increased quantity of material. Assuming the first peak in Fig.31 to be due to unsplit bovine serum albumin, partial separation into five components appeared to have been achieved.

Twenty milligrams of cyanogen bromide treated protein were applied to a short (60 x 2cm) column of

Sephadex G 200 with the result shown in Fig.32. A separation of some sort had obviously been obtained. In an attempt to reduce the complexity of the applied material and to improve the separation, fraction 11 from Fig.31 was applied to the long column of Sephadex G 200. The result is shown in Fig.33. This indicates that fraction 11 was probably a mixture of several components. The cleavage products of bovine serum albumin therefore appeared to be of greater complexity than at first supposed.

The method of separation on Sephadex G 75, followed by combination of the tubes containing the desired fraction, lyophilisation and reapplication to G 200 meant that any partial separation obtained on G 75 was largely lost because of the necessity for combination of the tubes containing the fraction prior to lyophilisation. In an effort to retain any such separation, the columns of G 200 and G 75 were run in series. The effluent from the G 200 column was led by narrow diameter teflon tubing (0.7 mm bore) into the gel at the top of the G 75 column. The results of one such separation are shown in Fig.34. The curve of ultraviolet absorption was relatively simple, however the upper curve of ninhydrin colour after alkaline hydrolysis showed a considerable complexity for the mixture. The presence of 15-20 components did not seem improbable. Although the ninhydrin estimation after alkaline hydrolysis

does tend to give rather variable results, the same type of pattern was obtained in two other separations using the coupled columns which suggested that the large estimate of the number of components was in fact correct. The separation obtained in Fig. 34 illustrates the benefit of using the two columns in a coupled manner. The separation was better than would be obtained from individual use of the two columns. The last two peaks in the curve of ultraviolet absorption in Fig. 34 correspond to the fractions 3 and 4 in Figs. 30 and 31.

Purification of fraction 3 was carried out on the G 75 bead column with the result shown in Fig. 35. Fraction 3_1 was used for amino acid analysis and fingerprinting after tryptic digestion.*

Fraction G from the tubes indicated in Fig. 34 was rerun on the coupled columns to give the graph shown in Fig. 36. Fraction G_1 was used for amino acid analysis and fingerprinting.

Fraction 4 was initially purified on G 75 bead (Fig. 37). Fraction 4_1, when rerun on G 50 gave only one, symmetrical peak and was used for amino acid analysis and fingerprinting.

---

* Discussion of the significance of these results and those pertaining to other fractions is reserved for the section headed Discussion (p. 86.).

As will be shown in the Discussion (p.105), deductions from the amino acid content and fingerprint pattern of fraction 4 would suggest for a single polypeptide chain a molecular weight of 30,000. The equivalent deductions for fraction 3 give a value of 15,000. The elution of fraction 4 after fraction 3, on gel filtration is therefore incompatible with the representation of fraction 4 as a single polypeptide chain. The experimental observations are best explained by the presence of three components in fraction 4. A further fractionation was attempted by repeating the separation on G 75 (Fig.37). The skewness of the peak obtained suggested that there was a partial separation on the left hand side. The main part of the peak contained in tubes 372-377 was isolated to give fraction 4D, a much narrower fraction than 4, which was rerun on G 75 to give $4D_1$ (Fig.38). This was used for amino acid analysis and fingerprinting.

The two peaks shown in the area IJ in Fig.34 were rerun separately on the coupled columns of G 200 and G 75. The results are shown in Figs.39 and 40. In neither case was there any indication that an adequate separation had been achieved. Since the material contained in each part of the effluent was insufficient in itself to allow further purification and since in the initial separation shown in Fig.34 the peaks I and J were in close proximity, the vague similarity in the two curves obtained

was taken to suggest identity of the two fractions I_1 and J_1. They were therefore combined and used for tryptic digestion and fingerprinting.

Katchalski, Benjamin and Gross (1957) reported a pH dependence in the extent of reduction of bovine serum albumin. It seemed possible that cyanogen bromide fragmentation of a partially reduced protein would yield a simpler mixture of products since some of the fragments might be held together by unreduced disulphide bonds. In one experiment, reaction of native bovine serum albumin with cyanogen bromide was carried out in 0.1 N HCl which was 8 M with respect to urea, with the addition of 0.3 ml ethanethiol to give reduction of the disulphide bonds. After reaction for 24 hr. the pH was adjusted to 8.6 with sodium hydroxide and 0.380 g of iodoacetate added as before and allowed to react with the protein. The products were applied to the column of G 75 and gave the results shown in Fig. 41. The same graph was obtained when reduction of the protein was carried out separately using 0.1 N HCl in place of the pH 8.6 tris buffer and the reduced protein carboxymethylated (pH 8.6) and then treated with cyanogen bromide (0.1 N HCl). Amino acid analysis of the partially reduced protein is compared with that of the fully reduced material in Tables 9 and 10. Fraction S1 obtained from the fractionation illustrated in Fig.41 was separated on

Sephadex G 200 as shown on Fig.42. Fraction S2 (Fig.41) when rechromatographed on G 75 appeared to separate into two main fractions, S2B and S2C (Fig.43). Fractions S2B and S2C were purified by further chromatography on G 75 (Figs.44 and 45). Fractions $S2B_1$ and $S2C_1$ were used for amino acid analysis and fingerprinting. The significance of these fractions is considered in the Discussion .

The amino acid analyses of the fractions G, 3, 4, 4B, S2B and S2C are shown in Table 11 and the values to the nearest integer in Table 12.

Fig.46 shows the position of all the fingerprint spots obtained from bovine serum albumin and its fractions. The numbers allocated to each spot in Fig.46 will be used to represent the spot in the Discussion. Tables 13 a,b,c,dshow the spots obtained after tryptic digestion and fingerprinting of reduced alkylated albumin, reduced alkylated split albumin, native albumin, and fractions G, 3, 4, 4B, S2C and IJ.

N-terminal determination of the reduced alkylated split protein showed the presence of glutamic acid, serine, lysine and arginine as new N-terminal residues together with aspartic acid, the N-terminal residue of the intact protein. N-terminal determinations have not been carried out with the individual fractions.

TABLE 9.

Amino acid analyses of bovine serum albumin.

	24 hr. hyd.	72 hr. hyd.	Reduced and Alkylated		Red. in 0.1N HCl	RA and CNBr treated	
Cys.A.	-	-	-	-	-	-	-
CM Cys.	-	-	41.2	38.9	31.8	18.6	32.9
Asp.	57.7	55.1	50.7	52.3	48.7	50.4	51.2
Thr.	27.3	17.5	30.7	30.1	26.8	29.5	32.0
Ser.	17.3	6.1	21.4	18.6	18.2	20.1	23.2
Glu.	74.6	78.9	+	+	+	+	+
Pro.	28.3	29.1	30.9	30.9	29.4	30.3	31.5
Gly.	15.8	16.3	17.2	17.3	16.0	17.8	17.3
Ala.	42.3	44.7	41.2	41.5	38.7	41.3	48.7
Cys.	23.9	19.1	-	-	3.2	12.0	6.3
Val.	37.4	37.8	36.9	36.3	40.8	36.0	35.9
Met.	4.9	4.1	3.8	3.5	3.4	1.6	1.7
Ileu.	18.6	13.65	+	+	+	+	+
Leu.	66.5	52.4	+	+	+	+	+
Tyr.	20.2	17.8	20.7	20.1	+	14.2	17.8
Phe.	26.3	26.2	27.9	27.2	+	27.9	26.4
Lys.	50.4	52.4	+	+	+	+	+
His.	15.9	16.8	17.3	+	+	18.8	+
Arg.	23.7	22.9	+	25.5	25.0	+	24.2

+ indicates that the amino acid was present but not calculated for the reasons given in the text.

- indicates that the amino acid was absent.

TABLE 10.

Amino acid analyses of BSA expressed as integers. (TABLE 9).

	24 hr hyd.	72 hr. hyd.	Reduced and alkylated		Red in 0.1N HCl	RA and CNBr treated	
Cys.A.	-	-	-	-	-	-	-
CMCys.	-	-	41	39	32	19	33
Asp.	58	55	51	52	49	50	51
Thr.	27	18	31	30	27	30	32
Ser.	17	6	21	19	18	20	23
Glu.	75	79	+	+	+	+	+
Pro.	28	29	31	31	29	30	32
Gly.	16	16	17	17	16	18	17
Ala.	42	45	41	42	39	41	49
Cys.	24	19	-	-	3	12	6
Val.	37	38	37	36	41	36	36
Met.	5	4	4	3	3	2	2
Ileu.	19	14	+	+	+	+	+
Leu.	67	52	+	+	+	+	+
Tyr.	20	18	21	20	+	14	18
Phe.	26	26	28	27	+	28	26
Lys.	50	52	+	+	+	+	+
His.	16	16	17	+	+	19	+
Arg.	24	23	+	26	25	+	24

+ indicates that the amino acid was present but not calculated for the reasons given in the text.

- indicates that the amino acid was absent.

FIGURE 30.

Separation of the products of cyanogen bromide action on reduced alkylated bovine serum albumin by chromatography on Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

1←—1—→1 indicates the tubes from which the appropriate fraction (1 etc.) described in the text was isolated.

FIGURE 30.

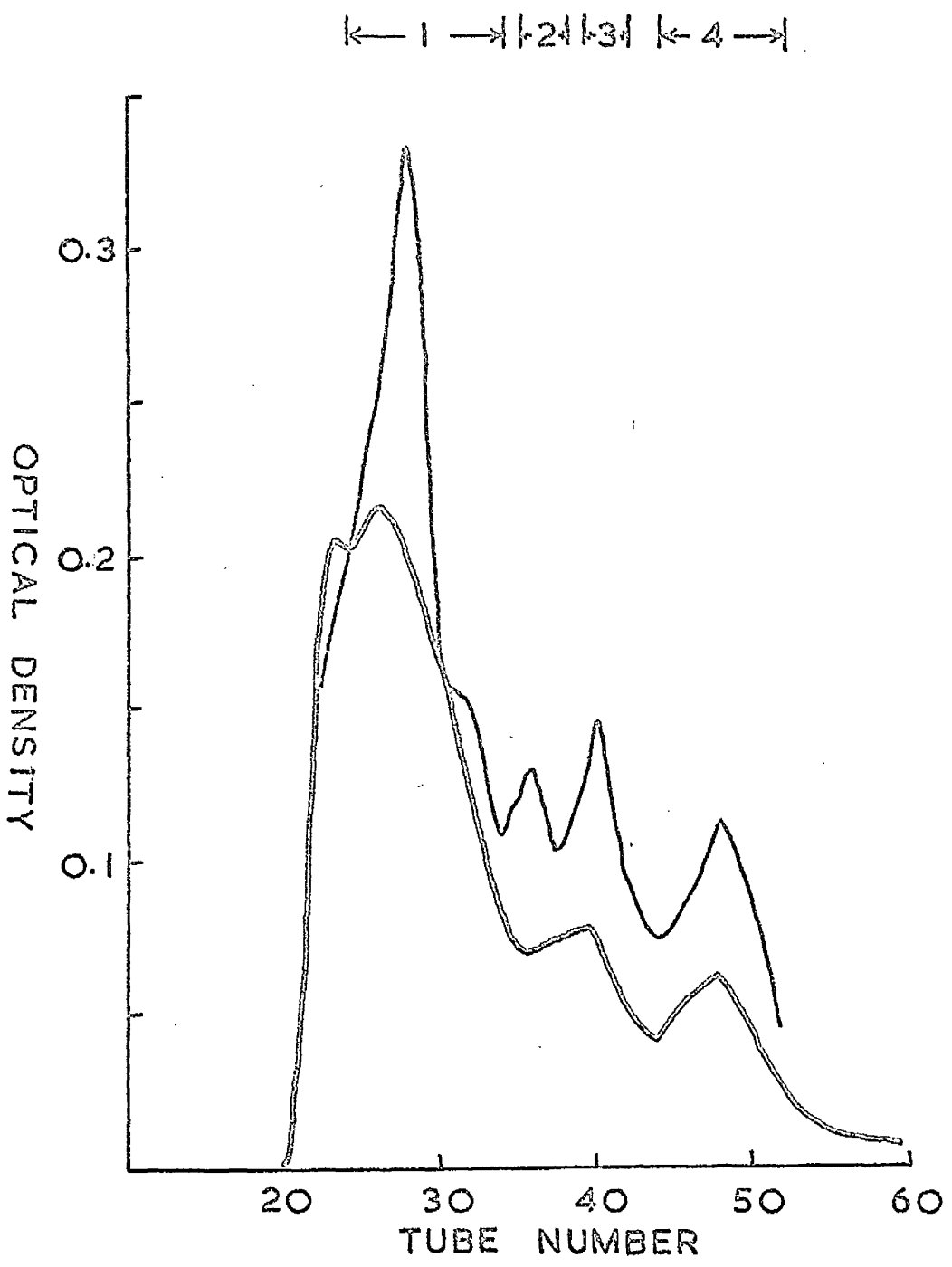


FIGURE 31.

Separation of the products of cyanogen bromide action on reduced alkylated bovine serum albumin by chromatography on the bead form of Sephadex G 75. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

1←1→1 indicates the tubes from which the appropriate fraction (1 etc.) described in the text was isolated.

FIGURE 31.

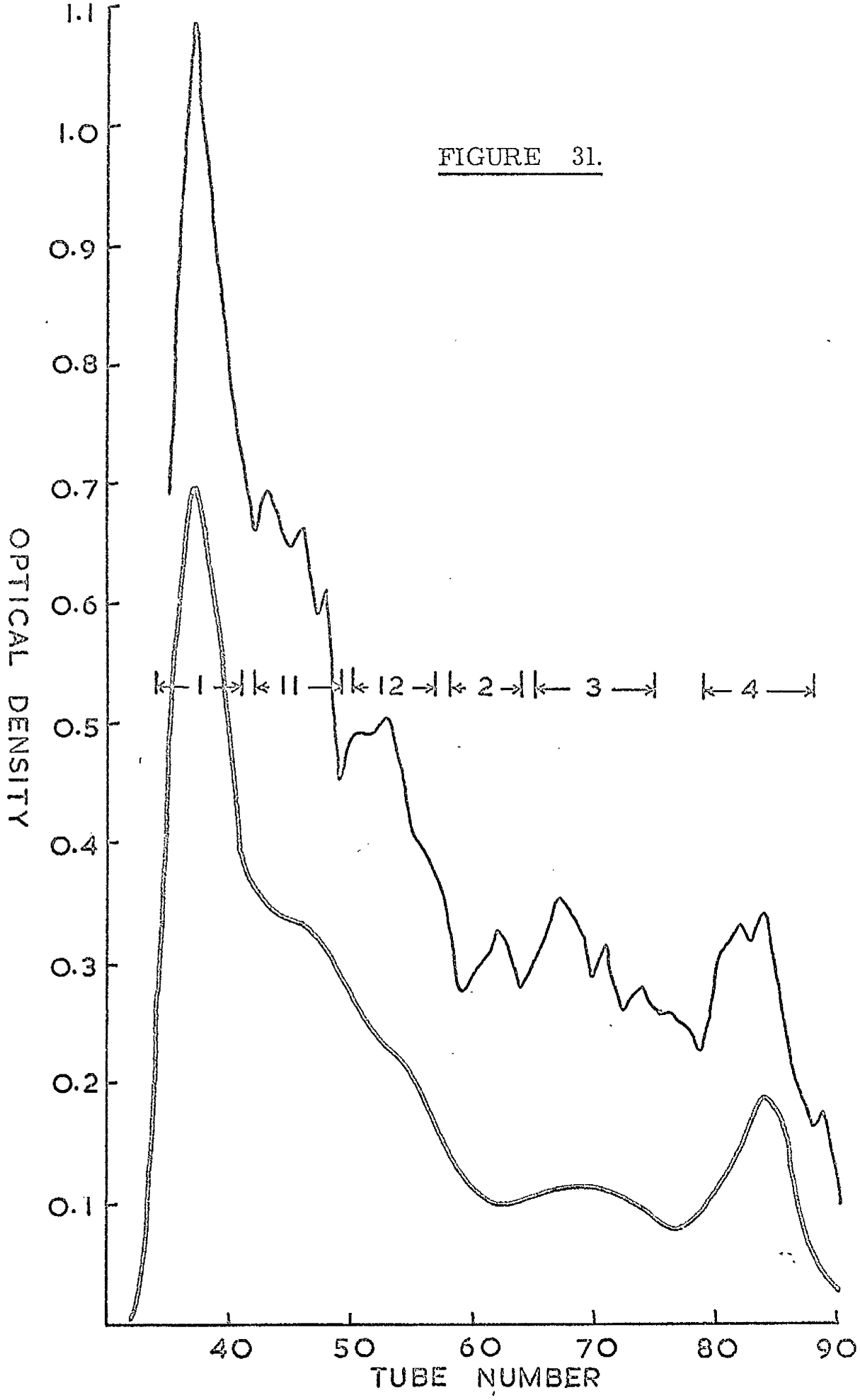

1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
OPTICAL DENSITY
1
11
12
2
3
4
40
50
60
70
80
90
TUBE NUMBER

FIGURE 32.

Separation of the products of cyanogen bromide action on reduced alkylated bovine serum albumin by chromatography on a short column of Sephadex G 200.

——— Optical density at 280 mμ.

FIGURE 32.

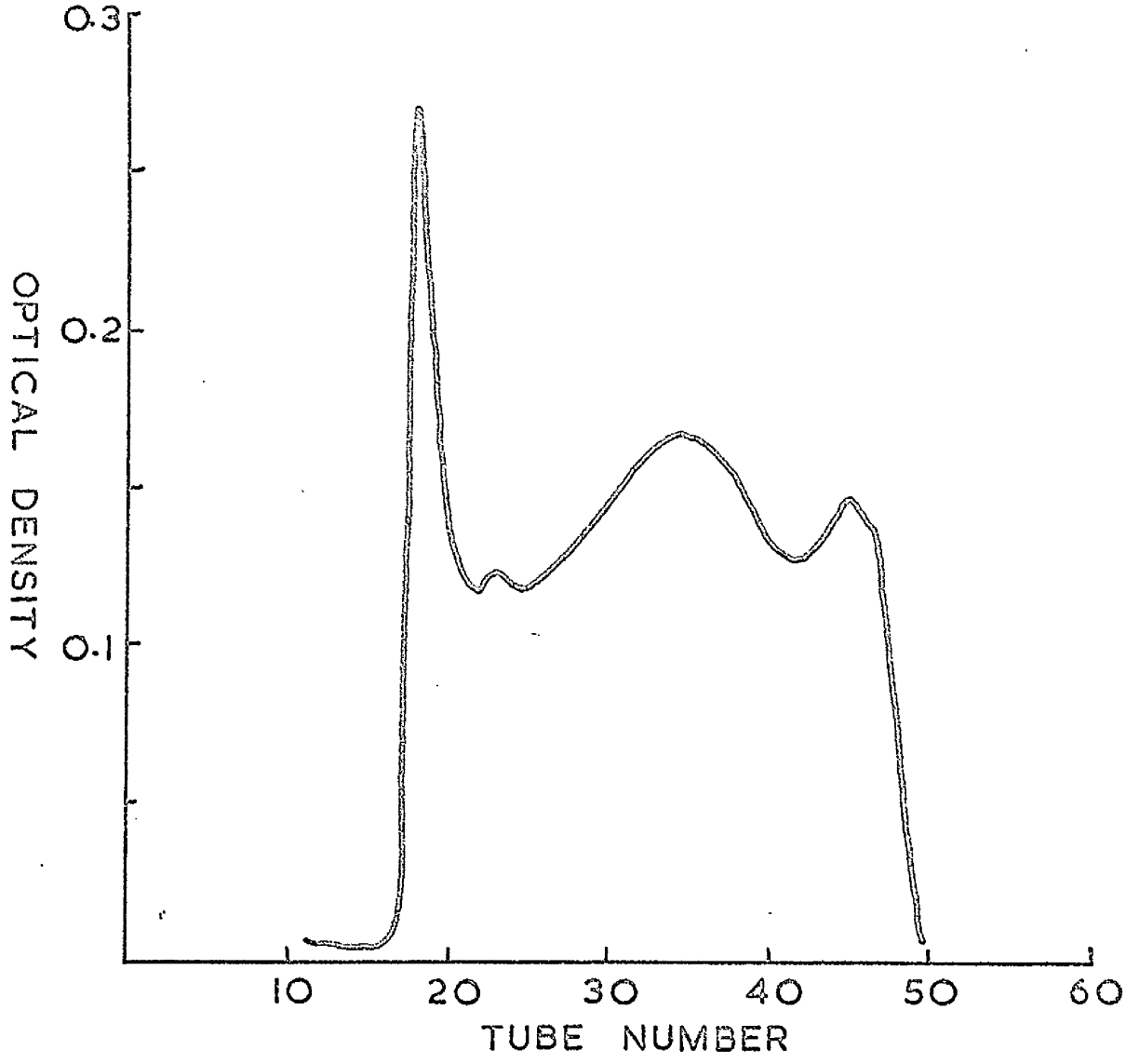


FIGURE 33.

Separation of fraction 11 (Fig. 31) by chromatography on a long column of Sephadex G 200. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

FIGURE 33.

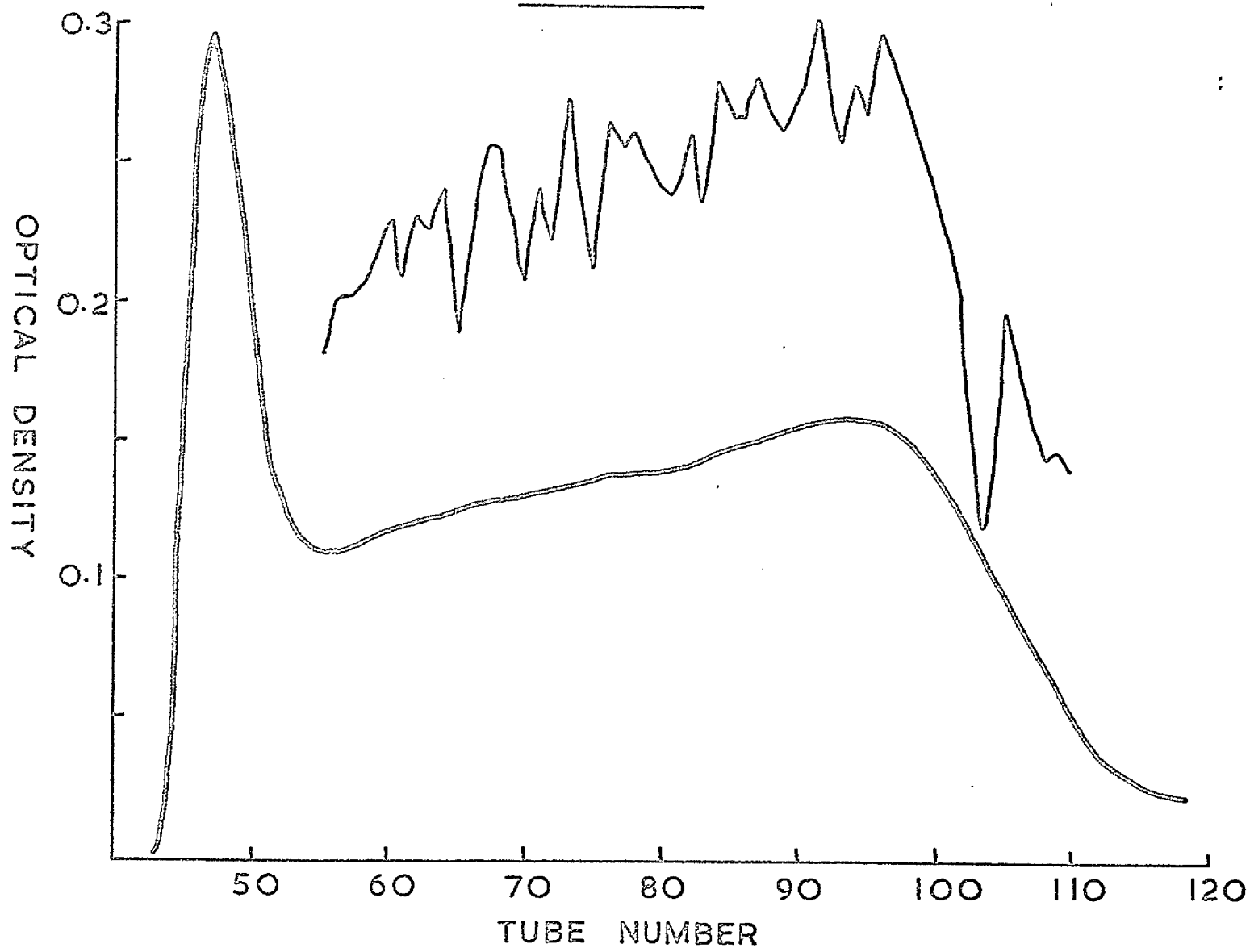


FIGURE 34.

Separation of products of cyanogen bromide action on reduced alkylated bovine serum albumin by chromatography on columns of Sephadex G 200 and G 75 bead run in series. Chromatographic conditions as described in the Experimental and Results Sections.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

1←G→1 indicates the tubes from which the appropriate fraction (G etc.) described in the text was isolated.

FIGURE 34.

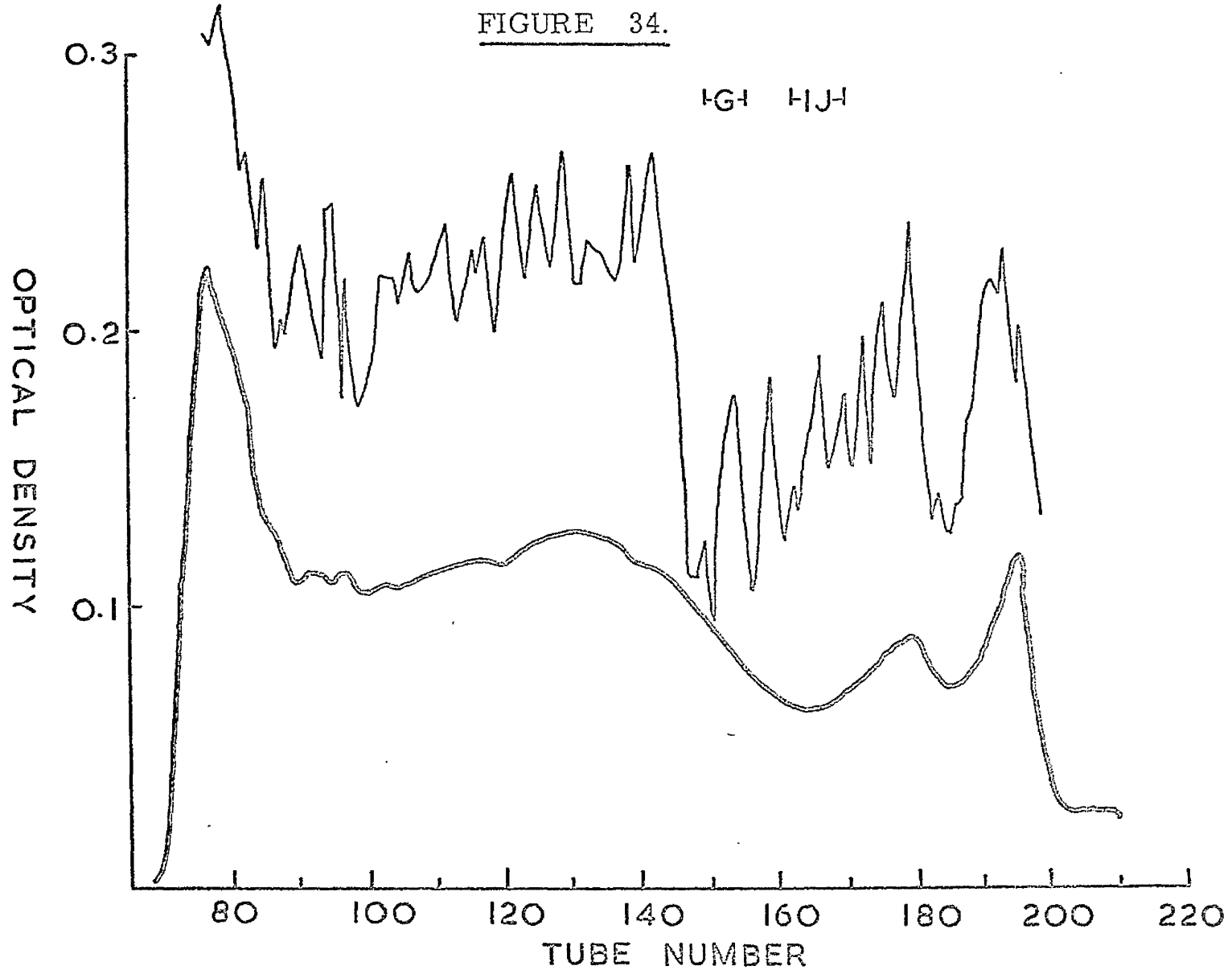

0.3
0.2
0.1
OPTICAL DENSITY
80
100
120
140
160
180
200
220
TUBE NUMBER

FIGURE 35.

Purification of fraction 3 (Figs. 30, 31, 34) on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion. Fraction 3_1 isolated from the tubes indicated.

FIGURE 36.

Purification of fraction G (Fig. 34) on coupled columns of Sephadex G 200 and G 75. Chromatographic conditions as described in the Experimental and Results Sections.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

FIGURE 35.

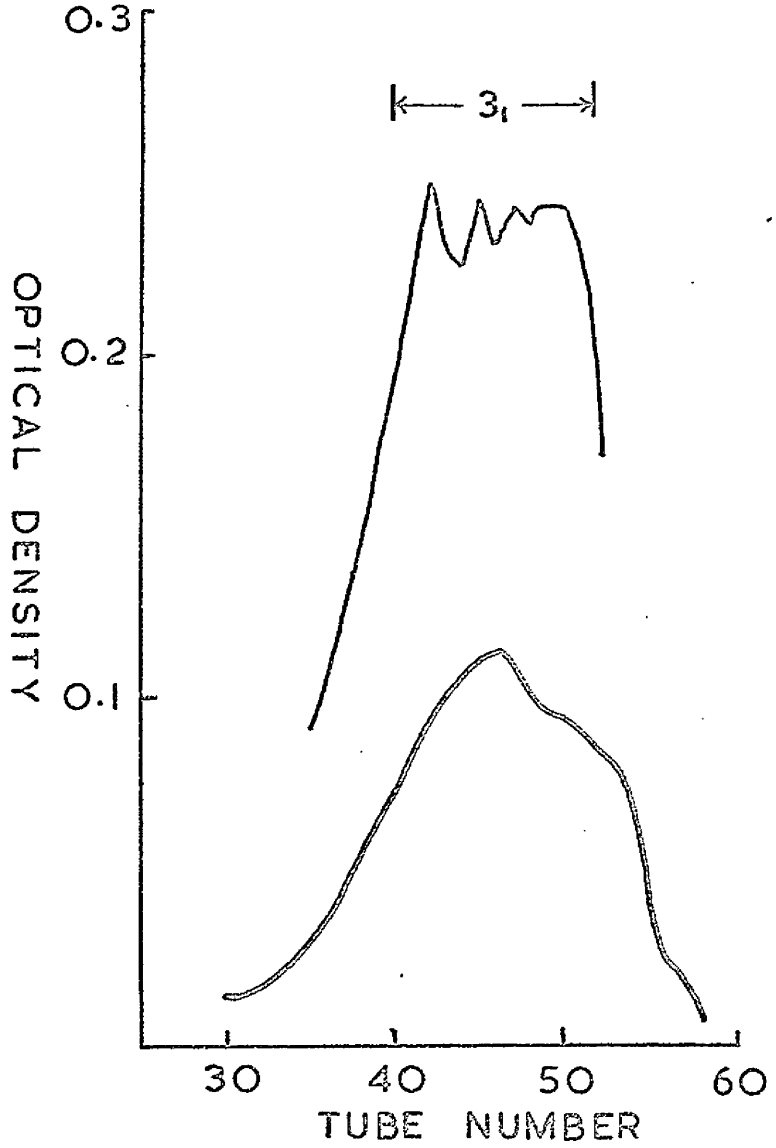


FIGURE 36.

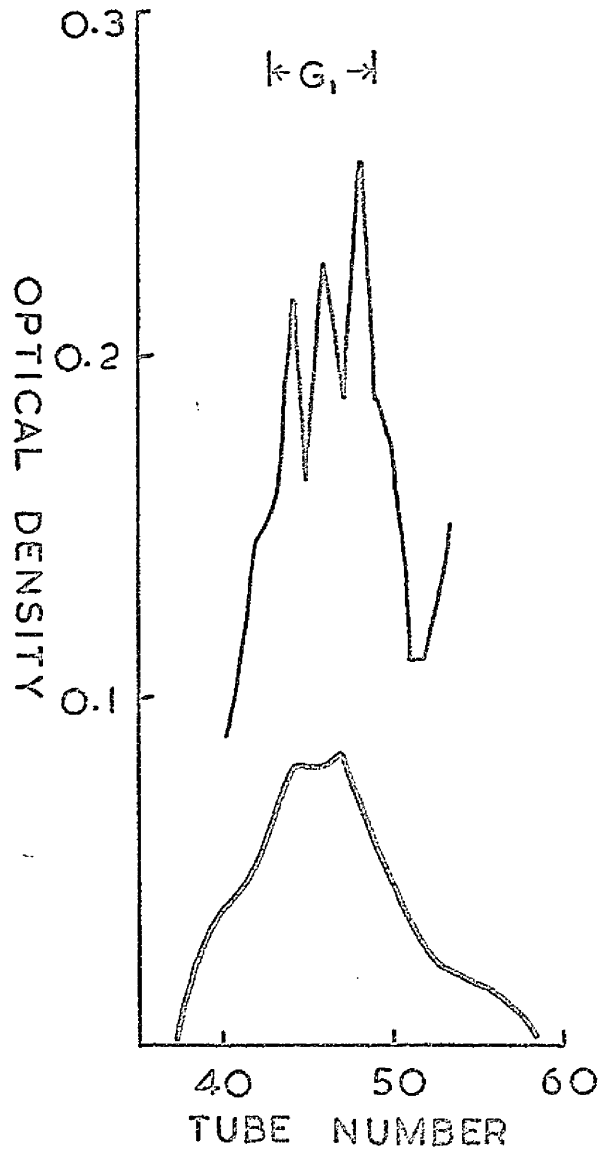


FIGURE 37.

Purification of fraction 4 (Figs. 30, 31, 34) on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion. Fraction 4_1 isolated from the tubes indicated. In a further attempt at purification the separation of fraction 4 was repeated and the much smaller fraction, 4B , obtained from the part of the peak indicated.

FIGURE 38.

Purification of fraction 4B (Fig. 37) on Sephadex G 75 bead. Chromatographic conditions as for the previous separation.

——— Optical density at 280 mμ.

Fraction $4B_1$ isolated from the tubes indicated.

FIGURE 37.

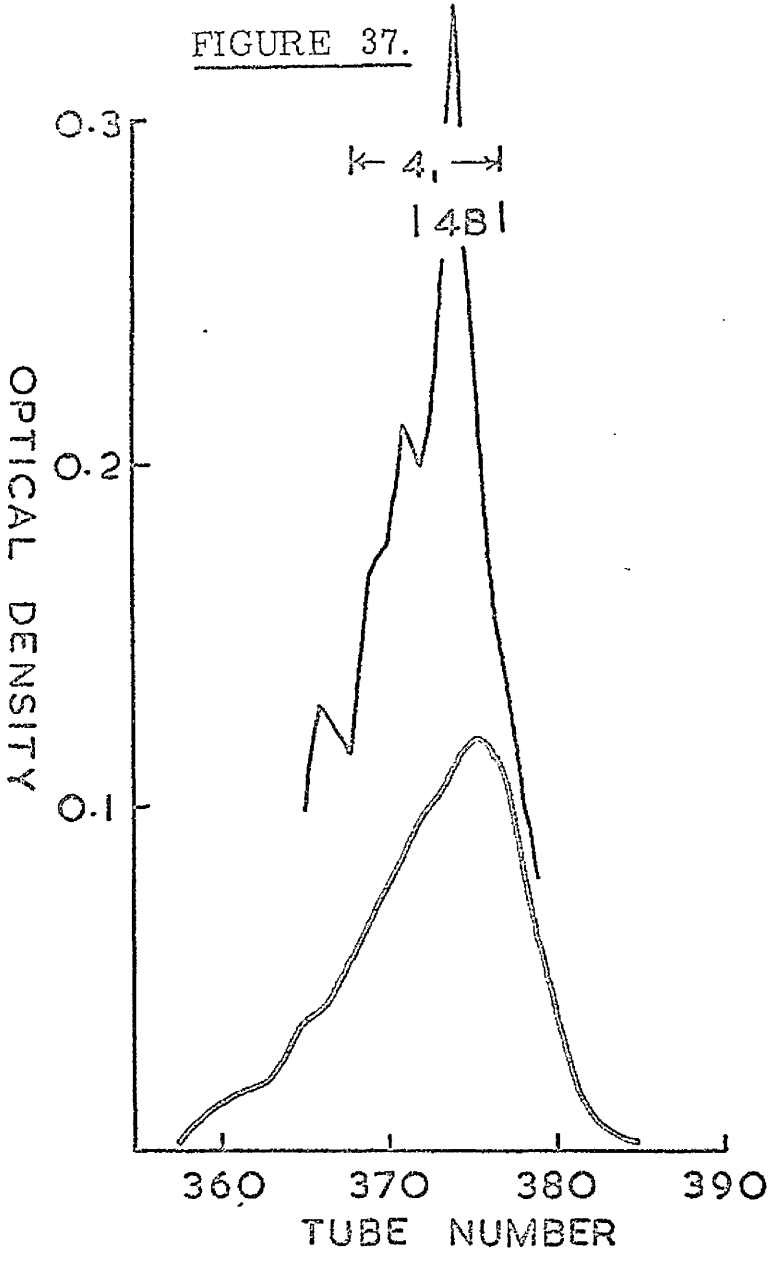


FIGURE 38.

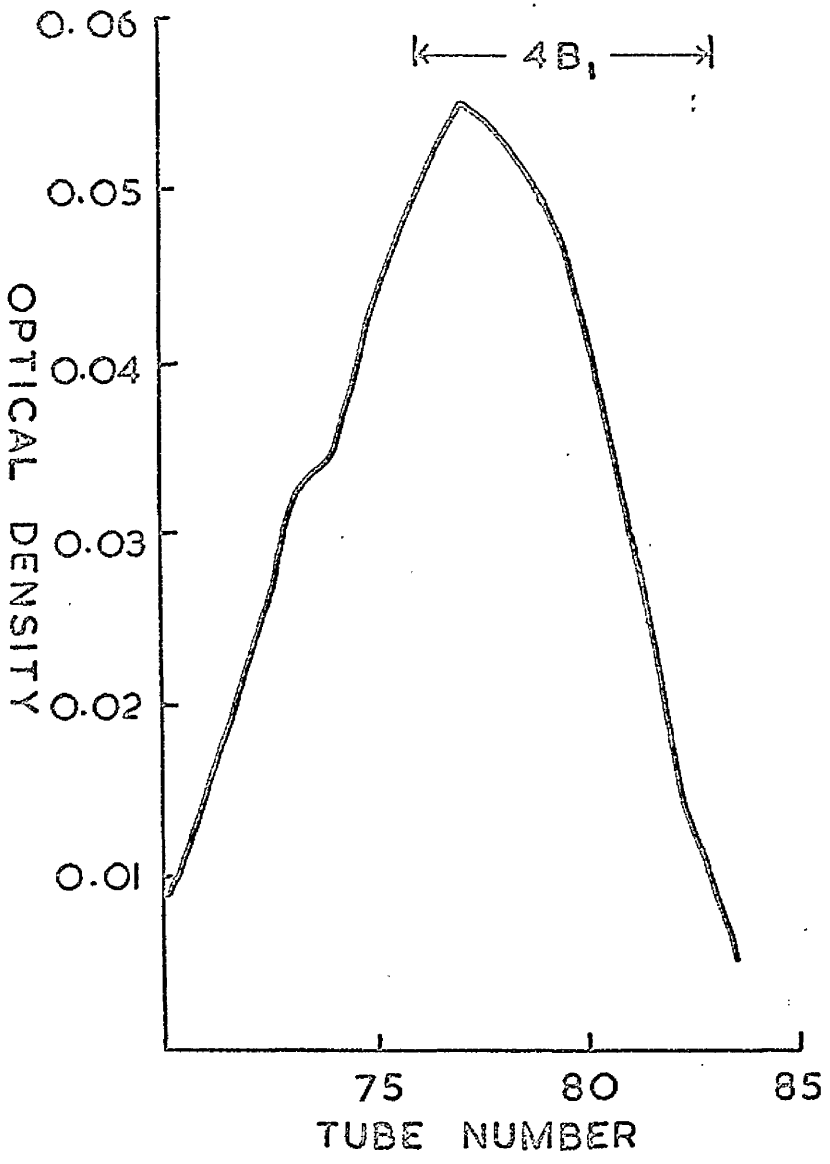


FIGURE 39.

Separation of fraction I (Fig. 34) on coupled columns of Sephadex G 200 and G 75 bead. Chromatographic conditions as described in the Experimental and Results Sections.

——— Optical density at 280 mμ.

Fraction I_1 isolated from the tubes marked.

FIGURE 40.

Separation of fraction J (Fig. 34) on coupled columns of Sephadex G 200 and G 75 bead. Chromatographic conditions as described in the Experimental and Results Sections.

——— Optical density at 280 mμ.

Fraction J_1 obtained from the tubes indicated.

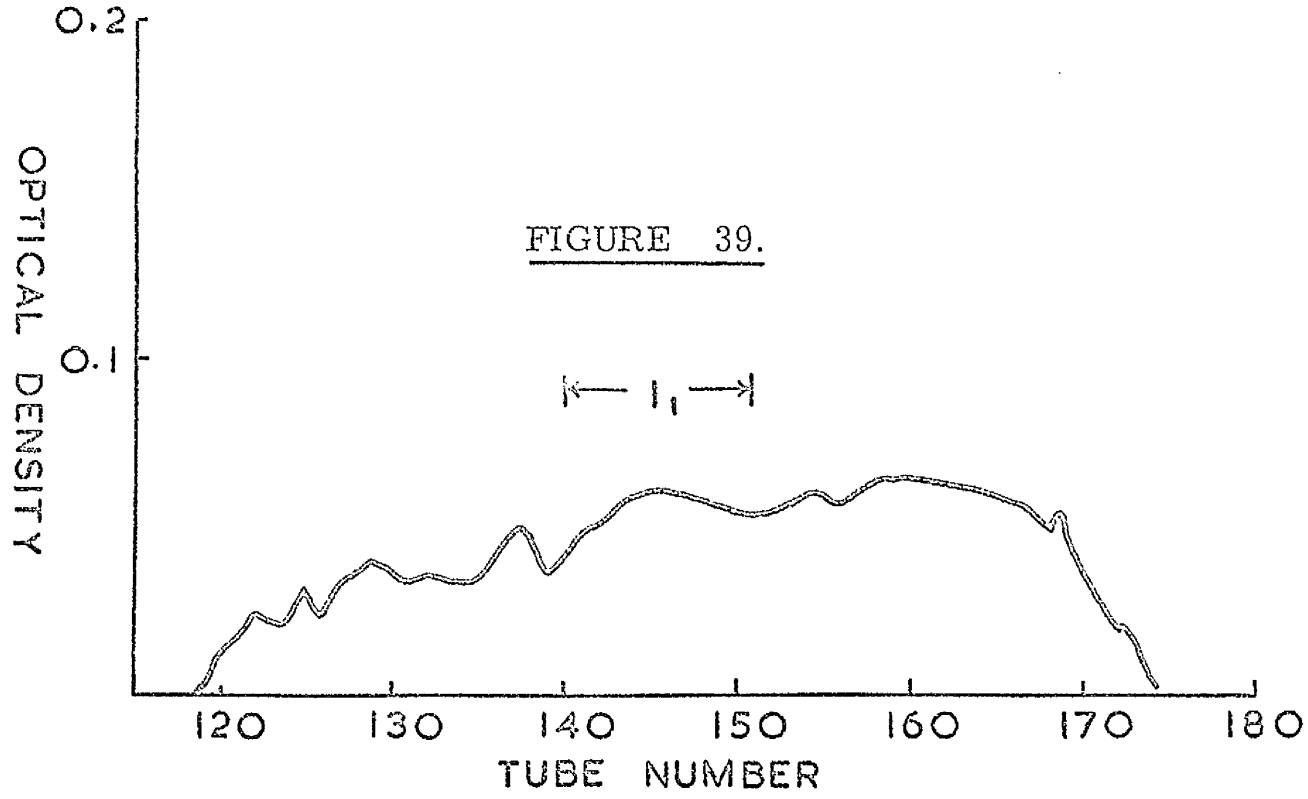


FIGURE 39.

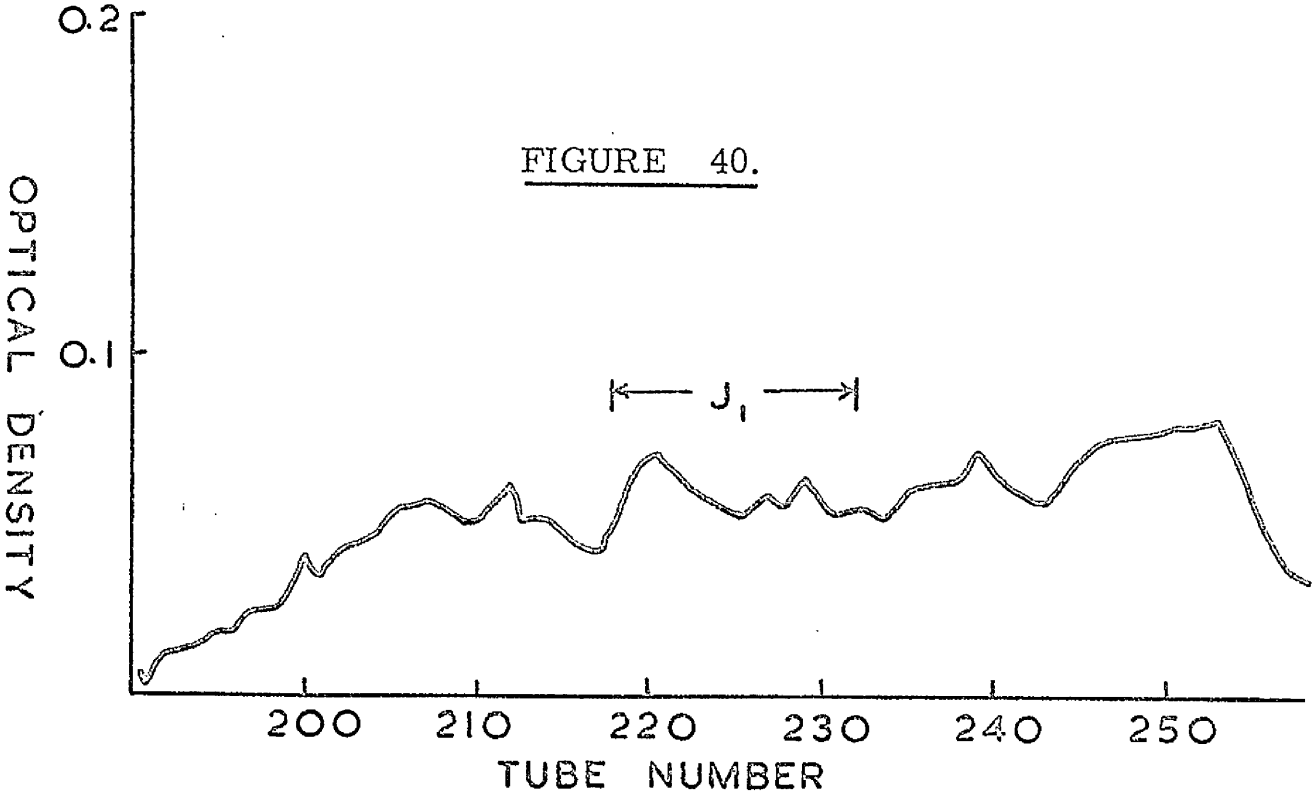


FIGURE 40.

FIGURE 41.

Separation of the products of cyanogen bromide action on partially reduced, alkylated bovine serum albumin by chromatography on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

1←S1→1 indicates the tubes from which the appropriate fraction (S1 etc.) described in the text was isolated.

FIGURE 41.

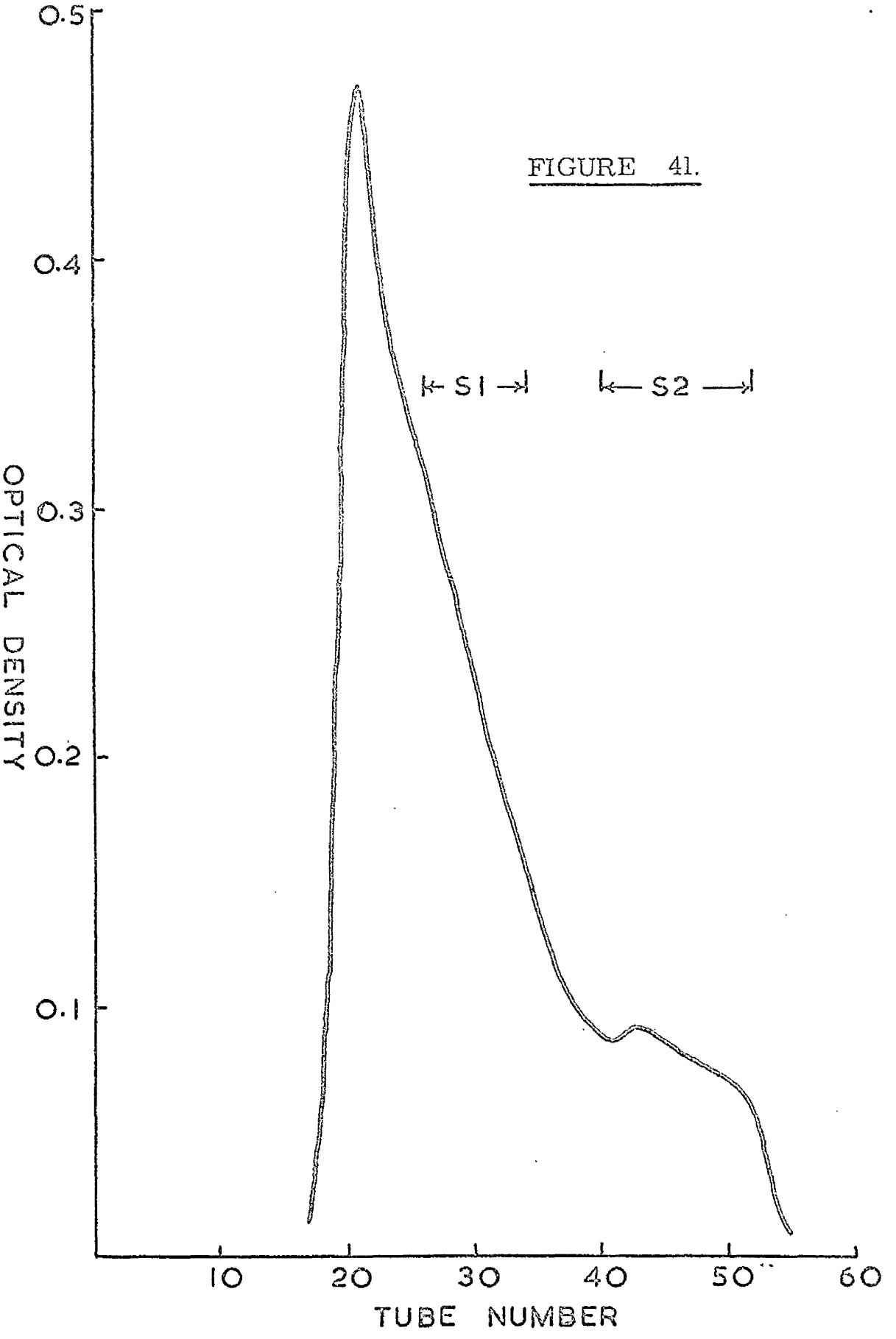

0.5
0.4
0.3
0.2
0.1
OPTICAL DENSITY
S1
S2
10
20
30
40
50
60
TUBE NUMBER

FIGURE 42.

Separation of fraction 81 (Fig. 41) on Sephadex G 200. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

FIGURE 42.

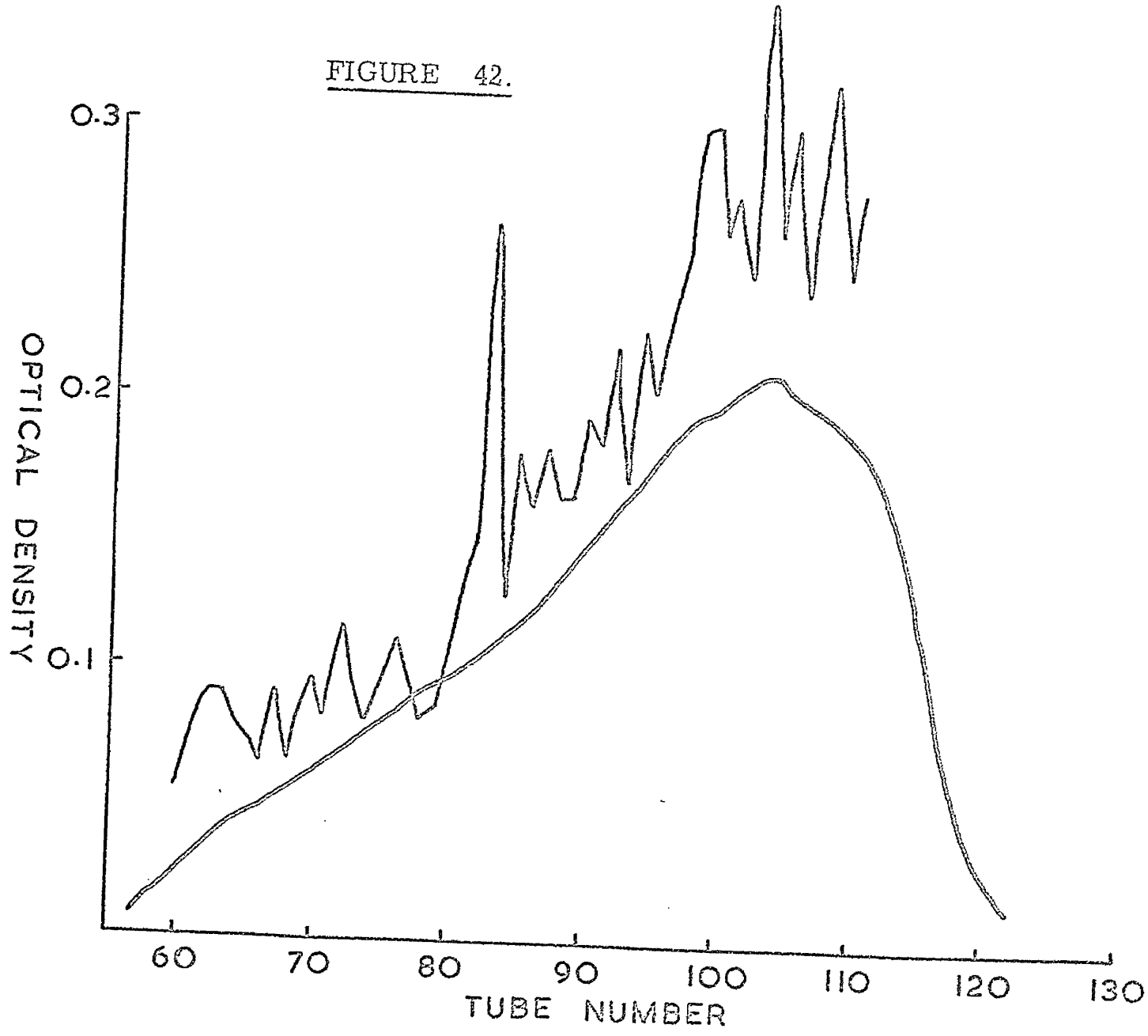

0.3
0.2
0.1
OPTICAL DENSITY
60
70
80
90
100
110
120
130
TUBE NUMBER

FIGURE 43.

Separation of fraction S2 (Fig. 41) on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

——— Optical density at 570 mμ of the ninhydrin colour developed with an aliquot after alkaline digestion.

|←S2I→| indicates the tubes from which the appropriate fraction (S2B etc.) described in the text was isolated.

FIGURE 43.

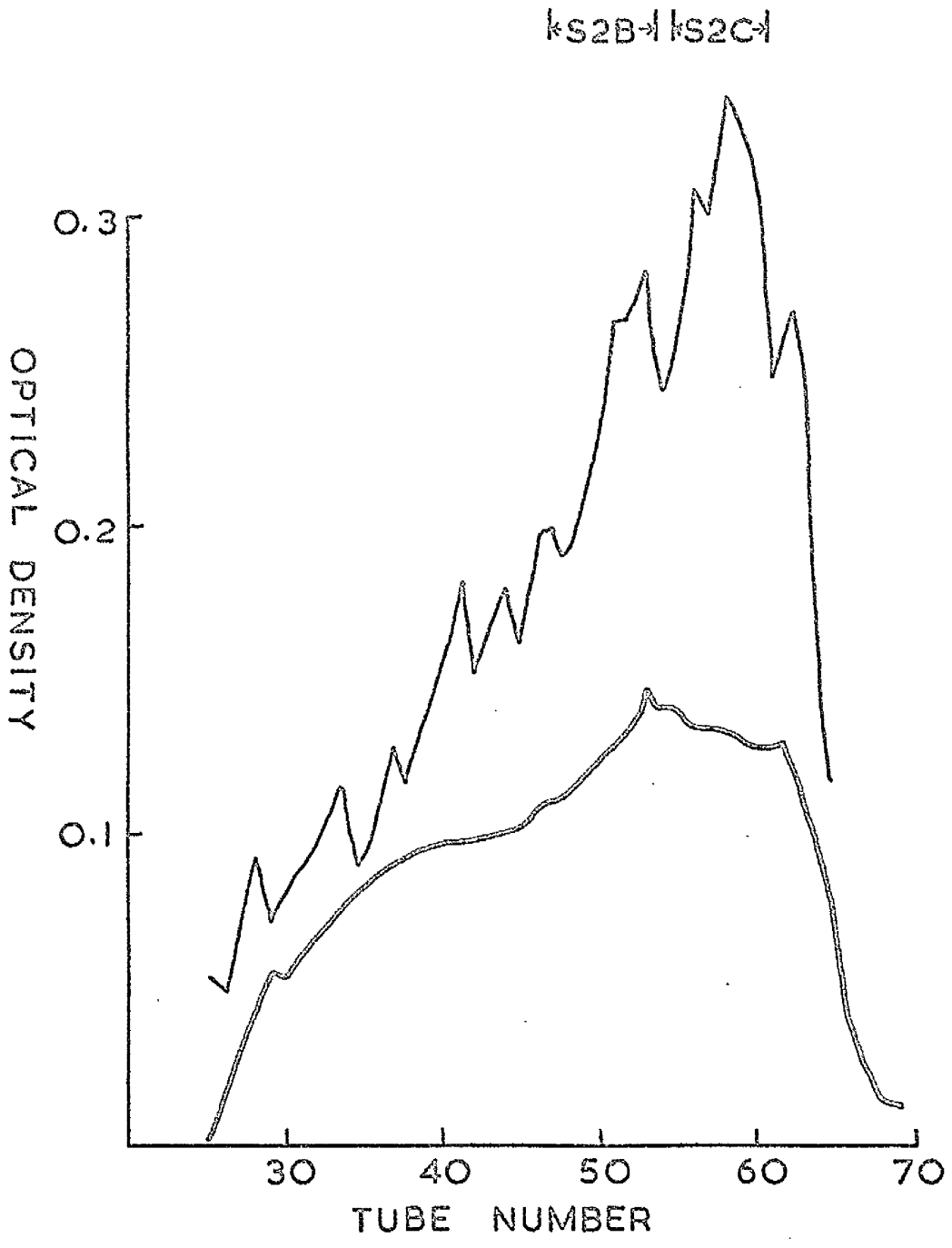


FIGURE 44.

Purification of fraction S2B (Fig. 43) on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction $S2D_1$ isolated from the tubes indicated.

FIGURE 45.

Purification of fraction S2C (Fig. 43) on Sephadex G 75 bead. Chromatographic conditions as described in the Experimental.

——— Optical density at 280 mμ.

Fraction $S2C_1$ isolated from the tubes indicated.

FIGURE 44.

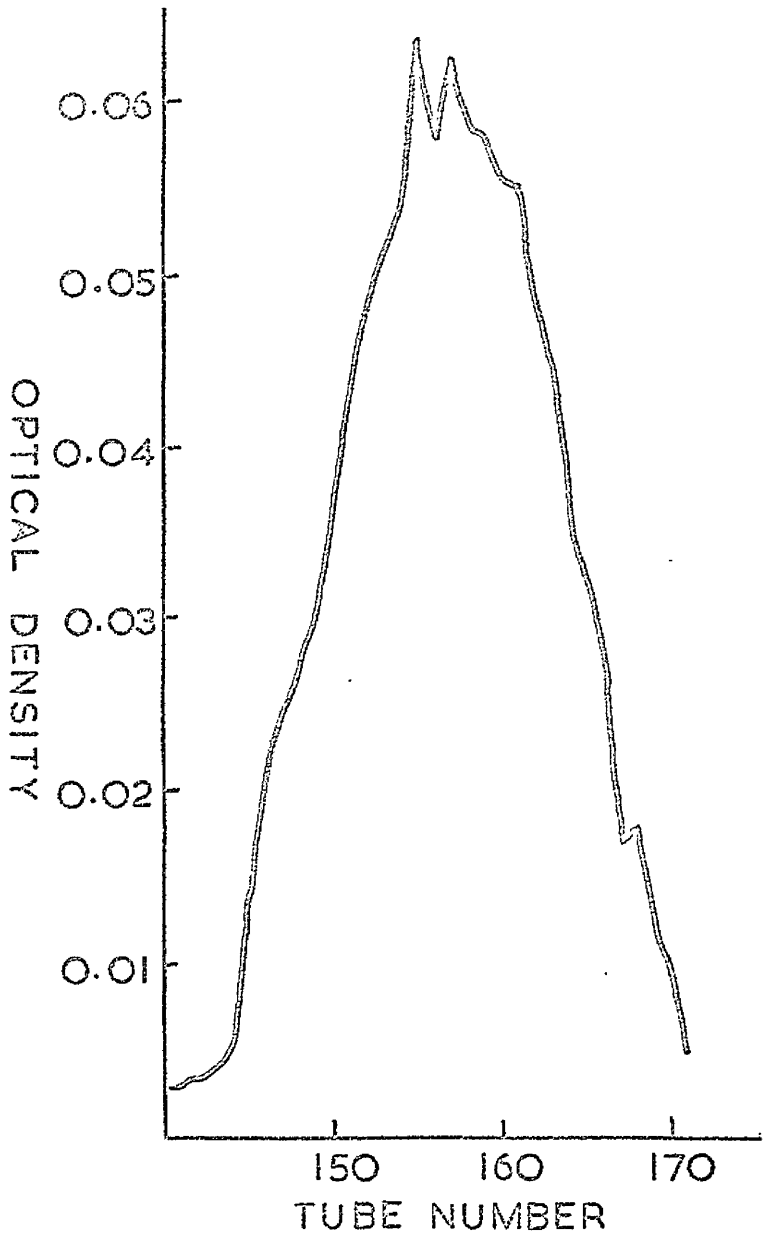


FIGURE 45.

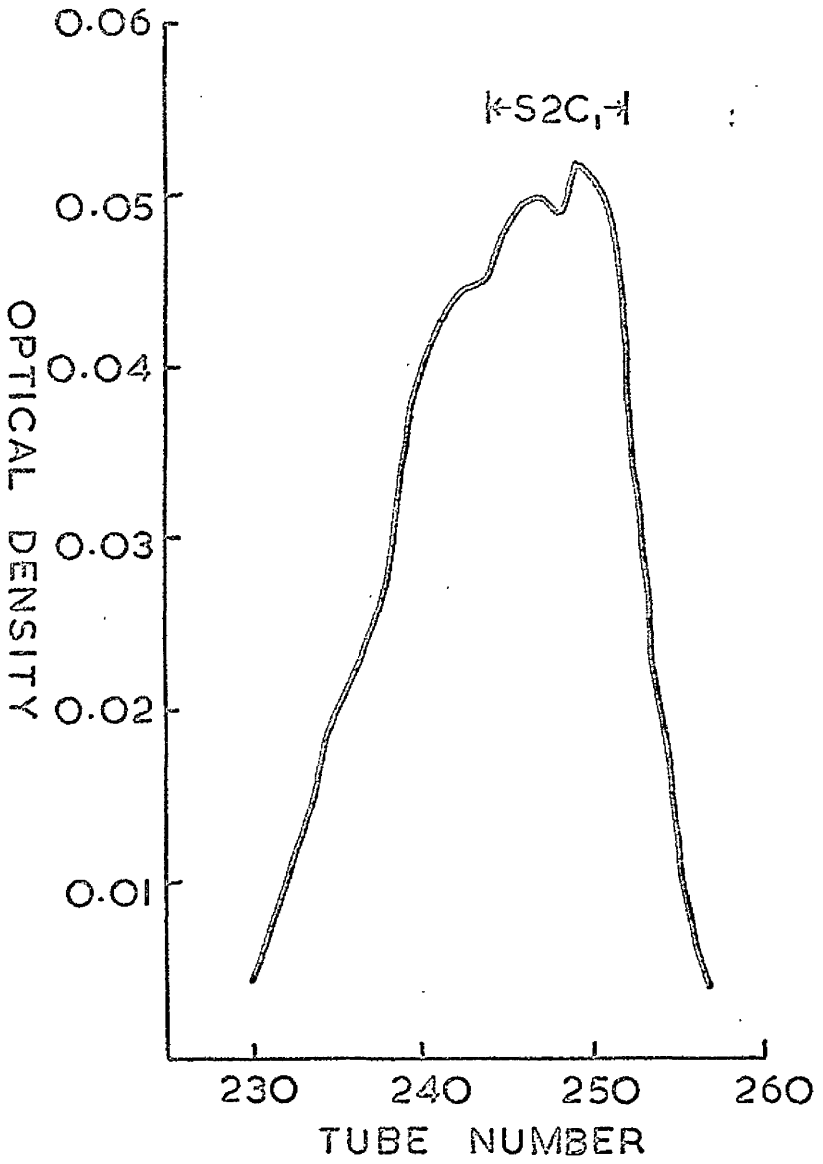


TABLE 11.

Amino acid analyses of the fractions obtained from BSA.

	G		3		4		4B	S2B	S2C
Asp.	19.6	19.8	11.3	13.0	25.5	24.7	29.3	15.0	12.0
Thr.	8.4	8.3	11.1	10.6	10.3	9.9	11.7	9.3	6.1
Ser.	7.0	7.3	5.0	4.5	8.4	7.8	8.1	4.5	3.6
Glu.	29.6	27.2	19.0	20.4	39.0	39.0	46.3	21.1	19.0
Pro.	9.24	+	8.8	8.9	12.1	9.8	10.4	10.2	+
Gly.	6.1	6.2	2.0	2.6	12.9	12.9	14.8	4.0	5.1
Ala.	15.6	15.0	8.5	9.9	22.9	18.3	20.9	11.9	9.1
Cys.*	9.4	10.3	2.8	7.6	4.5	9.4	9.9	4.5	4.4
Val.	13.8	12.9	8.8	9.1	12.9	13.3	13.6	11.3	7.7
Ileu.	5.8	5.3	3.9	4.2	8.0	6.6	6.4	5.0	2.6
Leu.	21.9	18.4	16.1	16.3	33.8	27.7	28.9	16.8	12.3
Tyr.	5.8	6.0	1.6	3.4	8.1	5.7	6.0	2.5	2.5
Phe.	7.9	7.1	6.1	7.0	17.1	15.9	17.0	6.8	6.9
Lys.	20.0	19.3	15.4	15.4	24.6	24.8	30.8	+	10.5
His.	5.0	4.7	5.1	4.7	13.1	14.3	15.3	3.4	5.1
Arg.	10.2	9.7	3.9	4.5	7.8	6.7	8.2	4.6	3.1

* includes the values for cysteic acid and carboxymethyl cysteine;

+ indicates that the amino acid was present but not calculated for the reasons given in the text.

TABLE 12.

Amino acid analyses of the fractions from BSA expressed as integers (TABLE 11).

	G		3		4		4B	S2B	S2C
Asp.	20	20	11	13	26	25	29	15	12
Thr.	8	8	11	11	10	10	12	9	6
Ser.	7	7	5	4	8	8	8	5	4
Glu.	30	27	19	20	39	39	46	21	19
Pro.	9	+	9	9	12	10	10	10	+
Gly.	6	6	2	3	13	13	15	4	5
Ala.	16	15	9	10	23	18	21	12	9
Cys.*	9	10	3	8	4	9	10	5	4
Val.	14	13	9	9	13	13	14	11	8
Ileu.	6	5	4	4	8	7	6	5	3
Leu.	22	18	16	16	34	28	29	17	12
Tyr.	6	6	2	3	8	6	6	2	2
Phe.	8	7	6	7	17	16	17	7	7
Lys.	20	19	15	15	25	25	31	+	10
His.	5	5	5	5	13	14	15	3	5
Arg.	10	10	4	4	8	7	8	5	3

* includes the values for cysteic acid and carboxymethyl cysteine.

+ indicates that the amino acid was present but not calculated for the reasons given in the text.

FIGURE 46.

Fingerprint spots obtained from tryptic digests of bovine serum albumin, its products and fragments. The numbers allocated to each spot are used to indicate the appropriate spot in the Discussion. Fingerprinting was carried out by the method of Anfinsen et al (1958) with slight modifications.

FIGURE 46.

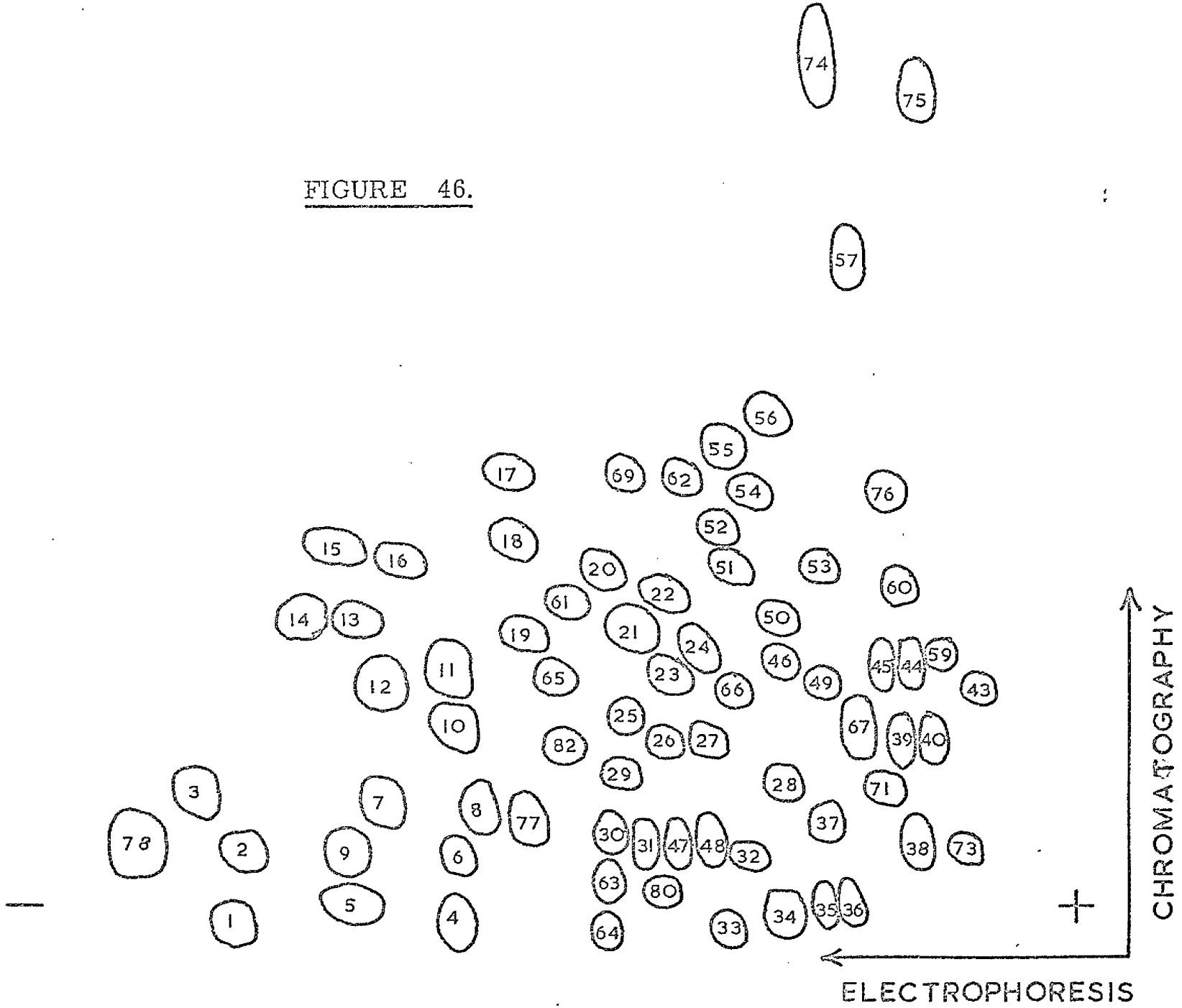

74
75
57
56
55
17
69
62
54
76
52
18
15
16
20
51
53
60
61
22
50
14
13
21
24
19
46
45
44
59
11
23
65
49
43
12
66
25
67
10
26
27
39
40
82
29
28
71
3
7
8
77
37
78
2
9
30
31
47
48
32
38
73
6
63
80
5
35
36
1
4
34
33
64
CHROMATOGRAPHY
ELECTROPHORESIS

TABLE 13 a.

Fingerprint spots obtained from bovine serum albumin and its cleavage products.

RA BSA		RA Split BSA		NAT BSA	G	3	4	4B	S2B	S2C	IJ
1	-	1	1	1	1	-	-	-	-	-	1
2	-	2	2	2	-	-	2	2	-	2	2
3	-	3	3	-	-	-	3	3	-	3	3
4	4	4	4	4	-	4	4	4	4	4	4
5	5	5	5	5	-	-	5	5	-	-	-
6	6	6	6	6	6	6	6	6	-	6	6
7	7	7	7	7	7	7	-	-	7	-	7
8	8	8	8	8	8	-	8	8	8	-	8
-	-	9	-	9	-	-	-	-	-	-	9
10	10	10	10	10	10	-	-	-	-	-	10
11	11	11	11	11	11	-	-	-	11	-	11
12	12	12	12	12	12	12	-	-	12	-	12
13	13	13	13	13	13	-	-	-	13	-	13
14	14	14	14	14	14	-	14	14	14	-	14
15	15	15	15	15	15	-	-	-	15	-	15
16	16	16	16	16	16	-	-	-	-	-	16
17	17	17	17	-	-	-	17	-	-	-	-
18	18	18	18	-	-	18	18	18	18	-	18
-	-	19	19	-	-	19	-	-	19	-	19

TABLE 13 b.

RA BSA		RA Split BSA		NAT BSA	G	3	4	4B	S2B	S2C	IJ
20	20	20	20	-	-	-	-	-	-	-	20
21	21	21	21	-	21	21	21	21	21	-	21
22	-	22	22	-	22	-	-	-	-	-	22
23	23	23	23	-	23	-	-	-	-	-	23
24	24	24	24	-	-	24	24	24	24	24	24
25	-	25	25	25	25	-	-	25	-	-	25
26	26	26	26	-	-	26	26	26	-	-	26
27	27	27	27	-	-	27	27	27	27	27	27
28	28	28	28	-	28	-	-	-	28	-	28
29	29	29	29	29	-	-	-	-	-	-	-
30	30	30	30	30	-	30	-	-	30	30	30
31	31	31	31	-	31	-	-	-	-	-	31
-	-	32	32	32	-	-	32	32	-	-	32
-	-	33	33	-	-	-	-	33	-	-	-
34	34	34	34	-	34	-	-	34	-	-	-
35	35	35	35	-	35	35	35	35	-	35	-
36	36	36	36	-	-	36	36	36	-	-	36
37	37	37	37	-	37	-	-	-	-	-	-
38	38	38	38	-	38	-	38	38	-	-	-
39	39	39	39	-	-	-	39	39	-	-	39
40	40	40	40	-	40	-	-	-	-	-	-
43	43	43	43	-	43	-	43	-	-	-	-
44	44	44	44	-	44	44	-	-	-	-	-

TABLE 13 c.

RA BSA		RA Split BSA		NAT BSA	G	3	4	4B	S2B	S2C	IJ
45	45	45	45	-	45	-	-	-	45	-	45
46	46	46	46	-	-	-	46	46	-	-	46
47	47	47	47	47	47	-	47	47	-	-	47
48	48	48	48	-	48	48	48	48	-	-	48
49	49	49	49	49	-	-	49	49	-	-	-
50	50	50	50	50	-	50	-	-	50	-	50
51	51	51	51	-	51	-	-	51	-	-	51
-	-	52	52	-	-	-	52	52	-	-	52
-	-	53	53	-	-	53	-	-	-	-	53
54	54	54	54	-	54	54	54	54	-	-	54
55	55	55	55	55	55	-	55	55	-	-	55
56	56	56	56	56	-	-	-	-	-	-	-
57	57	57	57	57	57	-	-	57	-	-	57
-	-	59	-	-	-	-	59	59	-	-	59
-	-	60	60	60	-	-	-	-	-	60	60
-	-	61	61	-	-	-	-	-	-	-	61
-	62	62	62	62	-	62	-	-	62	-	62
-	-	63	63	63	-	-	-	63	-	-	-
64	64	64	64	64	-	-	-	64	-	-	-
-	-	65	65	-	-	-	-	-	-	-	65
-	-	66	66	-	-	-	66	66	-	-	-

TABLE 13 d.

RA BSA		RA Split BSA		NAT BSA	G	3	4	4B	S2B	S2C	IJ
67	67	67	67	-	67	-	-	-	67	-	67
71	71	71	-	71	-	-	71	-	-	-	-
-	-	73	73	-	-	-	-	-	-	-	-
74	74	74	-	74	-	-	-	-	-	-	74
75	75	75	-	75	-	-	-	75	-	-	75
-	-	76	76	76	-	-	76	76	-	-	-
77	77	77	77	77	-	77	77	77	77	-	77
78	-	-	78	-	-	-	-	78	78	78	78
-	-	-	79	-	79	-	-	-	-	-	-
80	-	-	80	-	-	-	-	80	-	-	-
-	-	-	-	85	-	-	-	-	-	-	-
-	-	-	-	86	-	-	-	-	-	-	-
-	-	-	-	87	-	-	-	-	-	-	-

RA BSA - Reduced and alkylated bovine serum albumin

RA Split BSA - Reduced, alkylated, cyanogen bromide cleaved bovine serum albumin.

NAT BSA - Native bovine serum albumin.

G, 3, 4, etc. - The fragments described in the text obtained by cyanogen bromide action on reduced alkylated bovine serum albumin.

DISCUSSION

Discussion.

Before the individual results are discussed, some consideration must be given to their accuracy and therefore relevance to the problem under examination.

In all work on protein structure the initial requirement is for a pure protein in order that the results obtained may be representative of only one structure. The cytochrome c used would appear to satisfy all normal criteria of purity. Various reports have however appeared in the literature concerning the heterogeneity of myglobin and albumin preperations.

The myoglobins of several species have been fractionated during the purification procedures into multiple components. Thus seal myoglobin gave five components (Rumen, 1959), horse myoglobin three (Åkeson and Theorell, 1960) and sperm whale myoglobin four components (Edmunson and Hirs, 1962). In a recent study Atassi (1964) has obtained twelve components from sperm whale myoglobin. In all cases the components obtained were indistinguishable on the basis of chemical properties or amino acid composition. Separation was achieved by electrophoresis or chromatography on ion exchange celluloses, methods which depend on differences in the electrical charge on the components of the mixture. There is at present no evidence for differences in the primary sequence within the components obtained from each species. The isolation of different components appears to be due to differences in their amide content. This may represent a natural difference but

is more probably an artifact of the isolation procedure. The heterogeneity of the myoglobin preparation therefore would not be expected to cause any ambiguity or difficulty in the current investigation. Electrophoretic heterogeneity of serum albumin preparations was first shown by Luetscher (1939). This heterogeneity is due to isomerisation between the N form which exists above pH 4.5 and the F form which is produced when the pH is lowered below 4.5 (Aoki and Foster, 1956). The transformation is reversible. The mechanism of the N-F transition is obscure but appears to be due to slight differences in the secondary and tertiary structure of the protein (Sogami and Foster, 1963). The confusion caused by the apparent existence of two forms of albumin - albumin and mercaptalbumin - suggested by repeated determination of a sulphydryl content of 0.4-0.6 mole/mole of albumin, has been clarified by the observations of King (1961) that non-mercaptalbumin contains either an additional half cystine residue or a glutathione residue. The heterogeneity detected in ultracentrifugation and gel filtration is due to association of the monomer to give dimers and higher aggregates. In the absence of evidence to the contrary the preparation of albumin used can therefore be taken to contain a unique primary structure.

The next major consideration is the purity of the peptides obtained from the separation procedures following cleavage of the protein with cyanogen bromide. Separation has been

attempted in the present investigation by gel filtration. Gel filtration separates solutes on the basis of their molecular size. Small molecules have a greater access to the gel matrix and so are retarded. Larger molecules having a lesser access are eluted first. Successful separation of two components in a mixture therefore depends on a difference in their molecular size. The greater the difference the more probable is a complete separation. If complete separation is not achieved in the first instance due to insufficient difference in size, isolation of the enriched portion of the eluate followed by a repeated gel filtration yields a further purification which in favourable circumstances may be complete. Repetition of this procedure gives portions of eluate which contain increasing percentages of the desired component. The limit to the number of repetitions possible is imposed by the loss of yield at each step due to rejection of eluate containing the contaminant plus some of the desired component. In a number of instances this limit has operated in the present investigation to make analysis of a slightly impure peptide necessary. The columns used in this work were of the greatest length consistent with practical operation. Comparison of Figs.32 and 33 showing separations obtained on short and long columns respectively of Sephadex G 200 illustrates the desirability of using such columns in work involving complicated

mixtures of peptides where the optimum separation must be obtained. The technique of recycling chromatography introduced by Porath and Bennich (1962) would seem to be particularly applicable to the separation problem posed by the bovine serum albumin cleavage products. This technique eliminates the combination of the tubes containing the desired fraction which is necessary before lyophilisation and reapplication to the column in the procedure used in this work. This means that the eventual yield is increased and the separation at each step is improved.

In any work involving separation of a large number of components from a complex mixture a simple form of analysis to check the separation obtained at each step is very desirable. It serves as an indication of the utility of each procedure introduced into the isolation sequence and indicates the purity or otherwise of the final product. The cleavage products of cytochrome c and myoglobin were relatively simple and so did not pose major problems of component separation and identification. Nevertheless the application of cellulose acetate electrophoresis to the cytochrome c products was of great value in proving the non-identity of peptide 1A and cytochrome c. The failure to obtain such a simple electrophoretic technique for the products of bovine serum albumin degradation has greatly hindered their separation and purification. The complexity of the mixture only became obvious as attempted separation

proceeded. The purification could only be approached by selection of a peak in the initial separation followed by reapplication of gel filtration to give as symmetrical a peak as could be obtained in a practicable yield. In most cases this approach has appeared to give a reasonably pure product, however fraction 4 was only shown to consist of more than one component after amino acid analysis and fingerprinting of the supposedly purified material. Although partial separation of some larger components has been obtained on G 200 as shown in Fig. 33 no precise idea can be obtained of their number or location in the eluate.

The peptides obtained from the gel filtration procedures were first characterised by amino acid analysis.

This involves determination of the concentration of each amino acid present in an acid hydrolysate of the peptide, followed by calculation of the amino acid content of the peptide in terms of residues. Where accurate figures are not available for the molecular weight of the peptide, the number of residues of each amino acid present is obtained by division of the amino acid composition by a constant. The constant is selected to give values which most closely approximate to integers for the amino acids known to give reliable results on analysis. The method gave satisfactory results for the peptides obtained from cytochrome c and myoglobin. However with some of the larger peptides obtained from bovine serum albumin there is a possibility of error in the calculated

amino acid composition. This is a result of the inaccuracy of the amino acid analysis. The limits of accuracy achieved by the instrument in use are ±3%. As the number of residues increases so the limits of accuracy of the analysis become more significant. Thus the limits for determination of 1, 10 and 100 residues are 0.97 - 1.03, 9.7 - 10.3, 97 - 103. Any value within the first set of limits is close to an integer. It is still possible to obtain the correct integral value from the second set of limits, however this is no longer possible for an analysis of 100 residues. With twenty residues the limits are 19.4 - 20.6 and a method based on determination of integral values must depend on a Gaussian distribution of the results about the mean. The presence of small quantities of impurities in the peptides being analysed greatly increases the difficulty of selecting the correct constant.

A further difficulty in interpreting the results of amino acid analysis of peptides derives from the fact that homoserine, the product of cyanogen bromide reaction on methionine residues, was found to elute from the ion exchange column of the analyser together with glutamic acid. This means that many of the glutamic acid estimations are higher than they should be. Failure to obtain a separate peak for homoserine also means that the absence of homoserine can not be used to indicate the peptide arising from the C-terminal

end of the protein (Edmundson 1963, Hofmann 1964).

The fact that the amino acid composition of the peptides obtained accounts for the amino acid composition of the complete molecule and the correlation between the amino acid analysis and fingerprint data of the respective peptides suggests the absence of any serious error in the calculations.

The fingerprinting technique provided a useful way of distinguishing between different peptides and correlation of the number of peptides separated on the fingerprint with the determined lysine and arginine content provided some check on homogeneity. (However see later discussion of this point with reference to fraction 4 from bovine serum albumin). The method of fingerprinting.....p.93.

The method of fingerprinting does not give exactly reproducible results, mainly due to variations in the wetness of the paper and therefore of the current passing during the electrophoretic separation. In the comparison of different fingerprints it has been necessary to identify the spots by their positions relative to the general pattern obtained. This can be carried out most successfully where there are a large number of spots. With some of the smaller peptides identification of the component spots was particularly difficult and the selections made may well be erroneous.

Additional chance of error is introduced if the peptides used for tryptic digestion are impure. This would not be so important if tryptic peptides all gave the same colour with ninhydrin, however the amount of colour varies greatly with the structure of the peptide. It is thus possible for an impurity to give tryptic peptides some of which yield spots of an intensity similar to that obtained for some of the peptides from the major component. They would therefore be assumed to be tryptic peptides of the major component.

With these points in mind it is now possible to make a detailed examination of the results obtained. The amino acid sequence of cytochrome c is known from the work of Margoliash et al (1961). Cyanogen bromide cleavage at the two methionine residues should give three peptides of different size and of known amino acid composition. There is slight

variation among literature reports for the amino acid composition of horse myoglobin. These are shown in Table 14 together with the results of analysis of the myoglobin used in this work. They all show a common methionine content of two and so cyanogen bromide cleavage of myoglobin would be expected to yield a maximum of three peptides. Of the various analyses reported for bovine serum albumin, that of Spahr and Edsall (1964) is taken as being representative of the best values obtainable by the present analysis methods. Their figures are shown in Table 15 together with the amino acid contents of cytochrome c and myoglobin. The value of four methionine residues indicates that a maximum of five peptides should be obtained from complete cleavage of albumin.

TABLE 14.

Comparison of various reported amino acid compositions for horse heart myoglobin.

	(1)	(2)	(3)	(4)	(5)
Asp.	10	10	10	10	10
Thr.	7	7	6	7	6
Ser.	5	6	5	5	4
Glu.	19	19	19	19	20
Pro.	5	5	4	4	4
Gly.	15	13	13	16	15
Ala.	15	15	14	15	14
Val.	7	6	6	7	7
Met.	2	2	1 or 2	2	2
Ileu.	9	22	22	9	8
Leu.	17			17	17
Tyr.	2	2	2	2	1
Phe.	7	5	6	7	7
Lys.	18	18	17	19	+
His.	10	9	9	11	11
Arg.	2	2	2	2	2

(1) Boardman and Adair (1956).

(2) Tristam (1949).

(3) Holleman and Biserte (1959).

(4) Åkeson and Theorell (1960).

(5) Values obtained for the preparation used in this work.

TABLE 15.

Amino acid analyses of the proteins used in this work.

	Bovine serum albumin	cytochrome c	myoglobin
Asp.	54	8	10
Thr.	34	10	6
Ser.	26	-	4
Glu.	78	12	20
Pro.	30	4	4
Gly.	16	12	15
Ala.	46	6	14
Cys.	36	2	-
Val.	37	3	7
Met.	4	2	2
Ileu.	14	6	8
Leu.	62	6	17
Tyr.	20	5	1
Phe.	27	4	7
Lys.	62	19	+
His.	17	3	11
Arg.	22	2	2

Cytochrome c.

The amino acid content of the three peptides obtained from cytochrome c has been compared in the results section with the values to be expected from the sequence of Margoliash et al (1961).

Table 5 shows the comparison of the results for peptides 2 and 3 with the sequences 80 - 104 and 65 - 80 from the proposed amino acid sequence. The agreement is almost complete. The values obtained for tyrosine are the only major point of difference. The explanation for the low results with tyrosine is not clear since similar effects have not been noted with the other proteins. The observation of a reaction of cyanuric halides with bound tyrosine in proteins by Kurihara, Horinishi and Shibata (1963) offers a possible explanation since cyanuric bromide could arise by decomposition of cyanogen bromide. They do not however propose a mechanism for the reaction.

Table 6 shows the comparison of peptide 1A and the sequences 1 - 65, which is the peptide expected from complete fission of cytochrome c, and 1 - 80, which would arise from cleavage only at the methionine next to the C-terminal end of the protein. Since the porphyrin is attached to the 1 - 65 sequence both of these peptides would be coloured. The presence of three coloured bands on cellulose acetate electrophoresis of the unfractionated cleavage products suggests

that both peptides are present. The purified peptide 1A gave a single band on electrophoresis which was clearly distinguishable from that of cytochrome c. The lack of better agreement between peptide 1A and one of the sequences is rather disappointing and probably reflects insufficient purification of the peptide. However from the figures obtained peptide 1A appears to correspond to the 1 - 80 sequence.

The expected sequence 1 - 65 must be present in the unfractionated cleavage products since sequences from residue 65 onwards have been obtained but it has never been isolated and purified.

Cyanogen bromide has therefore split cytochrome c selectively into four peptides, three of which are obtained by complete cleavage at the two methionines of cytochrome c. The other arises from a more rapid cleavage of the met-ileu bond at 80-81 than the met-glu bond at 65-66.

Myoglobin.

The amino acid analyses taken to the nearest integer and the N-terminal amino acid of myoglobin and the fractions 1B, 2 and 3 are shown in Table 16. It can be seen that the addition of 1B, 2 and 3 gives an amino acid composition very similar to that of myoglobin. The difference is most probably due to incomplete purification of fraction 1B since the separation shown in Fig. 25 is by no means complete. The results do however show that myoglobin has been selectively cleaved into the three peptides 1B, 2 and 3. An additional peptide, 1A, was also obtained from the cleavage products which, like the peptide 1A from cytochrome c, arises from cleavage at only one of the methionyl bonds. Table 17 shows how the amino acid composition of peptide 1A can be accounted for by addition of peptides 1B and 2. Once again discrepancies are probably due to incomplete peptide purification. From the N-terminal determinations of the four peptides the order in the intact protein is seen to be 2, 1B and 3, with 2 being the N-terminal peptide. With myoglobin the met-lys bond has split faster than the met-ser bond to give rise to peptide 1A.

It is difficult to obtain a relationship between the constituents of the susceptible bond and the extent of cleavage since Gross and Witkop (1962) found that in ribonuclease met-ser was split to a greater extent than met-lys, in direct contradiction to the results obtained with myoglobin.

At the present time insufficient evidence is available to make any existing relationship clear.

Edmundson (1963) obtained three fragments from sperm whale myoglobin which accounted for all of the amino acids of the complete protein. Table 18 shows a comparison of the amino acid compositions of the two proteins and of their N-terminal, middle and C-terminal peptides. The respective peptides are of the same order of size. The similarity in the compositions of the two proteins is also evident in the peptides. However similarity is not greater in one of the peptides than in any of the others. In other words, if parts of the primary sequences are common to both proteins they do not appear to be restricted to any one region of the molecules. It remains to be seen whether the similarity of methionine distribution is coincidental or whether it is indicative of a greater similarity between the structures of the two proteins.

TABLE 16.

Comparison of the amino acid compositions of myoglobin and its fractions 1B, 2 and 3 which together account for the composition of myoglobin.

	Myoglobin	1B	2	3
Asp.	10	4	4	3
Thr.	6	3	2	1
Ser.	4	3	1	1
Glu.	20	7	9	4
Pro.	4	3	1	-
Gly.	15	7	6	3
Ala.	14	8	4	4
Val.	7	3	4	1
Ileu.	8	5	2	1
Leu.	17	8	6	4
Tyr.	1	1	-	1
Phe.	7	2	3	2
His.	11	7	3	1
Arg.	2	-	1	1
N-terminal	Gly.	Ser.	Gly.	Lys.

TABLE 17.

Comparison of the amino acid compositions of fractions 1A, 1B and 2 from myoglobin showing that 1A contains the 1B plus 2.

	1A	1B	2
Asp.	9	4	4
Thr.	6	3	2
Ser.	3	3	1
Glu.	16	7	9
Pro.	4	3	1
Gly.	12	7	6
Ala.	12	8	4
Val.	7	3	4
Ileu.	7	5	2
Leu.	13	8	6
Tyr.	1	1	-
Phe.	5	2	3
His.	10	7	3
Arg.	1	-	1
N-terminal	Gly.	Ser.	Gly.

TABLE 18.

Comparison of the peptides obtained from horse heart myoglobin with the corresponding peptides from sperm whale myoglobin.

	Whole protein		N-term.		middle		C-term.	
	SW	HH	SW	HH	SW	HH	SW	HH
Asp.	8	10	3	4	3	4	2	3
Thr.	5	6	2	2	3	3	0	1
Ser.	6	4	2	1	4	3	0	1
Glu.	19	20	8	9	7	7	2	3
Pro.	4	4	1	1	3	3	-	-
Gly.	11	15	3	6	6	7	2	3
Ala.	17	14	4	4	10	8	3	4
Val.	8	7	4	4	3	3	0	1
Ileu.	9	8	2	2	6	5	1	1
Leu.	18	17	7	6	8	8	3	4
Tyr.	3	1	0	0	1	1	2	1
Phe.	6	7	3	3	2	2	1	2
His.	12	11	4	3	8	7	0	1
Arg.	4	2	2	1	1	0	1	1
Total	134	128	47	46	65	61	18	27

SW sperm whale myoglobin.

HH horse heart myoglobin.

Bovin serum albumin.

The amino acid analyses shown in Table 10 are all of intact albumin or its unfractionated cleavage products.

The first two columns show the results obtained after 24 and 72 hr. hydrolyses of the native protein. The values for threonine, serine and cysteine show the expected reduction with increase in hydrolysis time. The other values are similar except for those for isoleucine and leucine which are lower after 72 hr. instead of higher as would be expected due to the slower acid hydrolysis of peptide bonds involving these amino acids.

The efficiency of the reduction and alkylation procedure was checked by amino acid analysis of the reduced alkylated protein. Results are shown in columns 3 and 4 of Table 10. Values for several of the amino acids were not obtained since the column was overloaded in an attempt to detect small quantities of cysteine. The analysis figures show that all of the cystine and cysteine have been successfully converted to carboxymethyl cysteine. There is considerable disagreement between the carboxymethyl cysteine content of the reduced alkylated protein and the determined cysteine content of the native protein. However it is well known that cysteine undergoes destruction during

acid hydrolysis (Light and Smith, 1963) so the carboxymethyl cysteine value is probably more reliable. This conclusion is supported by the half-cystine value of 36 given by Spahr and Edsall (1964) for the native protein as determined by the cysteic acid method of Moore (1963).

Column 5 of Table 10 shows the values obtained when reduction was carried out in 0.1 N HCl followed by alkylation at pH 8.6. Since the amino acid analysis of the fully reduced alkylated protein shows by the absence of cysteine that alkylation is complete under these experimental conditions, incomplete reduction of cystine would be expected to give rise to an even number of cysteine residues in the hydrolysate. The probable value is therefore four, which would mean that two cystine residues are still intact following reduction in the acid medium. This result follows from the work of Katchalski, Benjamin and Gross (1957) who showed that the degree of reduction varied with pH.

The last two columns of Table 10 give the amino acid analyses obtained with the reduced alkylated albumin after treatment with cyanogen bromide. These analyses were carried out in an attempt to determine whether peptide material was being lost at the step of acid acetone precipitation in the isolation of the split products from the 8 M urea solution. There is

no obvious difference between these results and those obtained without cyanogen bromide treatment which would be suggestive of such a loss. Two points however arise from the figures. After cyanogen bromide treatment there is a reduction in the quantity of carboxymethyl cysteine which takes place with a corresponding appearance of cysteine. The possible significance of this observation will be discussed later. The methionine value of 2 shows that only 50% reaction with cyanogen bromide has taken place.

Bovine serum albumin was chosen as a readily available protein which would pose problems representative of the sequence determination of large, single polypeptide chains. The difficulty of separation and isolation of the components of the complex cyanogen bromide degradation products became immediately obvious. The subsequent discussion should show that some success has been obtained with the smaller fragments although a considerable number of the larger fragments defy resolution.

Some explanation should perhaps be offered for the divergence between the expected number of fragments and the considerably larger number apparently obtained. It has been seen from the work on cytochrome c and myoglobin that unexpected peptides can be obtained from

partial cleavage of the susceptible bonds in the protein. Albumin with four susceptible bonds offers greater opportunity for partial cleavage. If none of the susceptible bonds are adjacent, partial cleavage at each of these bonds could give rise to fifteen different products including complete unsplit albumin. The irregularity of the upper curve in Figure 34 suggests that there may indeed be this many components produced.

A correlation between the arginine and lysine content of a protein or peptide and the number of spots obtained after tryptic digestion and fingerprinting generally indicates purity. This technique has been used in this work to characterise the peptides obtained from serum albumin. The technique has been invaluable but in this context has been shown to suffer from one limitation. An equivalent result is obtained with a single component or an equimolar mixture of components. Whilst the occurrence of such a mixture would be extremely unlikely in a protein preparation it has been found to occur in these peptide fractions. The possible errors involved in amino acid analysis and fingerprinting have already been discussed.

The amino acid analyses of fractions 6, 3 and 4 are shown in Table 12. The values obtained give arginine plus lysine values of 30, 19, and 33 for 6, 3 and 4

TABLE 19.

Calculation of molecular weight of fraction G.

Amino Acid	No. of residues	Corres. weight
Asp.	20	2660
Thr.	8	952
Ser.	7	735
Glu.	28	4120
Pro.	9	1035
Gly.	6	450
Ala.	15	1336
Cys.	10	1210
Val.	13	1520
Ileu.	6	786
Leu.	20	2620
Tyr.	6	1086
Phe.	7	1155
Lys.	20	2920
His.	5	775
Arg.	10	1740
Total	190	25,100

25,100 less 190 molecules of water formed during peptide bond synthesis gives 21,680, as the approximate molecular weight of fraction G.

TABLE 20.

Comparison of the amino acid composition of bovine serum albumin and the combined amino acid compositions of its fission products.

	BSA	G+3+4.
Asp.	54	59
Thr.	34	30
Ser.	26	20
Glu.	78	89
Pro.	30	29
Gly.	16	22
Ala.	46	45
Cys.	36	23
Val.	37	35
Ileu.	14	17
Leu.	62	66
Tyr.	20	16
Phe.	27	31
Lys.	62	62
His.	17	24
Arg.	22	22
Total	581	590

respectively. Table 13, a, b,c,d lists the fingerprint spots of each of these fractions. Fraction G gives 33 spots, 3 gives 20 and fraction 4 gives 31. The agreement is therefore well within the limits of experimental accuracy. On this basis fractions G, 3 and 4 contain either pure peptides or an equimolar mixture of pure peptides. It is impossible to make any distinction between the two possibilities from these results alone.

The calculation of the molecular weight from the amino acid composition is shown for fraction G in Table 19. The molecular weights calculated in this way from the amino acid compositions of the fractions are 21,680 for G, 15, 540 for 3 and 30,120 for 4. Addition of these values gives 67,340 which can be compared with the value of 65,560 calculated for bovine serum albumin in a similar fashion from the amino acid composition given by Spahr and Edsall (1964). Within the experimental limits fractions G, 3 and 4 account for the complete amino acid composition of bovine serum albumin. Table 20 shows a comparison of the amino acid composition of bovine serum albumin with the values obtained by addition of the compositions of fractions G, 3 and 4. Agreement is not complete. However when the errors involved are taken into consideration the similarity

between the two sets of values is remarkable. The major point of deviation is in the values for glutamic acid. This will partially be due to inclusion of homoserine in the glutamic acid peaks of some of the fractions.

Comparison of the fingerprint spots of the three fractions shows that they have many spots in common. This explains why the reduced alkylated protein with a lysine plus arginine content of 84 only gives 59 fingerprint spots. This cannot be taken to indicate that the fractions have sequences in common but rather that the common spots represent different peptides which have similar properties under the conditions of fingerprinting and are therefore indistinguishable. The total number of spots contained in the three fractions is 84 which is in better agreement with the lysine and arginine content of the complete protein.

The methionine content of bovine serum albumin gives a theoretical expectancy of five fragments, two more than the number obtained which apparently account for the complete albumin molecule. This could mean that only three peptides have been obtained since three of the methionines either exist together or almost together in the sequence or are near to the N or C-terminal ends of the protein in which case the product of cyanogen bromide treatment would either be homoserine

or a small peptide which would escape detection. The other possibility is that some of the fractions contain more than one peptide. The calculated molecular weight of fraction 4 (30,120) is not consistent with its position in the elution patterns of gel filtration. It elutes after fraction 3 which has a calculated molecular weight of 15,540. Since there is an agreement between the lysine and arginine content and the number of fingerprint spots of fraction 4 any components it contains must be present in equimolar amounts. Since fraction 4 is obtained from a limited area of the elution curve any components must be of a similar molecular weight. Two components would require molecular weights of 15,000 which would make separation from fraction 3 improbable on gel filtration. Three components of approximate molecular weight 10,000 would fit better with the known facts. This would also account for the expected five fragments from albumin.

Fraction 4B was isolated in an attempt at separation of the components. However the amino acid compositions and fingerprint results of 4B and 4 do not show any significant difference (Table 12 and Table 13, a, b, c, d). No success has therefore been achieved in attempts to separate the components of fraction 4.

The analysis of the albumin reduced in the acid medium

suggested that it contains two intact cystine residues. Cyanogen bromide treatment of this material might give a slightly less complicated mixture of products since it is conceivable that some of the components obtained with the completely reduced albumin be joined by the cystine residues and so behave as single fragments. This would of course only be true if the partial reduction was selective. Random partial reduction would give additional complexity.

Comparison of Figs. 31 and 41 of separation of the fully and partially reduced material after cyanogen bromide treatment shows a simpler elution pattern for the partially reduced material. This supports the view that variation in the extent of reduction by pH alteration is a selective process. The fraction 4 indicated in Fig 18 appears either to be entirely absent or greatly reduced in the products of the partially reduced material.

Figs 33 and 42 are of separations on G200 of the equivalent fractions 11, from the fully reduced, and S1, from the partially reduced material. The graphs of ultraviolet absorption suggest a simpler composition for fraction S1. However the curve of ninhydrin estimation after alkaline hydrolysis shows that S1 still represents a complex mixture. Partial reduction would appear to alter the proportions of the constituents. It may be that the continuing complexity represents contamination of

the partially reduced albumin by some of the completely reduced material.

Fraction S2B is one of the smaller components of the partially reduced albumin. The fingerprint and amino acid analysis of S2B suggest that S2B and 3 are due to the same peptide. It must however be noted that the fingerprint of S2B shows a number of spots which are found in fraction G and not in fraction 3. This indicates contamination with fraction G but raises the question of the existence or otherwise of this fragment in the products of the partially reduced protein. No evidence can be presented towards a solution.

Fraction S2C, the smallest component isolated, gives 10 fingerprint spots and has an arginine and lysine content of 13. The amino acid composition of S2C gives an approximate molecular weight of 12,800. From the previous discussion it can be seen that this is the order of size expected for the components of fraction 4. The identity of S2C with one of the components of fraction 4 is supported by comparison of the fingerprint spots of the two fractions. The majority of the spots found for S2C are also present in fraction 4.

Returning to the products of the fully reduced protein, the fingerprint data obtained with fraction IJ, from the area shown in Fig 34 between G and the peak of 3, when

compared with the data for fractions G and 3 show that fraction IJ can not represent a pure peptide. This result is not unexpected since the separations obtained for fractions I_1 and J_1 were not complete (Figs 39 and 40) and combination of I_1 and J_1 was not based on any real evidence for their similarity. Examination of the spots obtained shows the presence of several spots which are only present in fraction 4 and are absent from fractions G and 3. This suggests that fraction IJ contains some of the components of fraction 4 which in view of its elution volume could only be possible if two components were present as a single polypeptide chain. If the three components of fraction 4 are of a similar size a polypeptide containing two of them would have a molecular weight of the order of 20,000. This would mean that such a polypeptide would elute between fractions G and 3 on gel filtration, as seems to be the case. Since virtually all of the spots of fraction S2C appear in the fingerprint of IJ, the fraction IJ must contain the peptide S2C linked in a single polypeptide chain with one of the other components of fraction 4.

Some additional information can be obtained from the results of the fingerprinting. Tryptic digestion of reduced alkylated albumin would be expected to yield four peptides containing methionine. On cyanogen bromide

treatment of the reduced alkylated protein followed by tryptic digestion two new peptides should be obtained from each of the methionine containing peptides present in the original digest. If the cyanogen bromide cleavage is incomplete at each of the methionine bonds, fingerprinting after tryptic digestion of the reduced, alkylated, cyanogen bromide treated protein should give the four originals plus the eight new spots. The fingerprints of the fractions should not contain any of the original methionine containing spots.

From Table 13, a, b,c,d, of the fingerprint spots obtained for the reduced alkylated protein, the reduced alkylated split protein and the derived fractions it is found that spots 20, 29, 56 and possibly 64 or 74 are only present in the proteins, where methionyl bonds are still intact. They are absent from the fractions with the exception of spot 20 which is found in fraction IJ. This would be expected if fraction IJ represents a single polypeptide chain containing two of the components of fraction 4 with the methionyl bond intact.

None of the spots 19, 32, 33, 52, 53, 59, 60, 61, 63, 65, 66, 73, 76 or 79 are present in the fingerprint of the reduced alkylated protein. If spots of questionable significance are eliminated 19, 32, 52, 53, 59, 66 and 79 can be considered with reasonable certainty to arise from

splitting of the methionyl bonds. Spots 19 and 53 are found in fraction 3, spots 32, 52, 59 and 66 in fraction 4 and spot 79 in fraction 6. The isolation of more than two spots arising from methionyl bond fission in fraction 4 is additional evidence that it contains more than one component. Fragments from the N and C-terminal ends of a protein should contain only one of the spots arising from methionyl cleavage. Fragments from within the molecule should contain two such spots. The results obtained in the present work are insufficiently accurate to allow this principle to be applied with certainty; however fraction 3 would appear to be within the molecule and either two of the fragments from fraction 4 or fraction 6 to be at the N and C-terminal ends of the molecule.

In summary, it has been shown that reduced alkylated bovine serum albumin is split by cyanogen bromide into five fragments, which account for the amino acid composition and tryptic peptides of the intact protein. Three of these fragments have been isolated. Evidence has been obtained for the existence of a peptide composed of two of the smaller fragments joined together. Other partial degradation products are present but have not been adequately separated.

The assumption made at the beginning of this work was that bovine serum albumin exists as a single polypeptide chain. The only report which has appeared in contradiction to this is

by McLure, Schieler and Dunn (1953) as a result of their N-terminal determinations by the phenylthiohydantoin method. No opposing evidence has been obtained in the present work to the single chain theory. Circumstantial evidence for some form of subunit structure has been presented by the results of Reichman and Colvin (1955), Porter (1957), Richard and Kegeles (1959) and Adkins and Foster (1964). Foster (1960) on the results of the titration behaviour and other properties proposed a structure containing four "subunits" composed of large areas of ordered structure joined together by parts of the polypeptide chain having a more random order. The results obtained from the fingerprinting of the tryptic digestion products of native bovine serum albumin might possibly have some bearing on this proposed structure.

Very little is known of the configurational freedom necessary before tryptic digestion of a polypeptide chain can be achieved. The fact that native proteins are often completely resistant to tryptic attack or are only partially digested to leave a tryptic core has been known for some time. Denaturation generally renders a protein susceptible to tryptic digestion. A random configuration therefore seems the most probable requirement for attack. The α helix is the most obvious example of an ordered polypeptide chain arrangement and the

results of Kendrew et al (1961) indicate that the tryptic core of myoglobin arises from helical regions of the molecule. However other helical regions of the molecule are digested by trypsin, and chymotrypsinogen, which has not been shown to contain any α helix, is notably resistant to trypsin in the native form. Nevertheless attack at random regions seems more probable than at regions possessing a higher order. It therefore seemed worthwhile to examine the origin of the fingerprint spots obtained from native bovine serum albumin which are compared with those of the fractions 6, 3 and 4 in Table 13, a, b, c, d. If examination is restricted to spots which are present only in one or other of the fractions and therefore not open to ambiguity, 8 of the spots arise from fraction 6, 4 from fraction 3 and 4 from fraction 4. In a similar manner the tryptic core contains 9 spots from fraction 6, 2 from fraction 3 and 4 from fraction 4 neglecting spots only obtained after methionyl bond cleavage. From these results it appears that the fractions contribute equally to the soluble digest and to the tryptic core. This tends to support a "subunit" structure of the type proposed by Foster (1960). If the protein was in a form analogous to a ball of wool, digestion would be restricted to the loose end or ends of the chain. The tryptic peptides would then arise from one, or at the

most two, of the fragments obtained by cyanogen bromide fission. The fact that this is not so suggests that the protein does contain ordered segments interspersed with parts of a more random nature. Isolation and fractionation of the tryptic core would help to clarify this point. The large number of cystine residues present in bovine serum albumin of course complicate/any structural consideration of the molecule.

Cyanogen bromide has thus been shown to be a successful and useful reagent for the specific cleavage of proteins. It has been shown to give fragments of a size suitable for primary structural determination from the 560 residue-containing bovine serum albumin molecule, making determination of the primary structure of the complete protein feasible. However improvements are indicated before such work is attempted. The incomplete cleavage is an undesirable feature of the reagent under the conditions used in these experiments. While the isolation of partial cleavage products can be useful in aligning the products of complete fission in their correct sequence, as has been shown with the myoglobin results, their presence does complicate the isolation of the desired fragments. This has been particularly so with cytochrome c where the expected haem-containing peptide has yet to be obtained from the mixture of cleavage products.

It also means that the yields of the desired fragments are considerably lower than would otherwise be obtained. This is an important consideration since many proteins are available only in limited quantities. The yield of each purified peptide obtained from 100 mg of bovine serum albumin was of the order of only 2-3 mg, as compared with the theoretical yield of 20 mg. Considerable quantities of each peptide would be required before detailed structural studies could be continued. There is therefore a need for a more complete cleavage to increase the value of cyanogen bromide in the specific cleavage of proteins.

The effect of the reaction conditions on the extent of cleavage has not been studied in the present work. Hofmann (1964) found only 50% cleavage of trypsinogen when reaction was carried out in 0.1 N HCl for 24 hr. at room temperature. The methionine content of reduced alkylated albumin after cyanogen bromide cleavage shows the same percentage of fission under the identical conditions. He however obtained 92-95% cleavage by increasing the reaction temperature to 30° and the time of reaction to 30 hr. A similar extent of cleavage was obtained by carrying out the reaction in 50% formic acid. Examination of the effect of these measures on the percentage cleavage of bovine serum albumin is desirable.

quantitative N-terminal determination would serve as the best measure of the cleavage obtained during the varying conditions. In the work on trypsinogen a non-specific peptide bond cleavage was obtained under the acidic conditions of the cyanogen bromide reaction. No similar non-specific cleavage has been detected in the present work.

The work on the products of cleavage of bovine serum albumin in particular has shown up the lack of really effective methods for the preparative separation of large peptides. Recycling gel filtration is an advance on the existing technique which could yield results. The present work has shown that even for the larger cleavage products of bovine serum albumin a partial separation is obtained on Sephadex. However the components of fraction 4 from bovine serum albumin are in need of a different approach. Most of the other methods of peptide separation depend on charge differences on the peptide. It is for this reason that the failure to obtain a separation of the albumin products on electrophoresis is rather disappointing, since this indicates a parity in the charge carried at the pHs of the attempted separation. Further investigation of this field might help to indicate the direction of future separation and purification attempts. Study of the solubilities of the albumin products would also be useful

since in the present work it has been necessary to use 50% acetic acid to effect solution of the lyophilised material. Few of the separation techniques are capable of dealing with such a solvent.

Hofmann (1964) split the cystine bonds in trypsinogen after cyanogen bromide cleavage of the protein. In the case of bovine serum albumin it was felt that access of the reagent to the susceptible linkages would be improved by reduction of the cystine residues before cyanogen bromide treatment. The alkylation of cysteine to give carboxymethyl cysteine might however introduce a new functional group into the protein. It is possible that the cyanogen bromide react with the carboxymethyl cysteine sulphur in an analogous manner to its reaction with methionine sulphur (Fig.8). However complete reaction to give cleavage seems improbable since carboxymethylthiocyanate is not a good leaving group and cleavage would require the formation of a four membered lactone which is an extremely unlikely structure. Nevertheless the difference between the carboxymethyl cysteine content of the reduced alkylated protein before and after cyanogen bromide treatment (Table 10) shows that a reaction of some type has taken place. It is possible that the initial formation of an intermediate at room temperature is followed by further reaction at the temperature of hydrolysis to

give cysteine which appears in the hydrolysate. Cleavage at the thirty six carboxymethyl cysteine residues produced after alkylation would give products of an even greater complexity than those actually obtained. There is no evidence from the present work of any cleavage by cyanogen bromide at peptide bonds involving residues other than methionine.

SUMMARY

SUMMARY.

Fulfilment of the need for specific agents to cleave the polypeptide chain at highly specific locations is one of the most important prerequisites for further advances in the primary structural determination of proteins. Cyanogen bromide is a chemical agent known to cause selective fission of protein molecules at methionyl bonds. In the present work its action on the three proteins, horse heart cytochrome c, horse heart myoglobin, and bovine serum albumin has been investigated. The products of reaction were fractionated by chromatography on Sephadex gels.

1. Cytochrome c, which contains two methionine residues, yielded three peptides which were subjected to amino acid analysis. The peptides 2 and 3, except for low tyrosine values, had the composition to be expected from the published work for the sequences 65-80 and 81-104. The third peptide, 1A, did not give figures in agreement with the 1-65 sequence. It appeared to represent a pure product, free from cytochrome c contamination, and on further inspection was shown to have a composition similar to that for the sequence 1-80. Peptide 1A therefore arises from fission at only one methionine locus, that nearest the C-terminus of the peptide. The presence of the peptide corresponding to the sequence 1-65 is supported by the results of electrophoresis of the

reaction products, but it has not been isolated.

2. Myoglobin, which also contains two methionine residues, gave four peptides on cyanogen bromide treatment followed by gel filtration. The peptides 1B, 2 and 3 account for the amino acid content of myoglobin and are therefore the products of complete fission of the molecule. The compositions of peptides 1B and 2 together give the composition of the other peptide, 1A, which is therefore a partial degradation product. N-terminal determination has allowed the peptides to be placed in the order 2,1B and 3, with 2 containing the N-terminal residue of the intact molecule. The peptides obtained from complete fission of the protein molecule are similar in size to the corresponding peptides which have been obtained from sperm whale myoglobin.

3. Since bovine serum albumin contains four methionine residues it would be expected to yield five fragments when treated with cyanogen bromide. The protein was first reduced and alkylated, and after cyanogen bromide treatment the initial separations on Sephadex G 75 appeared to indicate the presence of five components. Further separation of one of the fractions on G 200 however showed that it had a complex composition. Separation on coupled columns of Sephadex G 200 and G 75 indicated that approximately fifteen components might be present

which would be the case if partial fission was taking place at each of the four methionine residues in the albumin molecule. From the Sephadex separations several fractions were isolated and examined by amino acid analysis and fingerprinting after tryptic digestion. Three of the fractions, 6, 3 and 4, account for the amino acid composition and fingerprint spots of the intact albumin molecule. Fraction 6 has a calculated molecular weight of 20,000 and fraction 3, 15,000. The amino acid content and elution behaviour of fraction 4 suggest that it contains three components of approximate molecular weight 10,000. This would account for the five fragments expected from bovine serum albumin.

Bovine serum albumin was also partially reduced by performing the reduction in 0.1 N HCl. Cyanogen bromide treatment of the partially reduced protein gave a simpler mixture of products. Two of the components have been isolated. Amino acid analysis and fingerprinting of one, S2B, indicates that it is identical with fraction 3 obtained from the fully reduced protein. The other component has a calculated molecular weight of 12,000, and appears to be one of the components contained in fraction 4 from the fully reduced protein.

Evidence has also been obtained that a fraction eluting between 6 and 3 on gel filtration of the products of the fully reduced protein contains the peptide S2C

joined by an intact methionyl bond to one of the other components of fraction 4.

By use of the fingerprinting technique after tryptic digestion it has been possible to identify some of the tryptic peptides which contain methionine and so represent sequences joining the cyanogen bromide produced fragments in the intact polypeptide chain. Tryptic peptides have also been identified after cyanogen bromide treatment which represent the N-terminal and C-terminal sequences of the fragments produced by methionyl bond fission.

The tryptic peptides obtained from native bovine serum albumin have been compared with those contained in the three fractions 6, 3 and 4, and by their distribution in these fractions support a form of "subunit" structure for the native albumin molecule.

BIBLIOGRAPHY

BIBLIOGRAPHY.

Adkins, B.J. and Foster, J.F. (1964) Fed. Proc., 23, 474.

Akabori, S., Ohno, K. and Narita K. (1952)
Bull. Chem. Soc. Japan, 25, 214.

Åkeson, Å. and Theorell, H. (1960)
Arch. Biochem. Biophys., 91, 319.

Ambler, R.P. (1963) Biochem. J., 89, 349.

Anderer, F.A. (1959) Z. Naturforsch, 14b, 642.

Anderer, F.A., Uhlig, H., Weber, E. and Schramm, G. (1960)
Nature, Lond., 186, 922.

Anfinsen, C.B., Sela, M., and Tritch, H. (1956)
Arch. Biochem. Biophys., 65, 156.

Anfinsen, C.B., Åqvist, S.E.G., Cooke, J.P. and Jönsson, B.
(1958) J. biol. Chem., 234, 1118.

Ansevin, A.T. and Lauffer, M.A. (1959) Nature, Lond., 183, 1601.

Aoki, K. and Foster, J.F. (1956)
J. Amer. Chem. Soc., 78, 3538.

Atassi, M.Z. (1964) Nature, Lond., 202, 496.

Avey, H.P., Carlisle, C.H. and Shukla, P.D. (1962)
Brookhaven Symposia in Biology, 15, 199.

Beaven, G.H. and Holiday, E.R. (1952)
in Adv. in Prot. Chem., Vol. 7, 320.

Benson, J.V. and Patterson J.A. (1964)
Fed. Proc., 23, 371.

Berger, A., Kurtz, J. and Noguchi, J. (1958)
in "Recent Advances in Gelatin and Glue Research",

p.271, (Ed. by Stainsby, G.) Permagon Press, New York.

Bernal, J.D., (1958) Discuss. Faraday Soc., 25, 7

Bishop, J., Leahy, J. and Schweet, R. (1960)
Proc. nat. Acad. Sci., Wash., 46, 1030.

Blaedel, W.J.and Todd, W.J. (1961)
Analyt. Chem., 33, 205.

Blow, D.M. (1964)
Abstr. Fed. Europ. Biochem. Soc., 135.

Boardman, N.K. and Adair, G.S. (1956)
Nature, Lond., 177, 1079.

Braunitzer, G., Gehring-Muller, R., Hilschmann, N., Hilse, K.,
Hobom, G., Rudloff, V. and Wittmann-Liebold, B. (1961)
Hoppe-Seyl. Z. 325, 283.

Brown, J.R., Cox, J.D., Greenshields, R.N., Walsh, K.A.,
Yamasaki, M. and Neurath, H. (1961)
Proc. nat. Acad. Sci., Wash., 47, 1554.

Canfield, R.E. and Anfinsen, C.B. (1963 a).
Biochemistry, 2, 1073.

Canfield, R.E. and Anfinsen, C.B. (1963 b)
in The Proteins 2nd. Edn. Vol. 1, p311 (Ed. by Neurath, H)
Academic Press, Ney York.

Carlton, B.C. and Yanofsky, C. (1963)
Fed. Proc., 22, 656.

Corfield, M.C. and Robson, A. (1962)
Biochem. J., 84, 146.

Craven, R.G., Steers, E. and Anfinsen, C.B. (1964)
Fed. Proc., 23, 263.

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

Crick, F.H.C. (1958)
Symp. Soc. Exptl. Biol., 12, 138.

Deal, W.C. and van Holde, K.E. (1962)
Fed. Proc., 21, 254.

Dintzis, H.M. (1961)
Proc. nat. Acad. Sci., Wash., 47, 247.

Dixon, G.H. and Wardlaw, A.C. (1960)
Nature, Lond., 188, 721.

Doolittle, R. and Blomback, B. (1964)
Nature, Lond., 202, 147.

Doscher, M.S. and Richards, F.M. (1963)
J. biol. Chem., 238, 2399.

Durrum, E.L. (1950)
J. Amer. Chem. Soc., 72, 2943.

Dus, K. and Kamen, M.D, (1963) Biochem. Zeit., 338, 364.

Eastoe, J.E. (1960) Biochem. J., 74, 8P.

Eastoe, J.E. (1961) Biochem. J., 79, 653.

Ebata, M. and Akabori, S. (1960) Seikagaku, 31, 588.

Ebata, M. (1961) J. Biochem.(Tokyo), 49, 110.

Edman, P. (1950) Acta chem. Scand., 4, 283.

Edmundson, A.B. and Hirs, C.H.W. (1962)
J. molec. Biol., 5, 663.

Edmundson, A.B. (1963) Nature, Lond., 198, 354.

Einstein, J.R., M^{c}Gavin, A.S. and Low, B.W. (1963)
Proc. nat. Acad. Sci., Wash., 49, 74.

Elmore, D.T., Baird, J.B., Roberts, D.V. and Smyth, J.J.
(1964) Abstr. Fed. Europ. Biochem. Soc., 50.

Epsedin, C.J. and Anfinsen, C.B. (1962)
J. biol. Chem., 237, 2175.

Epstein, C.J. and Goldberger, R.F. (1963)
J. biol. Chem., 238, 1380.

Fanshier, D.W. and Kun E. (1962)
Biochem. biophys. Acta, 58, 226.

Fischer, E. and Abderhalden, E. (1907) Ber., 40, 3544.

Foster, J.F. (1960) in The Plasma Proteins, Vol. 1,
p.179 (Ed. by Putman, F.W.), Academic Press, New York.

Fraser, R.D.B., and Inglis, A.S., and Miller, A. (1964)
Analyt. Biochem., 7, 247.

Frattali, V., Steiner, R.F., Millar, D.B.S. and Edelhoch, H.
(1963) Nature, Lond., 199, 1186.

Goldberger, R.F. and Anfinsen, C.B. (1962)
Biochemistry, 1, 401.

Goldberger, R.F., Epstein, C.J. and Anfinsen, C.B. (1963)
J. biol. Chem., 238, 628.

Goldstein, A. and Brown, B.J. (1961)
Biochem. biophys. Acta, 53, 438.

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

Grazi, E. Cheng, T. and Horecker, B.L. (1962)
Biochem. biophys. Res. Commun., 7, 250.

Green, D.E. (1962) Comp. Biochem. Physiol., 4. 81.

Gros, P. and Labouesse, B. (1960)
Bull. Soc. chim. Biol., 42, 559.

Gross, E. and Witkop, B. (1961)
J. Amer. Chem. Soc., 83, 1509.

Gross, E. and Witkop, B. (1962) J. biol. Chem., 237, 1856.

Gundlach, H.G., Stein, W.H. and Moore, S. (1960)
J. biol. Chem., 235, 3177.

Haavik, A.G. and Hatefi, Y. (1961) Fed. Proc., 20, 43.

Habeeb, A.F.S.A. (1961)
Can. J. Biochem. Physiol., 39, 31.

Hamilton, P.B. (1958) Analyt. Chem., 30, 914.

Hamilton, P.B. (1960) Analyt. Chem., 32, 1779.

Hamilton, P.B. (1963) Analyt. Chem., 35, 2055.

Hartley, B.S. (1964) Abstr. Fed. Europ. Biochem. Soc., 129.

Hill, R.L., Buettner-Janusch, J. and Buettner-Janusch, V.
(1963) Proc. nat. Acad. Sci., Wash., 50, 885.

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

Hirs, C.H.W. (1957) Fed. Proc., 16, 196.

Hirs, C.H.W. (1960) Ann. N.Y. Acad. Sci., 88, 611.

Hofmann, T. (1964) Biochemistry, 3. 356.

Holleman, J.W. and Biserte, G. (1959)
Biochem. biophys. Acta, 33, 143.

Hubbard, R.W. and Kremen, D. (1964) Fed. Proc., 23, 372.

Hvidt, A., Johansen, G., Linderstøm-Lang, K., and Vaslow, F. (1954) Compt. rend. trav. lab. Carlsberg, Ser. chim., 29, 129.

Inglis, A.S. (1964) Analyt. Biochem., 7, 168.

Ingram, V.M. (1956) Nature, Lond., 178, 792.

Ingram, V.M. (1958) Biochem. biophys. Acta, 28, 539.

Isemura, T., Takagi, T., Maeda, Y. and Imai, K. (1961)
Biochem. biophys. Res. Commun., 5, 373.

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

Iwasaki, H., Cohen, L.A., and Witkop, B. (1963)
J.Amer.Chem.Soc., 85, 3701.

Jirgensons, B., Ikenaka, T., (1959)
Die Makromolekulare Chem., 31, 112.

Katchalski, E., Benjamin, G.S. and Gross, V. (1957)
J.Amer.Chem.Soc., 79, 4096.

Katsoyannis, P.G., Fukuda, K., Tometsko, A., Suzuki, K. and Tilak, M. (1964)
J.Amer.Chem.Soc., 86, 930.

Keil, B., Sorm, F. (1964)
Abstr.Fed.Europ.Biochem.Soc., 126.

Kendrew, J.C., Dickerson, R.E., Standberg, B.E., Hart, R.G., Davies, D.R., Phillips, D.C. and Shore, V.C.

(1960), Nature, Lond. 185, 422.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G. Wyckoff, H. and Phillips, D.C. (1958) Nature, Lond. 181, 662.

Kendrew, J.C., Watson, H.C., Standberg, B.E., Dickerson, R.E. Phillips, D.C., and Shore, V.C., (1961) Nature, Lond. 190, 666.

King, T.P., (1961), J.biol.Chem., 236, PC5

Kirsten, E. and Kirston, R. (1962) Biochem.Biophys.Res.Commun., 7, 76.

Koch, G. and Weidel, W. (1956) Hoppe-Seyl.Z. 303, 213.

Koike, M. and Reed, L.J., (1961), J.biol.Chem., 236, PC34.

Konigsberg, W. and Hill, R.J., (1962) J.biol.Chem., 237, 2547.

Koshland, D.E., Strumeyer, D.H. and Ray, W.J. (1962) Brookhaven Symposia in Biology, 15, 101.

Krampitz, G. and Wieneke, J. (1963) Z.für Tierphysiol, Tiernahrung und Futtermittelkunde 18, 147.

Kraut, J. (1964), Abstr.Fed.Europ.Biochem.Soc., 135.

Kurihara, K., Morinishi, H. and Shibata, K. (1963) Biochim.biophys.Acta, 74, 678.

Lawson, W.B., Gross, E., Foltz, C.M. and Witkop, B. (1961) J.Amer.Chem.Soc., 83, 1509.

Leibig, J.von. in "Researches on the Chemistry of Food" (1847) (trans. Gregory, W.) Taylor Walton, London.

Lenard, J., Schally, A.V. and Hess, G.P. (1964) Biochem.biophys.Res.Commun., 14, 498.

Levinthal, C., Garen, A. and Rothman, F. (1961) Proc.Vth.Intern.Congress of Biochemistry, Moscow, Vol.1, 196.

Levy, A.L. (1954), Nature, Lond. 174, 126.

Li, C.H. and Bertsch, L. (1960), J.biol.Chem., 235, 2638.

Light, A. and Smith, E.L. (1963) in "The Proteins" Vol.1, p.1. (ed. by Neurath, H.), Academic Press, New York.

Linderstrøm-Lang, K.U. (1952), Lane Medical Lectures, Vol.VI, 58, Stanford University Press, Stanford, Calif.

Lindley, H. (1956), Nature, Lond. 178, 647.

Ludwig, M.L. and Byrne, R. (1962) J.Amer.Chem.Soc., 84, 4160.

Luetscher, J. (1939), J.Amer.Chem.Soc., 61, 2888.

Margoliash, E., Smith, E., Kreil, G. and Tuppy, G. (1961) Nature, Lond. 192, 1125.

Margoliash, E. (1963), Proc.Nat.Acad.Sci., Wash. 50, 672.

Matsubara, H. and Smith, E.L. (1962), J.biol.Chem., 237, 3575.

McLaren, W.B. (1963), Ph.D. Thesis, Glasgow University.

McLure, L.E., Schieler, L. and Dunn, M.S. (1953)
J.Amer.Chem.Soc., 75, 1980.

Meedom, B. (1956), Acta chem.Scand. 10, 881.

Merigan, T.C., Dreyer, W.J. and Berger, A. (1962)
Biochim.biophys.Acta, 62, 122.

Mills, G.L. (1952), Biochem.J., 50, 707.

Moore, S. and Stein, W.H. (1948), J. biol. Chem., 176, 367

Moore, S. and Stein, W.H. (1951), J.biol.Chem., 192, 663.

Moore, S. and Stein, W.H. (1954), J.biol.Chem., 211, 893.

Moore, S. (1963), J.biol.Chem., 238, 235.

Mosolov, V.V., and Loginova, M.D. (1962)
Dokl.Akad.Nauk SSSR, 146, 1209.

Mulder, G.J. (1838), Natuur en Scheikundig Archief.
6, 146.

Neel, J.V. (1949), Science, 110, 64.

Nirenberg, M.W., Matthaei, J.H., Jones, O.W., Martin, R.G.
and Barondes, S.H. (1963), Fed.Proc., 22, 55.

Ochoa, S. (1963), Fed.Proc., 22, 62.

Patchornik, A., Lawson, W.B., Gross, E. and Witkop, B.
(1960), J.Amer.Chem.Soc., 82, 5923.

Patchornik, A. and Sokolovsky, M. (1964)
J.Amer.Chem.Soc., 86, 1206.

Pauling, L., Corey, R.B. (1951), Proc.nat.Acad.Sci., Wash.
37, 251, 729.

Pauling, L. Corey, R.B. and Branson, H.R. (1951)
Proc.nat.Acad.Sci., Wash. 37, 205.

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H. and Will, G. (1960), Nature, Lond. 185, 416.

Peterson, E.A. and Sober, H.A. (1959), Analyt.Chem. 31, 857.

Piez, K.A. and Morris, L., (1960), Analyt.Biochem., 1, 187.

Porath, J. and Flodin, P. (1959), Nature, Lond. 183, 1657.

Porath, J. and Dennich, H. (1962)
Arch.Biochem.Biophys. Suppl.1, 152.

Porter, R.R. (1957), Biochem.J., 66, 677.

Porter, R.R. (1962) in "Basic Problems in Neoplastic Disease" p.177 (Gellhorn, A. and Hirschberg, E., Eds.), Columbia University Press, New York.

Raftery, M.A. and Cole, R.D. (1963)
Biochem.biophys.Res.Commun., 10, 467.

Redfield, R.R. and Anfinsen, C.B. (1956)
J.biol.Chem., 221, 385.

Reichmann, M. and Colvin, J. (1955)
Can.J.Chem., 33, 163.

Richard, A.J. and Kegeles, G. (1959),
Arch.Biochem.Biophys., 80, 125.

Rosen, H., Berard, C.W. and Levenson, S.M. (1962)
Analyt.Biochem., 4, 213.

Rumen, N.M. (1959), Acta chem.Scand., 13, 1542.

Ryle, A.P., Sanger, F., Smith, L.F. and Kitai, R. (1955)
Biochem.J., 60, 541.

Sakakibara, S., Shin, K.H. and Hess, G.P. (1962)
J.Amer.Chem.Soc., 84, 4921.

Samejima, T. and Yang, J.T. (1963)
J.biol.Chem., 238, 3256, 3262.

Sanger, F. (1945), Biochem.J., 39, 507.

Sanger, F. (1949), Biochem.J., 44, 126.

Sasakawa, S. (1963), J.Biochem. (Tokyo), 53, 188.

Schachter, H. and Dixon, G.H. (1964)
J.biol.Chem., 239, 813.

Schaffer, N.K., May, S.C. and Summerson, W.H. (1954)
J.biol.Chem., 206, 201.

Schlesinger, M.J. (1964), Fed.Proc., 23, 263.

Schroeder, W.A., Shelton, J.R., Shelton, J.B. and Cormick, J.
(1962), Proc.nat.Acad.Sci., Wash., 48, 284.

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

Schwartz, J.H. and Lipmann, F. (1961)
Proc.nat.Acad.Sci., Wash., 47, 1996.

Sela, M., White, F.H. and Anfinsen, C.B. (1959)
Biochim.biophys.Acta, 31, 417.

Shalitin, Y. (1961), Abstr.Proc.4th.Congress Sci.Soc.
Bull Research Council, Israel, 25, 65.

Shaltiel, S. and Patchornik, A. (1963)
J. Amer. Chem. Soc., 85, 2799.

Siegel, F.L. and Roach, M.K. (1961) Analyt. Chem., 33, 1628.

Simmonds, D.H. (1958) Analyt. Chem., 30, 1043.

Simmonds, D.H. and Rowlands, R.J. (1960)
Analyt. Chem., 32, 256.

Smith, E.L. and Hill, R.L. (1960) J. biol. Chem., 235, 2332.

Sogami, M and Foster, J.F. (1963)
J. biol. Chem., 238, PC2245.

Sokolovsky, M., Sadeh, T. and Patchornik, A. (1964)
J. Amer. Chem. Soc., 86, 1212.

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

Spackman, D.H. Stein, W.H. and Moore, S. (1960)
J. biol. Chem., 235, 648.

Spackman, D.H. (1963) Fed. Proc., 22, 244.

Spahr, P.F. and Edsall, J.T. (1964)
J. biol. Chem., 239, 850.

Stein, W.H. and Moore, S. (1948)
J. biol. Chem., 176, 337.

Steinrauf, L.K., Reddy, J.M. and Dickerson, R.E. (1962)
Acta Cryst., 15, 423.

Stellwagen, E. and Schachman, H.K. (1962)
Biochemistry, 1, 1056.

Synge, R.L.M. (1944) Biochem. J. 38, 285.

Terminiello, L., Sri Ram, J., Bier, M. and Nord, F.F. (1955)
Arch. Biochem. Biophys., 57, 252.

Theorell, H. (1932) Biochem. Z. 252, 1.

Tieze, F., Gladner, J.A. and Folk, J.E. (1957)
Biochem. biophys. Acta., 26, 659.

Tokuyasu, K., Funatsu, M. (1962) J. Biochem. (Tokyo) 52, 103.

Trincas, L. and Rinetti, M. (1962) Minerva Dietol. 2, 137.

Tristam, G.R. (1949), in "Haemoglobin" p.109
Butterworth, London.

Tsugita, A., Gish, D.T., Young, J., Fraenkel-Conrat, H.
Knight, C.A. and Stanley, W.M. (1960)
Proc.nat.Acad.Sci., Wash., 46, 1463.

Venetianer, P., Krause, E.G. and Straub, F.B. (1964)
Abstr.Fed.Europ.Biochem.Soc., 125.

Walsh, K.A., Kauffman, D.L., Sampath Kumar, K.S.V. and
Neurath, H. (1964), Proc.nat.Acad.Sci., Wash., 51, 301.

Walsh, K.A. and Brown, J.R. (1962)
Biochim.biophys.Acta, 58, 596.

Weil, L. and Telka, M. (1957), Arch.Biochem.Biophys., 71, 473.

White, F.H. (1960), J.biol.Chem., 235, 383.

White, F.H. (1961), J.biol.Chem., 236, 1353.

Wilchek, M. and Patchornik, A. (1962)
J.Amer.Chem.Soc., 84, 4613.

Wilson, J.G. and Cohen, L.A. (1963)
J.Amer.Chem.Soc., 85, 560.

Witkop, B. (1961) in "Advances in Protein Chemistry" Vol.16,

Wittman, H.G. (1959), Experientia, 15, 174.

Wood, J.L. and Catsimpoolas, N. (1963) J.biol.Chem., 238, PC 2887.

Yang, J.T. and Doty, P. (1957), J.Amer.Chem.Soc., 79, 749, 761.

Yanofsky, C. (1962) in "Cytodifferentiation and Macromolecular Synthesis" 21st Symposium of the Society for the Study of Development and Growth, p.15, Asilomar, Calif.

Yemm, E.W. and Cocking, E.C. (1955), Analyst, 80, 209.

Yoshida, A. and Tobita, T. (1960), Biochim.biophys.Acta, 37, 513.