Revised

SUMMARY OF THE Ph.D. THESIS

"STUDIES ON THE AMINO ACID SEQUENCE OF CYTOCHROME C"

submitted by Wm. B. McLaren

The purpose of this work was to obtain information about the sequence of amino acid residues in the peptide chain of cytochrome c, and at the same time to study methods for obtainir such information.

Cytochrome c was isolated from horse heart and purified by chromatography on an ion exchange resin (Amberlite IRC 50) and on sephadex.

Attempts to identify the N-terminal amino acid residue yielded negative results, and it was concluded that the protein may not have a free terminal amino group. This was subsequentl confirmed by Kreil and Tuppy who showed that the N-terminal group is acetylglycine.

To investigate the internal sequence of the protein, cytochrome c was hydrolysed with trypsin to yield a mixture of peptides. Chromatography on carboxymethyl cellulose, brough about partial resolution of this mixture, and the application of high voltage paper electrophoresis permitted the isolation of some pure homogeneous peptides from the material thus obtained. These methods, however, seemed inadequate for the exhaustive analysis of the peptide mixture. An effort was made, therefore to reduce the severity of the problem of peptide fractionation, by reducing the complexity of the mixture to be separated.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

This was achieved by acylating the lysine amino groups of cytochrome c, and thus restricting the action of trypsin to linkages involving the carbonyl groups of the two arginine residues in the protein. Methoxy-carbonyl chloride was first tried as an acylating agent, but proved unsuitable. Ethyl thiotrifluoroacetate was satisfactory and, at pH 9, this ester reacted with all the amino groups of the protein. It was found also that the substituent trifluoroacetyl groups could subsequently be removed by means of 0.5 M ammonia, without affecting the peptide structure.

The action of trypsin on the trifluoroacetyl protein yielded a mixture of peptides which was rather more complex than had been anticipated. On the assumption that this was due to the action of traces of chymotrypsin in the trypsin preparation, the time of exposure to the protease was reduced, and it was eventually possible by means of chromatography on sephadex to isolate a small number of large peptides. The amino acid compositions of these peptides, and the nature of their N-terminal residues were determined. Also, one of them was hydrolysed with chymotrypsin, and the resulting mixture of small peptides fractionated on a column of ZeoKarb 225 x 2. The peptides thus separated were analysed and their sequence in the original peptide determined.

The results of the analysis carried out during the course of this work were internally consistent, and in general found agreement with the recently published amino acid sequence for cytochrome c.

STUDIES ON THE AMINO ACID SEQUENCE OF CYTOCHROME C

By

William Blackwood McLaren, B.Sc.

Thesis submitted for the Degree of
Doctor of Philosophy
in the University of Glasgow

Department of Biochemistry November, 1963

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Professor J.N. Davidson for giving me the opportunity of carrying out this work in his Department and to Dr. G. Leaf for his constant help, guidance and encouragement during the course of these studies.

I am also grateful to the Science Faculty of the University of Glasgow and the Medical Research Council for the respective awards of a studentship and a research assistantship.

CONTENTS Page

Page

GENERAL INTRODUCTION

Proteins constitute the main chemical foundations of all living matter. Without them, life and its maintenance would be impossible.

The functions of protein substances are many and varied. Some, such as keratins, perform structural functions and this type of protein is present in hair, feathers and nails and belongs to a special class of proteins called the scleroproteins. Other members of this class are the collagen of bones and cartilage, and fibroin, the protein of silk.

Other proteins, however, such as the enzymes, hormones and antibodies, have specific physiological functions. The enzymes catalyse chemical reactions in living systems, whilst the hormones occupy a central role in the regulation of metabolic processes. The antibodies on the other hand, help to protect the organism against infection by foreign organisms such as bacteria and the viruses. The latter as a group consist of apparently non living nucleo-protein matter which is, nevertheless, capable of reproduction in a suitable living host.

Now, the great diversification of the functions that proteins perform, is due to structural differences in the proteins themselves.

Protein structure can most conveniently be examined at four different levels. These are (1) primary (2) secondary (3) tertiary (Linderstrøm - Lang, 1952) and (4) quaternary (Bernal, 1938).

By primary is meant the sequence of the amino acid residues in the peptide chains, whereas secondary structure describes the simplest

ways in which these polypeptide chains may be folded. There is a good deal of evidence to support the view that the peptide chain in many proteins is arranged in a helical structure. In no case, however, is the structure described by a single and continuous helix, but the helical portions are folded into a more complex 3-dimensional structure--this is denoted as the tertiary level of structure e.g., myoglobin (Kendrew et al. 1958, 1960). As an extension to this, quaternary structure describes the linking of groups of tertiary molecules e.g. see haemoglobin (Peruts et al. 1960).

Now, biochemical activity may in some instances be directly dependent on a few amino acid residues, e.g. with papain (Smith et al. 1960) showed that it was associated with a sulphhydryl and a carboxyl group and that two-thirds of the protein could be split off, using leucine amino peptidase, without seriously impairing it. In the majority of cases, however, it is associated with one particular 3-dimensional structure. This is characterised as the native configuration and involves the secondary, tertiary and quaternary structures discussed above. These, however, are to a large extent determined by the primary structure, for the secondary linkages involved in secondary, tertiary and quaternary structures require the correct placement of suitable amino acid residues, e.g. cysteine for formation of - S - S- cross linkages. Crick has in fact suggested that primary structure determines the secondary, tertiary and quaternary during biosynthesis, and Anfinsen et al. (1961) have presented evidence suggesting that in

ribonuclease, the native configuration is determined by the amino acid sequence, and in this case appears to involve the most thermodynamically stable folding of the chain.

However, the secondary, tertiary and quaternary levels of structure are not dependent on the entire amino acid sequence of the protein as is evidenced by the appearance of slightly different proteins performing the same function in different species. Inter-species differences were first observed with the amino acid composition of haemoglobin in the 1930's and since then actual sequence differences have been recorded e.g. with insulin (Sanger et al. 1955) and cytochrome C (Tuppy and Paleus 1955).

Intra-species differences also occur, particularly with haemoglobin. Although with some of these, function is unaffected, others interfere considerably e.g. Sickle cell anaemia (Itano 1955).

It is obvious, therefore, that the primary structures of proteins is of fundamental importance. If the behaviour of different proteins is to be properly understood, the chemical structures which govern their properties must be known.

In the present work, a study has been undertaken of some methods of sequence determination with special reference to cytochrome C.

This protein is associated with a number of similar compounds, the cytochromes, which were first described by MacMunn in 1886. MacMunn discovered a class of pigments in all types of animals, which

had four characteristic absorption bands in the reduced state. These pigments he found in various organs and tissues and he discovered further that the absorption bands disappeared on oxidation. MacMunn rightly concluded that these substances were concerned in respiration in tissues and this introduced a new concept, as previously it had been believed that respiration took place in the blood. He named these substances myo- and-histohaematins and demonstrated that they were haemochromogen in nature. Although he made an oxalic acid extract of washed skeletal muscle in which he observed absorption bands at 554.5 - 548.5 mμ, he was unable to proceed further with the isolation of the substances responsible for this. It was unfortunate that MacMunn believed that these compounds replaced haemoglobin in muscle, in which they carried on a similar role to the latter in blood. This involved him in serious controversy with Hoppe-Seyler, who maintained that myo- and-histohaematins were merely modified forms of haemoglobin. Hoppe-Seyler was by far the more influential of the two and so MacMunn's work was disregarded.

This remained the position for many years, until Keilin in 1925 showed that if MacMunn's theories had been wide of the mark, his experimental observations were nevertheless sound. Keilin confirmed and extended the work of MacMunn, showing that these "myo-and--histohaematins" are found also in plants and unicellular organisms, as well as in animal tissues. In order to describe better their wide distribution, he called them by the general name of cytochromes.

Keilin found that the four-banded spectrum of these compounds varied little from tissue to tissue, and careful microspectroscopic studies of baker's yeast showed that the bands had maxima as follows:- a 604 mμ; b 566 mμ; c 550 mμ; and d, which appeared to display three maxima at 532, 528 and 521 mμ. This type of spectrum was not immediately identifiable, but Keilin (1925) proposed that the cytochrome spectrum was due to the combined spectra of three distinct substances, each contributing a pair of absorption bands as shown below:

Cytochrome	α (mμ)	β (mμ)
a	604	532
b	566	528
c	550	521

This gives the effect of four main absorption bands, as the d - band of cytochrome is made up of the three β - bands of cytochromes a, b and c.

Apart from the cytochromes mentioned above, many others have been detected and distinguished by differences in their absorption spectra. Some of these are detailed below.

Cytochrome a Related Pigments.

Keilin found that heart muscle preparations contain a cytochrome with absorption spectrum (α-band 600 mμ; Y-band 448 mμ) similar to that of cytochrome a, but differing in being sensitive to CO. This pigment was denoted cytochrome a_3 and appears to be identical with the cytochrome oxidase of heart muscle (Keilin and Hartree 1939).

Spectroscopic studies have also revealed that electron transfer from ferrocytochrome c(c^{2+}) to oxygen in cells containing cytochromes c (or c_1), a and a_3, involves the reversible oxidation reduction of the cytochrome a and a_3 systems. Some doubt exists, however, whether a and a_3 are separate molecules. Some workers e.g. Wainio (1961) think they represent two functions of one molecule.

Rigourous studies of cytochromes a and a_3, however, have been hindered by the difficulties encountered in their extraction from tissues in a soluble form.

Some bacteria have been found to contain cytochrome a - like pigments and these are designated cytochromes a_1 and a_2 (Smith 1954 (1) 1954 (2) and Barrett (1956).

In Acetobacter pasteurianum, cytochrome a_1 (α-band 588 mμ) has been shown to be the terminal respiratory enzyme. This pigment forms a CO - compound that is not dissociated by light. It does not, however, catalyse the oxidation of mammalian cytochrome c by oxygen.

Cytochrome b Related Pigments

The extraction of cytochrome b from the tissues in which it is found cannot be readily achieved. Consequently, little is known of its chemical nature, although it has been established that it does not react with CO or cyanide, but does appear to be autoxidizable. It has been shown also, by means of spectroscopic studies, that the cytochrome b of heart muscle preparations is reduced by succinate, and that cytochromes c, a and a_3 are involved in the oxidation of

ferrocytochrome b by oxygen.

Spectroscopic measurements have also indicated that in some cellular systems, the transfer of electrons from succinate or DPNH to ferricytochrome c appears to be mediated by heme proteins similar to the cytochrome b of heart muscle preparations (Chance and Williams 1955). Now, flavoprotein preparations have been obtained that catalyze rapid electron transfer from reduced pyridine nucleotides to ferricytochrome c. The possibility that in living cells, one or more additional electron carriers may be interposed between a flavin and cytochrome c (or c_1) was raised by the discovery that treatment of heart muscle preparations with naphthol quinones (Ball et al. 1947), or with 2, 3 - dimercaptopropanol (Slater, 1949), inhibits the reduction of cytochrome c by succinate or by DPNH, but does not affect electron transfer from succinate or DPNH to cytochrome b, or from cytochrome c to oxygen. A similar effect has been found with antimycin A (Potter and Reif, 1953). This antibiotic does not inhibit the purified pyridine nucleotide - cytochrome c reductases and study of its action has yielded the conclusion that the antibiotic blocks electron transfer from cytochrome b to cytochrome c (or c_1). A relation between the lipids apparently associated with cytochrome b, and this inhibition, is suggested by the report that the action of antimycin A is counteracted by vitamin E (Nasan and Lehman, 1956). The antimycin A - sensitive respiration of tissue preparations has been attributed therefore to the participation of cytochrome b in the

sequence of electron transfer from metabolites to oxygen, and this pigment has been assigned the role of an electron carrier between flavins and cytochrome c.

Many biological systems have been shown to contain haem pigments with properties similar to heart muscle cytochrome b. Amongst these are cytochromes found in the microsomal fraction of liver preparations (Strittmatter and Ball, 1952) cytochrome m and in insect tissues cytochrome b_5 (Pappenheimer and Williams, 1954). These two have very similar spectroscopic properties, and are considered to be very closely related. The absorption maxima for reduced cytochrome from rabbit liver microsomes are: α-band 557 mμ, β-band 527 mμ, Y-band 423 mμ. The corresponding values for the insect pigment are: α-band 557 mμ, β-band 526 mμ, Y-band 421 mμ. Like cytochrome b, these heme pigments do not combine with CO or with cyanide, and appear to be autoxidizable.

Prior to the identification of the cytochrome b_5 (or m) group of pigments, several other intercellular heme proteins had been reported to have properties similar to those of cytochrome b. Cytochrome b_1 (α-band 558 mμ) has been found in several bacteria Vernon, 1956) including Corynebacterium Diphtheriae. Cytochrome b_2 (α-band 557 mμ, β-band 528 mμ) was found in baker's yeast and is associated with preparations having dehydrogenase activity towards L-lactic acid (S.J. Bach et al. 1946) and (J. Yamashita et al. 1957). Cytochrome b_3 (α-band 560 mμ) has been identified in microsomes from

the tissues of higher plants (Hill and R. Scarisbrick, 1951 and Martin and Morton, 1955 & 1957) and cytochrome b_4 was reported to be present in some bacteria, but its properties suggest that it belongs more properly to the group of bacterial pigments related to cytochrome c.

Cytochrome b_6 (α-band 563 mμ) was found in the chloroplasts of some green plants, where it appears to be associated with cytochrome f (R. Hill, 1954).

Finally, a pigment resembling cytochrome b, occurs in the hepatopancreas of the snail Helix pomatia, and has been named "cytochrome h;" it appears to be converted in the animal to helicorubin, a hemochromogen found in the gastrointestinal tract (J. Keilin, 1956 & 1957).

Cytochrome c Related Pigments

Spectroscopic measurements have shown in heart muscle, also cytochrome c_1 (or e). This is less easily extracted (α-band 554 mμ, β-band 524 mμ, γ-band 418 mμ) and is considered by some to be an endogenous form of cytochrome c (Keilin and Hartree, 1955). Like cytochrome c_1 this component is not autoxidizable, and does not appear to react with CO or with cyanide; however, cytochrome c_1 is thermolabile in contrast to cytochrome c.

Cytochromes c_2 , c_3 , c_4 , and c_5 are found in several types of bacteria, and in their ferrous form exhibit absorption maxima similar to those of heart muscle cytochrome c and c_1. In general, these bacterial cytochromes differ from mammalian cytochrome c in

their chemical properties, (e.g. isoelectric point, and oxidation - - reduction potential).

A haem pigment similar to mammalian cytochrome c, and named cytochrome f, has also been identified in the leaves of higher plants; it appears to be confined to the chloroplasts (Davenport and Hill, 1952).

Now, although these cytochromes and many other proteins have the same, or at least similar prosthetic groups, great differences in biological activities are apparent. Representitive of these are (a) haemoglobin, the oxygen carrying pigment in the blood; (b) catalase, an enzyme of wide occurrence which catalyses the decomposition of hydrogen peroxide; (c) peroxidase, another enzyme which utilises hydrogen peroxide for the oxidation of a number of substances; (d) lactic dehydrogenase, a ferroflavoprotein with "built in" cytochrome b_2; (e) many other haemoproteins with less well defined functions e.g. as in photosynthesis and sulphate reduction and (f) cytochrome c, the function of which is the transfer of electrons during the oxidation of substrates in the cell.

Since the prosthetic groups of these proteins are essentially the same, therefore, their functional differences must stem from differences in the protein parts of their molecules and/or its manner of attachment to the prosthetic group.

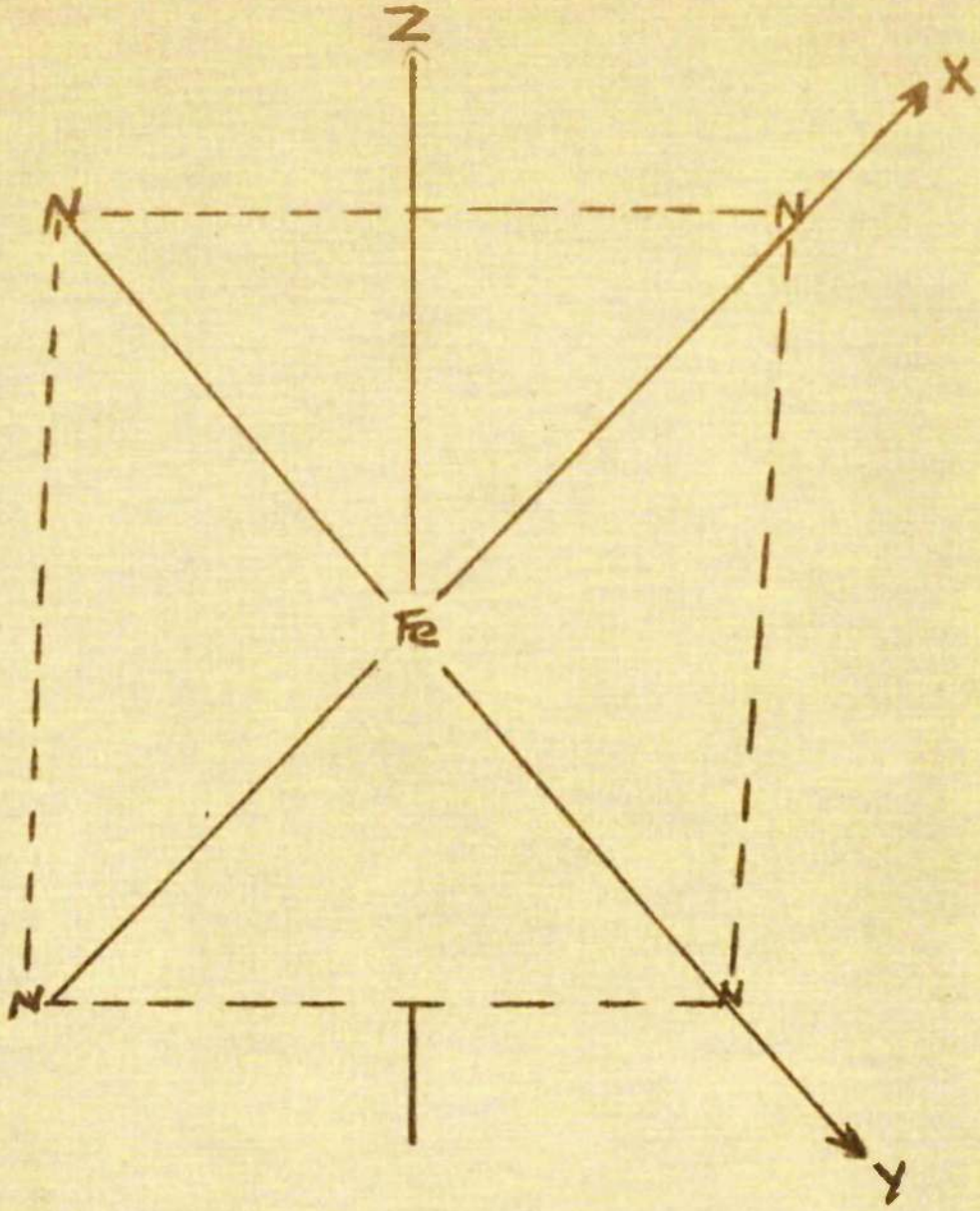
In fact, with haemoproteins, the relationship between the manner of attachment of the prosthetic group and the biological function of the protein is very strong. Indeed, the activities of all of them have been shown to depend on the mode and strength

of attachment of the iron in the porphyrin to the protein part of the molecule.

Now, the active site of all haemoproteins is the haem group, or perhaps even the iron atom, and it is apparent, therefore, that considerable modification of this is required to accommodate the variety of functions of these compounds.

In all of the haemoproteins, four of the co-ordination positions of the iron atom are taken up by the nitrogen atoms of the porphyrin; the two additional iron bonds are on a perpendicular to the plane of the porphyrin. These spatial relationships can be related most readily to the electronic configuration of iron by superimposing the octahedrally disposed iron bonds upon the Cartesian co-ordinate system as in figure. 1.

Now, in the case of cytochrome c, the structure and spatial arrangement of the protein moiety is such, that the two available co-ordination positions of the iron are bound to it through two functional groups which have been represented as imidazole groups. Therefore, this, plus the binding of the ethyl side chains of the porphyrin to the S atoms of two cysteine residues in the protein moiety, produces a very tightly knit structure for cytochrome c, and magnetic moment measurements (Theorell and Akesson, 1941a) have shown that this has resulted in the formation of co-ordinate bonds strong enough to bring about electron pairing in the t_{2g} level. Thus a structure of maximum stability is obtained, as is necessary to enable



GEOMETRICAL CONFIGURATION OF THE COVALENCES OF IRON (EICHORN 1960)

FIGURE 1.

the cytochrome c to carry out its biological function of electron transfer.

With catalase, peroxidase and haemoglobin, however, the biological functions of these haemoproteins are such that one of the available co-ordination positions of the haem iron must be free to co-ordinate with an extra molecular ligand. (In the above cases, the peroxide substrate and molecular oxygen respectively). This then results, with catalase and peroxidase possessing a desired degree of lability.

With haemoglobin, however, when the oxygen molecule is bound to the iron, the iron electrons become paired and extra strength is gained by the interaction of a π bond in the oxygen molecule with a t2g orbital of the iron.

It can thus be said that, with a haemoprotein, the electronic configuration of the iron is a true reflection of its biological function.

Recently, Perutz (1960), has shown from X-ray studies on the tertiary structure of haemoglobin, that the molecule is coiled in such a way as to hold the haem close to its surface, with one side attached to the protein strand, and the other side far enough away so as to accommodate an oxygen or a water molecule. This ideal positioning of the haem is thus largely due to the primary structure of the protein.

Similar spatial considerations must also be involved with cytochrome c, in order to position the prosthetic group so that the

biological function can be best carried out. As yet, however, no X-ray evidence is available to illustrate this. Considerable work has been done, nevertheless, on the properties (other than biological) and primary structure of the protein, and this is discussed in the next chapter.

The Structure of Cytochrome c

Cytochrome c is a remarkly stable substance, able to withstand exposure to mildly acid and alkaline conditions, and also considerable amounts of heat.

Apart from functional differences, its reactivity differs also in other ways from similar substances such as cytochrome oxidase and haemoglobin.

It does not, for example, react with carbon monoxide hydrogen sulphide or hydroxylamine, nor is it autoxidisable within the PH range of 4-11. Below PH4 and above PH11, however, reduced cytochrome c is modified to form an autoxidizable compound which does react with carbon monoxide (Keilin and Hartree, 1939; Theorell and Åkeson, 1941b; Tsou, 1951). Also, native cytochrome c will react with cyanide (Potter, 1941; Horecker and Kornberg, 1946) and with azide. Much of the early work done on the structure of cytochrome c was concerned with the linkages between the haem prosthetic group and the protein moiety. That this was much more powerful than those encountered in haemoglobin was evidenced by the fact that mild treatment with acid acetone failed to achieve separation of the iron porphyrin and the protein. Keilin,

however, succeeded in doing this by the more vigorous action of sulphur dioxide.

A clue as to the nature of the linkages was obtained by Theorell (1937), when he succeeded in preparing porphyrin c after hydrolysis of cytochrome c with HCl. This porphyrin he discovered contained 1-2 atoms of sulphur per molecule, and furthermore, L-cysteine could be split from it by the action of hydrogen bromide in glacial acetic acid. From these observations, Theorell (1938) was able to conclude that the vinyl side chains of the porphyrin were connected to the 2 cysteine residues in the protein by means of thio ether linkages.

Later work by Paul, (1950) lent support to Theorell's conclusions, when he split the porphyrin from the protein by the action of silver salts, known to cleave thio ether linkages.

Work done by Davenport (1952) has also lent support to Theorell's postulated linkage. He found that reductive cleavage by sodium amalgum of cytochrome c, yielded mesoporphyrin, whereas, similar treatment of protohaematin, horse liver catalase and ox haemoglobin gave rise to the unsaturated protoporphyrin. This suggests that sodium amalgum is incapable of reducing unmodified vinyl groups. This fits in with the idea that the vinyl groups in the porphyrin on cytochrome c are involved in thio ether linkages. If at the same time as reducing these, sodium amalgum removes the iron from the haem, then these findings go to confirm Theorell's views.

The above mentioned thio ether bonds, however, are not the only linkages between the prosthetic group and the protein. As

previously mentioned, bonds also exist between the iron atom of the haem and nitrogenous groups in the protein moiety. This is certain because of the haemochromagen type of spectrum of the whole molecule. Theorell and Åkeson, (1941, a, b & c) by means of titration, magnet--ometric and spectrophotometric measurements, showed that the probable points of attachment of the iron atom were two histidine residues in the protein molecule. A slightly inconclusive attempt was made by Paul (1951) to confirm these findings when he treated cytochrome c with fluorodinitrobenzene, and estimated the extent of labelling in the imidazole groups of histidine after hydrolysis. From these experiments, which were only partly successful, he concluded that at least one histidine molecule was linked, through its imidazole group to iron.

Ehrenberg and Theorell, (1955) found that if the peptide chain of the haem peptide is made into an α helix, either the histidine imidazole group, or the ε amino group of the lysine residue next to the cysteine, can be made to co-ordinate with the haem iron, but not simultaneously. On the basis of this, they supported the view that the haemochromogen forming groups of cytochrome c are two imidazole groups, one from the histidine present in the haem peptide, and the other from another histidine contained elsewhere in the molecule.

Margoliash (1959) pointed out, however, that the finding of a lysine or an arginine and a histidine adjacent to the two cysteines binding the protein to the haem in all the cytochromes c so far examined (at that time, Tuppy and Paleus, 1955 for ox, chicken

and salmon, and Tuppy and Bodo, 1954 for horse and pig heart) can lead to the alternative hypothesis that the imidazole and ε amino groups on these amino acids are in fact the haemochromogen forming groups in native cytochrome c. Moreover, an atomic model has shown that such an arrangement is sterically possible if the intervening chain of four amino acids was fully extended.

Which of the two hypothesis is correct, is not certain, but either way, it can be understood why the protein-prosthetic group complex in cytochrome c is so stable.

Work on the protein part of the cytochrome c molecule began, as would be expected, with attempts to determine the amounts of the various amino acids present therein.

In 1937, Keilin and Hartree estimated the amounts of the basic amino acids and cysteine present in an acid hydrolysed sample of cytochrome c containing 0.34% iron. Separation of these amino acids from the others present, was achieved by means of precipitation with phosphotungstic acid, and their estimation (with the exception of lysine) obtained by specific colour reactions.

The value for lysine was arrived at by subtracting the summed values of the others from the total base fraction.

Using a specific colour reaction, these same workers also obtained a value for the tryptophane present after digestion of the cytochrome c with trypsin. Enzymatic hydrolysis was necessary in this case because tryptophanecannot withstand acid hydrolysis and is destroyed.

In 1941, Theorell and Åkesson carried out more detailed determinations using similar techniques. In addition to the amino acids mentioned above, they obtained values for glutamic and aspartic acids and leucine.

The techniques used by both sets of workers in 1937 and 1941, however, were not very accurate and little reliance could be placed on the results. Nevertheless, their aquisition represented a considerable achievement and great credit was due to the workers who produced them.

It was many years, in fact, before accurate amino acid analysis of cytochrome c were obtained, and as is often the case in these matters, several sets of workers, working quite independently achieved the break-through, at about the same time.

Nunnikhoven (1958), Leaf, Gillies and Pirrie (1958), and Holleman and Biserte (1959) all published amino acid analysis of hydrolysed horse heart cytochrome c. These analysis are in agreement to within 1 residue per mole for the content of arginine, proline, methionine, isoleucine, tyrosine and phenylalanine, but vary by 1 residue or more for all of the other amino acids. See table.1.

This table is from Margoliash, Kimmel, Hill and Schmidt, 1962.
Table I

Amino Acid	Moles of amino acid per mole of protein			
	1	2	3	4
Tryptophan		1		1
Lysine	18	18	19	19
Histidine	3	3	4	3
NH_3	9	7		8
Arginine	2	2	2	2
Aspartic Ac	8	7	8	8
Threonine	10	9	10	10
Serine	0	0	0	0
Glutamic Acid	12	11	14	12
Proline	4	3	3	4
Glycine	12	11	11	12
Alanine	7	5	6	6
Valine	3	2	2	3
Cysteine		2		2
Methionine	2	2	2	2
Isoleucine	6	5	)	6
			) 11	
Leucine	6	5	)	6
Tyrosine	3	4	4	4
Phenyl-alanine	4	3	3	4

In this table, Column 1 gives the results of Nunnikoven; Column 2 those of Leaf et al; and Column 3, those of Holleman and Biserte. The analysis given in Columns 1 and 2 were obtained by column chromatography on ion exchange resins or starch and those given in Column 3 were obtained by dinitrophenylation of the amino acids followed by paper chromatographic separation of the dinitro--phenylated derivatives. In Column 4, data from an automatic amino acid analysis (Margoliash, Kimmel, Hill and Schmidt, 1962) on a complete enzymatic hydrolysis of the protein is given. This latter estimation, by reason of the accuracy of the tequnique is probably the correct analysis.

As regards the amino acid sequence of cytochrome c, very little information has been reported until last year, (1962). The only section of the chain on which information had been published was the portion immediately adjacent to the haem. Tuppy and Paleus, (1955), Leaf and Gillies, (1955) and Margoliash (1956) all published information on this section, but the most extensive communication on the matter was that of Paleus and Tuppy (1959). The haem peptide examined by them was the one resulting from tryptic hydrolysis of horse heart cytochrome c. The peptide, purified by means of adsorption on talc and by column partition chromatography, was sequentially investigated using the fluorodinitrobenzene technique (Sanger, 1945) to determine N-terminal residues, and the hydrazinolysis method of Akabori et al. (1952) to investigate C-terminal residues.

The tryptic peptide was degraded by means of chymetryptic hydrolysis and also subtilisin hydrolysis. The terminal amino acids of these sub-peptides were in turn discovered and in this way, by studying the small peptides and their overlaps, it was possible to obtain the sequence of the entire tryptic haem peptide of cytochrome c derived from horse heart.

Information has also been published concerning the N-terminal amino acid of horse heart cytochrome c. Margoliash (1955), using the fluorodinitrobenzene technique of Sanger, claimed two N-terminal residues of histidine.

Matsubara and Hagihara (1957), using the same technique could not confirm this work, and instead they identified arginine as the N-terminal residue.

Work carried out in this laboratory (see page 72) supported neither of these claims and indeed give no evidence for any free N-terminal residue. Since then, work by Tuppy (1961), contained in more detail in a later section, has shown that neither histidine or arginine is the N-terminal residue which is in fact acetyl-glycine.

Recently, however, almost two years after our own investigations began, the entire amino acid sequence of cytochrome c (horse heart) has been elucidated by Margoliash, Kimmel, Hill, Schmidt and Smith (1962) in conjunction with Kreil and Tuppy (1961). The techniques used, which are mentioned in more detail later,

involved enzymatic hydrolysis of the protein followed by purification of the peptides on ion exchange resins and by electrophoresis. Sequential analysis was carried out using the fluorodinitrobenzene technique, and Edman's phenyl isothiocyanate technique.

The value of the present work, therefore, is to confirm or otherwise, and also as new techniques have been used, to provide another approach to the general problem of protein chemistry.

Confirmatory Work

The value of confirmatory work can be well illustrated with respect to the protein Ribonuclease:-
In the derivation of the primary structure for this protein, there was one principal point of difference between the results of Spackman, Hirs, Moore and Stein (1960, a, b) and Anfinsen et al.(1956) and Redfield et al (1956) This concerned the identity of the amino acid residue in position 11, in the protein chain. Hirs, Moore and Stein having claimed it to be serine, whilst the work of Anfinsen indicated it to be glutamic acid (or glutamine).

Information on the subject from other authors adopting different experimental approaches, indicated that amino acid 11 was, in fact, glutamic acid. Allende and Richards (1962), for example, isolated a decapeptide which included residues 11-20 from tryptic digests of the S-peptide. The dinintrophenyl technique indicated N - terminal glutamic acid. Also, Gross and

Witkop, (1962) cleaved the protein at methionine residues by reaction with cyanogen bromide and obtained a peptide expected to contain residues from position 1 up to and including the position originally occupied by the first methionine residue in the chain. According to the formula originally proposed, that peptide would have contained glutamic acid and alanine in the ratio of 2 residues to 3. Gross and Witkop found equimolar amounts of these 2 amino acids in a hydrolysate of the peptide. This result supported the results of Anfinsen.

As a consequence to all this, Smyth, Stein and Moore, (1962) decided to make a thorough re-investigation of the sequence of amino acids in this part of the molecule. This time they obtained results which agreed with those of the other investigators and in addition, they found other errors in their originally proposed structure of this peptide.

It is obvious, therefore, that confirmatory work is of considerable value, particularly when the structures of large and very complicated molecules are involved.

The necessity for confirmatory work with cytochrome c is because of some inconsistency within the recent publications of Margoliash et al. Apart from the series of papers mentioned above, in which the entire sequence of the protein is elucidated, a publication also appears by Margoliash and Hill (1961). This describes the treatment of cytochrome c with leucine aminopeptidase,

which, it was claimed, removed thirty amino acid residues, leaving the remainder of the protein intact, and still containing the prosthetic group. Amongst the amino acids removed were the single tryptophane, one of the three histidines, one of the three valines and two serines which were hitherto undetected in cytochrome c. These findings are not in accordance with the sequence published in 1962, and although this came later, confirmation of one or the other, preferably using different techniques is desirable.

The means available for investigating the amino acid sequence of a protein have undergone rapid development over the past decade or so.

For example, in his studies on Insulin, Sanger relied on facts first pointed out by Synge et al. (1943) that while concentrated acid shows some side-chain specificity, in its catalysis of the hydrolysis of peptide bonds, it is more sensitive to chain length. The bonds of di and tri-peptides are hydrolysed relatively slowly as compared with those in the middle of longer peptide chains due to inhibition by the proximity of $-NH_3^+$.

Sanger therefore hoped to isolate di and tri-peptides representing every linkage in the peptide chain. He did in fact, isolate a substantial proportion of the expected di and tri-peptides from insulin A and B chains and was thus able to map out extensive sequences by this method. These sequences

did not overlap, however, for concentrated HCl does show some side chain specificity - in particular hydrolysing linkages involving amino groups of serine and threonine more rapidly than others. In order to discover how these sequences were linked, therefore, Sanger had to resort to the use of catalytic agents with specificity different from that of HCl - proteolytic enzymes.

Though the method of partial acid hydrolysis was very useful in elucidating the sequences in insulin, however, its use for longer peptide chains is precluded on the grounds:

(1) Quantitative analysis of mixtures of the complexity expected would be prohibitably difficult.

(2) The varying yields and possible recurrence of peptides would make interpretation difficult.

In the words of Katsoyannis (1961) therefore,

"Where peptide chemistry can make a contribution towards the proof of protein structure, is in the application and continued formulation of degradative techniques. Since fragmentation of the protein molecule to small peptides is probably the key reaction in degradative processes, the search for specific reagents for the selective cleavage of various peptide bonds, and the standardisation of existing techniques is of prime importance. "

Most of the proteins and large polypeptides studied, have been degraded with one or more proteolytic enzymes. Examples of these are trypsin which splits a protein chain at the carboxyl groups of lysine and arginine residues; chymotrypsin, which acts

at phenylalanine, tyrosine and sometimes histidine and tryptophane centres and pepsin, which has no particular specificity but which will considerably degrade most proteins. All three were used by Hirs, Moore and Stein in elucidating the structure of ribonuclease (1960), and the asignment of the fragments so produced, to their proper positions along the peptide chain was accomplished by taking advantage of the overlapping compositions of the peptides obtained from hydrolysates of the protein using different enzymes. Since the proper sequential alignment of peptides by this technique, however, from particularly long chains is likely to become quite difficult, it is obviously an advantage to be able to produce a limited number of large peptide fragments from the protein.

From this point of view, therefore, the selectivity of an enzyme as specific as trypsin is no longer adequate when large proteins are involved. Although in one of our own degradations of cytochrome c, it was used, even in such a case it would be convenient to have a cleaving agent that splits just one or two peptide bonds selectively, possibly next to an amino acid that occurs only twice in the protein molecule. So far, no enzymes are known that would perform such a task. Meanwhile, however, work is being done to modify the specificity of existing enzymes .

On the face of it, one possible approach in this direction could be the modification of the enzymes themselves. This seems

reasonable since it has been known for some time that enzyme activities depend on certain chemically defined groups, such as for example, the phenolic groups of tyrosine in pepsin (Herriot and Northop, 1934), or the indolyl or amide groups of trypsin (Fraenkel - Conrat, 1949). At the same time, it is obvious that all proteins possessessentially the same reactive groups as part of their structure. Therefore, the key to their specificity of action is found not in any particular groups, but in the spatial distribution of the latter, the general geometry of the molecule, the charge distribution on the surface and its fluctuation, or a possibility of resonance and chelation with the substrate etc. From all this, therefore, it would seem that the properties of an enzyme could certainly be modified by either blocking accessible reactive groups, or introducing new groups into its molecular structure.

The work done in this direction so far has, however, yielded no real change in enzymatic specificity although Nord, (1955) has shown that acetylation of the lysine groups of trypsin can be advantageous. The effects of this are to stabilise the enzyme to autodigestion, and at the same time increase the alkaline PH range over which trypsin is active. Katchalski, (1963) has also produced a water insoluble trypsin by coupling trypsin with a water insoluble polydiazonium salt, derived from a copolymer of p-amino-DL-phenylalanine and L-leucine (molar residue ratio 1:2). This, though less active than

native trypsin is more stable at the alkaline pH range 7.0 - 9.0 and can be used to prepare a column possessing tryptic activity.

Another approach has been the modification of the enzyme substrate. There are two ways of going about this, both of which have met with some success. (a) Creation of new sites in the molecule for enzymatic action and (b) the blocking of existing sites.

The creation of new lysine like sites for tryptic cleavage has been used by Lindley (1959). He achieved conversion of cysteine bonds to S -(β aminoethyl)- cysteine residues by reaction of the thiol groups with β-bromoethylamine. Enaymatic degradation of such modified proteins, e.g. reduced and β-aminoethylated insulin by the action of trypsin and carboxypeptidase β has been observed (Tietze et al. 1957).

Other techniques for producing "artificial" sites for tryptic cleavage have been devised by Ebata and Akabori (1960 &1961) and Shalitin (1961). The former showed that conversion of the free Y-glutamyl carboxy groups to Y-hydrazides produced tryptic sensitive modified glutamic acid residues, whilst the latter, by the introduction of glycyl residues into proteins with the leucine anhydrides of glycine, showed that peptide bonds next to serine were similarly sensitive. These techniques, however, have not yet been exploited and their value is difficult to assess.

With the other approach, i.e. the blocking of existing reaction sites, it is possible to remove arginine from tryptic attack through the use of sodium in liquid ammonia. This results in a protein containing ornithine instead of arginine, and the method has been used for the deguanidation of gelatin (Berger et al. 1958).

The first experiments reported, however, attempting to increase the specificity of trypsin, by blocking its reaction at the lysine residues in a protein, were done by Redfield and Anfinsen (1958). In these he substituted the reactive ε-NH_2 groups of the lysine residues in ribonuclease, using fluorodinitrobenzene. Thus, on tryptic hydrolysis, he obtained through hydrolysis only at arginine residues, a limited number of large peptides which he was able to separate and characterise. Unfortunately, the dinitrophenyl-lysine bond is very strong, and so no way of "de-blocking" the peptides could be found that did not at the same time completely hydrolyse them to their constituent amino acids. Thus, this technique was of limited application. He was soon able to improve on this, however, when he substituted the ε amino groups of the protein with benzyloxycarbonyl groups. This treatment appeared to be as efficient as a blocking agent, but had the enormous advantage that the blocking groups could be removed, after tryptic hydrolysis by treatment with anhydrous HBr. Unfortunately, this technique meets with difficulties concerning the

insolubility of its products and another disadvantage of it is the loss of basicity and of the hydrophilic character of the protein. Weil and Telka, 1957, attempted to overcome this by guanidating the free and accessible lysine ε - NH_2 groups. The resulting homoarginine peptide bonds they found were largely untouched by trypsin.

A modification of this approach by Merigan (1962), involves the conversion of the ε - NH_2 groups of lysine into dithiocarbamate derivatives, by reaction with CS_2 at alkaline pH. This method which has been tried out on an RNAase peptide, produced by digestion with subtilisin, and also on reduced egg-white lysozyme, has been shown to substantially block tryptic hydrolysis at lysine residues and the masking groups can be easily removed. The blocking, however, being far from complete and apparently reversible, the method will be of limited value when applied to large and more complicated proteins. This criticism applies also to the work of Weil and Telka (above).

In our own search for a suitable technique to modify the specificity of trypsin, we at first tried substitution with methyl chloroformate. This, however, was only a limited success and so it was with interest that we saw a paper by Schallenberg and Calvin (1955) on the trifluoroacetylation of amino acids and peptides in aqueous solution. Using ethyl thiotrifluoroacetate as the agent for this, they found that the intense electrophillic property of the trifluoroacetyl

group, and the unusual aminophilicity of the sulphur atom, combined in the thiol ester to produce considerable hydrolytic stability. This was in direct contrast to the highly reactive trifluoroacetic anhydride which has been previously used to prepare several trifluoroacetyl amino acid and simple peptide derivatives (Weygand et al.1952).

In addition, however, they found that although stable in acidic media, the trifluoroacetylamide bond undergoes hydrolysis at PH's greater than 10.

From reading this work, therefore, it seemed to us that the functional derivatives there described could be put to use to protect a protein from tryptic attack, in the same way Anfinsen used DNP and carboxybenzyl derivatives of proteins. Schallenberg and Calvin (1955) indeed had applied them successfully to peptide synthesis, in which the requirements of functional group masking and ultimate regeneration are much the same.

Consequently, we have thoroughly investigated this matter, with respect to its own values and also its application to the elucidation of the structure of cytochrome c.

It has been found possible to mask every lysine residue in the cytochrome c chain, and the peptides produced by tryptic hydrolysis have been purified and analysed. The results thus obtained have compared well with the structure of cytochrome c proposed by Tuppy and Margoliash.

The removal of the masking groups has also been accomplished, although to do this it has been found necessary to use either 0.1M NaOH or 0.5M NH_3. The latter technique is favoured, as it is likely to be the more gentle approach.

With all these techniques involving substrate modification, however, successful results can only be obtained if pure enzyme preparations are used. In the present work, steps were therefore taken to guard against interference of this sort in the trypsin used. These are described in the materials section.

MATERIALS SECTION

Cytochrome c

Some of the horse heart cytochrome c used in this work was prepared in this laboratory by means to be discussed, and solutions of this were checked for purity and concentration as follows:-

Solutionsof cytochrome c were diluted appropriately with 0.1 M phosphate buffer pH 7.3 and their optical densities measured at 280 mμ. A little sodium hydrosulphite was then added, and the optical densities measured again at 550 mμ.

The concentrations of the solutions could be calculated from the fact that

E 550 = 28,000 and if the solutions were pure

then $\frac{\text{O.D 550 red.}}{\text{280 mμ ox.}}$ = 1.28. Amino acid analysis were also done to check the purity of the preparations and the absence of serine in these was a good indication of purity.

The remainder of cytochrome c used was obtained from the Sigma Chemical Co., who claimed it to be 100% pure. This was checked by spectroscopic means and amino acid analysis. In interpretation of the latter.

Results

$$\frac{\text{O.D 550 mμ red.}}{\text{O.D. 280 mμ ox}} = 1.26$$

See also Table 2.

Amino Acid Analysis

Table 2.

Amino Acid	Residues Cytochrome c	Residues Cytochrome c (Margoliash et al.)
Asp	8	8
Thr	8	10
Glu	12	12
Pro	6	4
Gly	12	12
Al	6	6
Val	3	3
ILeu	6	6
Leu	6	6
Tyr	3	4
PhAl	4	4
Lys	18	19
Hist	3	3
Arg	2	2
Meth	2	2
Cyst	0.2	2

The absence of serine probably indicated purity, and the discrepancies seen in the table were almost certainly due to the fault of the analyser, (see Methods) and cysteic acid being present as cysteine.

Trypsin

Much of the work described in this thesis concerns a means of increasing the specificity of trypsin, thus causing it to fission proteins only at arginine centres.

The trypsin preparations in our possession, however, (from Worthington's) were found to be contaminated with chymotrypsin to the extent of 1% or less.

It was difficult to anticipate the importance of this, but what was obvious was, that any action which might derive from this enzyme would defeat the object of increasing the specificity of the trypsin as additional peptides would be produced in the hydrolysate. This was verified in practice when the fairly prolonged incubation required for the completion of a tryptic hydrolysis, allowed appreciable chymotryptic hydrolysis to occur on a sample of cytochrome c.

Attempts were, therefore, made to eliminate this chymotryptic activity, and the first was by incubating the trypsin preparation at 37°C, in $\frac{M}{10}$ HCl for a known time (24 hours or 40 hours). This treatment is known to render chymotrypsin unreactive, (Northrop, 1934) but no published information is available concerning its effect on trypsin. Experiments were therefore carried out (as described below) to determine this, and it was discovered that incubation in $\frac{M}{10}$ HCl destroys much of the activity of trypsin.

Thus, to achieve a satisfactory hydrolysis of a protein into peptides, using incubated trypsin, it would be necessary to use large amounts which could lead to confusion on analysis of the mixture.

Another approach was to limit the concentration of trypsin and the duration of the hydrolysis. In the present context, however, this sometimes leads to further difficulty because for achievement of the greatest benefit from the limitation of tryptic hydrolysis of the protein, it is desirable that such bonds as are attacked, are hydrolysed completely. (In this way, the simplest mixture of peptides will be obtained). With trifluoroacetyl cytochrome c, it was found a little difficult to obtain complete hydrolysis at the arginine residues and avoidance of significant chymotryptic action.

A third attempt was the separation of trypsin and chymotrypsin by electrophoresis as described by Grassmann(1961). In our hands, this technique gave only a partial separation, but as the overlap was small, and selective methods for the identification of the positions of the two enzymes were devised, it was possible to obtain some completely pure trypsin by these means.

This technique which is described below was only developed towards the end of the present work, at about the same time as B.D.H. marketed a chymotrypsin free preparation of trypsin. Although it was successful in itself, it was not used in the

cytochrome c structural investigations.

Determination of Tryptic Activities

In this experiment, the activity of trypsin before and after incubation in $\frac{M}{10}$ HCl is compared. The substrate used in this work was toluene sulphonyl-l-arginine methyl ester (Tsame) [123.7 mg in 50 ml water] Five ml of this solution and 2 ml HCl $\left(\frac{M}{10}\right)$ were used in every determination.

The trypsin solutions used were:-

(1) 0.706 mg per ml $\frac{M}{16}$ HCl (20 µl)

(2) As in (1), incubated for 40 hours at 37°C. (20 µl)

(3) 1% trypsin in $\frac{M}{16}$ HCl, incubated 24 hours at 37°C. (20 µl)

In all cases, the reactions were carried out in an automatic radiometer titrigraph which followed the reaction by titrating 0.194 M NaoH to maintain the pH of the reaction mixture at 7.5. At the same time, a graph was traced of alkali titrated versus time. See figure. 2.

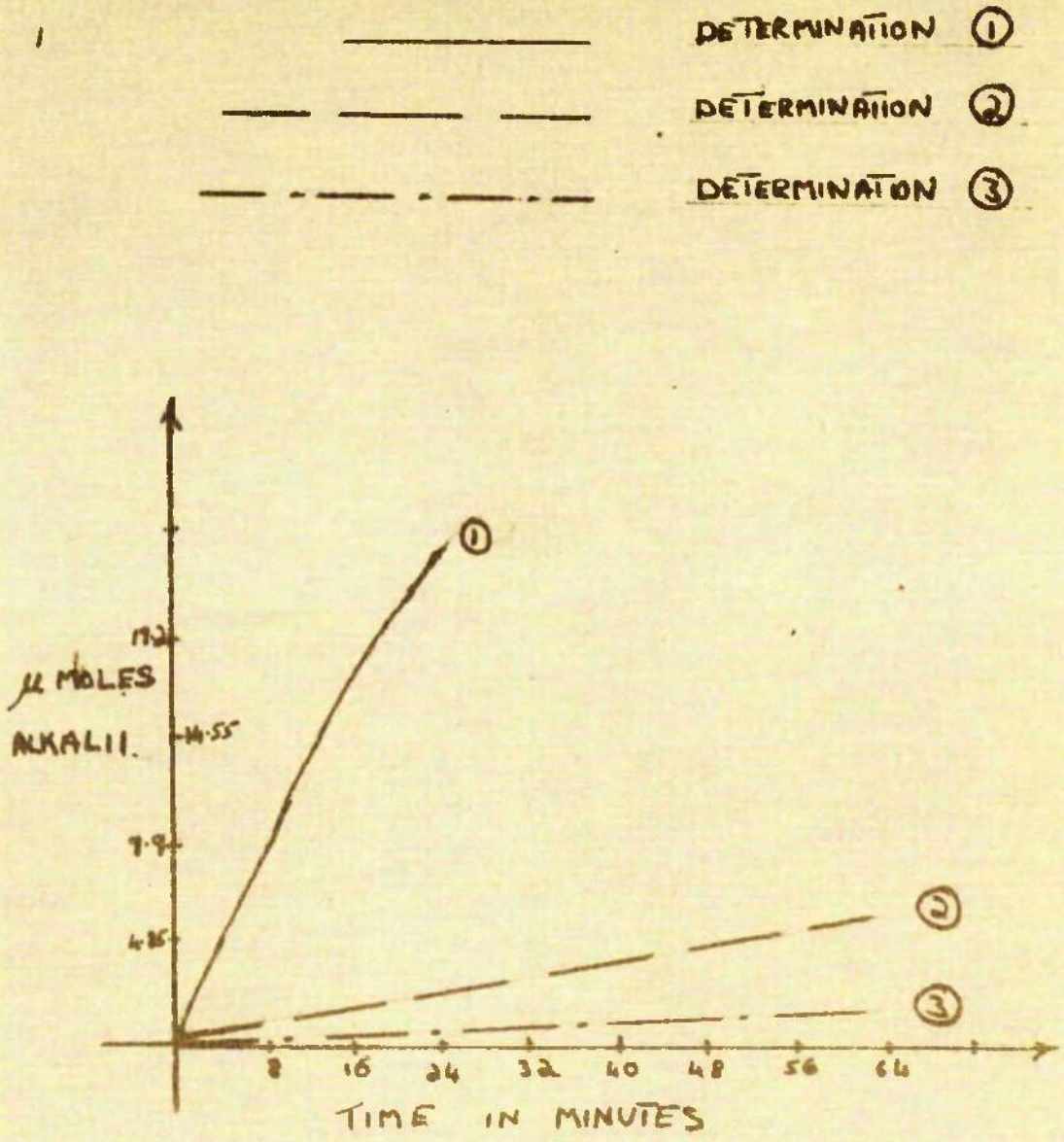

DETERMINATION ①
DETERMINATION ②
DETERMINATION ③
μ MOLES
ALKALI.
11.2
11.55
7.9
4.85
①
②
③
8
16
24
32
40
48
56
64
TIME IN MINUTES

FIGURE 2.

Results from the Graphs

The activity of the trypsin at any point on the curves was obtained by drawing a tangent to that point and calculating the gradient of this tangent. As very little spontaneous aqueous hydrolysis of the TSAME appeared to occur, it was not necessary to allow for this.

A more absolute evaluation of the activity of the trypsin in each case was obtained from the calculation of the reaction constants from the formula.

$$K = \frac{2.303}{t} \log \frac{Co}{Co - a}$$ where

Co = the initial concentration of substrate

and a = the amount hydrolysed in time t

(1) Activity = 1.82

Reaction Constant = 0.0197

(2) Activity = 0.02

Reaction Constant = 0.0025

(3) Activity = 0.09

Reaction Constant = 0.0028

Separation of Trypsin and Chymotrypsin by Electrophoresis

The trypsin preparation, in addition to about 1% chymotrypsin, also contained 50% $MgSO_4$. Before electrophoresis, therefore, it was necessary to desalt the preparation, both to prevent the electrophoretogram from being obscured and also

interference with the colour reactions described below for the selective detection of trypsin and chymotrypsin.

Experimental

10 mg of a Worthington trypsin preparation were desalted on a G25 sephadex column (11 cm long x 1.5 diameter) equilibrated with 0.05 M HCl. The trypsin was applied to the column in 1 ml solution and four ml fractions were collected. The emergence of the trypsin in the eluate was detected by optical density readings at 280 mμ and the $MgSO_4$ position was determined by the addition of βa Cl2 and HCl to aliquots from the fractions.

The trypsin containing fractions were then bulked, lyophilised and the trypsin applied to Whatman No.I paper for electrophoresis in a three-inch strip across the electrophoresis paper as shown. Fig. 3.

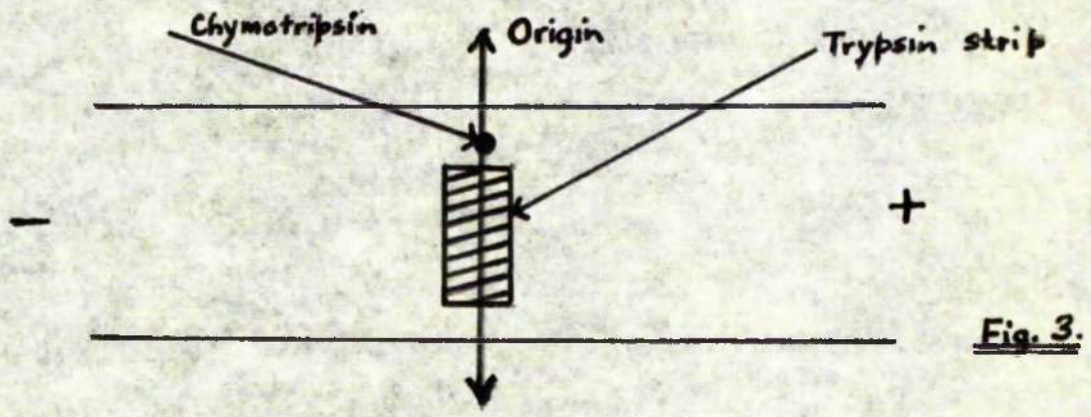


Fig. 3.

A spot of chymotrypsin was also applied as shown. Electrophoresis was carried out in pyridine -glacial acetic acid - water buffer (30 :2:1000) pH 6.4 with a potential gradient of 20 volts/ cm for 60 minutes.

After electrophoresis, the chymotrypsin and trypsin were located by a detection method using synthetic substrates:

Two strips of filter paper ½" x 22.5" were dipped in

(1) 0.01 M TSAME containing cresol red and adjusted to pH 8.5

(2) Similar to (1) except that TSAME was replaced by phenyl alanine ethyl ester

A ½" strip from the edge of the paper on which electrophoresis of the trypsin had been carried out was then sprayed with cresol red and exposed to NH_3 until it was just pink. This strip was then sandwiched between the test strips thus: Fig. 4.

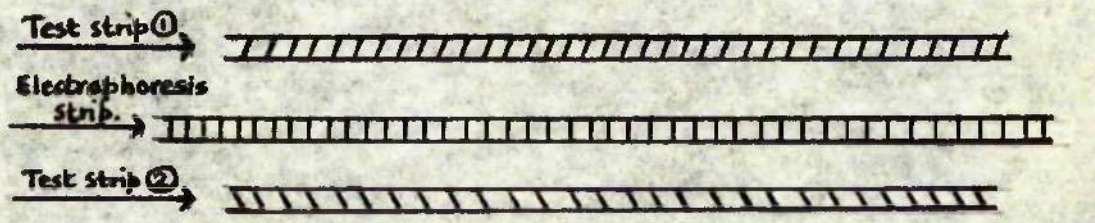


Fig 4

The effect in this sandwich was hydrolysis of the TSAME by trypsin and hydrolysis of the phenyl alanine ethyl ester by chymotrypsin. These hydrolysis were accompanied by a lowering in pH and consequent decolorisation of the cresol red. Thus the positions of trypsin and chymotrypsin can be determined by the appearance of yellow patches on the red indicator background of the test strips.

The experiment showed that trypsin and chymotrypsin were not entirely separated. Nevertheless, it was possible to elute pure trypsin from an appropriately cut strip. Fig. 5.

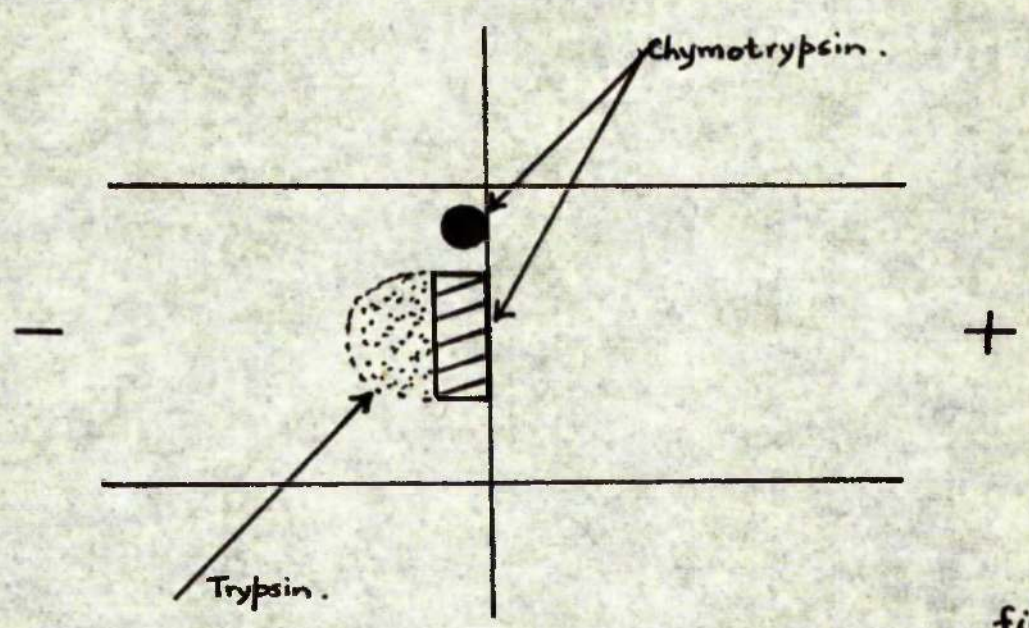


fig 5

Preparation of ethyl thiol trifluoroacetate (Hauptschein, Stokes and Nodiff, 1952).

8 gm ethyl mercaptan were chilled on an ice bath and 32 gm trifluoroacetic anhydride were added dropwise over a period of 1 hour with intermittent chilling and shaking. The reaction vessel was fitted with a reflux condenser fitted with a calcium chloride tube, permitted to stand at room temperature for 1 hour and then heated at 100°C for 3 hours. The excess anhydride and trifluoroacetic acid were removed by washing twice with 150 ml portions of cold 5% KOH, and twice with water. The material was dried over anhydrous magnesium sulphate and the ester was obtained by fractional distillation (boiling point 90.5°C at 760 mm) yield 50 - 60% based on ethyl mercaptan.

METHODS SECTION

Determination of Total Nitrogen

Samples of the material under investigation were digested for three hours in H_2SO_4 containing SeO_2 (50% w/v and 1% w/v) and then the digestion mixtures were made up to 10 ml with water. Two ml samples of these were subjected to Nesslerisation together with two controls of NH_4Cl.

Nesslerisation

To the samples, 3 ml 2N NaOH and 2 ml Nessler's reagent were added. The solutions were allowed to stand for 15 minutes and then the optical densities read at 420 mμ.

Amino Acid Analysis

(a) Qualitative

The qualitative analysis of amino acids was achieved by means of two dimensional chromatography on sheets of Whatman No.I filter paper. The solvents used to effect the separations were butanol/ acetic acid/ water (4:1:5) and phenol water, and the amino acids were detected and identified by dipping the papers in a 0.1% ninhydrin solution of composition, 95% ethanol - glacial acetic acid - collidine (50:15:2) and heating them at 90°C for five minutes. See Figure 6.

On some occasions, however, when interest was focussed on the basic or acidic amino acids, the amino acids were separated by means of high voltage electrophoresis. The apparatus used in these experiments was made by Shandon and had water cooled plates capable of accomodating paper 50 cm long/12.5 cm broad. The material to be electrophoresed was thus applied to Whatman No.I papers of these dimensions, which were then wetted with pyridine - acetic acid - water buffer pH 6.4 (100:90:890), blotted, covered both sides by polythene strip and placed between the plates of the apparatus. These were water cooled metal plates which were then pressed together with a pressure of 15 lbs. per square inch and contact between the paper and the buffer surrounding the electrodes was made by paper wicks. The required potential difference (50 volts/cm) was applied across the paper

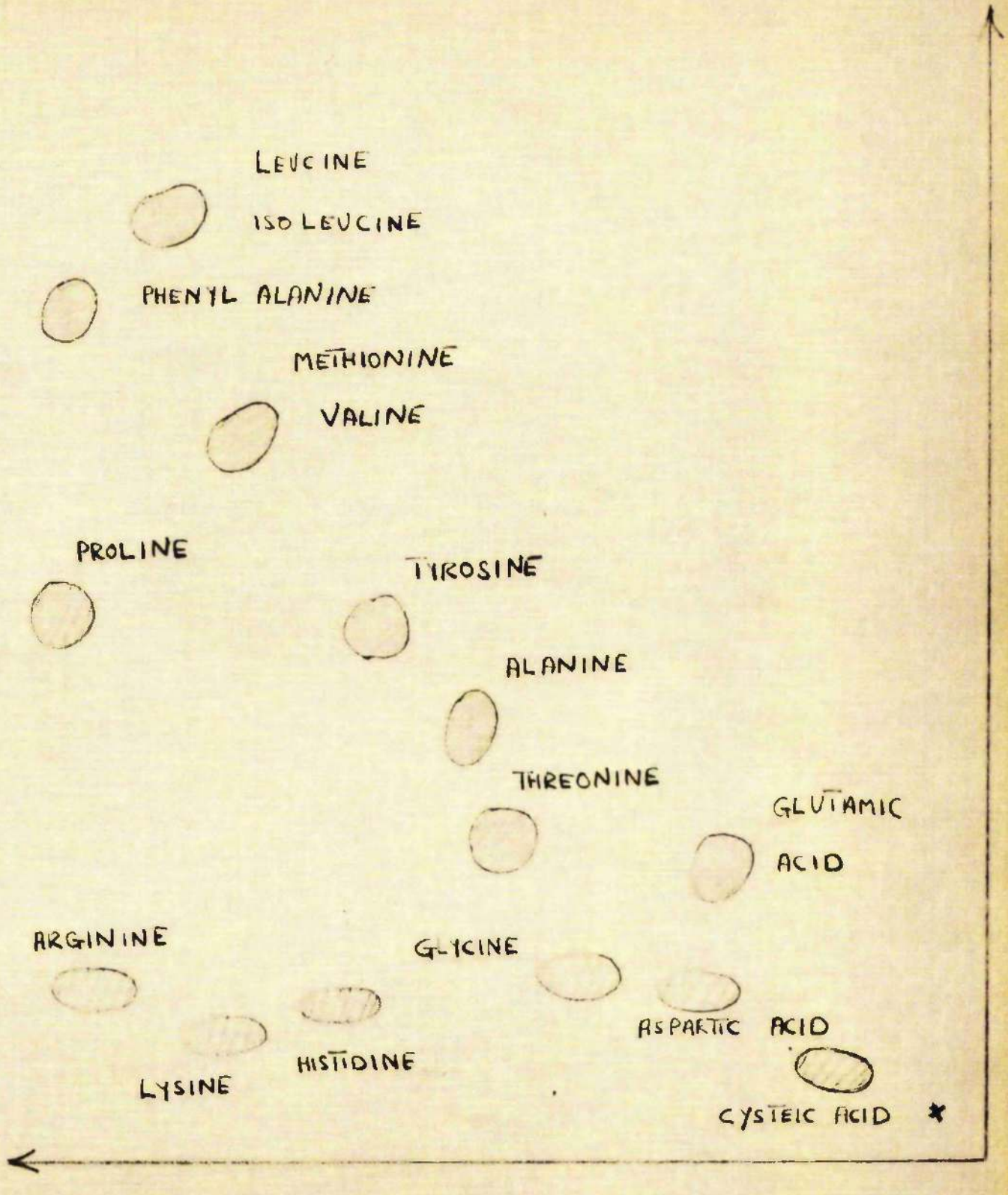

LEUCINE
ISO LEUCINE
PHENYL ALANINE
METHIONINE
VALINE
PROLINE
TYROSINE
ALANINE
THREONINE
GLUTAMIC
ACID
ARGININE
GLYCINE
ASPARTIC ACID
HISTIDINE
LYSINE
CYSTEIC ACID

FIGURE 6.

for 40 minutes and the detection of the amino acids was as described above. This technique separated only aspartic acid, glutamic acid, histidine, lysine and arginine. The mono amino monocarboxylic acids remained near the origin.

Quantitative Amino Acid Analysis

During the course of this work, three different techniques were employed. These are detailed below.

(i) The amino acids were subjected to two dimensional paper chromatography and estimated quantitatively using the special dipping solution of Heilmann, Barrollier and Watzke (1957). This consists of 100 mg cadmium acetate, 10 ml water, 5 ml acetic acid, 100 ml acetone and 1 gm ninhydrin, mixed in that order. The chromatograms were left to develop for 24 hours in a large desiccator over concentrated sulphuric acid, the resulting spots eluted with methanol, and the optical density of the eluate read at 500 mμ.

The values so obtained were converted to amounts of amino acid using the extinction coefficients given by Heilmann et al. Whilst the method, as employed, was not capable of great precision, experience suggested that the values obtained in any one set of analysis were sufficiently consistent to give a semi-quantitative measure of the ratios of the amino acids present. Glycine and proline, however, could not be estimated this way, neither could leucine or isoleucine be distinguished from each other.

The prosecution of this technique required considerable amounts of time and its use was sometimes precluded when the solutions to be analysed contained more than a certain amount of salts. On such occasions, the DNP technique, Levy, (1954), Kook and Weidel (1956), as described below, was used.

Preparation, Extraction and Quantitative Estimation of DNP Amino Acids

(ii) This was done according to the method of Schroeder (1957) whereby the amino acids were dissolved in 60% aqueous ethanol to which an excess of the $NaHCO_3$ necessary to maintain the pH at 9 had been added. Excess fluorodinitrobenzene was then added and the mixture shaken gently for three hours at 37°C.

The excess FDNB was first extracted from the above mixture, by peroxide free ether (four times). The solution was then acidified to pH3 and extracted three times with 4.5 ml ethyl acetate. The extracts were all washed with the same one ml of water, mixed, and the ethyl acetate evapourated off in a sublimation tube.

Most of the dinitrophenol present was then removed by vacuum sublimation (at 70°C for 10 minutes) to leave the residue ready for chromatography.

The aqueous phase was next extracted with secondary butanol. In order to prevent formation of a homogeneous solution, it was found necessary to add NaCl. The washing of

the three secondary butanol extracts was similarly useless, as it was necessary to salt out each time.

The secondary butanol extracts were bulked and the solvent evaporated, leaving a residue containing a considerable amount of salt. The DNP material was removed by extraction with acid acetone, (100 cc acetone + 1 cc NHCl) which on evaporation left the DNP amino acids ready for chromatography.

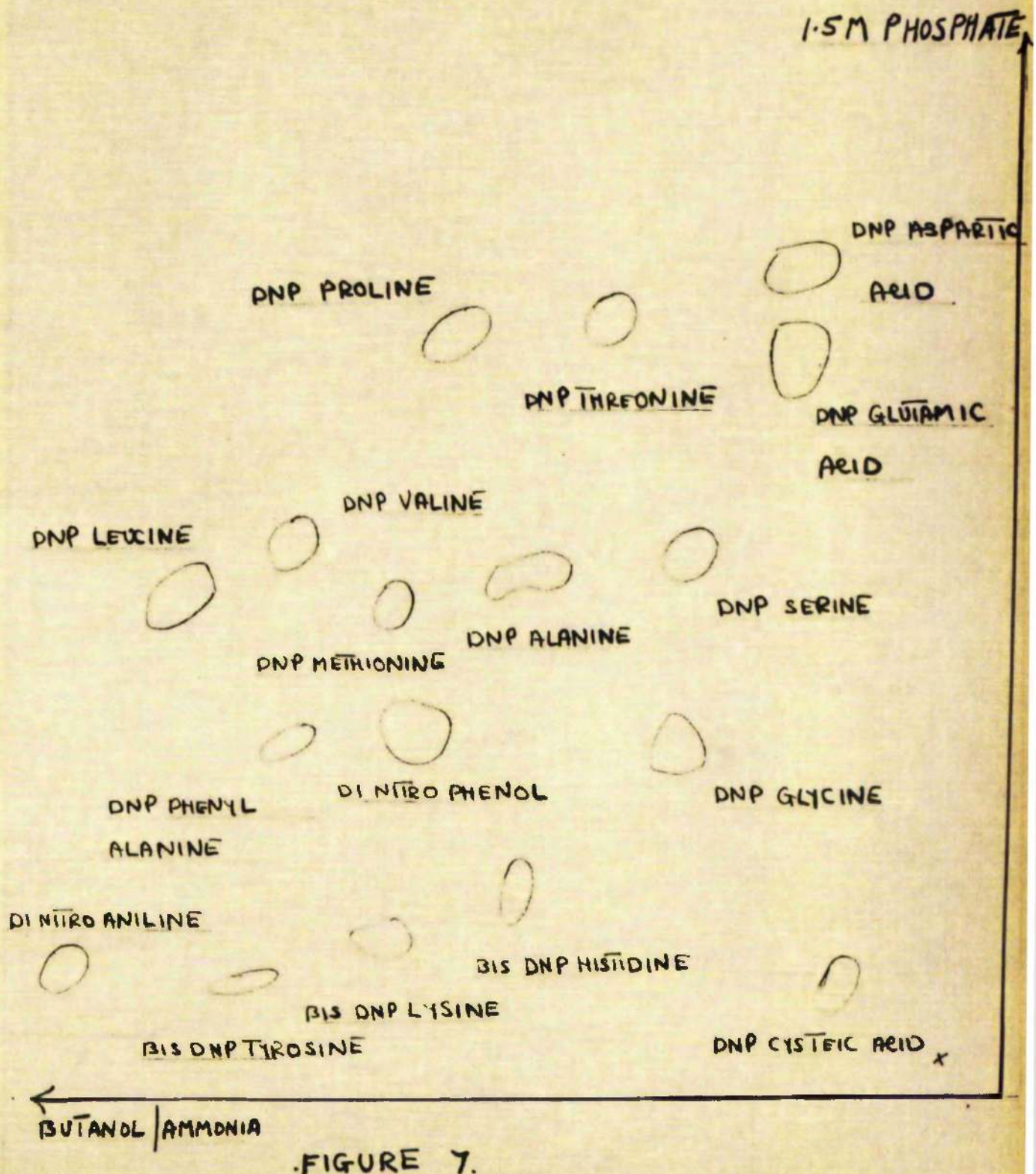
This was done two dimensionally using n-butanol/0.1% ammonia (1:1) and 1.5 M phosphate buffer pH7. The figures 7 & 8 show typical chromatograms.

To obtain quantitative results, the DNP amino acid spots were cut from the paper and then eluted from it with 5 ml H_2O. The optical densities of the resulting solutions were then measured at 360 mμ.

This technique has been used extensively in the present work with considerable success. Its immunity to salt effects, for example, makes it in some instances a superior tool to the method described under (i). It is superior also in that it makes possible the measurement of amounts of amino acid, ranging from 0.01 μmoles to 0.5 μmoles. It does not, however, distinguish between leucine and iso leucine.

(iii) Automatic Amino Acid Analysis

This procedure, although complicated, provides probably the quickest and most efficient means of characterising an amino acid mixture.


1·5M PHOSPHATE
DNP ASPARTIC ACID
DNP PROLINE
DNP THREONINE
DNP GLUTAMIC ACID
DNP VALINE
DNP LEUCINE
DNP SERINE
DNP ALANINE
DNP METHIONINE
DI NITRO PHENOL
DNP GLYCINE
DNP PHENYL ALANINE
DI NITRO ANILINE
BIS DNP HISTIDINE
BIS DNP LYSINE
BIS DNP TYROSINE
DNP CYSTEIC ACID
BUTANOL | AMMONIA

.FIGURE 7.

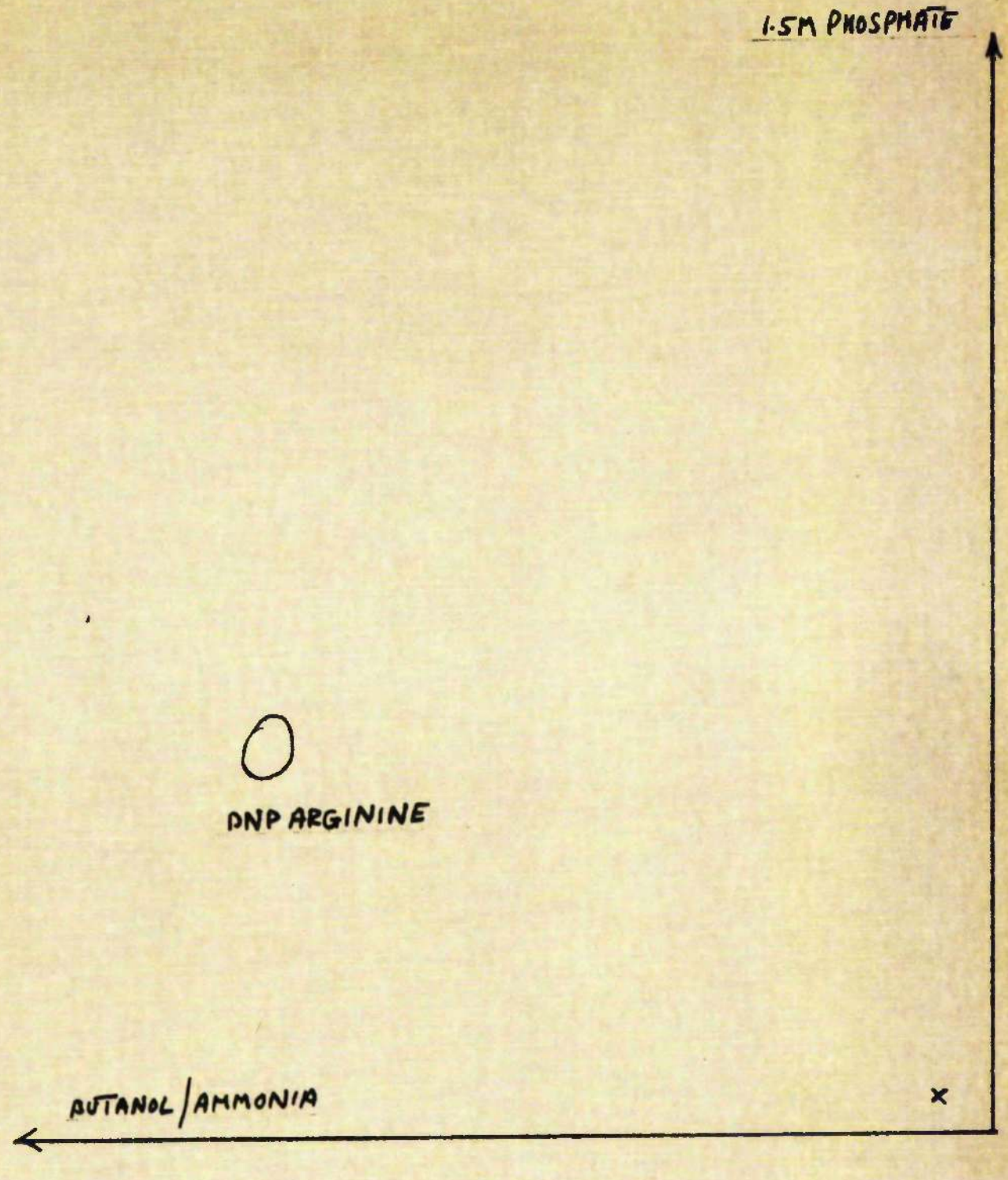


FIGURE 8

The assembly now available in this laboratory was designed and built by Dr. George Leaf, who based his design on earlier models by Spackman, Moore and Stein (1958), and Pies and Morris (1960).

The amino acid separation involved in this procedure is accomplished on an ion exchange resin column, the elution of which is controlled by a "varigrad" or mixing vessel (consisting of nine inter-connected chambers), that alters the ionic strength and pH of the eluate in a very specific manner. The inter-connections of the chambers, which are indicated by the small shaded circles can be temporarily blocked whilst the vessel is charged with its buffers; each chamber receives an equal weight of its required buffer. Figure 9.

The buffer to the column is pumped from chamber I and as its level consequently falls, so the levels in all the other chambers adjust themselves, through the various connections, in order to preserve the hydrostatic equilibrium. Thus, the desired gradients of pH and ionic strength are provided.

From the column, the buffer, as it emerges, is caused to mix with a special ninhydrin reagent. These two together then pass through a "reaction coil" immersed in a boiling water bath, and so the ninhydrin colour is developed, and the intensity of this depends on the quantity of amino acid (if any) present.

From the reaction coil, the solution passes on through a tubular cuvette in a photometer where the colour intensity

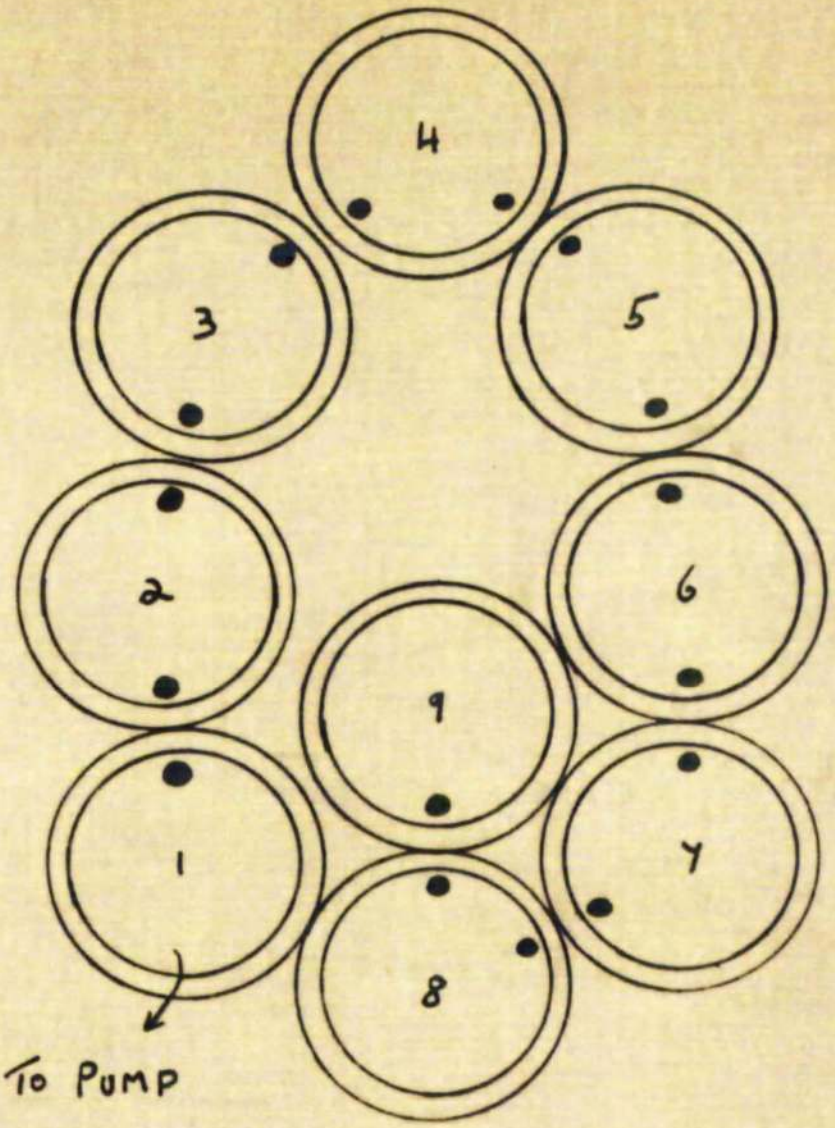


FIGURE 9

is accurately measured. This is recorded by means of an electronic recording device and a graph is traced vs time, which indicates the presence of the amino acids in the form of peaks. From the area under these peaks, it is possible to calculate the quantity of amino acid present. Individual amino acids are identified by their position on the graph, each amino acid always assuming the same position due to the very accurate and standard pumping of the buffers and reagents.

The benefits of this technique over the others used are numerous. The amino acids themselves are analysed directly, and therefore no reliance is placed on the completion of a secondary reaction. Further, the salt effects are minimal and the entire analysis, which is capable of measuring amounts of amino acids ranging from 0.1 - 3.0 μmoles, with recoveries of 97 - 103%, is completed in 24 hours.

It must be mentioned, however, that difficulty has sometimes been experienced with the analysis of lysine and histidine because of ammonia in the buffers. In the early analysis too, difficulty was found in eluting arginine quantit--atively from the column. This has since been overcome, however, by raising the column temperature.

Nevertheless, in view of its other advantages, on its becoming available, this technique was adopted in favour of the others .

Peptide Separation

To determine the amino acid sequence of a protein, it is first necessary to degrade it to smaller simpler fragments, which must then be separated from each other and characterised. In this work, various means of doing this were used, which relied on different principles. These are described below.

(1) Sephadex

Sephadex separates peptides according to their molecular size. It is a three dimensional cross linked, porous, polysaccharide substance with a basically non ionic character but slight polar properties due to its high hydroxyl group content and the presence of a very few - COO^- groups where some terminal alcoholic groups have become oxidised.

In the presence of water, dry sephadex swells, and the amount of water taken up depends on its grade. This is expressed as "water regain" of the material in gms of water per gm of dry sephadex.

The water in a column of sephadex consists of two parts, the internal water (Vi) i.e. the water in the grains, and the external water (Vo) which is the water in the column surrounding the gel grains. The value for Vi can be calculated from the known weight of dry sephadex, and its water regain, and Vo (the void volume) is determined as described below.

Thus, the total volume of the packed gel bed is

Vt = Vo + Vi + Vg where Vg = the volume of the gel matrix.

The way in which sephadex effects its separation of molecules is:-

Small molecules are permitted to diffuse ~~through~~ into the sephadex gel grains and are eluted from the columns in a volume Vo + Vi.

Large molecules, on the other hand, are not permitted to diffuse through the grains and therefore are eluted in a volume Vo.

Intermediate cases also exist where molecules are partly able to diffuse through the grains and these are eluted in volumes somewhere between Vo and Vo + Vi.

The different grades of sephadex determine whether or not a molecule is large enough to be excluded from the grains. These are:-

(1) G25 sephadex: a molecular weight of about 3,500 or greater will provide for complete exclusion.

(2) G50 sephadex: 8,000 and above

(3) G75 sephadex: 40,000 and above

Determination of Void Volume of a Column

This is carried out by passing through the column a solute with molecular weight large enough to ensure complete exclusion and measuring the elution volume. Diluted indian ink or haemoglobin are two materials which are suitable for the purpose.

The columns of sephadex used in this work showed that the resolving power of the technique is limited. Good separation of peptides were only obtained when a large size difference existed.

Method

The sephadex was equilibrated in water, or a selected buffer, overnight. The fines present were decanted off and the column then poured, in sections, from a fairly concentrated slurry.

The material to be separated was applied in as small a volume as possible, and the column eluted with the buffer used for equilibration. The eluate was divided into convenient fractions and these were analysed as described on page 62

(2) Paper Electrophoresis

This technique, to produce a successful separation, requires that the peptides concerned have a difference in net charge. With cytochrome c, which contains a large number of basic and acidic amino acids, the peptides obtained were generally quite amenable to this condition. The scale of this method, however, is limited, and it was found to be unsuitable for very complex mixtures.

The peptide material was applied to an electrophoresis paper in an 8 cm strip and run at a potential of 80 volts per cm for 30 minutes in the pyridine-acetic acid-water buffer previously described.

The detection of the peptides on the dried paper was achieved by means described below, on a small "marker strip" cut from the electrophoretogram. They were eluted with 0.05 M ammonia.

A two dimensional extension of this technique was also used on some occasions, to produce "peptide maps". This was done according to Ingram (1958), except that the electrophoresis was carried out at a higher voltage (see above). The second dimension was completed using ascending chromatography in butanol/acetic acid/water.

The Techniques used to Detect Peptide Positions on Paper

(1) Ninhydrin Dip

The papers are dipped in a 0.1% solution of ninhydrin, in 95% ethanol-glacial acetic acid-collidine (50:15:2), and heated at 90° for five minutes. The peptides appear as dark blue stains.

(2) Chlorine Method (Rundel and Hoppe, 1954)

The paper is air dried and any phenol completely removed with an ethanol-ether mixture (1:1 v/v). It is then moistened with ethanol-acetone (1:1 v/v) and blotted.

Next, it is enclosed in a desiccator above a solution of 10 ml $\frac{N}{10}$ $KMnO_4$ and 10 ml 10% HCl, and left for five minutes in the resulting atmosphere of chlorine (and oxides of chlorine). It is then immediately transferred to a solution of equal volumes of $\frac{N}{20}$ KI and saturated O-tolidene in 2% acetic acid. Maximum colouring appears after about two minutes, and the peptides are shown as dark blue patches on a light blue background.

(3) Ehrlichs Reagent for the Detection of Tryptophane Peptides (Jepson and Smith, 1955)

The paper is dipped in a 1% dimethylaminobenzaldehyde solution in acetone : chloroform 9:1. A positive test is indicated by the appearance of a purple spot after two minutes.

(4) Sakaguchi Test for the Detection of Arginine Peptides

(Jepson and Smith, 1955)

The paper is dipped in a solution of 0.1% oxine in acetone and dried in the air. It is next dipped in a solution of 0.2 ml bromine in 100 ml 0.5 M NaOH. A positive test is indicated by an orange red spot.

N.B. Methods (1), (3) and (4) can be carried out on the same paper, in that order.

(5) The Method of Phillips (1948)

The paper is heated at a temperature of 110°C for 10 minutes. The peptides thus react with the paper and can be detected under U.V. light.

Recovery of Peptides by Elution from Filter Paper after Electrophoresis and Chromatography

Low yields of large peptides obtained in some experiments indicated that substantial losses were occuring during fractionation. It was suspected that one source of loss was incomplete recovery of peptides from filter paper after electrophoresis and or chromatography. The efficiency of the elution technique was therefore investigated.

A sample of a peptide mixture, containing 744 µg N and known to consist predominantly of one large peptide, was applied to Whatman No.I filter paper and subjected to electrophoresis and chromatography (see page 54). The section containing the large peptide, remaining near the origin, was dissected, and the peptide eluted with 0.05 M NH_3 in the usual manner. The section of paper was then freed of NH_3 by saturation with 0.1 M Na_2CO_3 and drying in a vacuum dessicator over H_2SO_4. The paper was then homogenised in 10 ml ammonia free water and the suspension lyophilised. Weighed samples were then digested for determination of total N, together with a control containing 0.1 mg N.

Total weight of paper	=	225 mg
Weight of samples	(1)	49.5 mg
	(2)	45.5 mg

The digestion mixtures were made up to 10 ml, and 2 ml were subjected to Nesslerisation together with the two controls of NH_4Cl. The results are as shown.

Sample	Optical Density	N present in Sample
Control	0.820	100 µg
(1)	0.285	34.8 µg
(2)	0.270	33.0 µg

. . Mean total N in paper section = 161 µg.

Hence, some 25% of the total peptide N remained on this one section of paper. As four other peptides were also present on other parts of the paper, the recovery of peptide I may have been much lower.

(3) Ion Exchangers

The separation of amino acids or peptides or ion exchangers, depends not only on charge interactions, but also on the Van der Waal's forces existing between the exchanger and the side chains of the amino acids or peptides. Between synthetic resins, which are hydrocarbon in nature, and hydrophilic substances such as hydroxyamino acids, little Van der Waal's attraction exists. These substances, therefore, may pass fairly quickly through columns of the resins. Conversely, hydrophobic compounds such as histidine are more strongly bound and consequently pass through the columns more slowly.

The use of ion exchange resins for peptide separations, came after their development for amino acids. It was found, however, that the closely packed resin bed and the large numbers of side chains on peptides, were liable to cause the peptides to stick on the columns, unless resins with low cross linkage were used. Hirs, Moore and Stein and their collaborators were the major pioneers with this work and they were successful in using Dowex 50 - x2 and other resins with sodium and ammonium buffers to separate fairly complex mixtures of peptides deriving from ribonuclease (e.g. 1952, 1956).

Wittmann and Braunitzer (1959) also used Dowex 1 - x2 to separate the tryptic peptides from TMV. Using buffers consisting of mixtures of 2, 4, 6-collidine, pyridine, acetic acid

and water, the task of desalting the peptides was reduced to one step.

The technique used by us, however, in this work, was developed by Konisberg and Hill (1962). They used Dowex 5o - x2 and pyridine buffers to successfully separate the tryptic peptides of a-globin. This method has all the advantages of those methods previously mentioned, i.e., ease of desalting and good resolution. In addition, however, it requires a shorter column and therefore separation of the peptides is achieved more quickly.

The technique, which was used only once, is fully described in the experimental section page 124 .

The use of cellulose exchangers was pioneered by Sober et al. (1956), who found that suitable adsorbents could be prepared from "a-cellulose" by treating cellulose powder with chloroacetic acid to form a cation exchanger (CM-cellulose), or with 2-chloro-N, N-diethylamine to form an anion exchanger.

Almost all that was said of the theory of ion exchange resins, applies in reverse to cellulose exchangers i.e. they retain hydrophilic substances whilst those which are hydrophobic are unaffected. Also, they form relatively loosely packed beds, and so proteins and peptides are not liable to be so strongly attached.

Although they have been used successfully to effect protein separations, little work appears to have been done to

apply them to the separation of peptides.

In this connection, we have used carboxymethylcellulose with very limited success to separate the tryptic peptides of cytochrome c. The technique used is fully described in the experimental section, page 79 .

Methods of Peptide Detection in the Eluate from a Chromatography Column.

(1) Aliquots from the various fractions were subjected to alkaline hydrolysis, and the resulting amino acids were then reacted with ninhydrin at pH5. The intensity of the colour produced was measured at 570 mμ and this gave an estimation of the quantity of amino acid material present.

Procedure

1 ml 2-5 M NaOH is used for alkaline hydrolysis over a period of 2 hours at 100°C. This is neutralised with 0.5 ml 6N acetic acid and the adjustment to pH5 was completed by the addition of 0.3 ml 0.5 M citrate buffer pH5. 0.6 ml of the Yemm and Cocking (1955) ninhydrin reagent was then added and the colour developed at 100°C for 15 minutes, and read at 570 mμ.

The ninhydrin solution consists of 0.5 gm ninhydrin, 49 ml methyl cellosolve and 1.0 ml 0.01 M KCN solution.

During these estimations, great care must be taken to exclude ammonia which gives a positive test. For this reason, the citrate buffer used was passed through a column of zeocarb 225 in the sodium form and the methyl cellosolve was purified on a column of amberlite IR 120. Also, the ninhydrin was re-crystallised according to the method of Hirs, Moore and Stein (1951).

(2) Optical Density Measurements

The optical densities of the solutions suspected of containing peptide material are read at 280 mμ. Peptides containing aromatic amino acid residues, such as tyrosine, absorb at this wavelength and can thus be detected.

Preparation of DNP Peptides Followed by Acid Hydrolysis and Isolation and Estimation of the Resulting DNP Amino Acids. (Lockhart and Abraham, 1954 and 1956).

The peptide mixture, say 2-3 mg, was dissolved in 1 ml 5% trimethylamine carbonate pH9. To this, 100 µl F.D.N.B. in 2 ml ethanol was added, and the mixture shaken for 2 hours at 40°C. At the end of this time, most of the ethanol was evapourated off in a cold air stream and to the aqueous solution, 2 ml of trimethylamine carbonate were added. It was then extracted three times with peroxide free ether and the aqueous phase evapourated to dryness before hydrolysis overnight in 6N HCl at 110°C (1 ml).

The acid hydrolysate was made up to 3 ml with water, and the DNP amino acids were extracted, chromatogrammed and estimated as previously described (page 47).

EXPERIMENTAL AND RESULTS SECTION

Isolation and Purification of Cytochrome c

The earliest attempts to isolate and purify cytochrome c were made by Keilin (1930 & 1937), Zeile and Reuter (1933), and Theorell (1935 & 1936). Although their techniques represented a considerable achievement, however, and extensive purification resulted, it was not until 1939 that cytochrome c was obtained in reasonably pure form. Theorell and Åkeson were responsible for this, using an electrophoresis apparatus. Preparation of cytochrome c on a large scale could be performed by this method but it was not suitable for most laboratories on account of the expensive equipment required.

In 1945, Keilin and Hartree modified their previous technique (1937) and by means of ammonium sulphate precipitation on dilute acid extracts, obtained an electrophoretically pure preparation of cytochrome c. Tint and Reiss (1950) however, used their original method with success, but in this laboratory it was found that although it did occasionally produce cytochrome c of high purity, it could not be relied upon to do so.

In 1950, Paleus and Neilands used the polycarboxylic cation exchange resin amberlite IRC-50 for purifying cytochrome c. They were in the forefront of workers using ion exchange resins in protein purification and with cytochrome c they were able to take advantage of its high isoelectric point, 10.05, which allowed the adsorption of concentrates of the protein on the

resin at pH9. Impurities were removed by washing with ammonium buffers, leaving the cytochrome c to be eluted at pH10.8 or fractionated into three bands at pH9. The first of these bands was reduced, and the others oxidised, whilst the second appeared to be the purest as it had an iron content of 0.466%. This is purer than the cytochrome c eluted by pH 10.8 ammonium hydroxide and its physical constants compare favourably with those calculated for an electrophoretically pure sample by Tint and Reiss (1950).

One disadvantage of this preparation, however, was that the material was collected in dilute solution, thus making it necessary to precipitate with T.C.A. in order to concentrate it. This was unfortunate, as it has been shown by Margoliash (1952), that T.C.A. modifies cytochrome c into an enzymatically inactive product, spectroscopically identical with native cytochrome c.

A modification by Neilands (1952), however, improved the technique when he applied the charge at a lower pH of 7 in 0.1 M sodium phosphate buffer. Washing off the impurities in this buffer, he then eluted the cytochrome c in one band with saturated ammonium acetate.

It was also found in this laboratory (Leaf, Gillies and Pirrie, 1958) that cytochrome c could, more conveniently, be eluted by use of a sodium phosphate buffer of the same pH, but higher molarity (0.25 M).

One of the most up to date preparations of cytochrome c relies considerably on the above technique. As might be expected, it embodies the work of many laboratories. Submitted by Hagihara, Monikawa, Tagawa and Okomuki (1958), it begins with the extraction of fat freed minced heart muscle with cold dilute acetic acid.

These extracts are then purified using a series of amberlite IRC-50 columns with ammonium buffers of varying pH's and molarities, ammonium sulphate precipitation similar to that of Keilin and Hartree (1945) and crystallisation of the cytochrome c from a solution of ascorbic acid, ammonia and ammonium sulphate. This crystallisation is probably the main contribution made by this method. The cytochrome c thus obtained by the authors had an iron content of 0.45% and an extinction ratio $\frac{(550\ m\mu)}{(280\ m\mu)}$ of 1.28. Further tests, such as electrophoresis and ultracentrifugation also indicated homogeneity and so the material was considered to be 100% pure.

A method using an extraction procedure which did not involve acid extractants, however, was put forward by Murray, Marinetti and Stotz (1960).

In this method, a neutral salt solution is employed for the extraction of the cytochrome c, and a preliminary purification is brought about on this extract by the addition of acetone which at pH 4.5 precipitates much extraneous protein material. Further

purification on the ion exchange resin XE 64, using phosphate buffers, follows this to give a product with the extinction ratio $\frac{550 \text{ m}\mu \text{ red.}}{280 \text{ m}\mu \text{ ox.}} = 1.25$, which compares well with the values reported by other workers for purified cytochrome c. Also, the spectral properties of this product compared favourably with those reported by others for highly purified preparations (Hagihara et al. 1958 & 1959; Margoliash, 1959).

In our own preparation of cytochrome c, the methods used were based on the work of Hagihara et al. and Stats et al.

Minced heart muscle (3 kgm) was extracted by a solution of sucrose, containing saponin (Stats et al. 1960) which was then discarded. This was followed by extraction of the residue twice with acetone and the extracts were again discarded. These washes were carried out (Stats et al. 1960) to disrupt the cells, particularly the erythrocytes present, while keeping the mitochondria intact. It was thus possible to wash out most of the soluble cytoplasmic proteins whilst the cytochrome c remained in the particulate residue. This particulate was disrupted by the acetone washes.

The cytochrome c was extracted from the tissue with a sodium chloride solution, and on this, the preliminary purification, using acetone (mentioned above) was carried out.

Purification was continued (Stats et al., 1960) using a batch procedure, and the ion exchange resin XE 64 buffered to pH6.

The cytochrome c adsorbed on this resin was washed with distilled water and then eluted off using 0.5 M $Na_2 HPO_4$.

This eluate, after dialysis was put through a column of the same resin and then eluted off using 0.1 M $Na_2 HPO_4$.

At this stage, according to Statz et al. a very pure product of cytochrome c should result. This, however, was not our experience, and so our preparation was subjected to the ammonium sulphate treatment described by Hagihara et al. (1958) and then, after dialysis, chromatogrammed on amberlite IRC 50 equilibrated with 0.25 M phosphate buffer pH 7 (Hagihara et al. 1958). The material was eluted from this column with the same buffer and spectral determinations carried out on the eluate indicated 260 mg cytochrome c and $\frac{550\ m\mu\ red.}{280\ m\mu\ ox.} = 1.06$

It was found, however, that purification could be achieved by fractionation on a column of G25 sephadex. Consequently, the total cytochrome c solution was dialysed for 6 hours against running water, and then overnight against 4 litres of distilled water. The resulting solution was freeze dried, taken up in 20 ml water and applied to a column of G25 sephadex, equilibrated in distilled water, 60 cm long x 4 cm in diameter. The eluate from this column was collected in fractions of size 6 ml.

The cytochrome c emerged from the column after 50 ml of eluate, and the table shows the extent of the purification, Table 3.

Table 3.

Fraction Number	Ratio $\frac{550\ m\mu\ red.}{280\ m\mu\ ox.}$
9	1.26
11	1.5
13	1.26
15	1.25
17	1.26
19	0.94
21	1.1
23	1.07
27	0.84
29	0.84
31	0.45
33	0.27

Very dilute fractions

Fractions 9 - 17 were bulked and taken as reasonably pure cytochrome c.

From the table, it would appear that the impurities in the preparation, follow the cytochrome c down the column, over-lapping slightly with it at its tail. This suggests therefore that the contaminating material is a protein, or proteins, with molecular weights slightly less than that of cytochrome c.

The N-Terminal Residue of Cytochrome c

The number of free amino groups possessed by cytochrome c was first reported on by Theorell and Åkeson in 1941 Using the nitrous acid method of Van Slyke, they obtained results to suggest that cytochrome c had 9-10 free amino groups in excess of the number of lysine residues.

This matter was reinvestigated later by Leaf (P.C) however, and he found by means of Fraenkel Conrat's Orange G method, that horse heart cytochrome c had a total of 24.6 equivalents of basic groups per mole.

Now, the best amino acid analysis indicates that in cytochrome c there are 19 lysine, 3 histidine and 2 arginine residues, in all 24 basic side chain groups. This, therefore, left open the possibility of, at most, one N-terminal residue. Attempts made previously in this laboratory, to identify this, had proved unsuccessful. Also, the results of Margoliash (1955) and

Matsubara and Hagihara (1957) were conflicting, and so it was decided to reinvestigate the matter.

The first approach to the problem was substantially the same as that used by Margoliash (1955) and Hagihara (1957). This was the preparation of DNP cytochrome c, followed by acid hydrolysis, and extraction and chromatography of the resulting amino acids. In this way, Margoliash identified histidine and Hagihara identified arginine as the N-terminal amino acid.

The preparation of DNP cytochrome c was accomplished according to the method of Schroeder, (1957) using 12 mg cytochrome c (prepared in this laboratory). The DNP cytochrome c was acid hydrolysed and the DNP amino acids extracted with ethyl acetate and secondary butanol as previously described. The residue from the ethyl acetate extract was chromatogrammed as usual, and the one from secondary butanol was subjected to high voltage electrophoresis in N ammonia solution for half an hour with a potential difference of 30 volts/cm.

RESULTS See figures 10 and 11

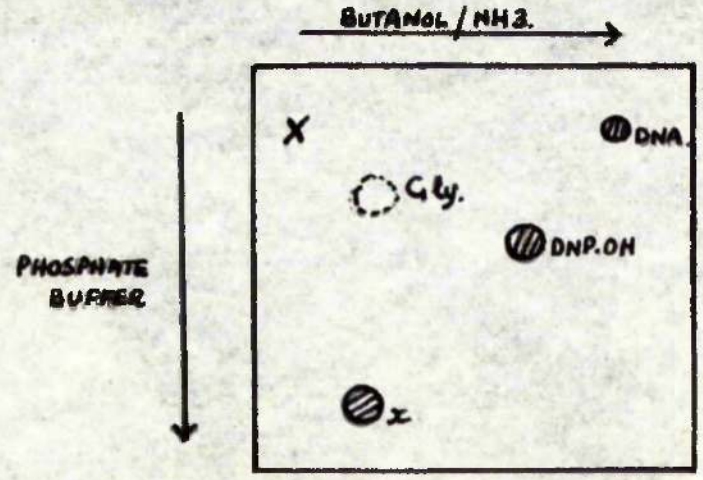


Fig 10.

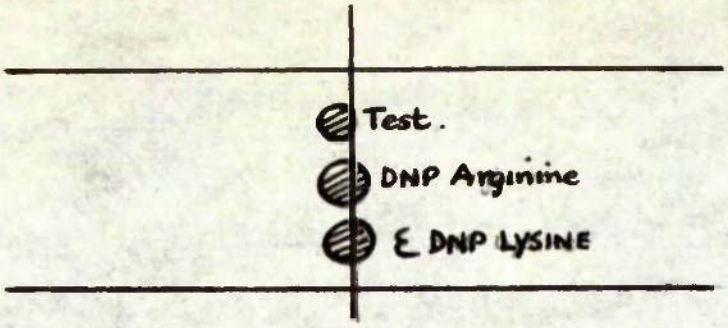


Fig. 11

Spot x, from the chromatogram resulting from the ethyl acetate extract, was further investigated. The material comprising the spot was extracted from the paper by water, which was extracted in turn by ethyl acetate. After evaporation of the ethyl acetate, the residue was then subjected once again to 6 N HCl hydrolysis. On completion, the hydrolysate was evaporated to dryness, and the residue taken up in water and chromatogrammed one dimensionally using butanol acetic acid. On development with ninhydrin solution, the chromatogram shown in the figure was obtained.

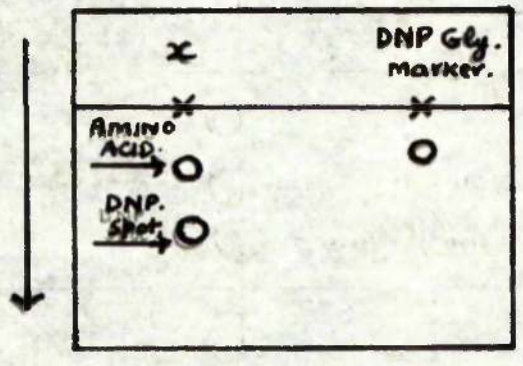


Fig. 12.

This indicates that substance x was the derivative of a small peptide, probably containing lysine. The electrophoretogram shown in figure, did not differentiate the test spot between E DNP lysine and DNP arginine. To do this, a Sakaguchi spray (see methods) for the detection of arginine was used, with negative results.

Therefore, these results supported neither Margoliash or Hagihara. It was possible, however, that some exceptional difficulty existed, and so two other approaches were tried.

Enzymatic Hydrolysis of Cytochrome c

The failure of the above experiment to give conclusive results, necessitated a different approach, and for this reason enzymatic hydrolysis was used. Bergmann(1941) has shown that the enzyme trypsin hydrolyses proteins at their lysine and arginine residues. Therefore, if the work of Hagihara was correct, and arginine is the N-terminal amino acid of cytochrome c hydrolysis of the DNP derivative of this protein with trypsin should yield DNP arginine.

The activity of a trypsin preparation in our possession was assessed, using DNP-l-arginine ethyl ester as substrates (see next page) and satisfactory results were obtained, showing no interference from the DNP group. 25 µl of a solution of this trypsin (1% w/v in 0.057 M HCl) were therefore added to 12 mg DNP cytochrome c, dissolved in 1.5 ml 0.1 M phosphate buffer pH 7.9, and the solution incubated overnight at 37°C. At the end of this time, the solution was carefully acidified to pH3 and then extracted with ethyl acetate and secondary butanol as previously described. The material thus extracted was chromatogrammed and electrophoresed (as above), and examination of the resulting chromatogram and electrophoretogram with the

Sakaguchi reagent, failed to indicate the presence of DNP arginine.

Determination of the Activity of Trypsin towards DNP-l-Arginine Ethyl Ester

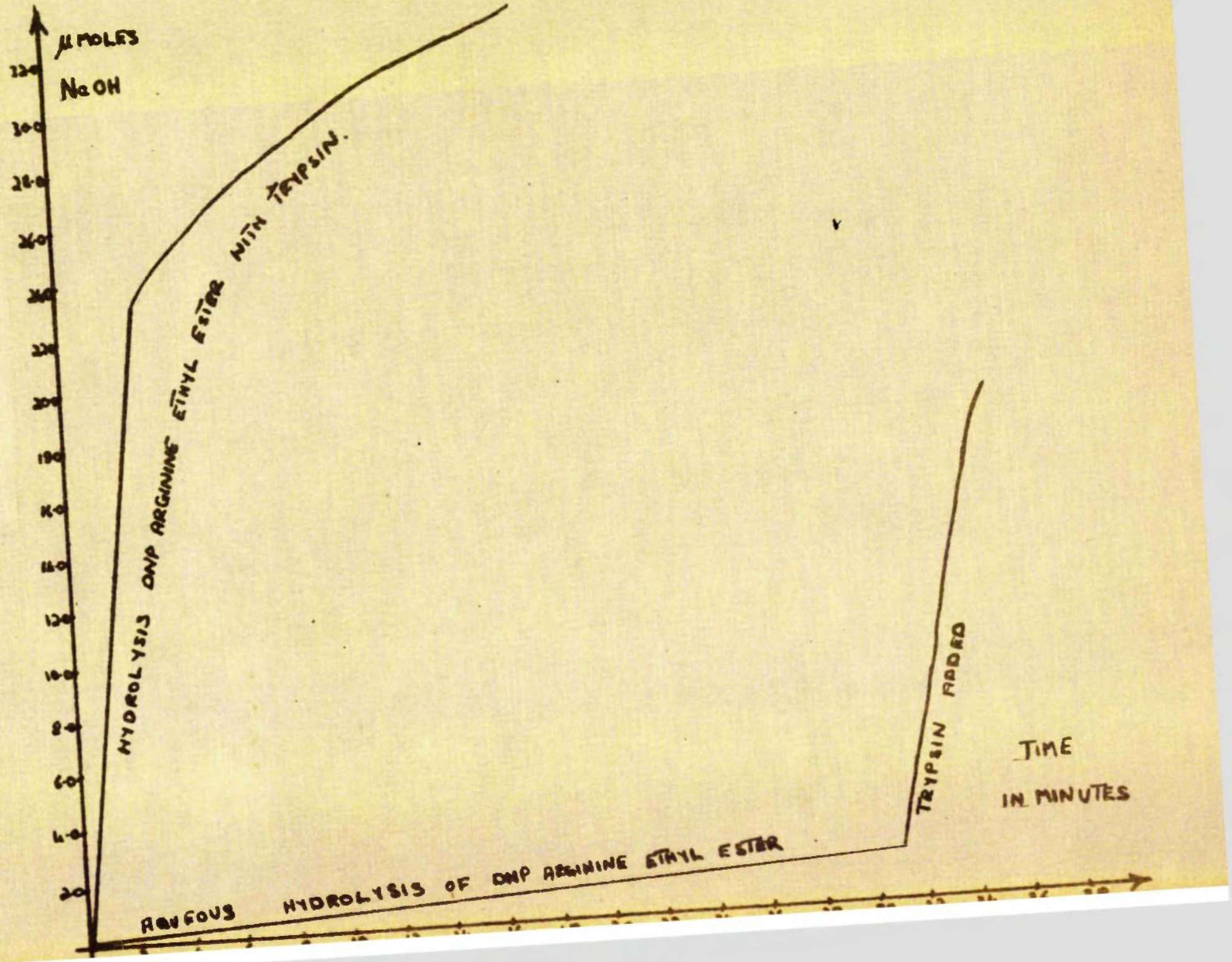
DNP-l-arginine was prepared according to the method of Sanger (1945) and converted to its ethyl ester using the conditions described by Bergmann et al. (1939). The DNP-l-arginine ethyl ester was precipitated from the reaction mixture with dry ether, in the form of a yellow oil. This was centrafuged off, allowed to stand under vacuum for twenty-four hours after which time it appeared as a yellow crystalline substance.

The compound was purified by passing it through an alumina column using alcohol-ether (1:1) as eluant. The material moved down the column in a single band, suggesting that it was already pure. This was confirmed by a one dimensional paper chromatogram, using butanol-ammonia as solvent, and having a reference spot of DNP-l-arginine present. The above preparation gave only one spot with a different R_f value from DNP arginine.

Tryptic Hydrolysis

A sample of DNP-l-arginine ethyl ester (9 mg) was dissolved in 4 ml of water and the pH adjusted to 7.5. 100 µl trypsin solution (1% w/v in 0.057 M HCl) was added, and a fairly rapid hydrolysis ensued. This was followed by titration of 0.1 M NaOH to pH 7.5 and a graph plotted of titration versus time. Figure 13.

FIGURE 13


μ MOLES
NaOH
HYDROLYSIS OF DNP ARGININE ETHYL ESTER WITH TRYPSIN.
TRYPSIN ADDED
TIME
IN MINUTES
AQUEOUS HYDROLYSIS OF DNP ARGININE ETHYL ESTER

A similar control experiment was carried out with no trypsin present, to find out the extent of aqueous hydrolysis. After 31 minutes of this, 100 µl trypsin were added in order to directly compare the two rates of hydrolysis. The results of both were plotted on a graph. Figure 13

Deductions from Graph

The activity of the trypsin is determined as the gradient of the tangent drawn to the reaction curve.

$\therefore$ Activity $= \frac{7.6}{1} = 7.6$

Control gradient $= \frac{0.7}{5} = 0.14$

$\therefore$ True activity $= 7.46$

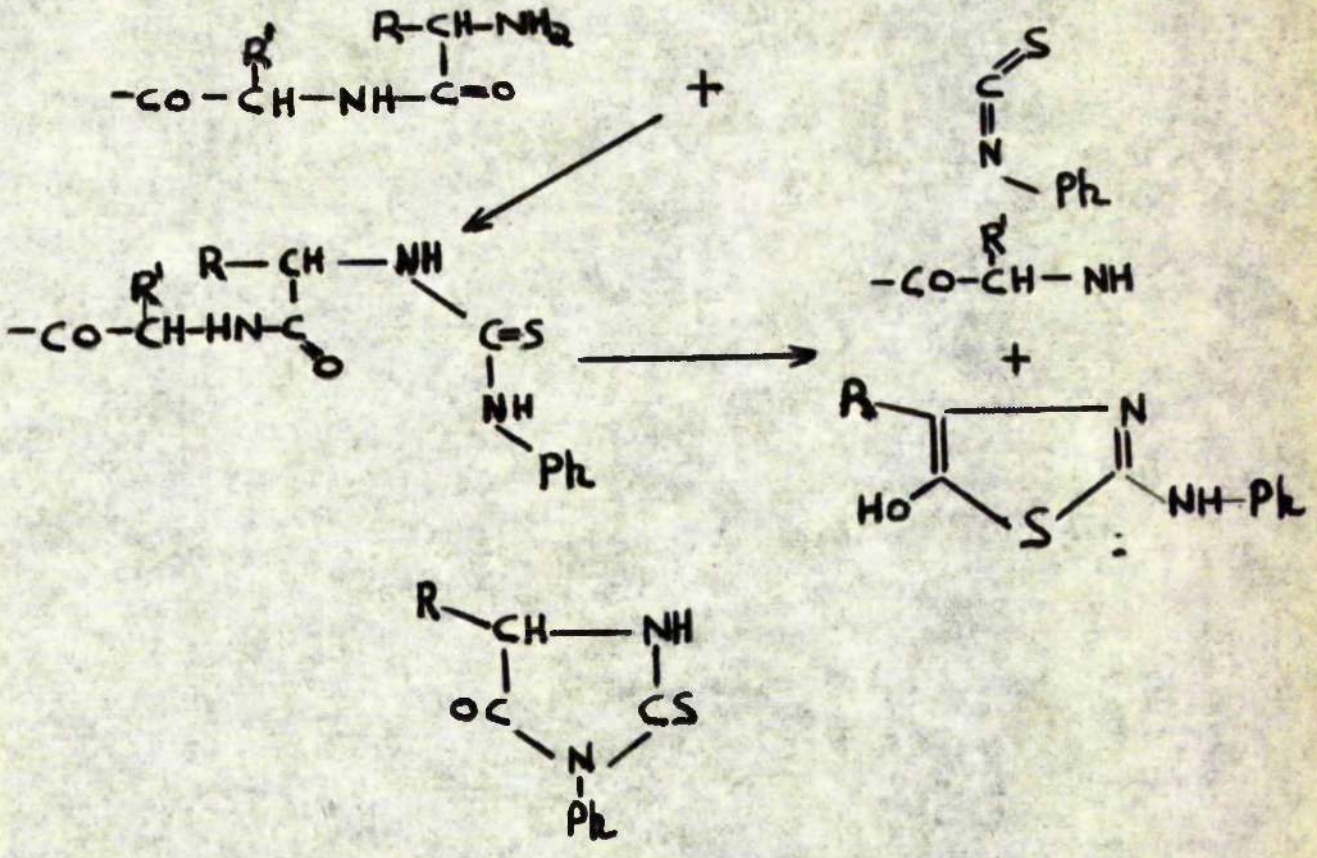
Reaction Constant $K = \frac{2.303}{5} \log \frac{23.4}{9.5} = 2.3$

The alkali uptake was 35 µmoles. This indicates that the hydrolyses was almost complete, as 36 µmoles DNP arginine ethyl ester were used.

The previous results established definitely that arginine was not the N-terminal amino acid of cytochrome c, despite the evidence published by Hagihara (1957). This conclusion, however, did not explain why no DNP end group had been obtained in the first experiment, when acid hydrolysis of DNP cytochrome c was carried out. This tended to suggest rather that cytochrome c had no N-terminal group, perhaps through being joined head to tail,

or that the N-terminal residue was such that methods involving substitutes with F^{D}NB were unsuitable for its detection.

A further attempt to detect an N-terminal residue was therefore made, using a method relying on rather different principles - the phenyl iso thiocyanate method of Edman for end group analysis (1950 and 1960)



The chemistry of the technique is outlined in Figure 14 and as can be seen, its effect is to strip off the N-terminal amino acid of the protein in the form of a phenyl thio hydantoin derivative. The presence of this derivative in the reaction mixture, is indicated by a U.V. maxima between 230 and 290 mμ, but this is not sufficient to identify the amino acid itself. This is done by hydrolysing the phenyl thio hydantoin amino acid with 6NHCl

thus producing the amino acid in its original state, and to identify this chromatographically. The experimental details are obtained from the reference.

This procedure was carried out on cytochrome c with negative results, whereas with the same conditions, myoglobin yielded glycine, as the N-terminal residue.

This therefore confirmed that the N-terminal amino acid of cytochrome c was not susceptible to the usual reactions.

It is now known that the N-terminal amino acid of cytochrome c is acetyl glycine (Tuppy, 1961). This explains the difficulty encountered in causing reaction, in that the amino acid is acetylated.

Carboxymethyl Cellulose and Peptide Fractionation

In order to investigate further the amino acid sequence of cytochrome c, the protein was degraded using the proteolytic enzyme trypsin. The separation of the resulting peptides was attempted unsuccessfully, using zeocarb 225 and so it was decided to use the cellulose ion exchanger, carboxymethyl cellulose to obtain pure samples of these.

Preparation of Column

Carboxymethyl cellulose was equilibrated overnight in 0.005 M acetate buffer, pH5. A column (50 cm long x 1.7 diameter) was then poured, the top surface being tapped down with a glass rod to give closer packing.

Meanwhile, 2 μmoles (24 mg) cytochrome c were subjected to tryptic hydrolysis in 5 cc 0.05 M phosphate buffer pH 7.4, for 4 hours at 37°C using 0.5 mg trypsin. At the end of this time, the solution was boiled to destroy tryptic activity , lyophilised and taken up in 2 ml water. The pH was adjusted to 5 with 2 M acetic acid and the solution was then applied to the column.

The elution from the column was begun with 0.005 M sodium acetate buffer pH 5.0. After the passage of 10 ml, the concentration of the eluting buffer was gradually increased by running 0.5 M sodium acetate into a 100 ml mixing flask (initially charged with the starting buffer) attached to the column. Fractions of three ml were collected, and the appearance of peptide material in these was detected by means of reaction with ninhydrin, after alkaline hydrolysis had been carried out on an aliquot of 20 μl, (see methods). The results of these were plotted on a graph. It was noted also that fractions 43-60 contained material with the characteristic red colour of haem, thus indicating the presence of the haem peptide of cytochrome c.

The absorption of the material in these fractions at 396 mμ was read and these results are also contained in the graph.

Calculation from these figures on the amount of haem peptides obtained, taking E = 28,000, show that 2.05 μmoles are in fact present. This agrees well with expectation.

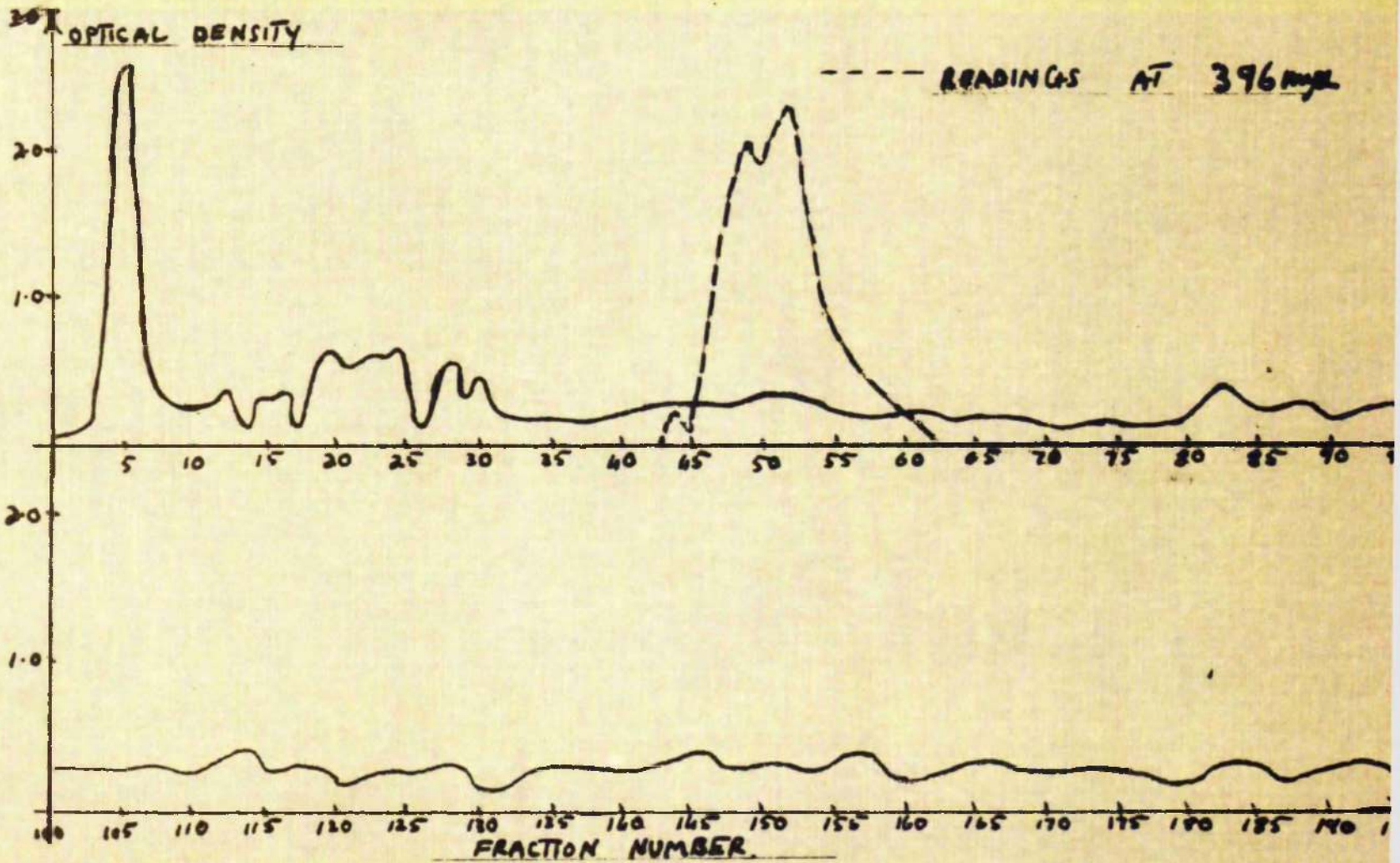
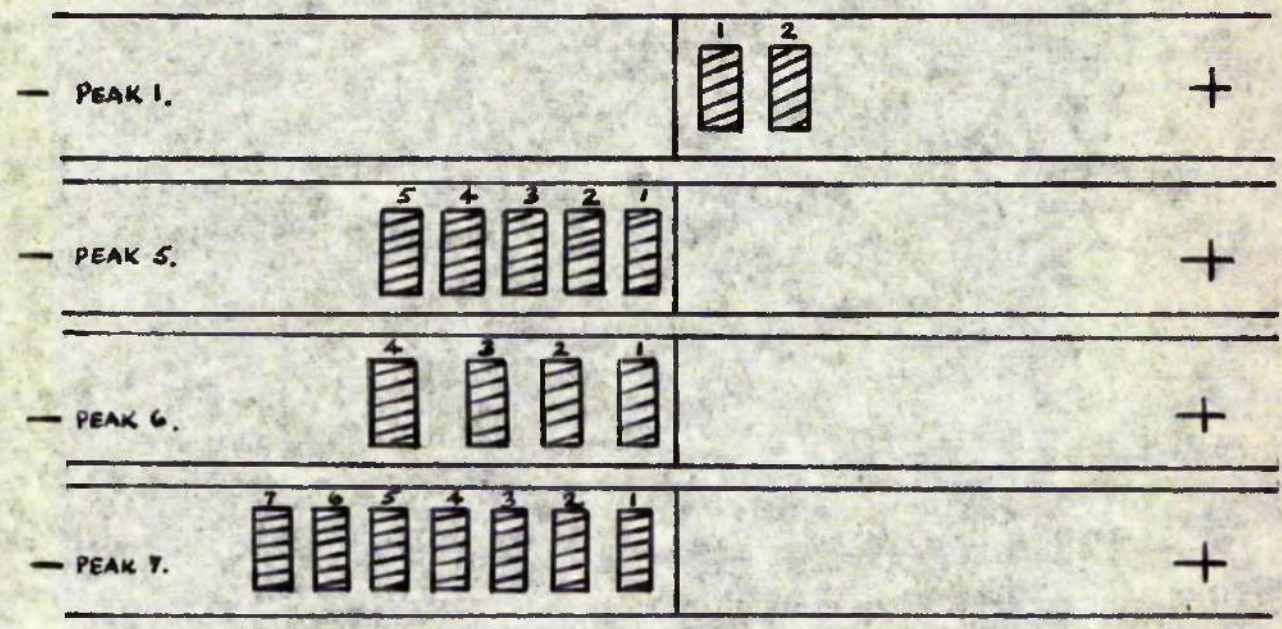

OPTICAL DENSITY
---- READINGS AT 396mμ
2.0
1.0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90
2.0
1.0
100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190
FRACTION NUMBER

FIGURE 15

As can be seen from the graph, very little clear cut separation of the peptides was obtained. In order to investigate this further, peaks I (tubes 4-7), 5, (18-20), 6 (21-23) and 7 (24-26) were lyophilised and taken up in 1 ml water. One tenth of the total material in each case was then subjected to high voltage paper electrophoresis in pyridine-acetic acid-water buffer pH 3.4 (10:100:890) at 50 volts per cm for 50 minutes. The resulting peptide positions were determined by means of ninhydrin dip, and the other techniques described in the methods section. Fig. 16



The separated peptides were *eluted, hydrolysed with HCl* and the amino acids liberated were detected by paper chromatography* with butanol-acetic acid-water vs phenol water.

* Details of these given in "methods."

In some cases, semiquantitative results were obtained using the method of Heilmann et al., but in the main, the results shown are merely qualitative.

RESULTS See Tables. 4 & 5

Table 4

Amino Acid	PI (1) Molar Ratios	PI (2)	P6 (1)	P6 (2)	P6 (3)
Asp	1.4	+	2+		1+
Glu	4	+	1+	1+	1+
Gly	not estimated by this technique		1+	1+	1+
Thr	1.53	+	2+	1+	
Al	1.1	+	1+		1+
Lys	2.2	+	1+	2+	2+
Meth	1.8		1+	1+	
Leu	3.4	+	2+	3+	
Tyr	0.72	+	1+	1+	
Phe	0.8		3+		
Pro	not estimated by this technique		1+	1+	

Table 5

Amino Acid	P5 (I)	P6 (4)	P5 (5)	P7 (2)	P7 (4)
Asp	2+			1+	
Glu	1+			1+	1+
Gly	2+	2+	1+		
Al	2+			1+	
Leu + ILeu	1+				
Phe	1+				
Tyr	1+				
Thr	4+			1+	1+
Pro	1+				
Lys	4+	+	2+	1+	
Arg					1+

Chromatograms derived from hydrolysates of P5 (2), P5 (3), P5 (4), P7 (1), P7 (3), P7 (5), P7 (6) and P7 (7) were all too faint to interpret.

Thus the peptide fractionation obtained on the column of carboxymethyl cellulose was not efficient. Only by the application of electrophoresis was it possible to obtain pure peptide samples. This technique was therefore abandoned in favour of a more sophisticated approach.

The Formation of Methoxycarbonyl Cytochrome c and its Hydrolysis by Trypsin

Preparation of Methoxycarbonyl Cytochrome c (Chibnall, 1958)

100 mg cytochrome c (extracted from horse heart in this laboratory) was taken up in 8 ml water and the pH adjusted to 9 using 2N NaOH. Nitrogen was used for stirring and 0.470 ml methyl chloroformate was added over two hours in 50 µl portions. The pH of the solution was maintained at 9 throughout the reaction by regular additions of the 2N NaOH.

The resulting material was dialysed against frequent changes of distilled water at 4°C and then precipitated on acidification of the solution. The precipitate was separated from the supernatant by centrifugation, but nitrogen determinations done on both indicated that although most of the material had been precipitated, a considerable amount (1/5th) still remained in solution. The nitrogen determinations were done using Nesslers reagent, after acid digestion of the samples.

Tryptic Hydrolysis of Methoxycarbonyl Cytochrome c

The precipitate and supernatant mentioned above were recombined and the pH adjusted to 8.2 thus once again giving a complete solution. An amount of this, corresponding to 1 μmole cytochrome c was withdrawn, the volume of the solution adjusted to 5 ml with H_2O and with the pH still at 8.2, 100 μl (1% w/v) trypsin solution added. The pH was maintained at 8.2 by regular additions of 0.1N NaOH and these additions were plotted against time on a graph. The graph shows the progress of the hydrolysis.(Fig.17)

From the graph, it can be seen that 4.46 μmoles NaOH were used per μmole cytochrome c. If masking had been complete, hydrolysis would have occurred only at the arginine residues and so only two μmoles NaOH would have been taken up.

Now, the excess uptake of 2-46 μmoles corresponds to the release of at least 2-46 equivalents of amino groups. Probably more in fact when it is remembered that the uptake of alkali depends on the relation between pH and pK of liberated groups. This is expressed in the Henderson equation

$$pH = pK + \log \frac{RNH_2}{RNH_3^+} - \frac{a}{1-a}$$

with respect to the reaction

$$R^1 - CO - NH - R \qquad R^1 - COO^- + (a)RNH_2 + (1-a)RNH_3^+ + a\,H^+$$

In order to check these results, the extent of the tryptic hydrolysis was further investigated by the fluorodinitrobenzene technique:-

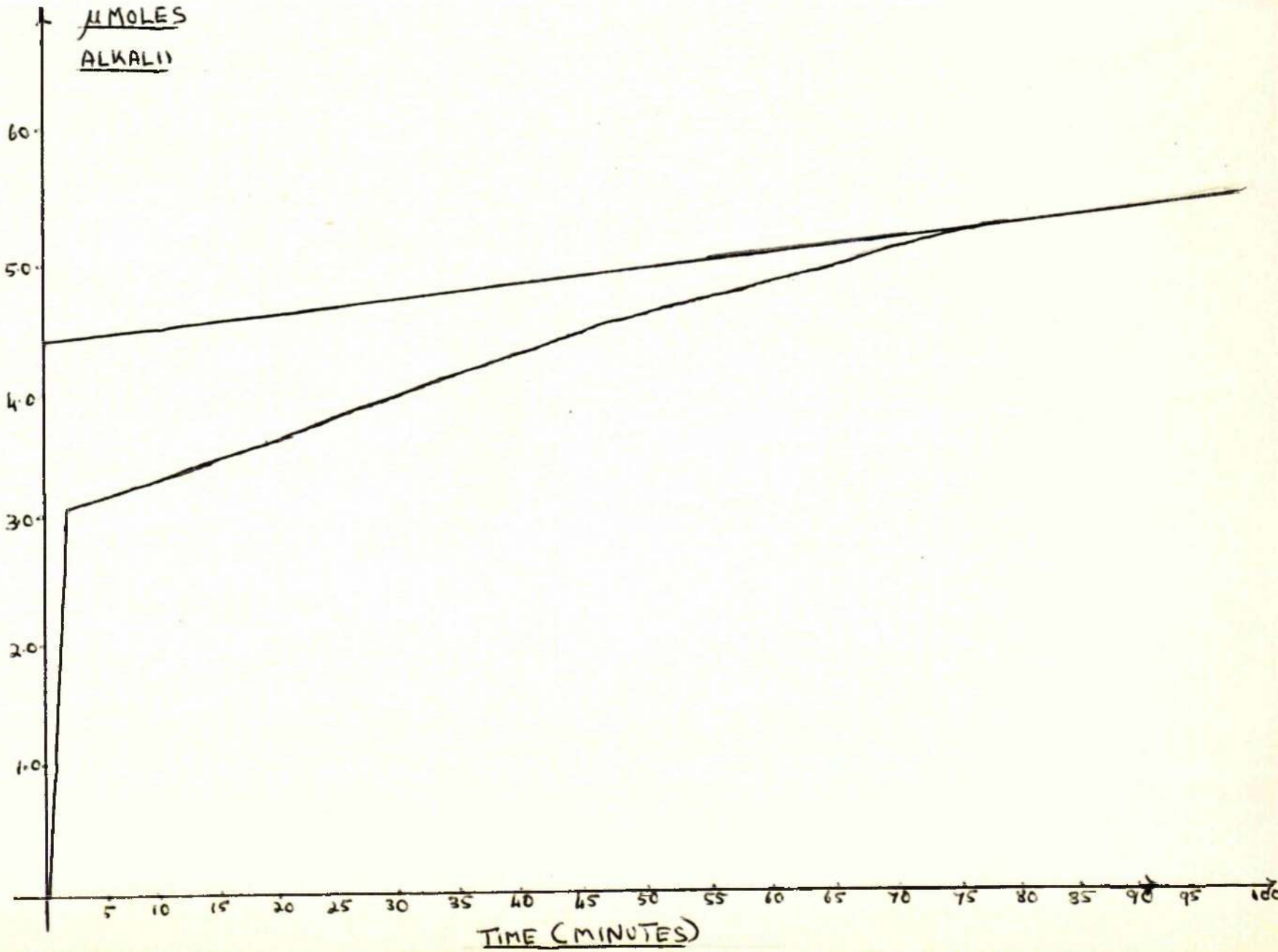

μMOLES
ALKALI
60
50
40
30
20
10
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
TIME (MINUTES)

FIGURE 17

The DNP derivatives of the N-terminal amino acids were prepared and chromatogrammed as described in the "methods section" (see figure 18). The DNP amino acids were identified as shown by reference to a control chromatogram run under identical conditions.

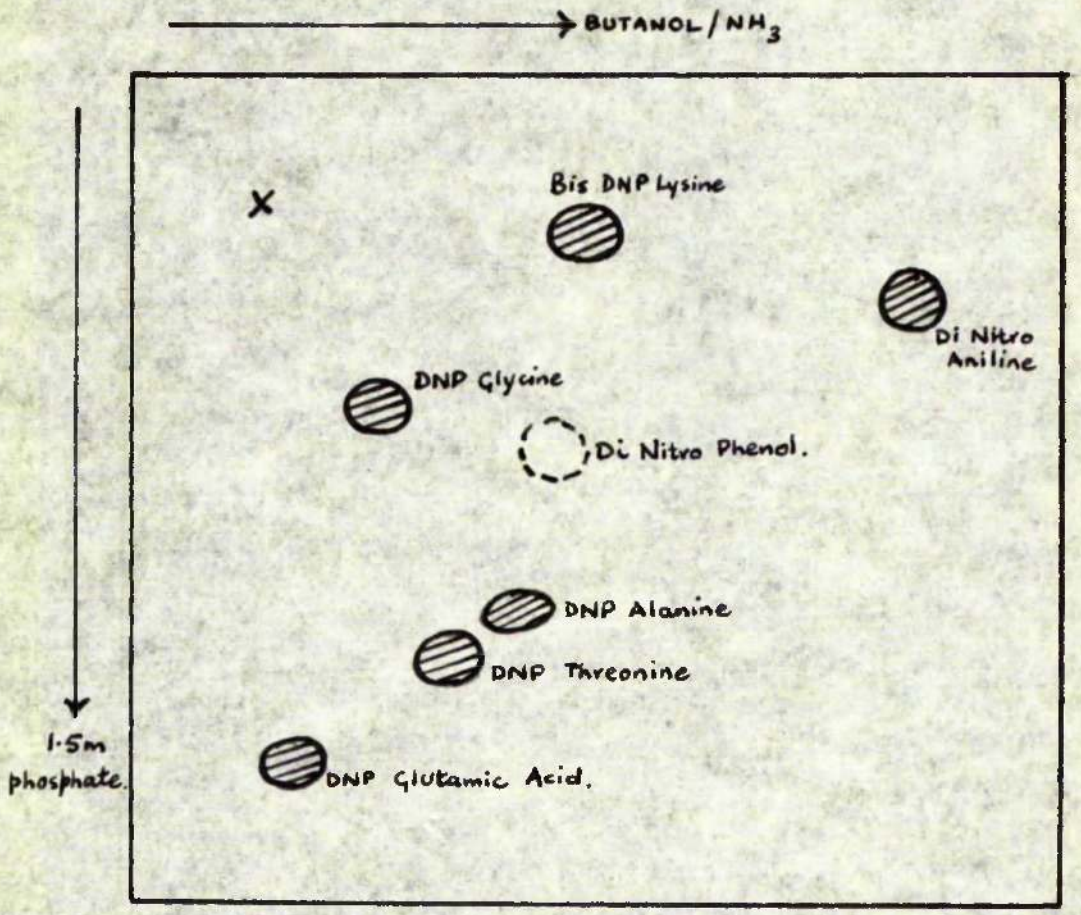


fig. 18.

The chromatogram suggests, therefore, that the tryptic hydrolysis has resulted in the creation of four or five new N-terminal end groups, as shown above. This would agree quite well with the results calculated from the uptake of alkali which indicated hydrolysis at 4 or 5 sites.

The Use of Trifluoracetylation to Increase the Specificity of Trypsin

The work by Weygand (1952) and subsequently Sahallenberg, and Calvin (1955) indicated that trifluoracetylation might be a convenient means of blocking tryptic hydrolysis at the lysine residues of a protein chain. The effectiveness of this technique, however, is dependent on the completeness of the reaction between the E amino groups of the lysine residues of the protein and CF_3 COS Et. A possible way to determine the extent of this reaction with respect to cytochrome c seemed to be to react the trifluoroacetylated cytochrome c with FDNB, and then after acid hydrolysis (6N HCl), estimate the ratio of free lysine: E DNP lysine. (The trifluoroacetyl groups being unstable to acid hydrolysis, whilst the DNP groups are stable).

To prove the validity of this method, however, it was necessary to establish that all the ϵ-lysine amino groups were able to react with FDNB, and define conditions under which complete substitution occurred.

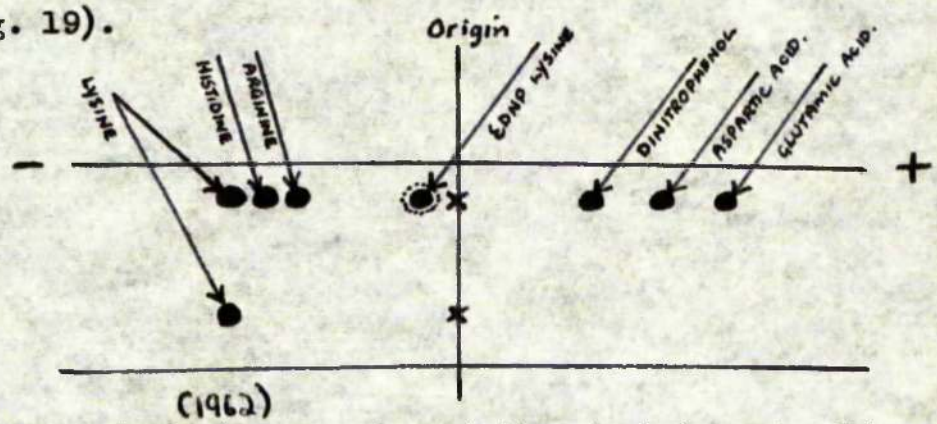
Obviously, with DNP cytochrome c, the ratio of free lysine: ϵ-DNP lysine once again gives a measure of the completeness of the reaction in this case with FDNB.

Technique for Separation and Estimation of ϵ DNP Lysine and Free Lysine

The ϵ DNP lysine and free lysine were separated from each other, and from the other constituents of the hydrolysate by high

voltage paper electrophoresis using pyridine-acetic acid-water buffer pH 6.4 and a potential of 100 volts per cm for 30 minutes. At the end of this time, the paper was thoroughly air dried to remove all pyridine and the material comprising the ε DNP lysine spot estimated as described in the methods section. The free lysine present was estimated according to the method of Heilmann et al. (1957).

A drawing of a typical electrophoretogram is shown below. (Fig. 19).



(1962)

Anfinsen et al. have used a similar technique to this one for the estimation of lysine.

Preparation of DNP Cytochrome c

(a) By the Method of Schroeder (1957) (see "methods")

20 mg cytochrome c were taken for this preparation after which the DNP protein was precipitated by acidification to pH3 using 0.5 M HCl. It was then centrifuged off, washed with dilute HCl (pH3) until the washings were no longer yellow and vacuum dried. A sample of this was taken (3.3 mg) and the extent of substitution estimated using the technique described above.

Results

Σ DNP lysine/3.3 mg DNP protein = 3.3 µmoles

lysine/3.3 mg DNP protein = 0.52 µmoles

. . % age Σ lysine amino groups free = 13.6

This corresponds to 2.54 amino groups/mole cytochrome c.

(b) Preparation of DNP Cytochrome c using an Automatic Titrator

12 mg cytochrome c was dissolved in 8 ml water and 30 µl FDNB added to the solution which was stirred vigorously at 37°C. The pH was adjusted to 9 and maintained there by automatic titration of 0.1 M NaoH for three and a half hours. At the end of this time, the DNP protein was precipitated by acidification of the solution to pH3 with 0.1 M HCl and after centrifugation it was washed with acidulated water pH3 (once), ethanol (once) and ether (three times).

Complete removal of the ether under vacuum followed this, and then a sample of the DNP protein was taken, (10 mg) and the degree of substitution determined as previously described.

Results

Σ DNP lysine/10 mg DNP protein = 12.5 µmoles

No free lysine was detected

. . substitution was complete.

A similar experiment to this was also carried out at room temperature, when it was found that complete substitution was not obtained.

Results

ϵ DNP lysine/5 mg DNP protein = 4.45 µmoles

lysine/5 mg DNP protein = 1.05 µmoles

. . % age ϵ lysine amino groups free = 20.0

This corresponds to 3.8 amino groups/mole cytochrome c.

Determination of the Extent to which Cytochrome c Becomes Trifluoroacetylated and the Time Taken

20 mg cytochrome c was dissolved in 1.5 ml carbonate--bicarbonate buffer (1 M pH9). 100 µl CF_3 COS Et was added and the mixture stirred at room temperature. More ester was added at regular intervals over five hours until in all 400 µl had been used. At the same time, samples were withdrawn from the reaction mixture and treated with FDNB to determine the degree of substitution as previously described. The ϵ DNP lysine produced was plotted against time. (Fig. 20).

Discussion

The results show that trifluoroacetylation is complete after three hours at pH9. It was assumed after this experiment that the pH requirements for the reaction were not critical in the absolute sense. For this reason, therefore, plus the fact that it was considered safer to operate as far as possible from pH11, at which value it was known that desubstitution was liable, subsequent blocking reactions were carried out at pH 8.5. Unfortunately, as was later discovered, 1.0 groups remain

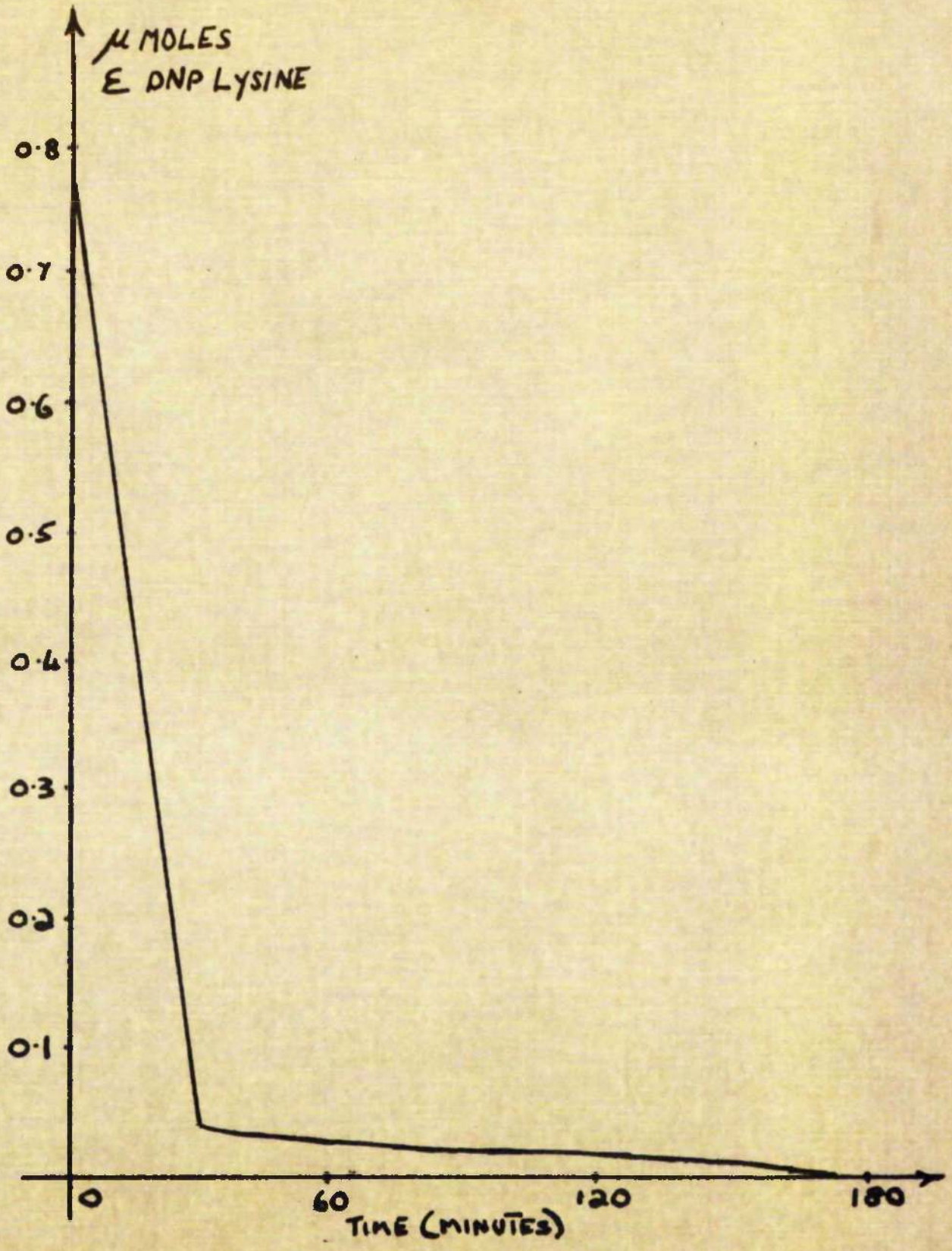

μ MOLES
ε DNP LYSINE
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
60
120
180
TIME (MINUTES)

FIGURE 20

unsubstituted at this pH which presumably lead on tryptic hydrolysis to a certain amount of undesired fragmentation of the protein. As only 1.0 groups were involved, however, this would be small.

Removal of Trifluoroacetyl Groups

In Schallenberg and Calvin's work, the removal of trifluoroacetyl groups (TFA) from amino acids was accomplished simply and easily, by raising the pH of their aqueous solution to 11.

In some of our earlier experiments, as will be seen, the removal of TFA groups from large peptides was attempted in a similar way. Treatment with FDNB, however, resulted in a low yield of E DNP lysine, thus showing that complete removal of the blocking groups was not achieved and so other methods were tried.

(a) A sample of a partially blocked peptide, on which an attempt had been made to remove the TFA groups at pH 11, was treated with FDNB in the usual way to ascertain the exact extent of desubstitution. The peptide was P(1 - 91) [Tuppy]* and 6 mg was taken for estimation.

Results

ε DNP lysine/6 mg DNP protein = 2.62 µmoles

lysine/6 mg DNP protein = 6.22 µmoles

∴ % lysine ε amino groups still substituted = ~~58~~ 68

This corresponds to ~~10.2~~ 12.0 amino groups/mole cytochrome c.

* 1961

An attempt was therefore made to find a method of quickly deblocking the remainder of this material (16 μmoles), which was therefore dissolved in a solution 0.1 M with respect to NaoH (volume 18 ml). Samples (10 μl) were withdrawn at 30 minutes and 60 minutes after which time the pH of the reaction mixture was returned to 7. These samples were subjected to Yemm and Cocking's ninhydrin reaction at the constant intensity of the colour obtained suggested that complete desubstitution had been achieved.

Results

	Optical Density at 570 mμ
30 minute sample	0.99
60 minute sample	0.99
0.2 ml $\frac{M}{2500}$ glycine (0.08 μmoles)	1.03

(b) A sample of cytochrome c (2 μmoles) which had been blocked at pH9 and stored at -10°C for three weeks was dissolved in 0.5 M ammonia and incubated at 37°C.

Samples (200) were withdrawn from this solution at zero time, 30 minutes, 60 minutes, 120 minutes and 16 hours (overnight), and treated with FDNB as usual to determine the number of ε amino lysine groups blocked and unblocked.

Results See Table 6.

Table 6.

Time of incubation (minutes)	ε DNP lysine μmoles	Free lysine μmoles	lysine groups blocked	lysine groups unblocked
0	2.88	5.46	11.9	6.1
30	6.3	4.86	8.0	10.0
60	7.2	4.92	7.0	11.0
120	9.0	3.0	4.2	15.8
overnight	10.8	0.096	0.1	17.9

The Effect of CF_3 COS Et and CF_3 COOH on the Activity of Trypsin

After the trifluoroacetylation of cytochrome c, considerable amounts of CF_3 COS Et and CF_3 COOH must remain in the reaction mixture. The following experiments were therefore carried out to measure their effect on the activity of trypsin and to determine whether or not their removal is necessary.

Two 1 ml 0.5% trypsin solutions were made up in 0.01 M phosphate buffer pH 9 and to one of these, 20 μl CF_3 COS Et was added. Both solutions were then shaken for three hours, at the end of which time, their activities were determined, using tolsene sulphonyl arginine ethyl ester (TSAME).

(a) Activity Determination using the Trifluoroacetylated Trypsin

This was done in the automatic titrator-radiometer at pH 7.5 using 20 mg TSAME dissolved in 8 ml water and 10 µl trypsin solution. The titrator was charged with 0.1 M NaOH, and the graph shows the rate of the hydrolysis. (Fig. 21).

A second experiment was carried out using the same conditions as above in a solution $\frac{M}{25}$ with respect to $CF_3.CooH$. This is roughly the situation after the trifluoroacetylation of 50 mg cytochrome c. The graph shows the rate of hydrolysis.(Fig. 21)

(b) Activity Determination using the Control Solution of Normal Trypsin

The conditions were exactly as in (a) except that 10 µl control trypsin solution was used. See graph. (Fig. 21).

Calculations from Graphs

The activities and reaction constants were calculated as previously described.

10 µl normal trypsin:

Activity = 4

Reaction constant = 0.02046

10 µl "blocked" trypsin

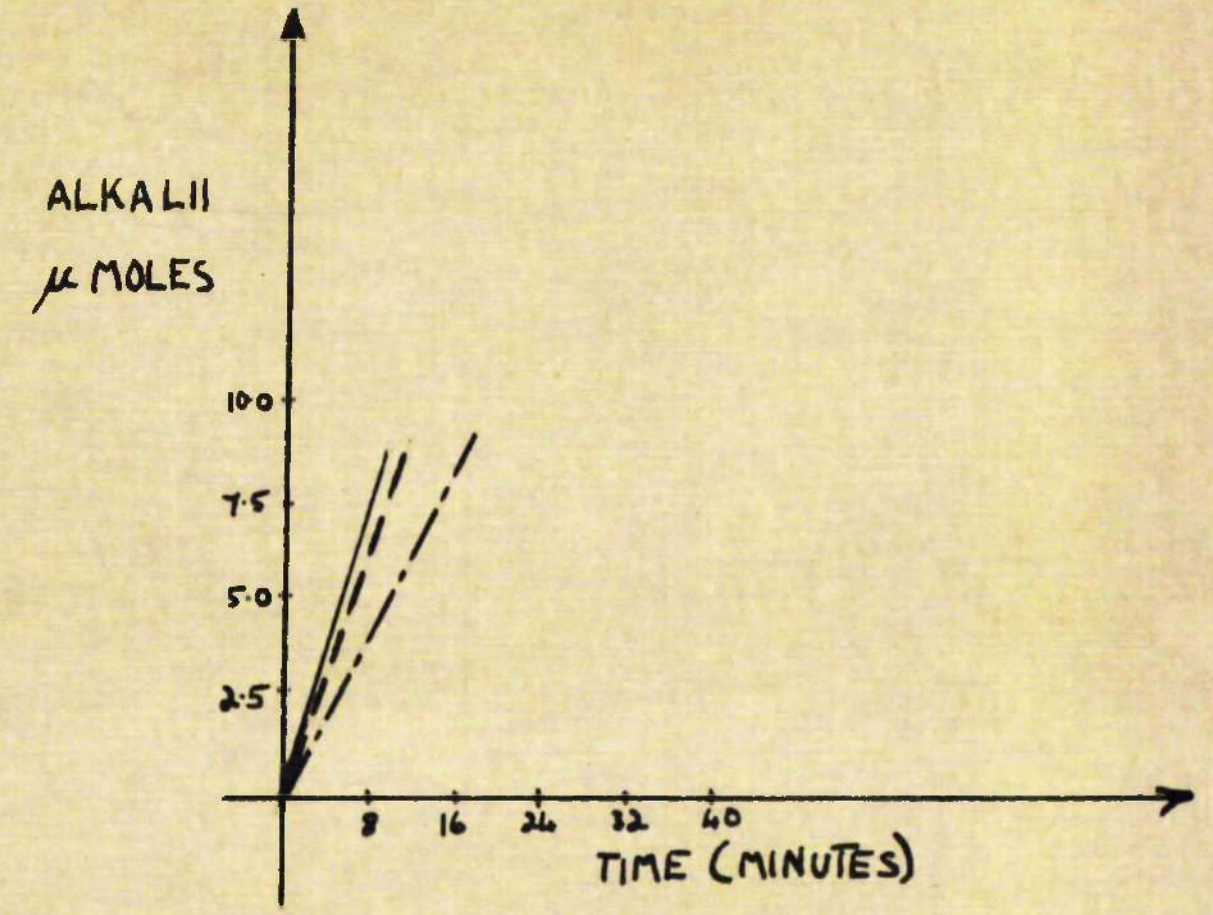
Activity = 3.28

Reaction constant = 0.01647

10 µl "blocked" trypsin + $CF_3.CooH$ (40 µl)

Activity = 2.16

Reaction constant = 0.01200



NORMAL TRYPSIN

TFA TRYPSIN

TFA TRYPSIN + $CF_3 \cdot COOH$

FIGURE 21

Discussion

The trifluoroacetylation of trypsin causes a reduction in its activity at pH 7.5, of 20 - 25%. The trifluoroacetylation of trypsin plus the presence of CF_3.COOH, reduces its activity by 40%. The removal of these substances from a tryptic hydrolysate is therefore desirable, but not essential.

Cytochrome c, its Trifluoroacetylation and Hydrolysis by Trypsin, followed by Investigation of the Peptides.

Experiment (1)

100 mg cytochrome c was dissolved in 8 ml water, and the pH of the solution was adjusted to 8.5. 100 µl CF_3.COS Et was added, and the pH maintained at 8.5 by regular additions of M NaOH from an automatic titrator. The reaction was continued for 3.5 hours and 400 µl more CF_3.COS Et was added over this period.

At the end of this time, the excess ester was extracted with chloroform (twice) - it was necessary to perform these extractions rapidly as aqueous hydrolysis of the ester with consequent pH reduction, causes precipitation of the blocked cytochrome c.

In order to hydrolyse it enzymatically, the blocked cytochrome c solution was returned to the automatic titrator and 1.86 mg trypsin was added. The pH was maintained at 7.5 for two and a half hours, after which time the solution was boiled for 10 minutes to deactivate the trypsin.

Preliminary Fractionation

After hydrolysis with trypsin, followed by boiling, a red precipitate which had formed was centrifuged off and labelled Fraction (1).

The pH of the remaining solution was reduced to 3, and a further coloured precipitate which appeared was also centrifuged off and labelled Fraction (2). Both precipitates (1) and (2) were washed with mildly acidulated water and the washings bulked with the final supernatant which was faintly yellow in colour and labelled Fraction (3).

Deblocking of Peptides

In this work, which was carried out before the control determinations described elsewhere, an attempt was made to deblock by raising the pH of each fraction to 11 with 0.25 M NaOH. The pH was maintained at this level for three hours at a temperature of 37°C, and at the end of this time, the pH's of the solutions were readjusted to 7.

Electrophoretic Separation of the Peptides

This was done as is described in the "methods" section on 10 % of each of the fractions. The buffer used was pyridine/HAC/H_2O pH 6.4 and the electrophoresis was carried out for 35 minutes at a potential difference of 80 volts per cm. (Fig.22).

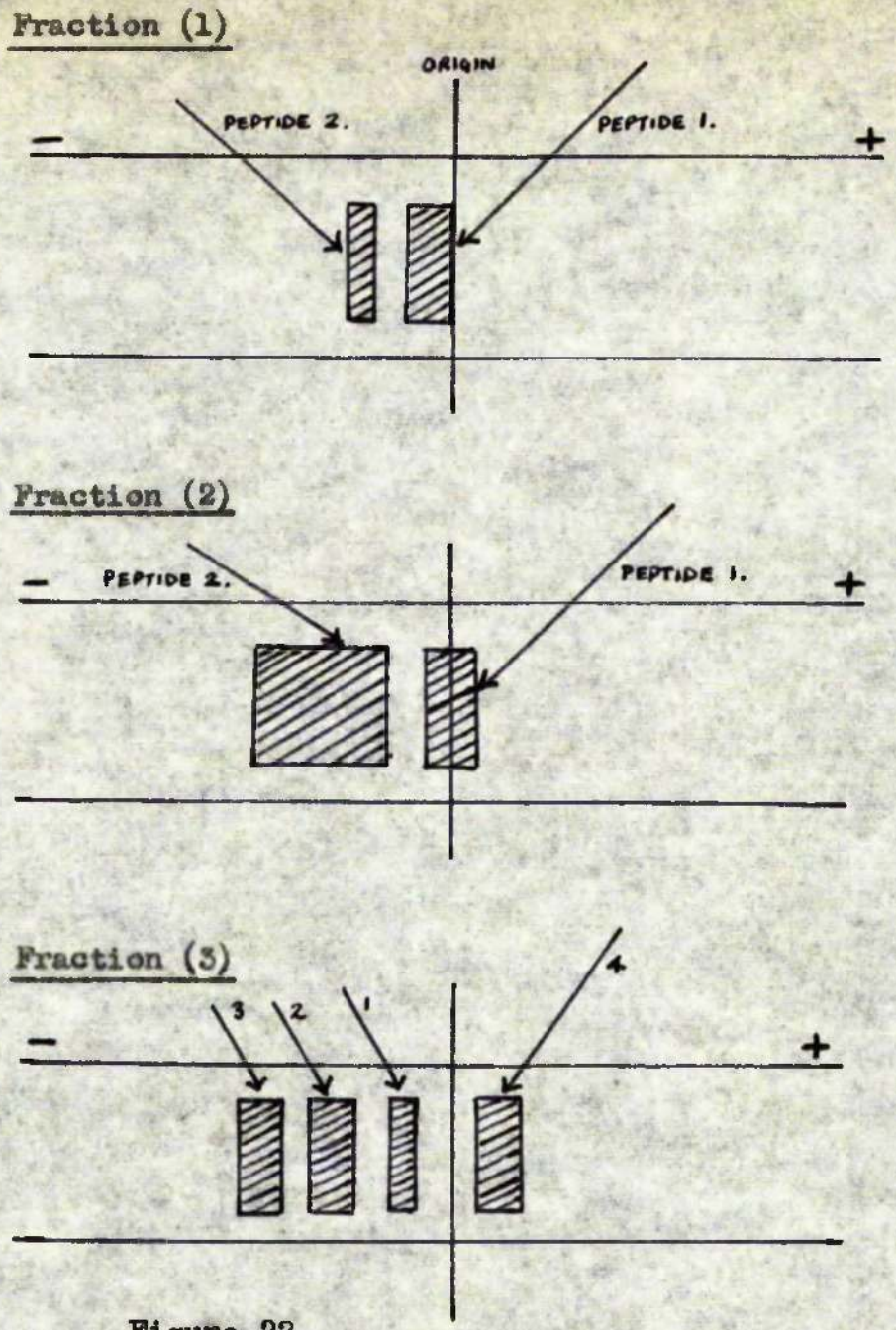


Figure 22

The detection of the peptides and their elution from the paper was achieved as described in "methods". The peptides were then acid hydrolysed and their amino acids estimated using either the DNP method or the method of Heilmann et al. as described in "methods."

Results

Fraction (1)

Here the amino acids were estimated from their DNP derivatives. See Tables 7 and 8.

pI (coloured)

Table 7.

DNP Amino Acid	Total µmoles in Fraction	Molar Ratios
Asp	1.3	2
Glu	3.125	4.9
Thr	1.35	2.16
Pro	0.6	0.97
Val	2.23	3.4
Al	0.6	0.97
Leu & ILeu	1.78	2.75
Gly	4.35	6.8
Phe	0.675	1.08
Hist	1.95	3
Lys	4.355	6.7
Cys	1.3	2
Arg	0.525	0.84

The yield of this peptide is 7 %, based on the amount of μmoles amino acid recovered, compared with expectation

P2 (coloured) Table 8

DNP Amino Acid	Total μmoles from this Peptide in Fraction	Molar Ratios
Asp	0.175	1.05
Glu	0.575	3.4
Thr	0.225	1.36
Val	0.510	3
Leu + ILeu	0.065	0.382
Al	0.175	1.05
Gly	0.775	4.58
Phe	0.15	0.89
Hist	0.25	1.48
Lys	1.012	6
Cys	0.338	2

Yield = 1.75%

Fraction (2) Amino acid analysis on DNP derivatives, see Tables 9 and 10

PI (colourless)

Table 9

DNP Amino Acid	Total μmoles from this Peptide in Fraction	Molar Ratios
Asp	3.21	4.1
Glu	4.98	6.2
Thr	4.86	6.1
Pro	1.53	2
Al	2.41	3.1
Gly	3.84	5
Leu	4.55	6.0
Phe	1.56	2
Lys	7.74	10
Arg	0.72	1
Meth	0.765	1
Tyr		
Try	Destroyed by acid hydrolysis	

Yield = 10%

P2 (coloured)

Table 10

DNP Amino Acid	Total μmoles from this Peptide in the Fraction	Molar Ratios
Asp	5.39	3
Glu	10.23	6
Thr	3.6	2
Pro	1.57	1
Gly	13.1	7.25
Ala	1.725	1
Val	4.35	2.5
Leu + ILeu	5.55	3
Phe	2.85	1.6
Hist	3.52	2
Lys	12.2	6.77
Cys	3.6	2
Arg	1.1	0.61

Yield = 20%

Fraction (3) See tables 11, 12, 13, 14.

PI (colourless) Amino Acid analysis on DNP Derivatives

Table 11

DNP Amino Acid	Total µmoles from this Peptide in Fraction	Molar Ratios
Asp	8.086	1.6
Glu	5.654	0.77
Thr	6.08	1.37
Pro	5.11	1.11
Al	5.9	1.27
Gly	8.33	2.01
Phe	5.02	0.72
Lys	8.95	1.97
Tyr		

Yield = 44%

P2 (colourless) Amino Acids Estimated According to Heilmann et al.

Table 12

Amino Acid	Total μmoles in Fraction from this Peptide	Molar Ratios
Asp	2.9	1.7
Glu	6.7	3.93
Gly	not estimated by this technique	
Thr	1.93	1.13
Meth	0.885	0.515
Lys	5.45	3.19
Leu	3.92	2.3
Pro	not estimated by this technique	
Tyr	0.716	0.42

Yield = 18%

P3 (colourless) Amino Acids Estimated by the Method of Heilmann

Table 13

Amino Acid	Total µmoles in Fraction from this Peptide	Molar Ratios
Glu	2.42	2.19
Gly	not estimated by this technique	
Thr	2.25	2
Al	0.92	0.84
Lys	3.46	3.15
Arg	0.359	0.325
Phe	1.62	1.46
Tyr	0.194	0.175
Leu	2.81	2.56
Meth	1.26	1
Pro	not estimated by this technique	

Yield = 12%

P4 (colourless) Amino Acids Estimated by the Method of Heilmann et al.

Table 14

Amino Acid	Total μmoles in Fraction from this Peptide	Molar Ratios
Asp	10.9	2.2
Glu	11.9	2.4
Thr	4.45	0.9
Al	9.4	1.9
Lys	15.9	3.2
Tyr	2.68	0.54
Leu + ILeu	14.8	3

Yield = 80%

The N-terminal end groups of the peptides were determined using the technique described in the "methods" section, and the results are shown in the table, (15).

The ϵ DNP lysine recoveries show the extent to which the TFA groups were removed from the various sections of the protein.

Table 15

Peptide	N-Terminal Amino Acid	As a %age of the total Amino Acid of its kind in the Peptide	E DNP Lysine as a %age of the total Lysine in the Peptide
Fraction I PI	None	Detected	None Detected
Fraction I P2	None	Detected	None Detected
Fraction 2 PI	Lysine	7.0	52.0
Fraction 2 P2	None	Detected	8.0
Fraction 3 PI	Lysine	42.2	None
Fraction 3 P2	Aspartic Acid	42.0	74.0
Fraction 3 P3	Tyrosine	790.0	85.4

2nd Approach to the Trifluoroacetylation of Cytochrome c, its Hydrolysis with Trypsin, Followed by Investigation of The Peptides.

In this experiment, a preparative attempt was made to isolate the large peptides of cytochrome c, produced by hydrolysis at its arginine residues.

In the previous experiment, the yields of some of these peptides were low, probably because of hydrolysis at centres other than arginine. On this occasion, therefore, less trypsin was used in an attempt to minimise this.

250 mg 100% commercial cytochrome c was subjected to the blocking technique and then hydrolysed by trypsin as described in the previous experiment (1 mg trypsin was used for the hydrolysis).

After boiling to deactivate the trypsin, the preliminary fractionation by pH adjustment resulted in only two fractions. The pH's of these fractions were then raised to 11 with 0.25 M NaoH for six hours at 37°C in order to deblock them. It was discovered, however, on reduction of the pH of the red precipitated fraction that a red precipitate developed at pH 6 leaving behind a deep red coloured solution. This solution plus the washings from the precipitate was labelled FRACTION 1, whilst the precipitate suspended in water and the original supernatant were labelled FRACTION 2 and FRACTION 3 respectively. It was noticed of Fraction 3, when returning its pH to neutral with 0.1 M HCl, that

its colour in acid solution was yellow, whilst at alkaline pH's it was pink (see Fig. 23).

The Total Nitrogen Content of the Fractions

The total nitrogen content of Fractions (1) and (2) was determined as is described in the "methods" section. The value for Fraction 2, which was found to be very insoluble, and showed a marked tendency to stick to the flask in which it was stored, was calculated by difference.

Results

Table 16

Fraction	Total Nitrogen	% Cytochrome c Nitrogen
1	7.9 mg	21
2	-.	58.5
3	7.64 mg	20.5

Comparison of the Extinction Coefficients of Fractions 1 and 2 and Cytochrome c.

Samples of Fraction 1 and cytochrome c containing 0.265 mg and 0.18 mg nitrogen respectively, were dissolved in 4 ml 0.1M phosphate buffer pH 7 and the iron in each reduced with a little sodium dithionate. The optical densities of these solutions were read at 550 mμ.

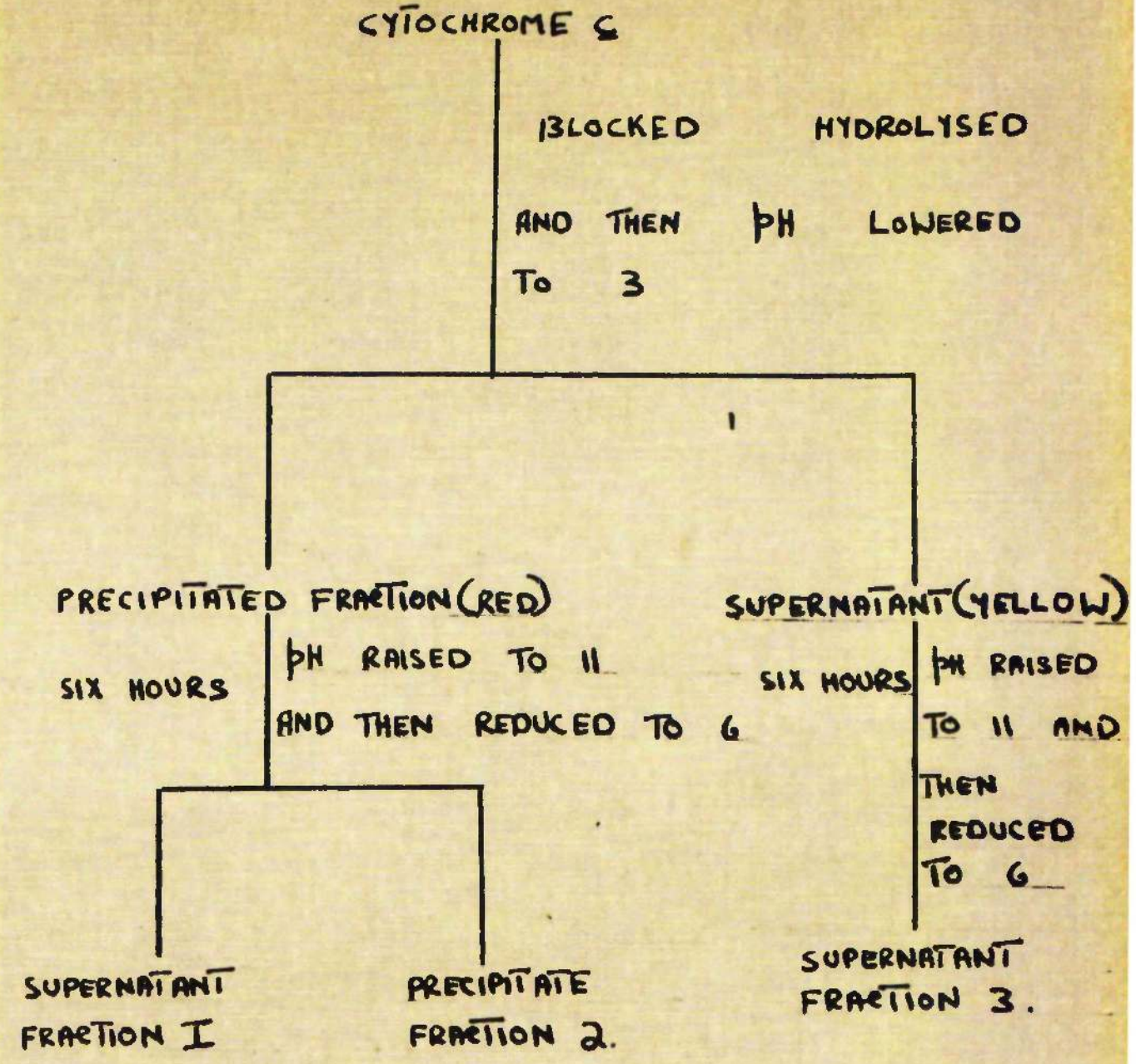

CYTOCHROME C
BLOCKED
HYDROLYSED
AND THEN pH LOWERED TO 3
PRECIPITATED FRACTION (RED)
SUPERNATANT (YELLOW)
SIX HOURS
pH RAISED TO 11 AND THEN REDUCED TO 6
SIX HOURS
pH RAISED TO 11 AND THEN REDUCED TO 6
SUPERNATANT FRACTION I
PRECIPITATE FRACTION 2.
SUPERNATANT FRACTION 3.

FIGURE 23

An attempt made to deal similarly with a sample of Fraction 2 (containing 2 mg N) was only partially successful as a considerable amount of this material did not enter into solution.

In each case, the extinction coeficient was calculated from the formula

Optical density = Extinction Coeficient x concentration

In this case, the concentration was in terms of 1 mg nitrogen.

Results

Table 17

Fraction	Extinction Coeficient
Cytochrome c	4.17
Fraction 1	0.457
Fraction 2	1.0

Peptide Maps of Fractions 1, 2 and 3 as an Indication of Their Composition

These were done according to the technique described in the "methods" section (see Fig. 23 a). The peptide maps were developed using some of the methods described in the "methods" section. The results are stated below.

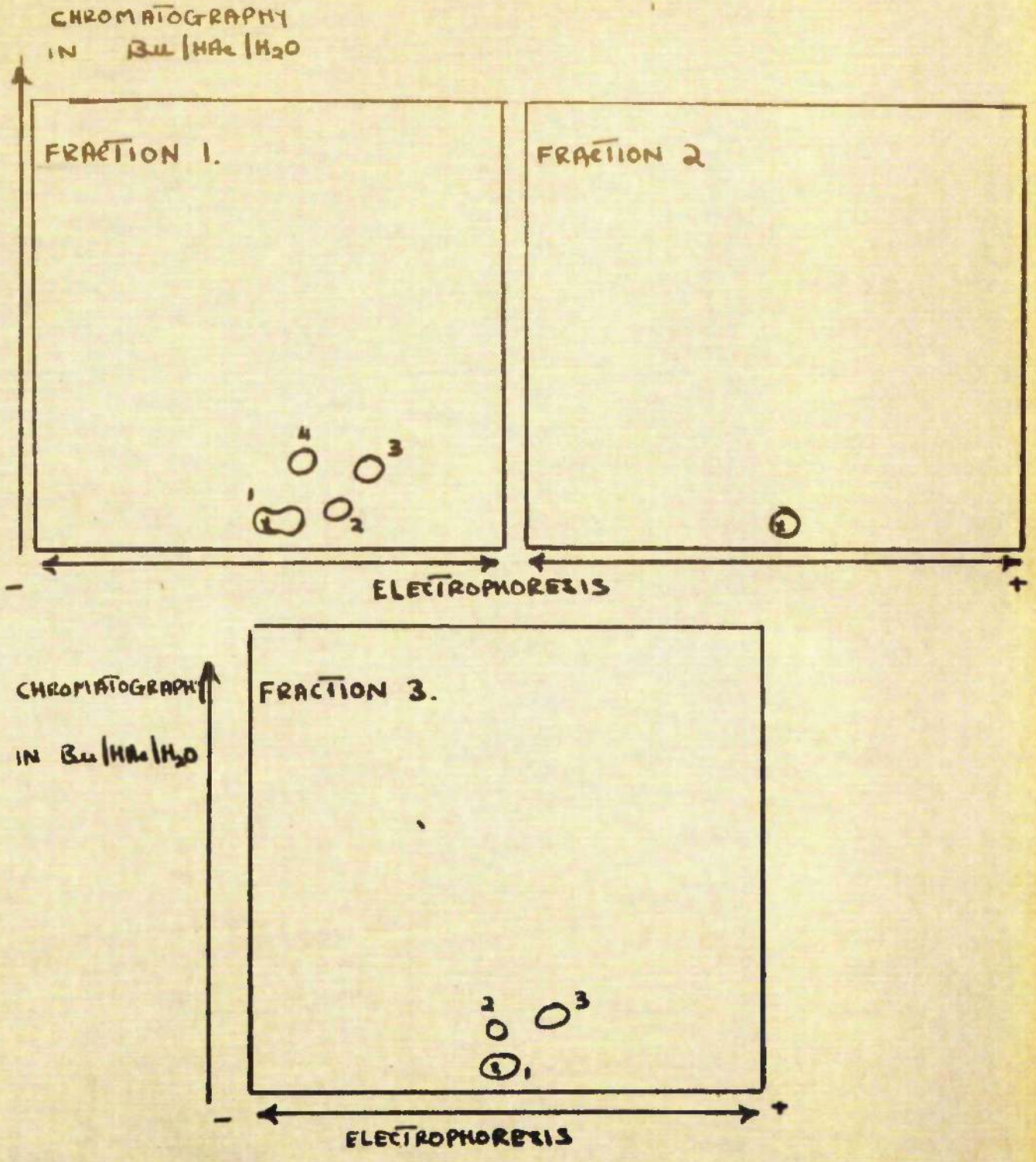

CHROMATOGRAPHY
IN BuOH/HAc/H2O
FRACTION 1.
FRACTION 2
4
3
1
2
1
-
ELECTROPHORESIS
+
CHROMATOGRAPHY
IN BuOH/HAc/H2O
FRACTION 3.
2
3
1
-
+
ELECTROPHORESIS

FIGURE 23(a)

FRACTION 1		Peptide 1, coloured
Ninhydrin	-	all peptides reacted positive
Sakaguchi	-	peptides 1 and 2 slight positive
Erlich's	-	all negative
FRACTION 2		Only one peptide, and it was coloured
Ninhydrin	-	positive
Erlich's	-	negative
Sakaguchi	-	negative, but the natural red colour of the peptide could have been obscured by a slight positive result.
FRACTION 3		Peptide 1, very slightly coloured
Ninhydrin	-	all positive
Sakaguchi	-	all negative
Erlich	-	all negative

Investigation of Peptides Separated from Fractions 1, 2 and 3

Fraction 1

One-tenth (779 µg N) of the total material in the fraction was applied to a peptide map and run under conditions identical to those already described. Then, using the previous map for Fraction 1 as a guide, peptide 1 was dissected and eluted with 0.05 M NH_3. This solution did not remove all the colour from the paper, however, and so the molarity was increased to 0.2 . This also failed to remove all the adsorbed material, and so that

which had been removed was dried under vacuum and then acid hydrolysed as previously described.

The amino acid content of the hydrolysate was estimated on the automatic amino acid analyser discussed in the "methods" section. The results are as shown.

Fraction 3 was dealt with in an identical fashion (760 µg N were applied to the peptide map.)

Fraction 2, which the peptide map technique had suggested was relatively pure, was hydrolysed directly (1/40th i.e. material containing 400 µg N) and the amino acid analysis obtained from the automatic analyser.

Results See Tables 18, 19 and 20

The absence of values for Pro, Arg and Cys in the above analysis, was due to inadequacies in the automatic analyser, which at this time was at an early stage in its development.

Table 18

FRACTION I, PI

Amino Acid	µMoles	Molar Ratios
Asp	0.488	2.72
Thr	0.710	3.96
Glu	1.11	6.12
Pro		
Gly	1.295	7.22
Al	0.6025	3.35
Val	0.1795	1
Meth	0.0656	0.387
ILeu	0.327	1.82
Leu	0.347	1.94
Tyr	0.0996	0.55
Phe	0.1965	1.095
Lys	1.15	6.5
Hist	0.1495	0.835
Arg		
Cys		

Table 19

FRACTION 2

Amino Acid	µMoles	Molar Ratios
Asp	0.471	6.02
Thr	0.489	6.02
Glu	0.854	10.8
Gly	0.966	12.2
Al	0.392	5
Val	0.283	3.6
Pro		
Meth		
ILeu	0.461	5.8
Leu	0.422	5.4
Tyr	0.1565	2
Phe	0.213	2.72
Lys	1.382	17.6
Hist	0.285	3.6
Arg		
Cys		

Table 20

FRACTION 3, P1

Amino Acid	µMoles	Molar Ratios
Asp	1.782	6.94
Thr	1.556	6.14
Glu	1.124	4.37
Pro		
Gly	1.655	6.44
Al	1.31	5.1
Val		0.484
Leu	0.435	1.68
ILeu	0.146	0.569
Tyr	0.257	1
Phe	0.565	2.2
Lys	1.910	7.44
Hist	0.15	0.584

Removal of the Trifluoroacetyl Groups from Fractions 1, 2 and 3

A check was made on the material in Fraction 2, using FDNB (see 'earlier' section), to ascertain whether or not incubation at pH 11 for six hours had completely removed the TFA groups. The result of this determination showed that some still remained and these were removed in the manner described, pp 92 Fractions 1 and 3 were then treated similarly.

Results

Fraction 1	Time	Optical Density at 570 mμ
	Zero	0.82
	15 min.	0.82
	30 min.	0.855
	45 min.	0.855
Fraction 3		
	Zero	0.898
	15 min.	0.94
	30 min.	0.94
	45 min.	0.94
	Blank	0.581

Despite the high blank, it can be seen with both fractions that there is an increase in the intensity of colour produced with ninhydrin. This indicates that the desubstitution was not quite complete before the above treatment, whilst the constant colour intensity ultimately obtained, suggests that desubstitution was achieved.

Pilot Run Purifications of Fractions 1, 2 and 3 on G25 Sephadex

G25 sephadex (medium grade) was equilibrated with 0.05 M phosphate buffer pH 9.9 and poured to form a column 58 cm long x 1.3 cm diameter (void volume 12-7 cc, determined using rat haemoglobin).

Attempts were made to purify samples of Fractions 1, (1/6th i.e. 1,500 µg N), 2, (1/9th i.e. 1,777 µg N) and 3, (1/6th i.e. 1,260 µg N) on this column. Each sample in its turn was applied to 2 ml solution pH 9.9 and the column was eluted with 0.05 M phosphate buffer pH 9.9. The positions of the peptides in the fractions collected were determined by taking aliquots, subjecting these to alkaline hydrolysis and then estimating the amino acids in the resulting hydrolysates with Yemm and Cockings ninhydrin solution (see "methods " section).

The results for all three fractions are shown in the graphs (Figs. 24,25 & 26)

As can be seen, the coloured peptide from Fraction 1 came through after 14 cc, the coloured peptide(s) comprising Fraction 2, after 16.5 cc and the main peptide in 3, after 15 cc. A sample of cytochrome c run through the same column emerged after 16 cc. Now, as the void volume of the column = 12.7 cc, all fractions and cytochrome c have been retained to some extent by the G25 sephadex. In the case of cytochrome c this is unexpected and must have been due to charge effects. With the peptides in the fractions,

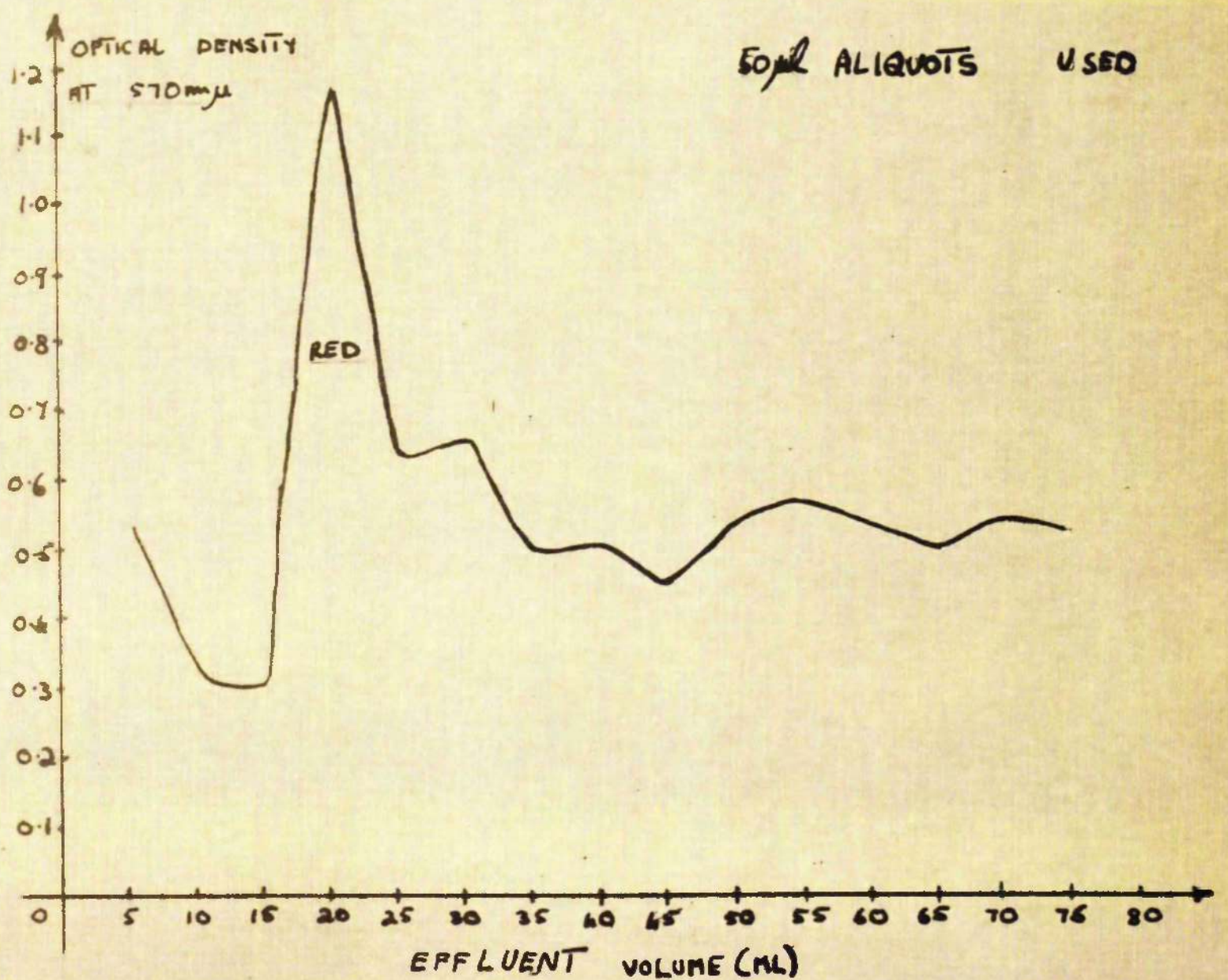


FIGURE 24 (FRACTION 1.)

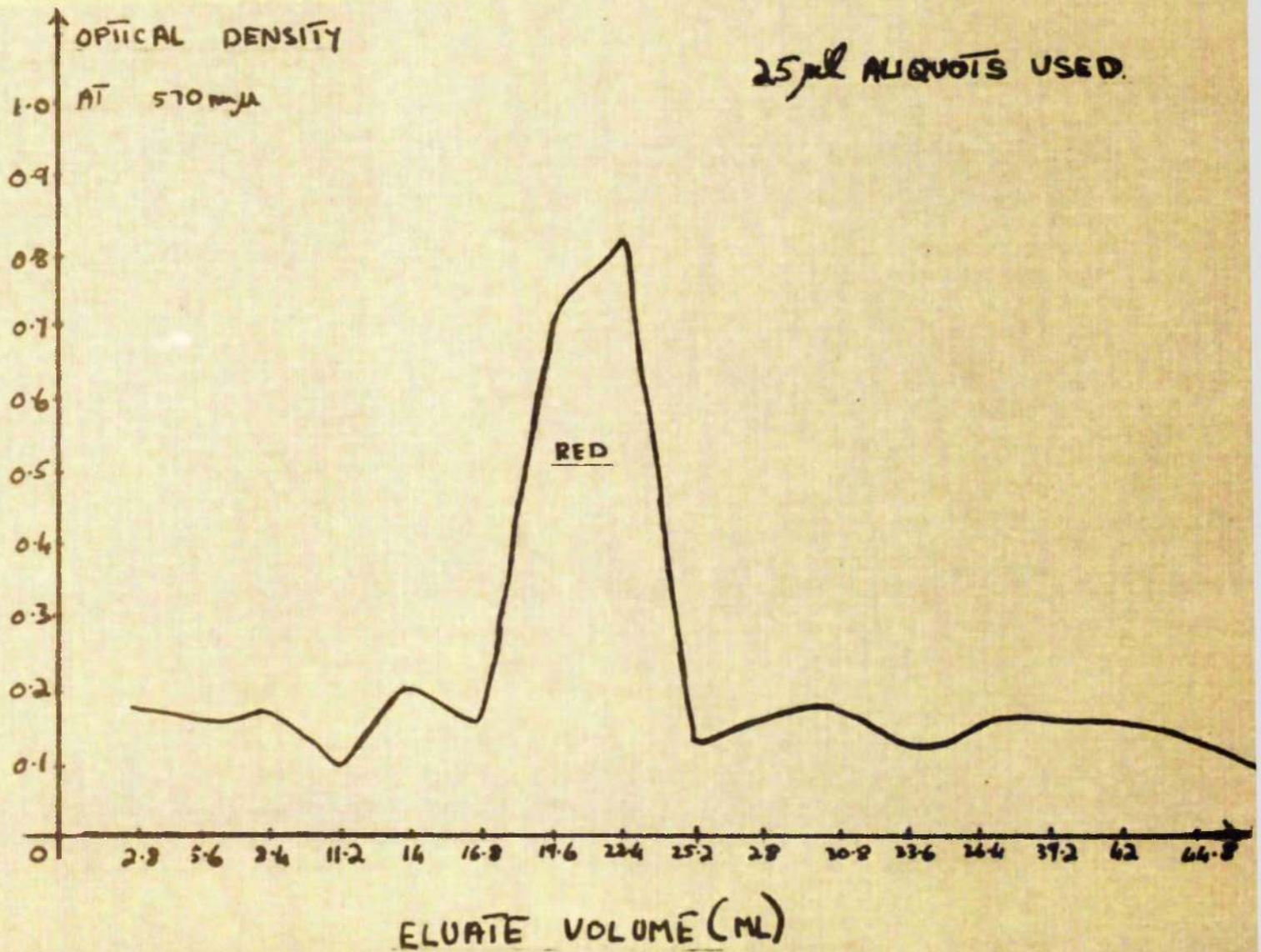


FIGURE 25 (FRACTION 2)

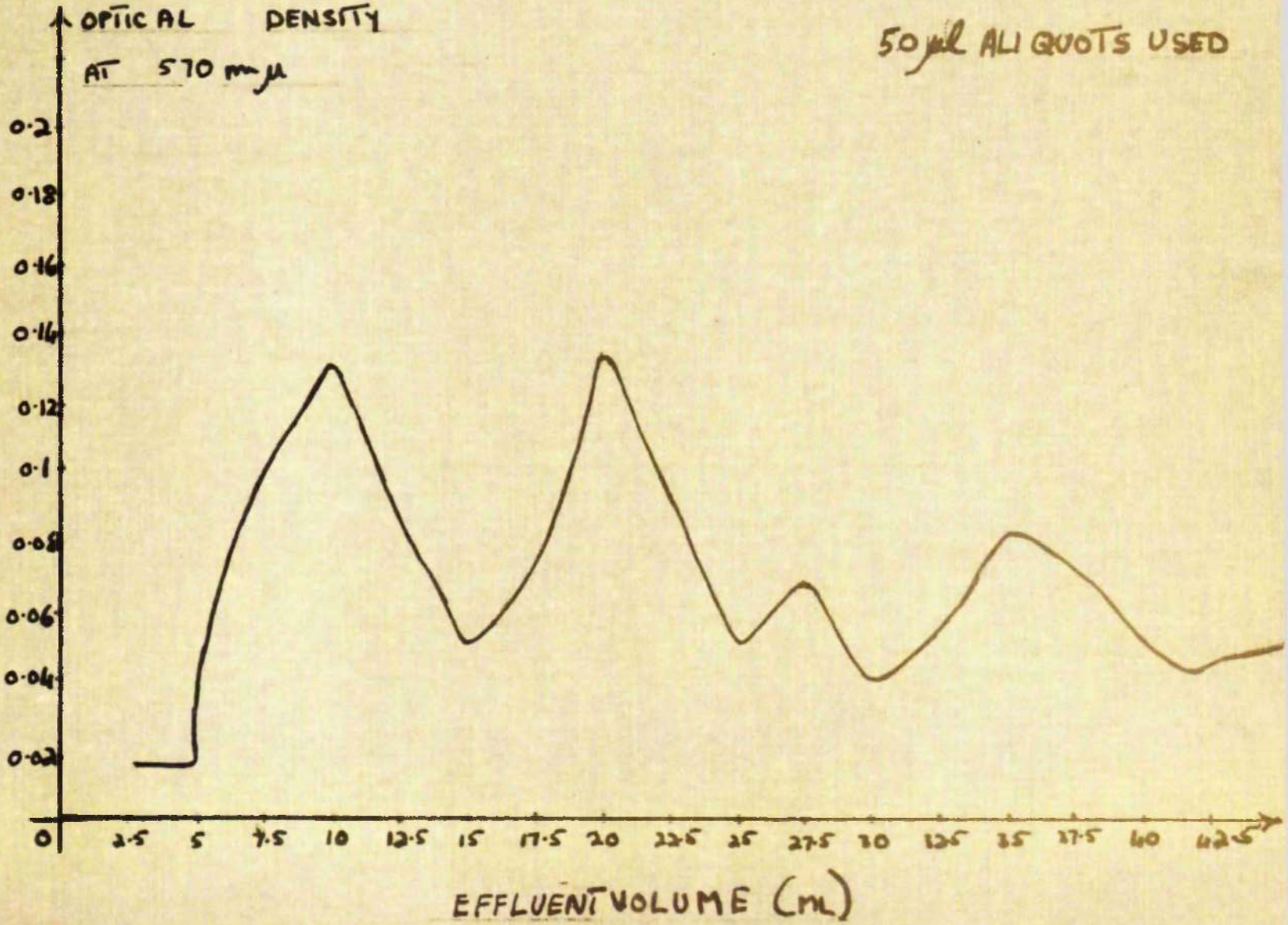


FIGURE 26 (FRACTION 3.)

therefore, it is difficult to draw conclusions concerning their size, as the charge effects operating between each and the sephadex are unknown.

Full Scale Purifications of the Fractions

These were done in a manner almost exactly similar to that used in the pilot runs, except that a fine grade of G25 sephadex was used, in a column 82 cm long x 2 cm in diameter (void volume 72.8 cc). Also, the material from each fraction was applied in its turn in 3 cc volume and in the Fraction 1 run, samples were collected of size 4.6 ml, in the Fraction 2 run, of size 5.2 ml and in the Fraction 3 run, of size 3.2 ml.

The results for all three runs are shown in the graphs (Figs. 27,28 & 29).

The samples comprising the peaks shown on the graphs were bulked and then lyophilised.

Peptide Maps of the Peaks

These were done in the usual way in order to investigate the homogeneity of the material contained in the bulked samples. 1% of the total material was used in each case (see Figs. 30,31 & 32)

Total Nitrogen Content of the Peaks

These were estimated as is described in the "methods " section.

Results see Table 21

The low yield of nitrogen in Fractions 1 and 2 indicates that some material must have been lost during the peptide

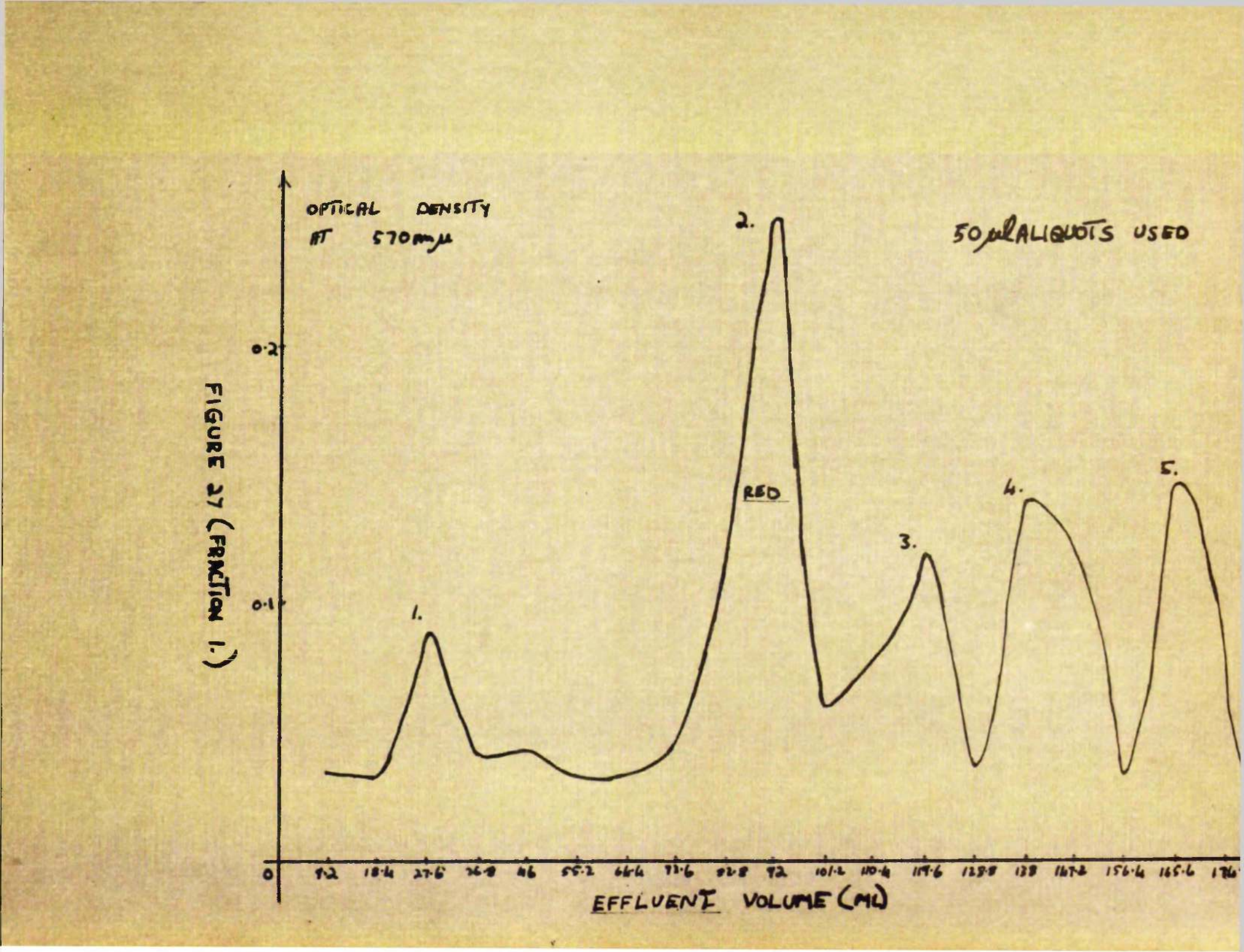


FIGURE 27 (FRACTION 1.)

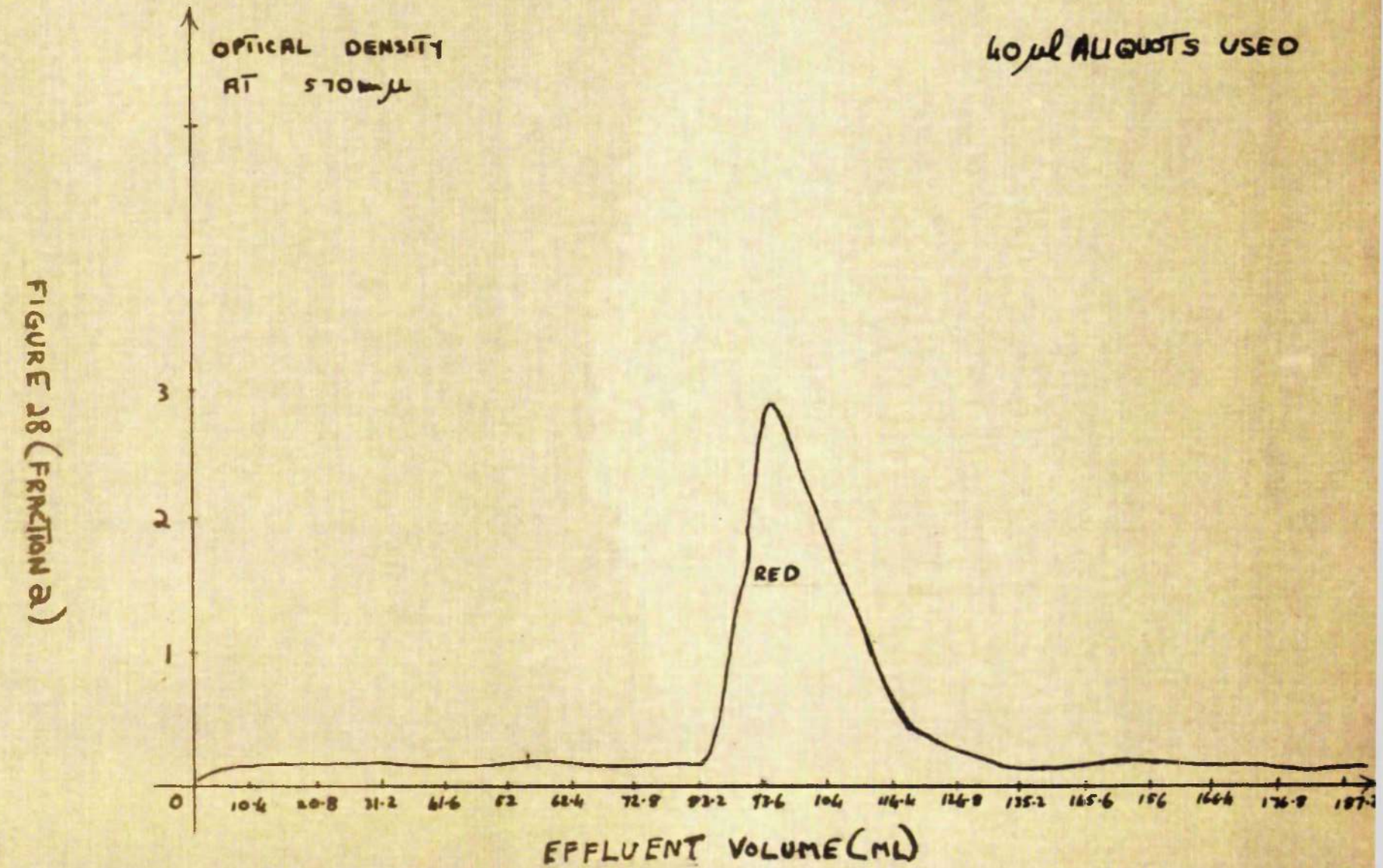
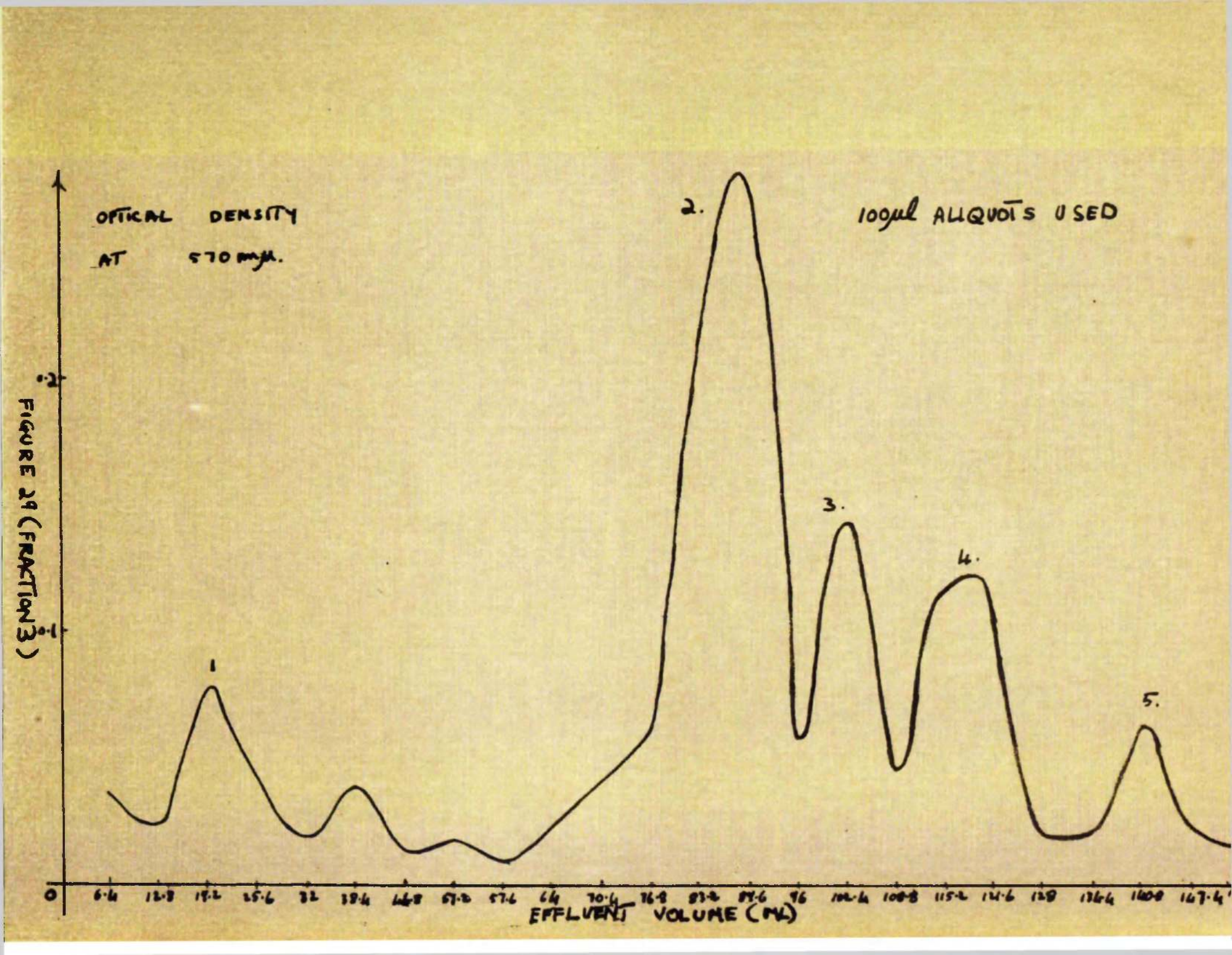
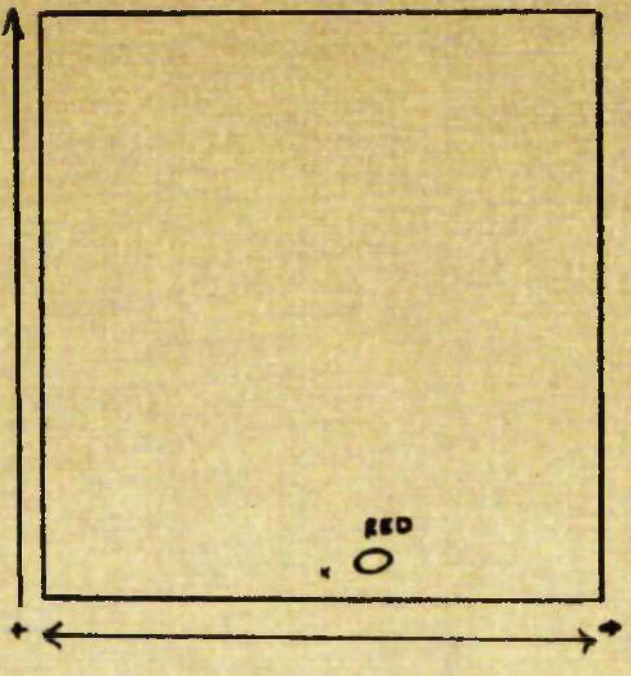


FIGURE 28 (FRACTION 2)

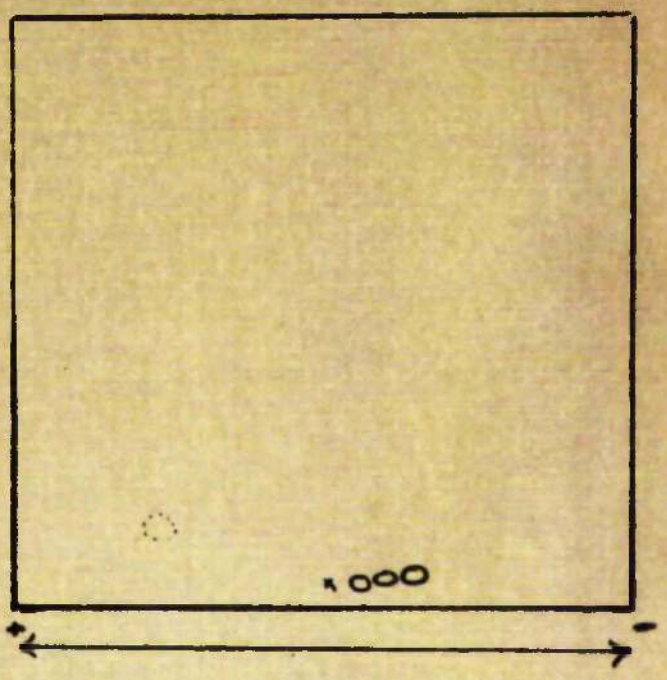

OPTICAL DENSITY
AT 570 mμ.
100μl ALIQUOTS USED
FIGURE 29 (FRACTION 3)
·2
·1
0
1
2.
3.
4.
5.
6·4
12·8
19·2
25·6
32
38·4
44·8
51·2
57·6
64
70·4
76·8
83·2
89·6
96
102·4
108·8
115·2
121·6
128
134·4
140·8
147·4
EFFLUENT VOLUME (ML)

PEAK 1. NOTHING SHOWN

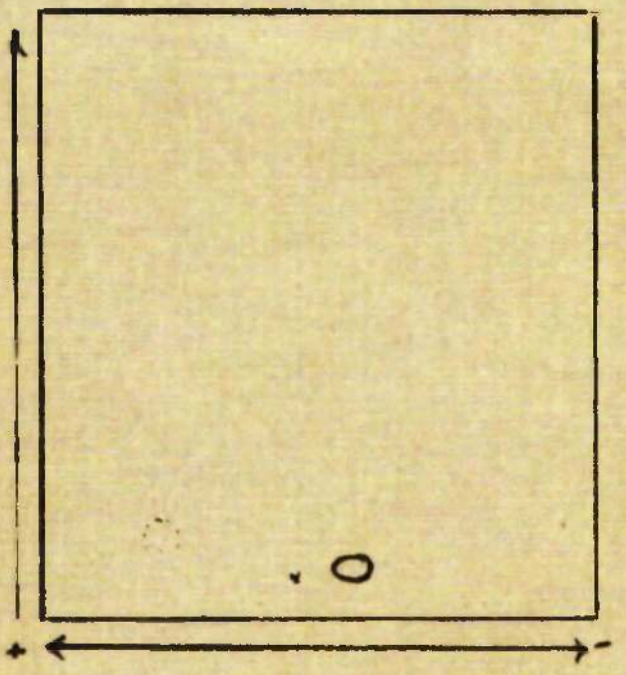
PEAK 2.



PEAK 3.



PEAK 4.



PEAK 5.

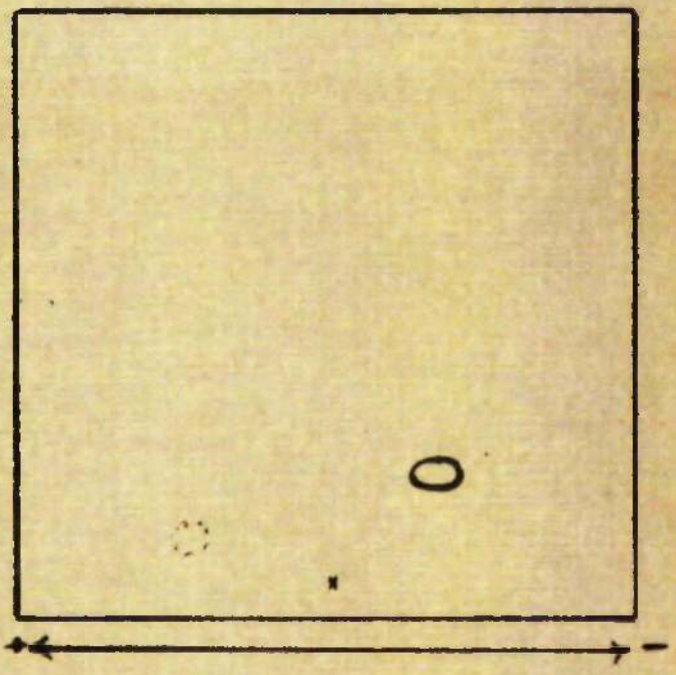


FIGURE 30 (FRACTION 1.)

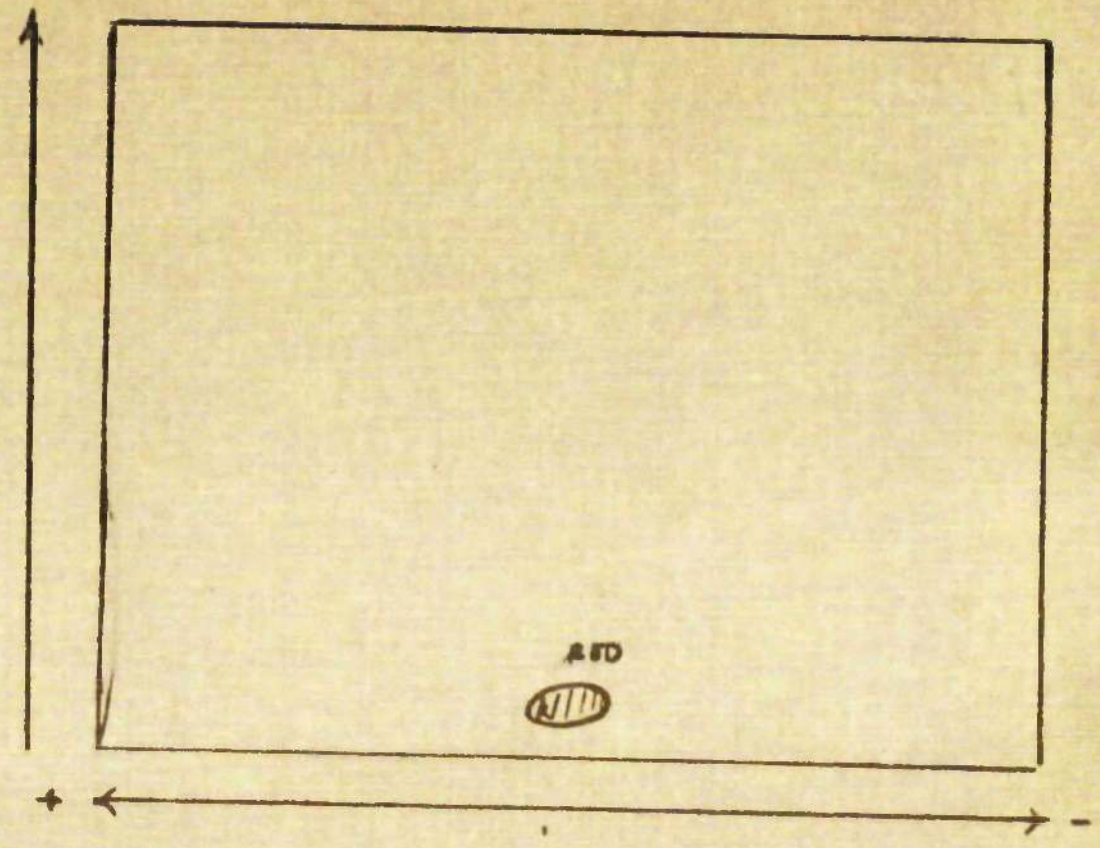
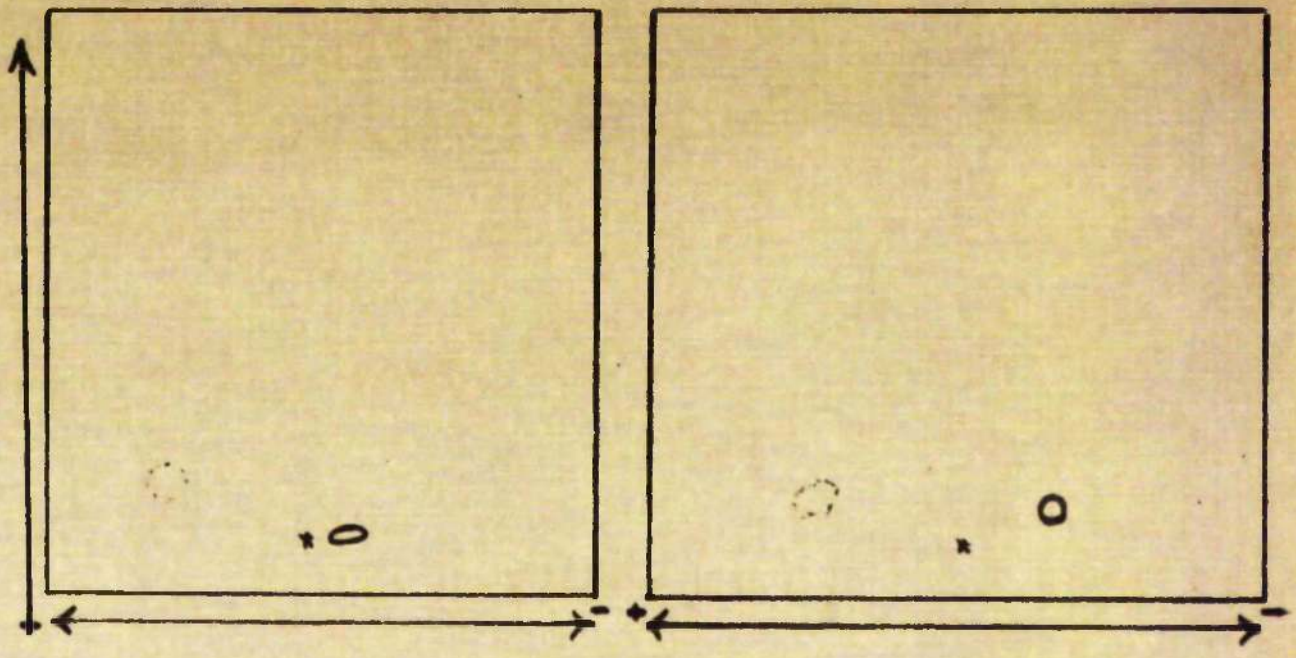


FIGURE 31. (FRACTION 2.)

PEAK 1. NOTHING SHOWN

PEAK 2.

PEAK 3



PEAK 4.

PEAK 5.

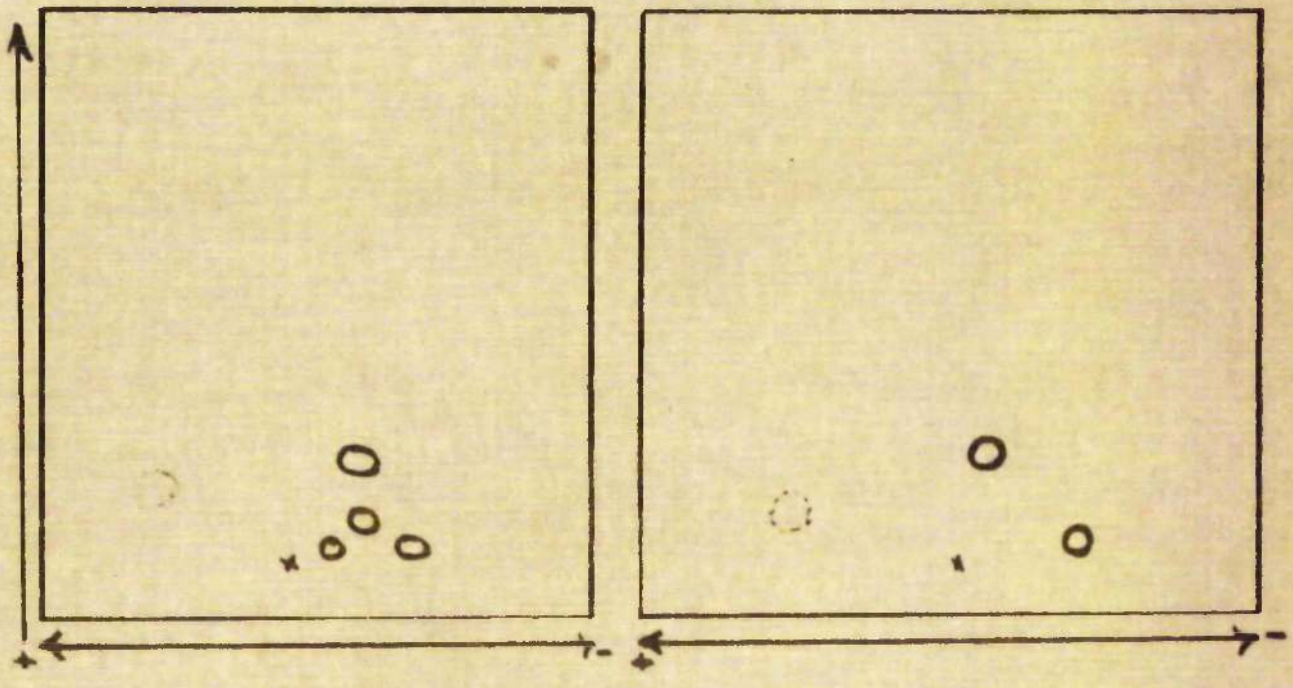


FIGURE 32 (FRACTION 3.)

fractionation. However, Fr1, P2, Fr2 and Fr P2 were all N-rich and peptide maps indicated only one component in each.

It was concluded, therefore, that these fractions probably contained the large peptides expected from the foregoing treatment.

Investigation of the Main Peaks from the Preceding Gel Filtration

(a) Fraction 3, Peak 2

Amino Acid Analysis

One-tenth of the total material was hydrolysed by 6N HCl overnight at 114°C and the amino acid content of the hydrolysate determined on the automatic analyser . See Table 22.

Table 21

Peak	Nitrogen Content (mgs)	% age N applied to the Column
FrI, PI	0.028	0.5
(coloured) FrI, P2	1.670	31.0
FrI, P3	0.394	7.3
FrI, P4	0.0135	0.25
FrI, P5	0.00994	0.18
(coloured) Fr2	9.3	72.0
Fr3, PI	0	0
Fr3, P2	1.485	38
Fr3, P3	0.099	1.84
Fr3, P4	0.206	3.8
Fr3, P5	0	0

Table 22

Amino Acid	µmoles	Molar Ratios
Asp	0.501	3.84
Thr	0.351	2.7
Glu	0.261	2.01
Pro	0.204	1.57
Gly	0.481	3.7
Al	0.428	3.29
Val	0.1063	0.82
ILeu	0.0331	0.255
Leu	0.0635	0.49
Tyr	0.1545	1.19
Phe	0.1602	1.23
Lys	0.5400	4.16
Hist	0.0209	0.161

Fraction 2

Approximately one-fifth of a μmole was acid hydrolysed and the amino acid content of the hydrolysate estimated. A similar amount of cytochrome c was similarly treated for purposes of comparison. See Table23

As the results of the amino acid analysis are rather inconclusive, it was decided to compare Fraction 2 with cytochrome c another way. To do this, they were both hydrolysed with trypsin and chymotrypsin and peptide maps were done of the hydrolysates. See Fig. 33

Experimental

21.4 mg of cytochrome c were dissolved in 3 ml 0.05 M phosphate buffer pH 7.8 and 50 μls (1% w/v) trypsin solution added. To this 0.5 mg chymotrypsin was added and the hydrolysis was carried out at 37°C for 24 hours.

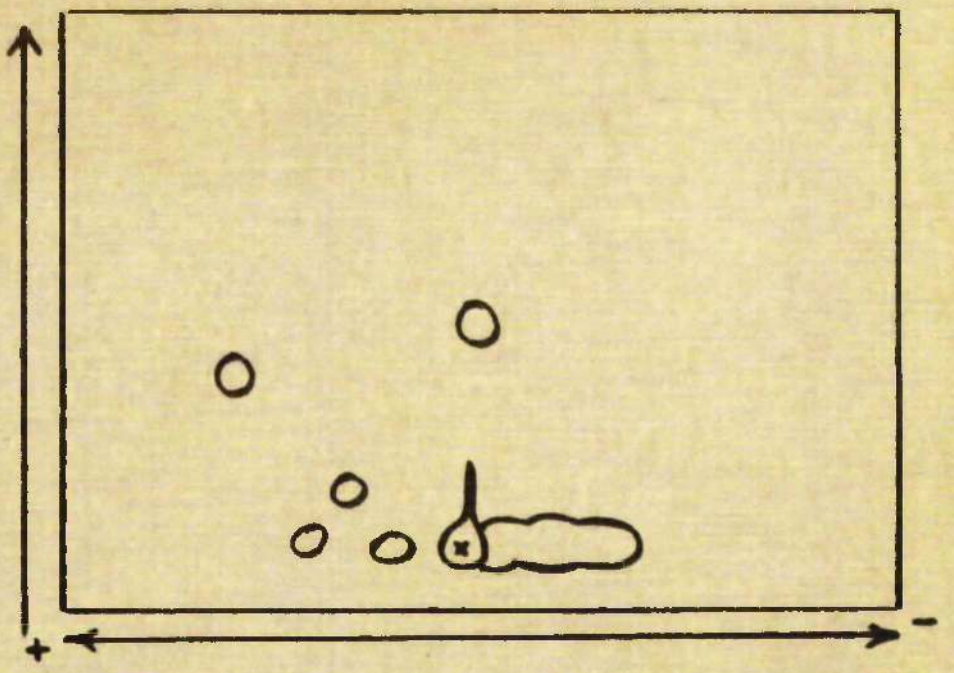
Fraction 2 was dealt with similarly.

The peptide maps were done in the usual way.

Fraction 1, P2

One-tenth of the total material was acid hydrolysed and the amino acids estimated as usual. See Table 24.

FRACTION 2.



CYTOCHROME C

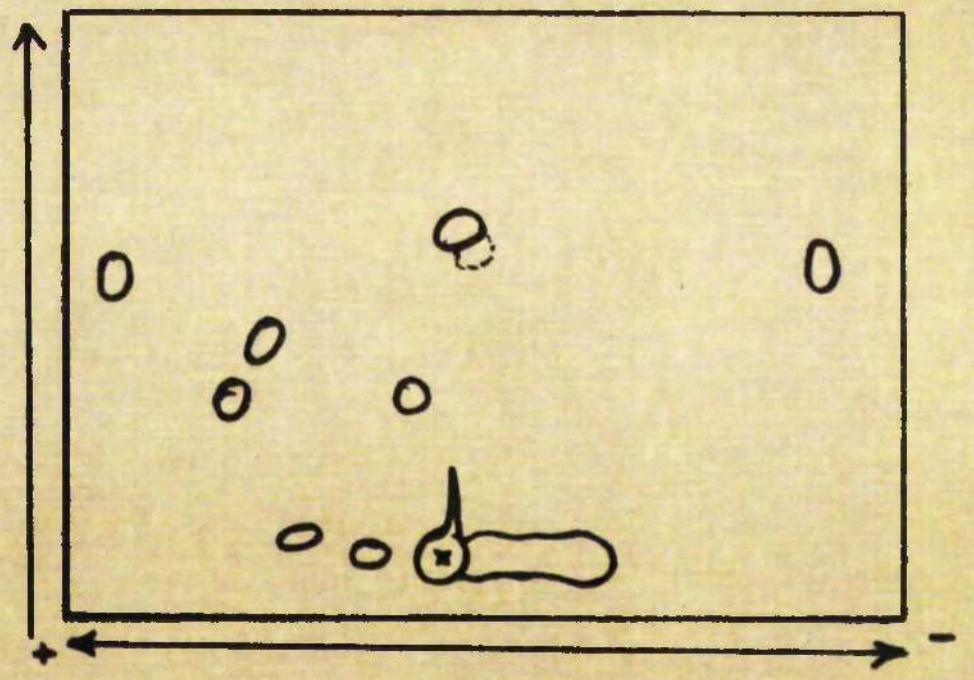


FIGURE 33

Table 23

Amino Acid	Molar Ratios Fr2	Molar Ratios Cytochrome c
Asp	4.95	8.12
Thr	6.06	8.18
Glu	11.9	12.46
Pro	4.25	6.25
Gly	11	11.12
Al	6.7	5.76
Val	3.77	5
ILeu	5	6.5
Leu	4	6.5
Tyr	2.82	3.0
Phe	4.05	5.56
Lys	16.8	17.6
Hist	3.3	3
Arg	1.92	2
Cys	0.64	0.2

Table 24

Amino Acid	µmoles	Molar Ratios
Cys	0.119	1
Asp	0.249	2.09
Thr	0.247	2.07
Glu	0.516	4.37
Pro	0.185	1.58
Gly	0.549	4.6
Ala	0.231	1.94
Val	0.229	1.92
ILeu	0.211	1.77
Leu	0.212	1.78
Tyr	0.0427	0.359
Phe	0.157	1.32
Lys	0.856	7.2
NH_3	0.672	
Hist	0.149	1.91
Arg	0.0427	0.85

Chymotryptic Hydrolysis of Fraction 3, P2 and Separation of the Resultant Peptides

All of the material in the fraction was hydrolysed with chymotrypsin for five hours at room temperature in the buffer solution in which it came off the sephadex column. The pH of this solution was 8.8. At the end of this time a sample was taken (1 /10th) and a peptide map done in the usual way. See Fig 34 . The remainder of the material was then lyophilised and taken up in 2 ml 25% acetic acid and applied to the ion exchange column described below. (Guidotti, Hill and Konigsberg, 1962). The column (50 cm long, 0.9 cm diameter) was composed of Zeocarb 225 x 2, which had been washed with 2N NaoH, 2N HCl, H_2O and then the starting buffer until equilibrium had been established.

The buffers used for elution were:

(a) 0.17 M pyridine acetic acid pH 4.7

(b) 0.4 M pyridine acetic acid pH 4.6

(c) 1.07 M pyridine acetic acid pH 5.4

(d) 2.25 M pyridine acetic acid pH 5.4

These buffers were introduced successively as is shown on the graph, making use of a gradient to effect the changes. The volume of the mixing chamber was 250 cc.

The effluent from the column was divided into 2.5 ml fractions, and 0.2 ml aliquots of these were analysed by alkaline hydrolysis and ninhydrin determinations as previously described. See graph (Fig. 35).

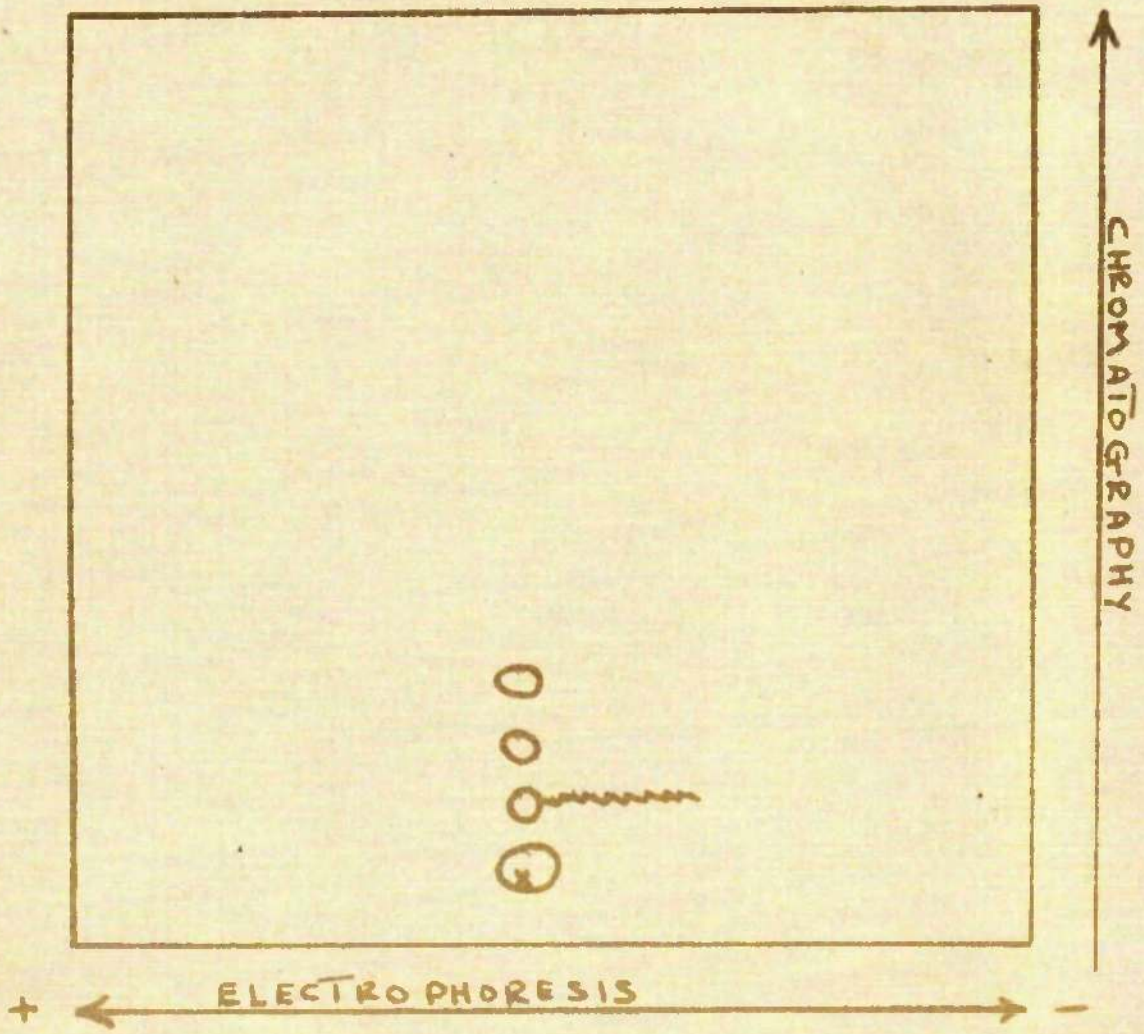


FIGURE 34

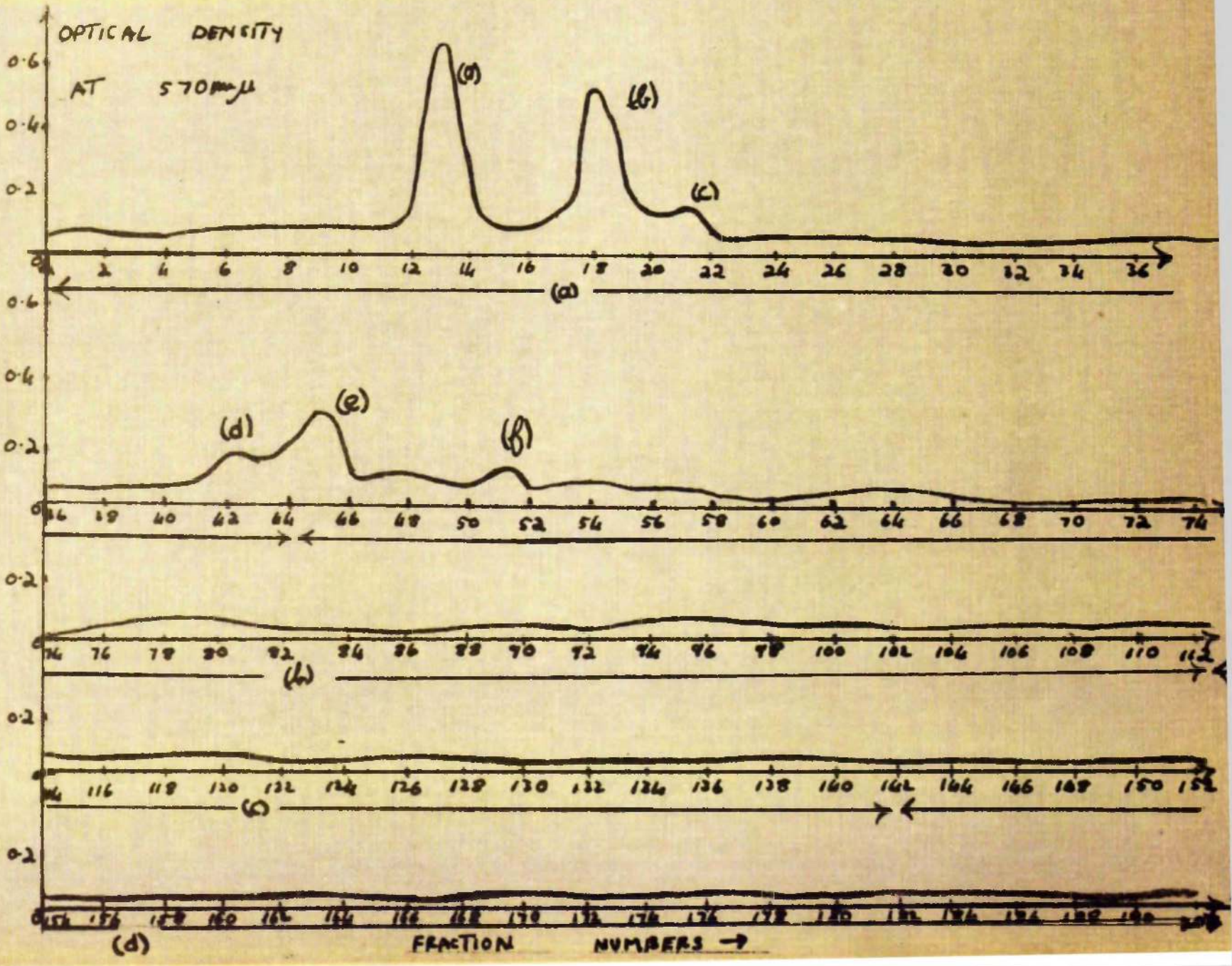


FIGURE 35

The fractions comprising the peaks shown on the graph were bulked:-

(a) 12, 13, 14, 15

(b) 17, 18, 19

(c) 22, 22

(d) 41, 42

(e) 44, 45, 46

(f) 50, 51, 52

These were then lyophilised, acid hydrolysed and the amino acids in the hydrolysates estimated as usual. See Table 25

Apart from (a), all these peptides are identifiable with fragments of P39-55 (Tuppy 1961). See discussion. (a) is probably a mixture, therefore, but this is not surprising as it was first to elute from the column. The occurrence of valine in (a) indicates that fragments of the impurity in Fr3, P2 (discussed above) are present. No other fragments of this were identified elsewhere.

Table 25

Amino Acid	Fr3, P2 CH (a) Residues	Fr3, P2 CH (b)	Fr3, P2 CH (c)	Fr3, P2 CH (d)	Fr3, P2 CH (e)	Fr3, P2 CH (f)
Asp	4.39	1.9				2.99
Thr	1.656	0.9	1	1.36	1.64	2.4
Glu	3.8		1.3	1.36	1	1
Pro	0.4		1.2	1.5	1.05	
Gly	1.5		2.1	1.8	2.12	1
Al	2.7	1.2	1.1	1	0.975	1.5
Val	1.6					
ILeu	0.8					
Leu	0.9					
Tyr	0.1				0.725	
Phe	0.3		1.1	1	1.07	
NH_3	7.4					
Lys	0.5	1		0.86		1.5
% yield	11	25.0	6.8	3.0	13.0	2.0

Cytochrome c, its Trifluoroacetylation and Hydrolysis by Trypsin, Followed by Investigation of the Peptides - Experiment 3.

Two μmoles cytochrome c were blocked with $CF_2.COS\ Et$ at pH 9.0 as previously described. The reaction mixture was extracted with chloroform, and then along with 2 mg trypsin, applied to a column (60 cm long x 1.2 cm diameter) of G50 sephadex equilibrated with 0.05 M NH_4AC pH 7.5.

The reason for performing the hydrolysis this way was because it was hoped that cytochrome c and trypsin would progress down the column together at the same rate, but as peptides were produced, they would be retained and thus separated from the trypsin. The flow rate of the column was controlled so that the hydrolysis lasted approximately 3 hours, and the eluate was collected in 5 ml fractions. These were analysed for protein material by reading their optical densities at 550 mμ and 280 mμ (see graph. Fig. 36). The results indicated a partial separation of trypsin.

Fractionation of the Tryptic Hydrolysate

Fractions 14-22 from the above sephadex treatment were bulked, their volume concentrated to 4 ml and applied to a G25 sephadex column (100 cm long x 1.8 cm diameter). The column was eluted with 0.05 M NH_4 AC pH 8.4 and 5 ml fractions were collected. See Fig. 37

Fractions 18-23 from this run were bulked, and the haem

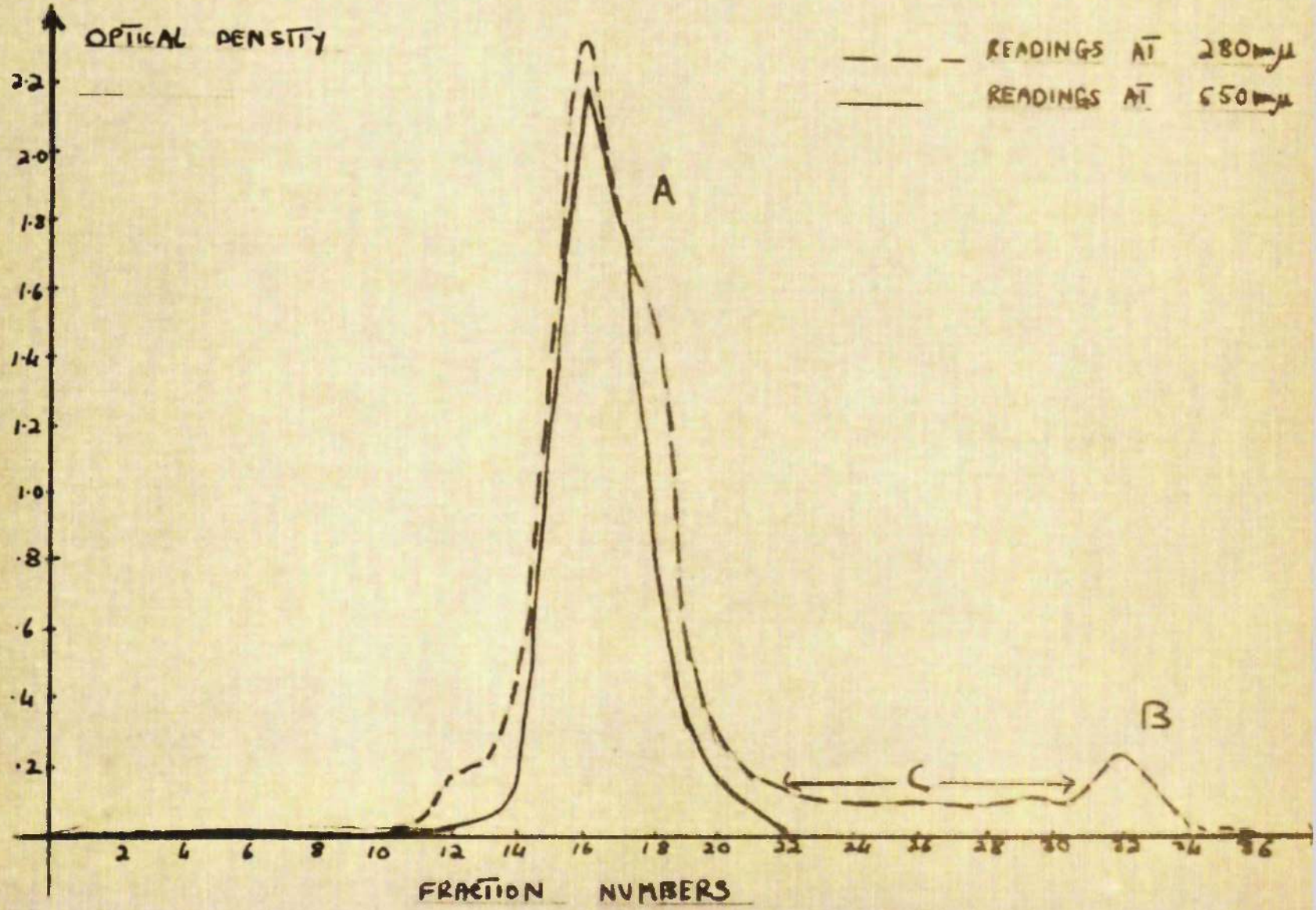

OPTICAL DENSITY
READINGS AT 280mμ
READINGS AT 650mμ
2·2
2·0
1·8
1·6
1·4
1·2
1·0
·8
·6
·4
·2
A
B
C
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
FRACTION NUMBERS

FIGURE 36

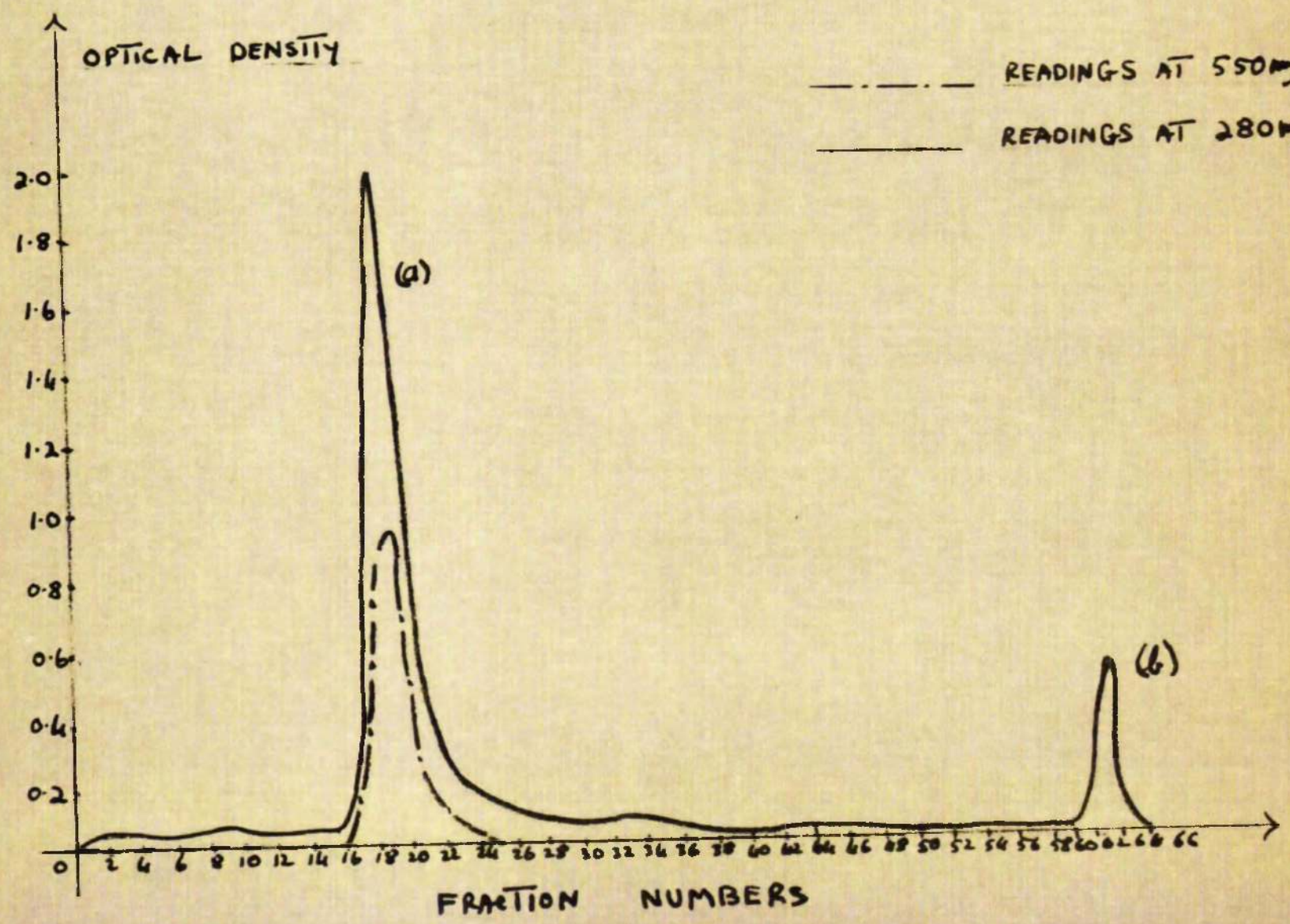

OPTICAL DENSITY
READINGS AT 550mμ
READINGS AT 280mμ
2·0
1·8
1·6
1·4
1·2
1·0
0·8
0·6
0·4
0·2
0
(a)
(b)
FRACTION NUMBERS

FIGURE 37

peptide precipitated by reducing the pH of the solution to 3 with acetic acid. The precipitate was washed with 0.05 M NH_4 Ac pH3, and the washings added to the supernatant.

Determination of the Protein Substitution Remaining after Tryptic Hydrolysis

After the tryptic hydrolysis on the G50 sephadex column, one-twentyth of the total hydrolysate was freed from ammonium acetate under vacuum and treated with FDNB in the automatic titrater as previously described. At the end of this treatment, the excess FDNB was removed from the acidified solution, by extraction with peroxide free ether, The material was reduced to dryness under vacuum, and then subjected to acid hydrolysis in 6N HCl. The estimation of the ϵ DNP lysine and free lysine in the hydrolysate was carried out as before.

Results

ϵ-DNP lysine recovered = 0.05 µmoles

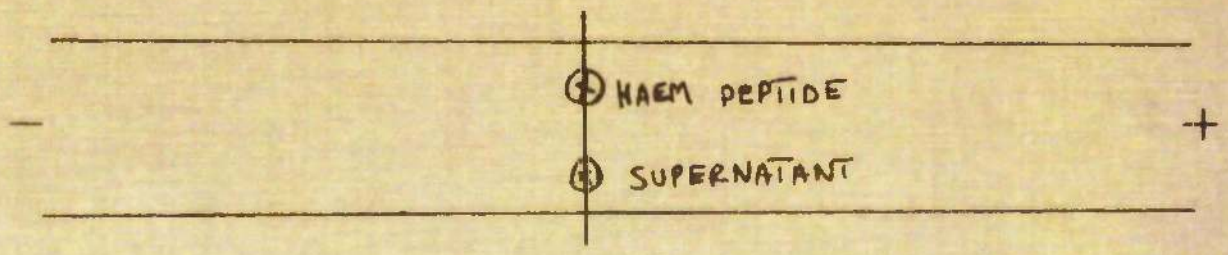
Free lysine recovered = 1.85 µmoles

This represents 0.55 lysine groups unsubstituted after tryptic hydrolysis.

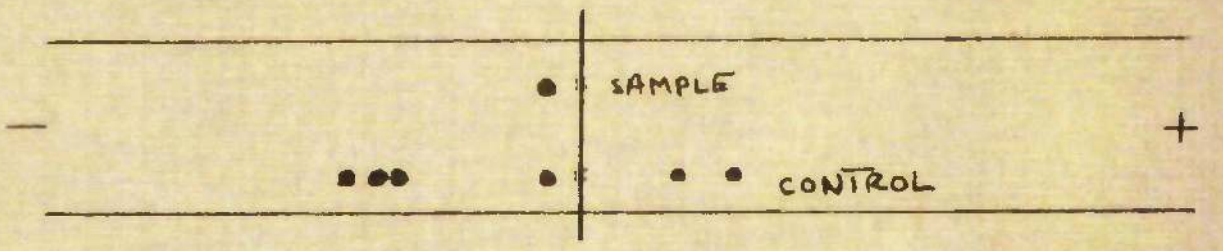
Electrophoresis of the Haem Fraction and the Supernatant

This was done in the usual way using pyridine-acetic acid--water buffer pH 6.4 and a potential of 70v/cm for 35 minutes. The paper was developed with ninhydrin dip. As can be seen in Fig. 38 no movement from the origin occurred in any case, and only one peptide was indicated in each sample.

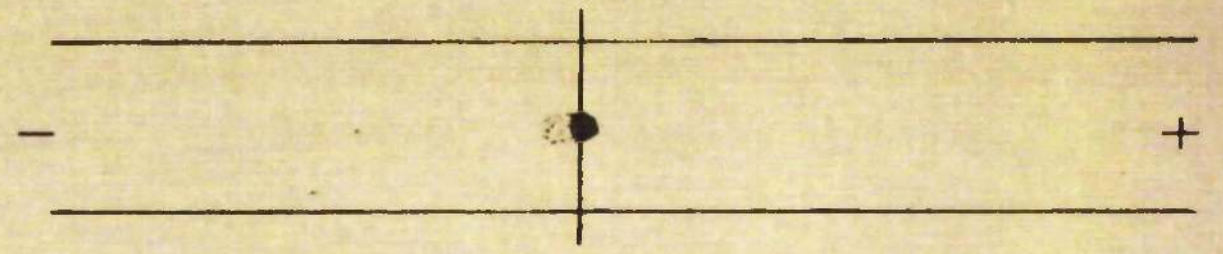
ELECTROPHORESIS OF THE HAEM FRACTION AND SUPERNATANT



INVESTIGATION OF PEAK (b) FROM THE G25 SEPHADEX FRACTIONATION



INVESTIGATION OF PEAK B FROM G50 SEPHADEX TRYPTIC HYDROLYSIS



INVESTIGATION OF PEAK C FROM G50 TRYPTIC HYDROLYSIS

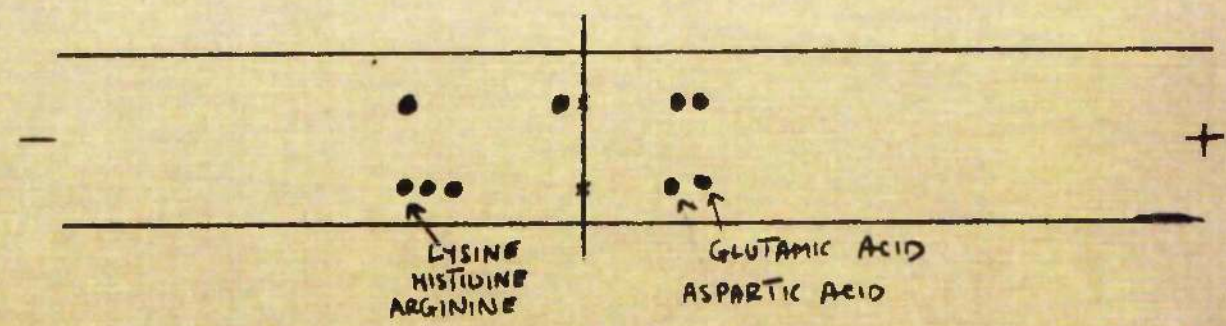


FIGURE 38

Amino Acid Analysis of the Fractions

Haem Fraction (1/12th of the total). See Table. 26

Table 26

Amino Acid	µmoles	Molar Ratios
Cys	0.205	1.3
Asp	0.8	5.0
Thr	0.86	5.38
Glu	1.5	9.4
Pro	.445	2.78
Gly	1.3	8.15
Al	0.62	3.87
Val	0.47	2.9
ILeu	0.36	2.25
Leu	0.32	2
Tyr	0.32	2.0
Phe	0.52	3.25
His		
Lys	1.7	10.6
Hist	0.49	3.06
Arg	.09	0.56

Yield = 100%

Supernatant (½). See Table 27

Table 27

Amino Acid	µmoles	Molar Ratios
Asp	0.0825	1.37
Thr	0.061	1
Glu	0.137	2
Pro	0.081	1.37
Gly	0.142	2.37
Al	0.067	1
Val	0.034	0.5
ILeu	0.0495	0.8
Leu	0.0655	1
Lys	0.286	4.5

Investigation of Peak (b) from the G25 Sephadex Fractionation

The amount of material contained in this peak is very small, in fact, 0.08 μmoles if E = 25,000.

A third of the material was acid hydrolysed as usual and then subjected to high voltage electrophoresis at pH 6.4 (70v/cm for 30 mins.) A control containing 5 μg lys, hist, arg, asp, gly, glu and phe was run also. See Fig. 38
The results showed that only neutral amino acids were present. This meant that this material did not constitute the "tail" peptide of cytochrome c and so in view of the limited amount of it available, it was discarded as being of little importance.

Investigation of Peak B from the G50 Sephadex Tryptic Hydrolysis

The amount of material contained in this peak is 0.09 μmoles, if E = 25,000. Electrophoresis done on one third indicated that this peak was composed of substantially one peptide (70v/cm 35 mins. buffer pH 6.4). See Fig. 38

One third of the material when subjected to N terminal end group analysis using the DNP technique (Schroeder 1957) yielded the results.

	500 mμ	μmoles	Total μmoles
Phe	0.11	.025	.075

This peak was also discarded as being of little importance.

Investigation of C from the G50 Sephadex Tryptic Hydrolysis

A very small amount of this was available. N-terminal end group analysis indicated the presence of Glu (0.02 μmoles)

and high voltage paper electrophoresis carried out as above on the acid hydrolysate of this material showed the presence of lysine, aspartic acid, glutamic acid and some neutral amino acids. See Fig.38.

These results indicate, therefore, that "C" is probably the tail peptide i.e. p 92-104 . Its position in the sephadex fractionation, i.e. behind the main peak tends to confirm this, as this is as would be expected from a smaller peptide.

DISCUSSION

Discussion

One of the earliest investigations made in this work of the amino acid sequence of cytochrome *c*, consisted of attempts to separate its peptides, produced by tryptic hydrolysis. In this connection, a column of Zeocarb 225 eluted with phosphate buffers was tried unsuccessfully, but better results were obtained using the ion exchanger carboxymethyl cellulose. Although with this, very little clear cut separation appeared to be obtained, the application of electrophoresis to some of the peaks showed that separation of the peptides into groups had occurred. Few of these peptides, however, gave integral molar ratios and so the technique was abandoned in favour of a more sophisticated approach. Nevertheless, in retrospect, some of them might well have been pure (see confirmation section).

The technique adopted, as discussed in the introductory section, involved the splitting of the protein into a few large peptides by limited hydrolysis with trypsin. The limitation of the hydrolysis was achieved by substituting the ε amino groups of the lysine residues with, in the first instance, methoxycarbonyl groups, and subsequently, with trifluoroacetyl groups. The peptides produced were separated on sephadex or by means of electrophoresis and their amino acid content, in some cases, their N-terminal end groups determined.

While this work was going on, however, the entire amino acid sequence of cytochrome *c* was published by Margoliash, et al (1962)

and Kreil & Tuppy (1961). Our work was continued, however, because of the desirability for confirmatory results, and also because of the necessity to test our new approach to sequence analysis.

Margoliash and his co-workers obtained the amino acid content of the protein by automatic analysis of a complete enzymatic hydrolysate.

Margoliash and Smith (1962) then degraded the protein, after ethanol-denaturation, into medium sized peptides by means of chymotryptic hydrolysis.

The peptides were separated chromatographically on Dowex 50-x2, using volatile buffers, and then purified where necessary by a mixture of paper electrophoresis and chromatography.

Samples of the purified peptides were acid hydrolysed using 6N HCl and their amino acid content estimated by automatic amino acid analysis.

In order to determine the amino acid sequence of the chymotryptic peptides, these in turn were hydrolysed enzymatically using a selection of enzymes. Trypsin, papain, carboxypeptidase and leucine amino peptidase were all used, the selection depending on the circumstances and the knowledge required.

The peptides produced were separated and purified by a combination of paper electrophoresis and chromatography, or, one or the other used separately.

The amino acid sequences of the peptides were determined by means of the fluorodinitrobenzene technique and also Edman's

(1950 and 1953) phenyl isothiocynate technique. This work, plus similar work done on the tryptic peptides by Kreil and Tuppy (1961) has permitted the deduction of a unique amino acid sequence for horse heart cytochrome c. This sequence is shown in Figure 39.

As can be seen, the N-terminal residue is acetyl-glycine. This was detected by Tuppy on hydrolysing the tryptic peptide Acetyl Gly-Asp-Val-Glu-Lys with pepsin to give Val-Glu-Lys and an acidic fragment which failed to react with ninhydrin. The latter on partial hydrolysis with 2N hydrochloric acid yielded free aspartic acid and a substance chromatographically indistinguishable from acetyl glycine; hydrazinolysis gave aspartic acid and two hydrazides behaving like glycyl hydrazide and acetyl-hydrazide. This was accepted as proof that the N-terminal residue of cytochrome c is acetyl glycine and explains some of the results obtained in this present work, as the acetylation of glycine would render it insusceptible to the usual reaction of an amino acid in this position.

The work done on the restricted tryptic hydrolysis of cytochrome c by the substitution of its lysine ϵ amino groups, using methyl chloroformate, showed that it was impossible to achieve complete substitution under the conditions described. From the knowledge of the N-terminal amino acids of the peptides produced, and the amino acid sequence of cytochrome c, however, it was possible to obtain some information concerning the identity of the lysine groups which remained unmasked.

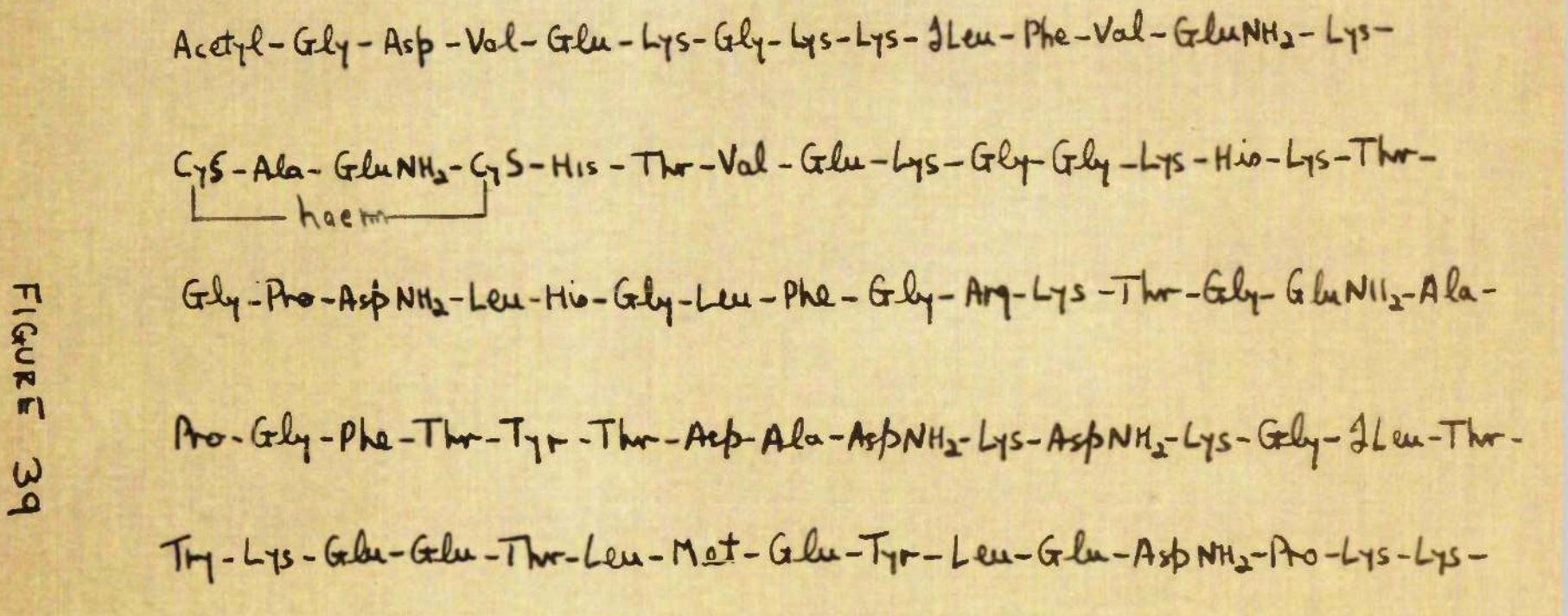

Acetyl-Gly-Asp-Val-Glu-Lys-Gly-Lys-Lys-ℓLeu-Phe-Val-GluNH2-Lys-
Cys-Ala-GluNH2-Cys-His-Thr-Val-Glu-Lys-Gly-Gly-Lys-His-Lys-Thr-
haem
Gly-Pro-AspNH2-Leu-His-Gly-Leu-Phe-Gly-Arg-Lys-Thr-Gly-GluNH2-Ala-
Pro-Gly-Phe-Thr-Tyr-Thr-Asp-Ala-AspNH2-Lys-AspNH2-Lys-Gly-ℓLeu-Thr-
Trp-Lys-Glu-Glu-Thr-Leu-Met-Glu-Tyr-Leu-Glu-AspNH2-Pro-Lys-Lys-
Tyr-ℓLeu-Pro-Gly-Thr-Lys-Met-ℓLeu-Phe-Ala-Gly-ℓLeu-Lys-Lys-Lys-Thr-
Glu-Arg-Glu-Asp-Leu-ℓLeu-Ala-Tyr-Leu-Lys-Lys-Ala-Thr-AspNH2-Glu·COOH

FIGURE 39

Although it cannot be assumed that the free amino groups belong to three or four particular lysine residues, the DNP results obtained did suggest that only a few of these were involved. The occurrence, for example, of DNP threonine and DNP alanine, as well as bis DNP lysine indicated lack of substitution at lysines 87, 88, 99 and 100 which could be explained by some sort of steric hindrance to reaction at multi lysine sites. It is possible, however, that some of the material identified as bis DNP lysine might in fact have been methoxycarbonyl DNP lysine.

The occurrence of DNP glycine, on the other hand, is most readily explained by a non specific lack of completeness in the reaction between methyl chloroformate and cytochrome c, as DNP glycine has a three times greater chance of appearing compared with those other amino acids not detected. Thus, the results obtained here showed that this technique required extensive further development. It was abandoned, however, in favour of the more promising method involving $CF_3.COS$ Et ethylthiotrifluoroacetate.

Using this substance, it has been demonstrated possible to substitute the E amino group of every lysine in cytochrome c at pH 9.0 and at pH 8.5, 94% substitution is obtained. Subsequent experiments showed also that although some spontaneous hydrolysis of these groups took place, this was very slow (see pp 130) and unlikely to affect the main issue, providing some efficient means of deactivating trypsin was available.

The critical nature of the pH requirements for the

substitution reaction was surprising, however, and not discovered until recently. Unfortunately, most of the previous reactions had been carried out at pH 8.5 because of the knowledge that desubstitution occurs at pH 11 and the conclusion, therefore, that at pH 8.5 spontaneous desubstitution would be less likely to occur than at pH 9. The damage resulting from this error, however, would be small.

The complete removal of the "blocking" groups has also been demonstrated, using either 0.1 M NaOH or 0.5 M ammonia solution. Although the NaOH method is obviously effective, it is possible that it may have some undesirable side-effects on the peptides, hydrolysing for example, the amide groups. It was for this reason that ammonia was tried. Despite the fact that it requires a higher concentration and a longer incubation time, it is more acceptable as it is unlikely to damage the peptides.

In experiment I, (pp 95) the peptides expected to result from hydrolysis only at arginine centres were obtained. Corresponding to the structure of cytochrome c published by Margoliash and Tuppy, these peptides were 1-38 (30%); 39-91 (10%) and 92-104 (80 %) and their yields are shown in brackets. With the exception of the latter, these yields are low and in addition, considerable amounts of other peptides were also obtained. Comparison of these, however, with the Tuppy Margoliash structure showed that they were fragments of P39-91, and this explained, to

some extent, the low yield of this peptide.

Nevertheless, the yields of the peptides comprising the cytochrome c fragment 1-91 were still low and ranged from 22%-54%.

One possible explanation for this could be the non detection of peptide material on the paper electrophoretograms. A more probable explanation, however, was incomplete elution of peptide material from paper, due to adsorption plus subsequent chromatographic losses. With the coloured peptide 1-38, indeed, we were unable to elute all the colour from the paper and considerable amounts were thus lost. In the "methods" section also, an experiment is described to show that in one particular case, at least 25% of a peptide was lost in this way.

This attraction between peptide and paper, however, is likely to vary in strength depending on the structure of the peptide. It is not surprising, therefore, that the yields of the peptides isolated, differed substantially,

The efficiency of the blocking technique in this experiment therefore was not entirely satisfactory. Although the first large peptide P1-38, appeared in an unfragmented form (apart from a little P1-25), the second large peptide P39-91 was considerably degraded. This fragmentation did not occur because of chymotryptic activity, (thus showing the success of the short hydrolysis) but as a result of hydrolysis at lysine residues, despite the earlier work showing almost complete trifluoroacetylation. It seemed

probable, therefore, that it occurred after the partial removal of the blocking groups at pH 11 because of continuing tryptic activity despite the efforts made to eliminate this. That P1-38 was not affected in this way is due to the fact that the blocking groups on this part of the protein are not removed at pH 11. This was shown by its treatment with FDNB which yielded no E DNP lysine.

In the second experiment (pp 107) in order to reduce this additional degradation, less trypsin was used, in the hope that the activity remaining after boiling would be negligible. Although reasonably successful in this direction, however, the intended hydrolysis was not full enough in that 60% of the starting material remained as P1-91. This was rather too close to the structure of cytochrome c to be very useful, but in other proteins with different dispositions of arginine residues, such incomplete hydrolysis might prove worthwhile.

In this experiment, an attempt was made to purify the fractions by means of gel filtration. This was quite successful, although considerable amounts of material appeared to get lost, particularly with Fractions 1 and 3. The explanation for this was probably failure of the ninhydrin technique to detect some material in the column eluates, plus errors in the ninhydrin determinations because of a high blank.

The failure of the ninhydrin technique to register was probably the explanation also for the low yields obtained from the chymotryptic hydrolysis of Fraction 3, P2. This sort of

thing has been noted by other workers also, e.g. Margoliash, (1962).

In the third experiment (pp 129 -), the problem of the continuing tryptic activity was tackled in the way described in the hope that the trypsin would be separated from the peptides. To a large extent, this was successful, but unfortunately, the intended hydrolysis did not materialise and no cleavage whatsoever occurred at arginine 38. Instead, all the material was hydrolysed at lysine 55 and possibly elsewhere in the section 56-104, although little of this was recovered and so it is impossible to say precisely. This hydrolysis of the protein at lysine 55, although even more pronounced on this occasion, occurred also, to a considerable extent in the other experiments. It seems probable, therefore, that this position is in some way specifically vulnerable to the tryptic attack.

The solution to this entire problem is therefore the use of some reagent which is capable of completely eliminating tryptic activity, e.g. di isopropylfluorophosphate. This material, however, is not readily available, and as yet we have been unable to obtain any. The next step in this work, however, will be an experiment as follows:-

(1) The blocking of cytochrome c with $CF_3.COS\ Et$ followed by the removal of excess ester with chloroform or on a column of sephadex. The latter would be better because it would also remove $CF_3.CooH$.

(2) Tryptic hydrolysis for 2 - 3 hours, followed by elimination of the enzymatic activity with di isopropylfluorophosphate.

(3) pH fractionation of the material.

(4) De-blocking with 0.5 M MnI_2.

(5) Separation of the peptides in the fractions on sephadex or electrophoretically.

The content of the fractions would presumably be

Fraction 1: P1 -38 and P39 -91

Fraction 2: P92-104

An alternative approach, however, would be to restrict the tryptic hydrolysis by substituting the lysine ϵ-amino groups of cytochrome c with t-butoxycarbonyl groups. Use has been made of this group in the field of protein synthesis by Schwyzer and Rittel (1961) and they have shown that it can be removed in cold trifluoroacetic acid or 1N HCl but not in aqueous or concentrated acetic acid.

It would seem, therefore, that as a "blocking" agent, this group has distinct possibilities. Investigation of its potential in this direction would thus be well worthwhile.

During the course of the present work, although the technique under development did not produce exactly the results hoped for, a good many peptides of cytochrome c were isolated in a pure form. The amino acid compositions of these peptides compared with the relevant section of the Margoliash-Tuppy structure of cytochrome c are shown in the tables 28 - 40.

Discussion of the Confirmatory Results

Peptide I i.e. P1-38 (Tuppy and Maryoliash)

Table 28

Amino Acid	Fraction 1 P1 1st Exper.	Fraction 2 P2 1st Exper.	Fr.1 P2 2nd Exper.	P1-38 (Tuppy
Asp	2.0	3.0	2.09	2
Glu	4.9	6.0	4.37	4
Thr	2.16	2.0	2.07	2
Pro	0.97	1.0	1.58	1
Val	3.4	2.5	1.92	3
Al	0.97	1.0	1.9	1
Leu)) ILeu)	2.75	3.0	1.78 1.77	1 2
Gly	6.8	7.25	4.6	7
Phe	1.08	1.6	1.32	2
Hist	3.0	2.0	1.91	3
Lys	6.7	8.77	7.2	7
Cys	2.0	2.0	1.0	2
Arg	0.84	0.61	0.83	1

Table 29

Amino Acid	Fr.1, P2 Exper. 1	P1-25 (Tuppy)
Asp	1.03	1
Glu	3.4	4
Thr	1.36	1
Val	5.0	5
Leu + ILeu	0.4	1
Al	1.05	1
Gly	4.58	4
Phe	0.89	1
Hist	1.48	2
Lys	6.0	6
Cys	2.0	2

Table 30

Amino Acid	Haem Peptide Exper. 3	Pl-55 (Tuppy)
Cys	1.3	2
Asp	5.0	5
Thr	5.38	5
Glu	9.4	5
Pro	2.78	2
Gly	8.15	9
Al	3.87	3
Val	2.9	3
Ileu	2.25	1
Leu	2.0	2
Tyr	2.0	1
Phe	3.25	3
Lys	10.6	10
Hist	3.06	3
Arg	0.56	1

Table 31

Amino Acid	Fr.2, P1 1st Exper.	P39 - 91 (Tuppy)
Asp	4.1	4
Glu	6.2	6
Thr	6.1	6
Pro	2.0	3
Al	3.1	3
Gly	5.0	5
Leu	6.0	6
Phe	2.0	2
Lys	10.0	10
Arg	1.0	1
Meth	1.0	2
Tyr	not detected	3
Try	destroyed by acid hydrolysis	1

Table 32

Amino Acid	Fr.5, P1 1st Exper.	P39 - 53 (Tuppy)
Asp	1.6	2
Glu	0.77	1
Thr	1.57	3
Pro	1.11	1
Al	1.27	1
Gly	2.01	2
Phe	0.72	1
Lys	1.97	2
Tyr	not detected	1

Amino Acid	Fr3, P2 C.H.(d) 2nd Exper.	P39-46 (Tuppy)	Amino Acid	Fr3, P2 C.H.(e) 2nd Exper.	P40-46 (Tuppy)	Amino Acid	Fr3, P2 C.H.(f) 2nd Exper.	P47-53 (Tuppy)
Thr	1.36	1	Thr	1.0	1	Thr	2.4	2
Glu	1.36	1	Glu	1.3	1	Glu	1.0	0
Pro	1.5	1	Pro	1.2	1	Asp	2.99	3
Gly	1.8	2	Gly	2.1	2	Gly	1.0	0
Al	1.0	1	Al	1.1	1	Al	1.5	1
Lys	0.86	1	Phe	1.0	1	Lys	1.5	1
Phe	1.0	1						

Table 34

Amino Acid	Fr3, P2 C.H.(3) 2nd Exper.	P40-48 (Tuppy)	Amino Acid	Fr3, P2 C.H.(b) 2nd Exper.	P49-53 (Tuppy)
Thr	1.64	2	Thr	0.9	1
Glu	1.0	1	Asp	1.9	2
Pro	1.05	1	Al	1.2	1
Gly	2.12	2	Lys	1.0	1
Al	0.975	1			
Tyr	0.725	1			
Phe	1.07	1			

Table 35

Amino Acid	Fr3, P2 1st Exper.	P54-73 (Tuppy)	Amino Acid	Fr3, P3 1st Exper.	P74-91 (Tuppy)
Asp	1.7	2	Al	0.84	1
Glu	3.93	4	Glu	2.19	1
Gly	not estimated by the technique used	1	Gly	not estimated by the technique used	2
Thr	1.15	1	Thr	2.25	2
Meth	0.515	1	Meth	1.0	1
Lys	3.19	4	Lys	3.15	4
Leu	2.3	3	Leu	2.56	3
Pro	not estimated by the technique used	1	Pro	not estimated	1
Tyr	0.42	1	Tyr	0.175	1
			Arg	0.325	1
			Phe	1.46	1

N-terminal end group Asp — Asp

N-terminal end group Tyr — Tyr

Table 36

Amino Acid	Fr3, P4 1st Exper.	P92-104 (Tuppy)
Asp	2.2	2
Glu	2.4	2
Thr	0.9	1
Al	1.9	2
Lys	3.2	2
Tyr	0.54	1
Leu + ILeu	3	3

N-Terminal end group Glu

Glu

Table 37

Amino Acid	Molar Ratios Fr5, P2 2nd Exper.	Corrected Molar Ratios	P39-55 (Tuppy)
Asp	3.84	3.3	3
Thr	2.7	2.16	3
Glu	2.01	0.95	1
Pro	1.57	1.3	1
Gly	3.7	2.08	2
Al	3.29	3.0	2
Val	0.82	0	0
ILeu	0.255	0	0
Leu	0.49	0	0
Tyr	1.19	1.19	1
Phe	1.25	0.7	1
Lys	4.16	2.3	3
Hist	0.761	0	0

Table 38

Amino Acid	Molar Ratios Fr2 Exper. 2	Molar Ratios Cytochrome c Exper. 2	Pl - 91 (Tuppy)	Cytochrome c (Margoliash)
Asp	4.95	8.12	6	8
Thr	6.06	8.18	8	9
Glu	11.9	12.46	10	12
Pro	4.25	6.25	4	4
Gly	11	11.12	12	12
Al	6.7	5.76	4	6
Val	3.77	5	3	3
ILeu	5	6.3	5	6
Leu	4	6.3	4	6
Tyr	2.82	3.0	3	4
Phe	4.05	3.56	4	4
Lys	16.8	17.6	17	19
Hist	3.3	3	3	3
Arg	1.92	2	2	2
Cys	0.64	0.2	2	2

Table 39

These Tables Contain the Generally Qualitative Results from the Fractionation of Cytochrome c Tryptic Peptides on Carboxymethyl Cellulose.

Amino Acid	Peak I Peptide I	P61-86 (Tuppy)
Asp	1.4	1
Glu	4	4
Gly	not estimated by this technique	
Thr	1.53	2
Al	1.1	1
Lys	2.2	3
Meth	1.8	2
Leu	3.4	4
Tyr	0.72	1
Phe	0.8	1
Pro	not estimated by this technique	2

Amino Acid	Peak 6 P4	P23-25 (Tuppy)
Gly	2+	2
Lys	1+	1

Amino Acid	Peak 5 PI	P39-60 (Tuppy)
Asp	2+	3
Glu	1+	1
Gly	2+	3
Thr	4+	4
Al	2+	2
Lys	4+	3
Try	destroyed	1
ILeu	1+	1
Tyr	1+	1
Phe	1+	1
Pro	1+	1

Table 40

Amino Acid	Peak 5 P5	P6 - 8 (Yuppy)	Amino Acid	Peak 7 P2	P100-104 (Yuppy)	Amino Acid	Peak 7 P4	P89-91 (Yuppy)
Gly	1+	1	Lys	+	1	Thr	+	1
Lys	2+	2	Al	+	1	Glu	+	1
			Thr	+	1	Asp	+	1
			Asp	+	1			
			Glu	+	1			

In Table 28, some slight discrepancies can be seen between the present work and P1-38 (Tuppy and Margoliash). However, none of these discrepancies is common to all three determinations and therefore, it would seem that the amino acid composition proposed by Tuppy and Margoliash for this part of cytochrome c is correct.

In Table 29, more or less complete agreement between our results and Tuppy and Margoliash's for P1-25 is shown. This at the same time, therefore, by difference confirms the composition for P26-28 as it has already been shown that their value for P1-38 is also correct.

In Table 30, some slight discrepancies also exist. These, however, are unlikely to be significant as it has been shown above that P1-38 (Tuppy and Margoliash) is correct, and later tables (32 and 37) show that their values for P39-55 are also correct.

Table 31 gives confirmation of Tuppy and Margoliash's amino acid composition for P39-91 except for the pro, tyr and meth residues against which discrepancies were found. Tables 32, 35 and 37 (see below) however, negate these and at the same time provide some evidence for the accuracy of the proposed sequential structure. Tables 33 and 34 which contain amino acid analysis of some small peptides also do this .

In Table 36, the analysis of the "tail" peptide is given and complete agreement exists between our results and Tuppy and Margoliash's, except for the lysine residues. The significance of this is open to speculation.

In Table 37 is contained the amino acid analysis of Fr.3, P2 (2nd experiment), which agrees with no part of the structure of cytochrome c as proposed by Tuppy and Margoliash. It will be remembered, however, that this material was subjected to chymotryptic hydrolysis and that the peptides when isolated, added up to P39-55 (Margoliash and Tuppy). Also, it was noticed that in the original analysis 0.82 residues of valine were present. Now, all the valine in cytochrome c is contained in the first twenty residues of the protein (starting from the N-terminal end), and these in turn are contained in the first large peptide produced by hydrolysis at arginine 38. It seemed possible, therefore, that this peptide might be a contaminating factor. Knowing the amino acid content of this peptide and the amount of valine present in the above analysis, it was possible to calculate the number of residues of the other amino acids present by proportion. From these, by subtraction, the amounts of amino acid residues due to the main peptide in Fr.3, P2 were obtained, and the peptide identified as P39-55.

In Table 38, the amino analysis of Fr.2, experiment 2, and cytochrome c are compared with the relevant values from the work of Tuppy and Margoliash. As can be seen, the results were not conclusive and so the peptide maps described in the experimental section pp 121 were done. As can be seen, the cytochrome c peptide map shows five spots not visible on the other one. These spots are shaded, two negatively changed, two neutral and one positive.

This is in accordance with expectation if Fraction 2 is P1-91 and the extra peptides from cytochrome c presumably are

(1) Glu, Asp, Leu (- ve)

(2) ILeu, Al, Tyr (neutral)

(3) Leu (neutral)

(4) 2 Lys (+ ve)

(5) Al, Thr, Asp NH_2 Glu.CooH (- ve)

In Tables 39 and 40 is presented some qualitative evidence for the structure of cytochrome c. Although this material does not carry the same weight as the quantitative results, in as far as it goes, it agrees with our own previous results and also those of Margoliash and Tuppy.

The results obtained, therefore, with the possible exception of the tail peptide, have provided strong confirmatory evidence for the structure of cytochrome c proposed by Tuppy and Margoliash. The substantial number of overlapping peptides of varying size which has been obtained, to cover the entire protein structure, shows that if there are errors, they must be small, and concerned with slight sequential differences.

SUMMARY

1. Cytochrome *c* was extracted from horse heart by a modification of the methods of Hagihara et al. (1958) and Stotz et al. (1960). The product thus yielded, however, was still impure, as indicated by spectral investigations, but complete purification was achieved by means of chromatography on a column of G25 sephadex.

2. In order to clarify the conflicting results of Matsubara et al. (1957) and Margoliash (1955), an investigation of the N-terminal residue of cytochrome *c* was made. The techniques used were:-

(a) A modification of Sanger's DNP technique (1945)

(b) Edman's phenyl isothiocyanate technique (1950)

(c) Tryptic hydrolysis of DNP cytochrome *c*

In all cases negative results were obtained, which indicated that with cytochrome *c*, the N-terminal residue is not susceptible to the reactions normally associated with an amino acid in this position.

3. An attempt was made to separate the tryptic peptides of cytochrome *c* on a column of carboxymethyl cellulose. The application of high voltage electrophoresis to some of the material thus obtained, enabled the isolation of some pure peptides, the amino acid content of which was determined.

4. In an attempt to limit the hydrolysis of cytochrome *c* by trypsin to the arginine residues in the protein, an attempt was

made to substitute the lysine ε-amine groups therein with methoxy carbonyl groups. Total substitution, however, was not achieved this way and so this method was abandoned in favour of another more promising one.

5. It was found possible to totally trifluoroacetylate cytochrome *c*, using the ester ethyl trifluoroacetate. The excess ester and its hydrolysis product, trifluoroacetic acid, were removed before tryptic hydrolysis, however, as it was discovered that they tended to reduce the activity of this enzyme. Electrophoretic investigation of the hydrolysate indicated the presence of the expected large peptides (in fairly low yield). Amino acid analysis of these compared well with the results of Tuppy and Margoliash et al. (1961 and 1962). Other peptides were also obtained, resulting from hydrolysis at lysine residues. The appearance of these was attributed to continuing tryptic activity after removal of the "blocking" groups, despite an attempt having been made, by boiling, to deactivate the enzyme. The removal of the blocking groups could be accomplished in either 0.1M NaOH or 0.5M NH_3.

In an effort to eliminate the additional hydrolysis at the lysine residues, a similar experiment to the above was carried out, using less trypsin. This time the peptides were separated on sephadex columns and it was found that insufficient trypsin had been used to cause a full hydrolysis of the cytochrome *c*.

Nevertheless, similar damage to that previously experienced did occur.

The amino acid analysis of the peptides obtained compared well with the results of Tuppy and Margoliash.

6. One section of cytochrome c i.e. peptide 39-53 (Tuppy-Margoliash structure) obtained in 5, was subjected to chymotryptic hydrolysis and the peptides separated on Zeocarb 225 x 2 (Konigsberg and Hill 1962). Amino acid and end group analysis carried out on these fitted exactly, the structure of this section of the protein proposed by Tuppy and Margoliash. The peptides obtained were P40-48, P49-53, P39-46, P40-46 and P47-53.

7. Because of the previous difficulty encountered, resulting from the continuing activity of trypsin, even after boiling, an experiment was carried out in which the "blocked" cytochrome c was hydrolysed on a column of sephadex. Thus the trypsin and cytochrome c peptides were separated as they progressed down the column and the danger of further damage after the "deblocking" of the peptides was eliminated. This technique, however, resulted in hydrolysis not at the arginine residues of cytochrome c, but at lysine 55.

The use, therefore, of some reliable means of deactivating trypsin is desirable, to render this technique of restricted tryptic

hydrolysis efficient. Di isopropylfluorophosphate would probably be suitable, but as yet we have been unable to obtain this substance.

Nevertheless, the amino acid analysis of the many peptides isolated during the course of this work, have provided strong confirmatory evidence for the structure of cytochrome o, proposed by Tuppy et al. and Margoliash et al.

REFERENCES

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

Allende and Richards ., Biochemistry in Press (1962)

Anfinsen C.B. Sela M. and Cooke J.P., J. Biol. Chem. 237, 1825 (1962).

Anfinsen C.B., Haber E., Sela M, and White F.H., Proc. Natl. Acad. Sci. U.S. 47, 1309 (1961)

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

Bach S.J., Dixen M and Zerfas L.G., Biochem.J., 40, 229 (1946)

Ball E.G. et al., J.Biol.Chem., 168, 257 (1947)

Barrett J., Biochem.J., 64, 626 (1956)

Berger A., Kurtz J and Noguchi J., in "Recent Advances in Gelatin and Glue Research" (G. Stainsby ed.)., pp.271 Pergamon Press New York, (1958)

Bergmann M., Fruton J.S. and Pollock H., J.Biol.Chem., 127, 646 (1939)

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

Brown H., Sanger F and Kitai R., Biochem.J., 60, 556 (1955)

Chance B. and Williams G.R., J.Biol.Chem., 217, 429 (1955)

Chibnall A.C. and Spahr P. F., Biochem.J., 68, 159 (1958)

Davenport H.E., Nature, Lond., 169, 75 (1952)

Davenport H.E. and Hill R., Proc.Roy.Soc., 139B, 327 (1952)

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

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

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

Edman P., Acta Chem. Scand. 7, 700 (1953)

Edman P., N.Y.Acad.Sci., 88, 602 (1960)

Ehrenberg A. and Theorell H., Acta Chem.Scand., 9, 1195 (1955)

Fraenkel-Conrat H., Bean R.S. and Lineweaver H., J.Biol.Chem., 177, 385 (1949)

Grassman W., Hannig K. and Schleyer M., Z.Phys.Chem., 322-325, 71 (1960-1961)

Gross E. and Witkop B., J.Biol.Chem., 237, 1856 (1962)
Guidotti G, Konigsberg W and Hill R.J., J. Biol Chem Aug 2184 (1962)
Hagihara B., Monikana J., Tagawa K. and Okonoki, K., Proc.Japan Acad., 34, 169 (1958)

Hagihara B., Tagawa K., Sezuki I., Morikawa J. and Okonoki, K., Biochem.J., 46, 11 (1959)

Hauptschein M., Stokes C.S. and Nodiff E.A., J.Am.Chem.Soc., 74, 4005 (1952)

Heilman J., Barrollier J. and Watzke E., Z.Phys.Chem., 309, 219 (1957)

Herriot R.M., Northrop J.H., J.Gen.Physiol. 18, 35 (1934)

Hill, R., Nature, 174, 501 (1954)

Hill, R. and Scarisbrick R., New Phytol., 50, 98 (1951)

Hirs C.H.W., Moore S. and Stein W.H., J.Biol.Chem., 190, 103 (1951)

Hirs C.H.W., Moore S. and Stein W.H., J.Biol.Chem., 195, 669 (1952)

Hirs C.H.W., Moore S. and Stein W.H., J.Biol.Chem., 221, 151 (1956)

Hirs C.H.W., Moore S. and Stein W.H., J.Biol.Chem., 219, 625 (1956)

Hirs C.H.W., Moore S. and Stein W.H., J.Biol.Chem., 235, 633 (1960)

Holleman J.W. and Biserte G., Bull.Soc.chim.biol., 41, 975 (1959)

Horecker B.L. and Kornberg A., J.Biol.Chem., 165, 11 (1946)

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

Itano H.A., A.M.A. Arch.Internal Med., 97, 145 (1955)

Jepsen J.B. and Smith, J., Nature, 172, 1100 (1953)

Katchalski E. and Bar-Eli E., J.Biol.Chem., 238, 1690 (1963)

Katsoyannis P.G., J.Polmer. Sci., 49, 65 (1961)

Keilin D., Proc.Roy.Soc.B., 98, 312 (1925)

Keilin D., Proc.Roy. Soc.B., 106, 418 (1930)

Keilin J., Biochem.J., 64, 663 (1956)

Keilin J., Nature, 180, 427 (1957)

Keilin D. and Hartree E.F., Proc.Roy.Soc.B., 122, 298 (1937)

Keilin D. and Hartree E.F., Proc.Roy.Soc.B., 127, 167 (1939)

Keilin D. and Hartree E.F., Biochem.J., 39, 289 (1945)

Keilin D. and Hartree E. F., Nature, 176, 200 (1955)

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

Kendrew J.C., Dickerson R.E., Strandberg B.E., Hart R.G., Davis D.R., Phillips D.C. and Shore V.C., Nature, London, 185, 422 (1960)

Koch G. and Weidel W., Z.Physiol.Chem., 305, 213 (1956)

Konisberg W. and Hill R.J., J.Biol.Chem., Aug.2547 (1962)

Kreil G. and Tuppy H., Nature, London, 192, 1125 (1961)

Leaf, G and Gillies N.E., Biochem.J., 61, vii (1955)

Leaf G., Gillies N.E. and Pirrie R., Biochem.J. 69, 605 (1958)

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

Linderstrøm-Lang K., Lane Lectures, 115 (Stanford University Press, Palo Alto, Calif., 1952)

Lindley H., Australian J.Chem., 12, 296 (1959)

Lockhart J.M. and Abraham E.P., Biochem.J., 58, 633 (1954)

Lockhart J.M. and Abraham E.P., Biochem.J., 62, 645 (1956)

MacMunn C.A., Phil.Trans., 177, 267 (1886)

Margoliash E., Nature, London, 170, 1014 (1952)

Margoliash E., Nature, London, 175, 293 (1955)

Margoliash E., Communications Third International Congress of Biochem., Brussels, 1955, Academic Press Inc., New York, 1956, pp.10

Margoliash E., Biochem.J., 71, 559 (1959)

Margoliash E., J.Biol.Chem., 237, 7, 2161 (1962)

Margoliash E. and Hill R., J.V.B. Symposium Series Vol.19 Pergamon Press, London, 1961, pp.383

Margoliash E., Kimmel J.R., Hill R.L. and Schmidt W.R., J.Biol.Chem., 237, 7, 2148 (1962)

Margoliash E. and Smith E., J.Biol.Chem., 237, 7, 2151 (1962)

Martin E.M. and Morton R.K., Nature, London, 176, 113 (1955)

Martin E.M. and Morton R.K., Biochem.J., 65, 404 (1957)

Matsubara H., Hagihara B., Horio T. and Okunuki K., Nature, London, 179, 250 (1957)

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

Murray R., Morrison M., Hollocher R., Marinetti, G., and Stotz E., Biochim.Biophys.Acta., 41, 334 (1960)

Nasan A. and Lehman J.R., J.Biol.Chem., 222, 511 (1956)

Neilands J.B., J.Biol.Chem., 197, 701 (1952)

Northrop J., J.Gen.Physiol., 17, 359 (1934)

Nunnikhoven R., Biochim.Biophys.Acta., 28, 108 (1958)

Paleus S. and Neilands J.B., Acta.Chem.Scand., 4, 1024 (1950)

Paleus S. and Tuppy H., Acta.Chem.Scand., 13, 641 (1959)

Pappenheimer A.M. and Williams C.M., J.Biol.Chem., 209, 915 (1954)

Paul K.G., Acta.Chem.Scand., 4, 259 (1950)

Paul K.G., Acta.Chem.Scand., 5, 379 (1951)

Perutz M.F., Rossman M.G., Cullis A.F., Muirhead H., Will G. and North A.C.T., Nature, London, 185, 416 (1960)

Phillips D.M.P., Nature, London, 161, 53 (1948)

Pies K.A. and Morris L., Anal.Biochem. 1, 187, 201 (1960)

Potter V.R., J.Biol.Chem., 137, 13 (1941)

Potter V.R. and Reif A.E., J.Biol.Chem., 194, 287 (1952)

Redfield R.R. and Anfinsen C.B., J.Biol.Chem., 221, 585 (1956)

Rundel F. and Hoppe W., Chem.Ber., 87, 1103 (1954)

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

Schallenberg E.E. and Calvin M., J.Am. Chem.Soc., 77_2, 2779

Schroeder W.A., J.Am.Chem.Soc., 79, 610 (1957)

Schwyzer R., "Protides of the biological fluids" pp.27 Elsevier Publishing Co., 1962

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

Slater E.C., Biochem.J., 45, 14 (1949)

Smith L (1)., Bact.Reviews, 18, 106 (1954)

Smith L (2)., Arch.Biochem.and Biophys., 50, 299 (1954)

Smith E.L. and Hill R.L., J.Biol.Chem., 235, 2352 (1960)

Smythe D.G., Stein W.H. and Moore S., J.Biol.Chem., 237, 1845 (1962)

Sober H.A., Gutter F.J., Wykoff M.M. and Peterson E.A., J.Am.Chem.Soc., 78, 756 (1956)

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

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

Strittmatter C.F. and Ball E.G., Proc.Natl.Acad.Sci., 38, 19 (1952)

Synge R.L.M., Chem.Revs., 32, 155 (1943)

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

Theorell H., Biochem.Z., 279, 463 (1935)

Theorell H., Biochem.Z., 285, 207 (1936)

Theorell H., Enzymologia, 4, 192 (1937)

Theorell H., Biochem.Z., 298, 242 (1938)

Theorell H. and Åkesson Å (a) J.Am.Chem.Soc., 63, 1804 (1941)

Theorell H. and Åkesson Å (b) J.Am.Chem.Soc., 63, 1812 (1941)

Theorell H. and Åkesson Å (c) J.Am.Chem.Soc., 63, 1818 (1941)

Theorell H. and Åkesson Å (d) J.Am.Chem.Soc., 63, 1820 (1941)

Tietze F., Gladner J.A. and Folk J.E., Biochim.Biophys.Acta., 26, 659 (1957)

Tint H. and Reiss W., J.Biol.Chem., 182, 385 (1950)

Tsou C.L., Biochem.J., 49, 302 (1951)

Tuppy H. and Bodo G., Monatsh.Chem., 85, 807, 1024, 1182 (1954)

Tuppy H. and Paleus S., Acta.Chem.Scand., 9, 353 (1955)

Vernon L.P., J. Biol.Chem., 222, 1035 (1956)

Wainio.W.W., "Haematin Enzymes" Ed. Falk, Pergamon Press, pp.28, 1961

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

Weygand F. and Csendes E., Angew Chem., 64, 136 (1952)

Witmann H.G. and Braunitzer G., Virology, Vol.9, No.4, 726 (1959)

Yamashita J., Nature, 179, 959 (1957)

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

Zeile K. and Reuter F., Z.Physiol.Chem., 221, 101 (1933)