

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE AMINO ACID SEQUENCE OF THE
TRYPTIC PEPTIDES OF THE f1
BACTERIOPHAGE COAT PROTEIN.

by

Barry Charles Richardson

A thesis submitted to Massey University of the
Manawatu in partial fulfilment of the requirements for
the degree of Master of Science in Biochemistry.

July, 1970

ABSTRACT

Five major peptides were isolated by paper electrophoresis from a tryptic digestion of purified f1 bacteriophage coat protein. The amino acid composition of the peptides was determined and shown to be:-

T ₁	Ala ₂ , Glu ₁ , Asp ₂ , Pro ₁ , Gly ₁ , Lys ₁ .
T ₂	Ala ₁ , Ser ₁ .
T ₃	Phe ₁ , Thr ₁ , Ser ₁ , Lys ₁ .
T ₄	Leu ₁ , Phe ₁ , Lys ₁ .
T ₅	Lys ₁ .

Sequential degradation of the intact f1 coat protein using the Edman technique showed the N-terminal sequence to be:-

Ala - Glu - Gly - Asp - Asp -

T₁:- The sequence of the tryptic peptide T₁ indicated it was derived from the N-terminal of the protein and was assigned the sequence:-

Ala - Glu - Gly - Asp - Asp - (Pro₁, Ala₁) - Lys.

T₂:- After two cycles of the Edman degradation reaction the sequence of T₂ was shown to be:-

Ala - Ser

Digestion of the intact f1 coat protein with carboxypeptidase A indicated T₂ was the C-terminal peptide since carboxypeptidase A showed the C-terminal sequence to be:-

-
- Lys - Ala - Ser

T₃:- This peptide was shown to have the sequence:-

Phe - Thr - Ser - Lys

T₄:- This peptide was assigned the sequence

Leu - Phe - Lys

T₅:- T₅ was shown to be a free Lys residue.

ACKNOWLEDGEMENTS

I am grateful to my supervisor, Dr G.G. Midwinter for his continual help and guidance throughout this study and also to Professor G.B. Petersen, Dept. of Biochemistry, University of Otago Medical School, Dunedin for the supply of the f1 phage.

I would like to thank my future wife, Joy for her continual assistance during the preparation of this manuscript, and for typing the draft copy. I also wish to thank Miss Wendy Nash of the N.Z. Dairy Research Institute for typing the final manuscript.

CONTENTS

	Page
<u>CHAPTER 1</u> INTRODUCTION	1
Bacteriophages and the study of their coat proteins.	10
<u>CHAPTER 2</u> MATERIALS AND METHODS	
(a) Purification of solvents for sequence analysis.	17
(b) Chromatography.	18
(c) Purification of the f1 phage coat protein.	19
(d) The amino acid analysis of the f1 phage coat protein.	20
(e) Tryptophan analysis of the f1 coat protein.	21
(f) Tryptic digestion of the coat protein.	21
(g) Separation of tryptic peptides.	22
(h) Separation of core and protein resulting from tryptic digestion of the f1 coat protein.	25
(i) Amino acid composition of the tryptic peptides.	25
(j) Edman degradation of the f1 coat protein.	26
(k) Carboxypeptidase digestion of the coat protein.	27
(l) Edman degradation of the tryptic peptides.	28

<u>CHAPTER 3</u>	RESULTS	30
(a)	Characterization of the f1 phage coat protein.	30
(b)	Amino acid analysis of the coat protein.	30
(c)	Tryptophan analysis.	31
(d)	Tryptic digestion of the coat protein.	32
(e)	Separation of the tryptic peptides.	33
(f)	Separation of the tryptic core and f1 coat protein.	36
(g)	Amino acid composition of the tryptic peptides.	36
(h)	N-Terminal sequence of coat protein.	38
(i)	Digestion of the coat protein with carboxy-peptidase A.	40
(j)	The Edman degradation of the tryptic peptides.	40
<u>CHAPTER 4</u>	DISCUSSION	44
(a)	Purification of the f1 coat protein.	44
(b)	Amino acid analysis of the f1 coat protein.	46
(c)	Tryptophan analysis.	51
(d)	Tryptic digestion of coat protein and separation of the tryptic peptides.	52
(e)	Edman degradation of the f1 coat protein and the tryptic peptides.	58
(f)	C-Terminal sequence of the coat protein.	63
(g)	General summary	65
	References	69

TABLES

Between pages

1. Amino Acid Composition of the Phage Coat Protein. 30-31
2. Values of S Calculated from the Spectra of Known Mixtures of Tyrosine and Tryptophan. 30-31
3. The Electrophoretic Mobility (E_m) of the Tryptic Peptides at pH 6.4 in Pyridine Acetate buffer. 34-35
4. Amino Acid Composition of the Tryptic Peptides of the f1 Phage Coat Protein. 36-37
5. Edman Degradation of the f1 Coat Protein. 38-39
6. Amino Acids Released During the Carboxypeptidase A Digestion of the f1 Coat Protein. 38-39
7. Edman Degradation of the Tryptic Peptides. 41-42
8. Amino Acid Composition of the f1, fd, M13 and ZJ-2 Coat Proteins. 48-49
9. Comparison of the C-Terminal Sequences of f1, fd, TMV vulgare and TMV dahlenense. 64-65

<u>FIGURES</u>	Between pages
1 Amino acid analysis of the f1 coat protein after hydrolysis in 6N HCl for 24 hours.	30-31
2a Ultra-violet spectra of the f1 coat protein obtained at pH 13.0.	31-32
2b Ultra-violet spectra of the f1 coat protein obtained at pH 12.5.	31-32
3 Value of S as a function of R for the standard amino acid mixtures used for estimating the tryptophan content of the protein.	31-32
4 Tryptic digestion of the f1 coat protein.	32-33
5 High voltage paper electrophoresis of the tryptic peptides, using pyridine : acetic acid : water buffer, pH 6.4.	34-35
6 Thin layer finger print map of a tryptic digestion after hydrolysis for 24 hours.	35-36
7 Separation of the tryptic peptides of the f1 coat protein on a AG 50W-X2 column.	35-36
8 Separation of the core and coat protein by chromatography on Sephadex G-50.	36-37
9 Edman degradation of the f1 coat protein and identification of the PTH-amino acids formed using solvent E.	38-39

- 10 Ultra-violet spectra of the N-terminal PTH-Ala
of the f1 coat protein in absolute methanol. 38-39
- 11 Edman degradation of the tryptic peptides. 41-42
- 12 Comparison of the amino acid sequence of the
f1 and fd coat proteins. 61-62

ABBREVIATIONS

Ala	Alanine
Asp	Aspartic acid
Dansylchloride	1-dimethylaminonaphthalene-5-sulphonyl chloride
DNA	Deoxyribonucleic acid
DNFB	2,4-Dinitro-1-fluorobenzene
DNP-amino acid	Dinitrophenyl-amino acid
DNS-peptide	Dansyl - peptide
DPTU	Diphenylthiourea
Glu	Glutamic acid
Gly	Glycine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
MPTU	Monophenythiourea
Phe	Phenylalanine
PITC	Phenylisothiocyanate
Pro	Proline
PTC-derivative	Phenylthiocarbonyl-derivative
PTH-derivative	3-Phenyl-2-thiohydantoin-derivative
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

INTRODUCTION

Proteins together with nucleic acids constitute the material basis of living matter and are the most complex chemical systems known. Proteins are composed of combinations of the twenty common amino acids, arranged in a certain order that determines their characteristics as well as determining their biological activity. Hence a knowledge of the primary structure of proteins (i.e. the arrangement of the amino acids) is necessary to gain an insight into the behaviour of these molecules.

The early workers in the field of protein chemistry defined a protein by its elementary composition in terms similar to that by which organic compounds are described. The first important step forward in the isolation and characterization of the common amino acids from protein hydrolysates was accomplished by precipitation, and later by ester-distillation methods (Fischer, 1906). The last common amino acid, threonine was discovered by McCoy, Meyer and Rose (1935) as a result of feeding experiments on rats.

These early methods for the estimation of amino acids in proteins were laborious and time consuming. Typical of the precipitation methods was that used for isolating and

estimating the basic amino acids (Kossel and Kutscher, 1900). This involved precipitation of the amino acids as their silver salts, separation of these due to their different solubilities in water and strong alkaline solutions, followed by precipitation with picric acid and phosphotungstic acid. This method required 25 - 50 grams of protein, the handling of large volumes of solution and frequent precipitation, filtration, washing of large amounts of barium sulphate and other lengthy procedures.

The ester-distillation method of Fischer (1906) required one kilogram of protein. After hydrolysis with hydrochloric acid, the amino acids were esterified with acidic ethanol and esters were separated from the hydrochlorides using sodium hydroxide, and extracted with ether. After removing the ether these derivatives were fractionally distilled in vacuo and converted to free amino acids by boiling in water. The separation of every component from this mixture required a special procedure. The time required for a single analysis was several months and the losses were large, yet these methods still laid the foundations of protein chemistry. These techniques were essentially macromethods and subsequent development of colorimetric methods reduced the level to the submicro scale.

An early method for analysis of tyrosine and tryptophan with phenol reagent was developed by Folin and Denis (1912). This was followed by the colorimetric methods for estimating the other amino acids: the Sakaguchi reaction for arginine (Sakaguchi, 1925), and the coupling of histidine with diazotized sulphanic acid (Hanke and Koessler, 1920). Microbiological assay of amino acids was also used to estimate the composition of protein hydrolysates (Barton-Wright, 1946; Barton-Wright and Curtis, 1948).

Separation of amino acids by paper chromatography was first achieved by Consden, Gordon and Martin (1944), by Dent (1948). Hanes and co-workers during an exhaustive study of the paper chromatography of amino acids described methods for their separation and quantitation using ninhydrin (Hanes, 1961; Wade, Matheson and Hanes, 1961; Matheson, Tigane and Hanes, 1961; Hanes, Harris, Moscarello and Tigane, 1961; Tigane, Wade, Wong and Hanes, 1961; Harris, Tigane and Hanes, 1961). A modification of the paper chromatographic technique employed high voltage electrophoresis to separate amino acids. The technique developed by Wieland and Fischer in 1948 (Wieland, 1948) used a field of 5 - 20 volts/cm. Although this enabled a useful separation of amino acids, fields of 50 - 200 volts/cm were found to be necessary for the highest resolution of complex mixtures.

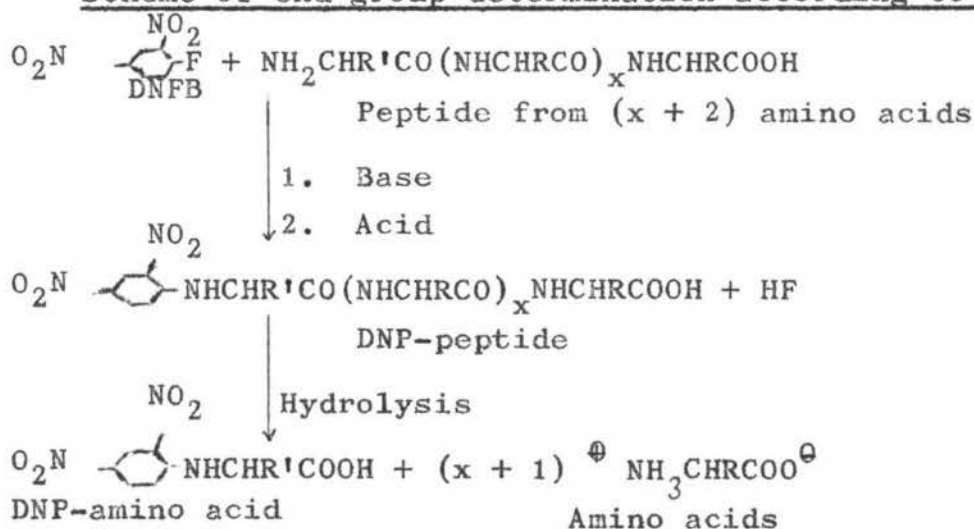
The composition of mixtures of amino acids has also been determined by separating and estimating their dinitrophenyl derivatives, phenylthiohydantoin derivatives, or dansyl derivatives by column chromatography (Sanger, 1945; Sjoquist, 1955; Sjoquist, 1957), paper chromatography (Blackburn and Lowther, 1951; Edman and Sjoquist, 1956; Sjoquist, 1960; Boulton and Bush, 1964), and high voltage electrophoresis (Gray and Hartley, 1963).

The precise amino acid composition of a peptide or protein must be known before its sequence determination can be undertaken. The development by Spackman, Stein and Moore at the Rockefeller Institute of a chromatographic column capable of separating all of the common amino acids was a remarkable development (Spackman, Stein and Moore, 1958). This kind of column forms the basis of the modern amino acid analyser which selectively elutes the amino acids from a strong cationic resin and analyses them ninhydrin, the elution pattern being continuously recorded. Five milligram samples of protein hydrolysate were sufficient to accurately determine the amino acid composition of a protein, with an analysis time of 24 hours. Today, using these instruments the analysis time for a protein hydrolysate is reduced to four hours, compared to several

months for the precipitation methods. These instruments are capable of accurately analysing samples that contain 0.1 μ M of each amino acid or less.

The first determination of the sequence of a protein was carried out using the amino specific reagent 2,4-dinitro-1-fluorobenzene, which under alkaline conditions couples with the free amino groups of peptides and proteins. Complete acid hydrolysis of the DNP-derivative yields the acid stable DNP-amino acid derived from the amino terminal, which can be identified quantitatively. This technique was employed extensively by Sanger in his investigation of the primary structure of bovine insulin (Sanger, 1945; Sanger, 1949; Sanger and Tuppy, 1951; Sanger and Thompson, 1953).

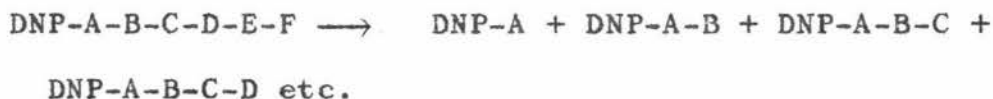
Scheme of end group determination according to Sanger



This investigation provided the first direct proof of the hypothesis that amino acids in a protein are joined

by an α -peptide linkage in a specific sequence. In addition this was the first demonstration of the specificity of the proteolytic enzyme trypsin in cleaving the peptide bond at the carboxyl terminal of the amino acids lysine and arginine. This study of insulin also shed light on the mode of action of other proteolytic enzymes such as chymotrypsin, pepsin and papain.

Sanger's major technique was to determine the amino acid sequences adjacent to the N-terminal derivative by partial acid hydrolysis as opposed to complete acid hydrolysis. Hence:



During the 1950's Edman developed a method for the sequential degradation of proteins and peptides (Edman, 1950, 1956 and 1960). The reagent phenylisothiocyanate was found to react with the N-terminal group of a protein or peptide under alkaline conditions to form the phenylthiocarbonyl derivative. On exposure to acid the terminal residue was cleaved as the 3-phenyl-2-thiohydantoin derivative of the terminal amino acid (PTH) with the release of the residual peptide. The residual peptide was then capable of being subjected to further degradation cycles.

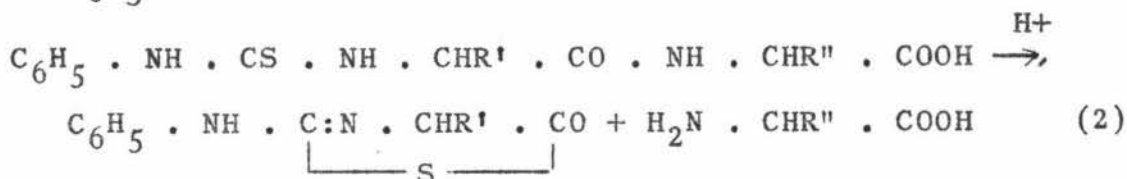
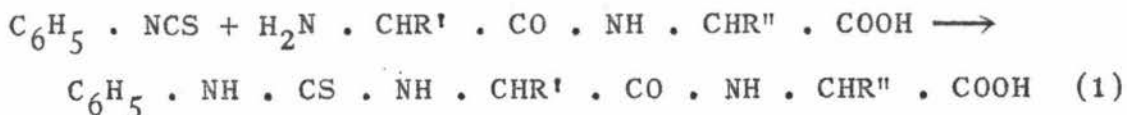
This technique has been used in almost every

investigation into the primary structure of proteins and as a result a number of variations of the method have been developed. These include methods where the degradation is carried out in solution (Sjoquist 1957, 1959; Guidotti, Hill and Konigsberg 1962; Sjoquist, Bloombach and Wallen 1961; Sjöholm 1964; Eriksson and Sjoquist 1960), using paper as a support for the degradation reaction (Schroeder, Shelton, Balog-Shelton and Cormick, 1963); and using silica gel as a support (Wieland and Gebert, 1963).

The PTH-derivative derived from the terminal amino acid may be identified directly by paper (Edman and Sjoquist, 1956) or thin layer chromatography (Edman and Begg, 1967). Alternatively the subtractive method of Hirs, Moore and Stein (1960) may be used which analyses a portion of the residual peptide after each degradation cycle and so identifies the N-terminal derivative removed.

The degradation reaction can be subdivided into three steps: coupling, cleavage and conversion (Edman and Begg, 1967).

Scheme of the Edman degradation reaction



and Begg (1967) showed that oxidative desulphuration of the phenylthiocarbonyl derivative blocked the degradation reaction. However despite these problems six and more amino acids have been removed from a peptide and successfully identified (Schroeder et al., 1963).

Gray and Hartley (1963) have used the dansyl-Edman method to sequence peptides. This is an Edman degradation followed by coupling a portion of the residual peptide with 1-dimethylaminonaphthalene-5-sulphonyl chloride, removing the acid-stable, terminal dansyl-derivative by acid hydrolysis and identifying it using paper electrophoresis. 10^{-10} M of the derivative is detectable, and only 20ngM of peptide is necessary to sequence a tripeptide. A sequenator has been developed recently by Edman and Begg (1967) which automatically degrades proteins and is capable of fifteen degradation cycles per 24 hours, and with a yield at each step in excess of 98%. This instrument which is still largely untried will undoubtedly prove a beneficial aid in the sequencing of proteins.

Recent years have seen the introduction of the mass spectrometer and the computer as a means of sequencing peptides (Senn and McLafferty, 1966). This method is still largely under development and success for the large part has been limited to small peptides of preselected amino acid composition. It remains to be seen whether this technique

will prove itself as a standard method in protein sequencing.

Bacteriophages and the study of their coat proteins

The bacteriophages were discovered by d'Herelle in 1917 (d'Herelle, 1917). They were not further characterized until the period 1930-36 when Schlesinger isolated a bacterial virus that consisted of protein and deoxyribonucleic acid, in roughly equal proportions (Schlesinger, 1936). Since this period a series of phages infecting Escherichia coli have been characterized. The nucleic acid content of these phages may be DNA (T phage, ϕ x174, f1) or RNA (MS2, R17, f2 and fr).

The DNA of a species is capable of coding for a protein. The composition of the protein depends on the order of the triplet nucleotide bases within the DNA. Each triplet can code for a specific amino acid and if the order of the nucleotide bases in the DNA is changed by the mutation of one base then we have a protein formed which differs from the original by one amino acid. This phenomenon is evident with bacteriophage coat proteins, especially within a group of phages which are all morphologically and serologically similar.

The nucleic acids of these phages direct coat protein synthesis in cell free extracts and studies along this line have contributed much to the elucidation of the

general mechanism of protein synthesis in vitro (Nathans, 1965). A study of "small" bacteriophages has an advantage since assuming a triplet genetic code these can only code for 5-10 proteins of molecular weight 20,000 - 30,000 (Sinsheimer, 1959b; Marvin and Hoffmann-Berling, 1963). By comparison the larger bacterial viruses such as the T even phages of E. coli contain sufficient information to code for several hundred proteins of a similar size (Rubenstein, Thomas and Hershey, 1961). Hence it is much easier to determine the functions of every gene within a small bacteriophage as it is likely that all the gene products can be isolated and characterized (Ray 1968).

ϕ x174 was the first of the "small" bacteriophages to be characterized (Sinsheimer, 1959a). This bacteriophage, which belongs to the class of isometric DNA phages, contains 25.5% weight/weight DNA, the molecular weight of this single stranded, circular DNA being $1.6-1.8 \times 10^6$ (Sinsheimer, 1959b). It is a regular icosohedra with a diameter of $250 \overset{\circ}{\text{A}}$ and has a coat consisting of 180 subunits each with a molecular weight of 25,000 (Hoffmann-Berling et. al., 1966).

The phages f2, fr, MS2, R17 and M12 which were first isolated by Loebb and Zinder (1961) are classified as isometric RNA phages. Zinder (1965) postulated that these phages are mutational derivatives of a common ancestor and it is on this basis that they have been extensively

studied. The first of these E. coli male specific bacteriophages, f2 was discovered by Loebb and Zinder (1961), followed by MS2 (Strauss and Sinsheimer, 1963), fr (Hoffmann-Berling, Marvin and Durwald, 1963), R17 (Enger, Stubbs, Mitra and Kaesburg, 1963), M12 (Hofschneider, 1963) and β (Nonoyama, Yuki and Ikeda, 1963). These bacteriophages which come under the classification of isometric RNA phages have a coat protein consisting of 180 subunits, each with a molecular weight of 14,700 (Lin, Tsung and Fraenkel-Conrat, 1967).

Characterization of the coat proteins of these phages is at such a stage that the amino acid sequences of R17 (Weber, 1967), f2 (Weber and Konigsberg, 1967), MS2 (Lin, Tsung and Fraenkel-Conrat, 1967) and fr (Wittmann-Liebold, 1966) are known. This work showed f2 and R17 to differ by two different single amino acid exchanges from MS2 and M12. The sequences of the coat proteins of MS2 and M12 were identical (Lin et. al., 1967). However fr differs by several more amino acid exchanges from the other coat proteins (Weber, 1967). f2 and R17 differ in the exchanges of leucine (position 88 in f2) for methionine (position 88 in R17). This change corresponds to a single nucleotides replacement (A \rightarrow C) when the codons for methionine (AUA, AUG) and leucine (CUA, CUG) are considered. The replacement of one amino acid for another within this

group of coat proteins can be explained by a single nucleotide base change except in the case of fr which appears to be a distinctly different protein (Enger and Kaesberg, 1965). These results show a good correlation to those results obtained from the seriological cross reactions observed by Scott, (1965).

f1, the bacteriophage with which this thesis is concerned was first isolated by Zinder, Valentine and Stockenius (1963), followed by fd (Marvin and Hoffman-Berling, 1963), M13 (Salivar, Tzagoloff and Pratt, 1964) and ZJ-2 (Bradley, 1964). These bacteriophages all belong to the class of small rod shaped DNA phages and have been shown to be serilogically related (Salivar et. al., 1964). Seriological cross reactions have indicated the close relationships that exist between viruses of the same type. (Zahler, 1958; Tessman and Shleser, 1963; Salivar et. al., 1964; and Arbor, 1966).

These phages which are specific for E. coli K12 F⁺ and HFr strains (Zinder, Valentine, Roger, Stoeckenius, 1963) are indistinguishable in electronmicrographs, all being flexible rods 8000 Å in length and 60 Å in diameter (Salivar et. al., 1964). The DNA of these phages is single stranded and circular (Petersen and Reeves, 1969; Marvin and Hoffmann-Berling, 1963). fd has been the most extensively studied phage of its class. A double helical base

paired structure for the DNA is only possible if non Watson-Crick pairing takes place (Day, 1969), as the base composition of the fd DNA is 35% thymidine and 24% adenosine.

Marvin (1966) proposes that the fd phage encloses the DNA in a tube, the wall of which has been calculated to be 11 Å thick. This wall is constructed supposedly from protein subunits with a high α -helical content; the α -helical axes follow a supercoil of 32.2 Å pitch around the DNA. Marvin (1966) also proposed that there are $9\frac{1}{2}$ protein subunits for every complete turn of the protein "screw". Day (1966) has measured the optical rotatory dispersion of the phage and estimates the α -helical content of its capsid to be as high 90-95%. This agrees with Marvin's model, however 40% of the amino acids present are not considered to be α -helix promoting (Braunitzer, 1964).

The amino acid sequence of fd has been elucidated by Braunitzer, Asbeck, Beyreuther, Kohler and von Wettstein (1967) and Asbeck *et. al.* (1969), and contains 49 amino acids per subunit with 1900 of these subunits making up the coat protein of the bacteriophage.

A study of the bacteriophage coat proteins of f1, M13 and ZJ-2 is necessary since our knowledge of the chemical structure is restricted to fd. In order to delineate the mechanism of protein synthesis and of

mutagenesis it is important to determine the chemical structure of the phage coat protein. The purpose of this investigation was to determine the amino acid sequence of the tryptic peptides of the f1 coat protein. Once complete studies of the bacteriophage coat proteins of this group, similar to the studies of the RNA phages f2, MS2, M12 and R17 have been achieved conclusions can be drawn between the differences in the proteins. An attempt could then be made to account for any amino acids replacements by mutagenesis, in terms of single base nucleotide changes within the codons in the genetic code.

CHAPTER 2MATERIALS AND METHODS

Amino acids were obtained from Mann Research Laboratories. Carboxypeptidase A was obtained from Sigma (Sigma Chemical Company Ltd.) 1-dimethyl aminonaphthalene-5-sulphonyl chloride (dansyl chloride) and ninhydrin were obtained from Pierce (Pierce Chemical Company). Trypsin, (Mann) was twice crystallized, lyophilized and substantially salt free.

N-Ethyl morpholine was obtained from Koch Light (Koch Light Laboratories Ltd.), diisopropylfluorophosphate (British Drug Houses Ltd.) and phenyl isothiocyanate from Pierce.

Solvents used were:- acetic acid (BDH), acetone and benzene (May and Baker), n-butyl acetate and chloroform from BDH. Diethyl ether was obtained from the Drug houses of Australia, ethyl acetate (May and Baker), ethyl alcohol and ethylene chloride from BDH. Formic acid, heptane and hydrochloric acid were all obtained from May and Baker.

Methanol was obtained from BDH, methyl ethyl ketone (May and Baker), propionic acid (Koch Light), pyridine and triethylamine from BDH.

Biogel P-2, exclusion limit (Molecular weight 100-2600) was obtained from Biorad (Biorad Laboratories).

Sephadex G-25 (fine) exclusion limit (1,000-5,000) and Sephadex G-50 (fine) exclusion limit (1,500-30,000) were obtained from Sigma. AG50W-X2 (200-400 mesh) ion exchange resin was obtained from Biorad and Dowex 50-X2 (20-50 mesh) was obtained from the Dow Chemical Company.

(a) Purification of solvents for sequence analysis

2.5 litres of solvent were purified, unless otherwise stated. Acetic acid was refluxed with chromic acid and distilled over ninhydrin. Acetone was refluxed for 3 hours with potassium permanganate (1 gm/l), and distilled over ninhydrin (1 gm/l), (Sjoquist, 1955). Benzene was purified according to the procedure of Edman and Begg, (1967). 2.5 l was stirred with concentrated sulphuric acid, with two acid changes being made. It was then washed with distilled water and shaken with 3% KMnO_4 (w/v) overnight. It was finally washed with distilled water and dried over KOH, the fraction distilling at 80°C being collected.

Chloroform was washed twice with water, dried over CaCl_2 and distilled over phosphorous pentoxide.

Ethylene chloride was purified by stirring 8 volumes of solvent with 1 volume of concentrated sulphuric acid and extracting twice, washed twice with water, 5% aqueous Na_2CO_3 , water, finally dried over anhydrous K_2CO_3 and distilled at $83 - 85^\circ\text{C}$ (Morgan and Henschen, 1969).

Methanol was refluxed with 2, 4-dinitrophenylhydrazine (1 gm/l) and distilled over ninhydrin. 10 ml of PITC was distilled in vacuo at 1 mm Hg pressure and the fraction boiling at 55°C was collected (Sjoquist, 1955). Pyridine was refluxed for 3 hours with KOH (5 gm/l) and then distilled over ninhydrin (1 gm/l), bp 115°C (Eriksson and Sjoquist, 1960). 500 mls of triethylamine was refluxed for 3 hours with phthalic anhydride (5 gm/100 ml) and distilled before use (bp 89°C), (Sjoquist, 1955).

All other solvents (reagent grade) were distilled before use.

(b) Chromatography

Thin layer chromatography was carried out using 0.25 mm thick layers of silica gel GF₂₅₄ (Merck). These layers fluoresced green under UV light of wavelength 254 mμ. 10 x 20 cm and 20 x 20 cm glass plates were used, the layers being spread with a Desaga spreader using a gel : water slurry of 1:2.

Thin layer plates were air dried and stored at 37°C until used. Plates were washed before being used by continuous elution with absolute methanol to remove a yellow band of contaminant which quenched strongly under UV light of wavelength 254 mμ.

Four solvent systems were used for the thin layer chromatography of PTH-amino acids. These were :- the

Solvent E of Sjoquist (1953). This solvent system consisted of 970 ml of n-butyl acetate, washed with water. Propionic acid was then added to a concentration of 3%, and the solution saturated with formamide, (approx. 40 ml/l). Solvent 1 : Chloroform containing 1½% ethyl alcohol (Brenner, Niederwieser and Pataki, 1961). Solvent 2 : Chloroform containing 1½% ethyl alcohol : methanol (9:1), (Brenner et al., 1961). Solvent 3 : Heptane : propionic acid : ethylene chloride (50:30:25) (Jeppsson and Sjoquist, 1967). Solvent 4 : Chloroform containing 1½% ethyl alcohol : formic acid (100:5) (Brenner et al., 1961).

(c) Purification of the f1 phage coat protein.

Purified f1 phage, grown on Escherichia coli K13 was obtained from Professor G.B. Petersen of the University of Otago Medical School, Dunedin. Purified coat protein was isolated by the phenol extraction method of Notani, Engelhardt, Konigsberg and Zinder (1965).

A 15 ml suspension of phage was centrifuged for 10 minutes at 2000 x g to remove any amorphous material. The opalescent material was mixed with an equal volume of warm phenol and incubated at 40°C for 30 minutes, after which it was cooled and centrifuged at 2000 x g for 10 minutes. The aqueous phase containing the phage DNA was discarded and the lower phenol layer was extracted

twice with two volumes of distilled water.

The protein was precipitated from the phenol by the addition of six volumes of cold ethanol : ether (1:1) and left to stand for 3 hours at 0°C. After centrifugation the precipitated protein was suspended in 15% acetic acid and dialysed against 50% acetic acid until a clear viscous solution was obtained. After dialysing the protein against 2 litres of distilled water for 24 hours it was stored in vials at -20°C.

The degree of contamination with nucleic acids was determined by the method of Warburg and Christian (1941) using the $E_{280/260}$ ratio. The protein concentration was estimated by the method of Lowry, Rosebrough, Carr and Randall (1951) with a bovine serum albumin standard.

(d) The amino acid analysis of the f1 phage coat protein.

Amino acid analyses were carried out using the standard procedure of Moore and Stein (1963). Samples containing approximately 1.5 mg of protein were placed in 16 x 125 mm heavy walled pyrex test tubes and freeze dried. 2 mls of 5.7N constant boiling HCl (three times distilled) was added to each tube, which was then sealed under a vacuum of 0.1 mm Hg while frozen.

Duplicate samples were hydrolysed for 24, 48, 72 and 96 hours at $110 \pm 1^\circ\text{C}$. After hydrolysis the HCl was

rapidly removed on a rotary evaporator and the hydrolysate washed with distilled water and re-evaporated. The samples were finally taken up in 0.6 ml of distilled water and analysed with a Beckman 120C Amino Acid Analyser using the standard 4 hour analysis procedure as outlined in the Beckman Procedures Manual.

(e) Tryptophan analysis of the f1 coat protein.

10 mM stock solutions of tyrosine and tryptophan were used to prepare 0.1mM solutions of each amino acid in 0.1N NaOH (freshly prepared). From these, solutions with a molar ratio tyrosine : tryptophan of 1.0, 1.5, 2.0, 2.5, were prepared (Bencze and Schmid, 1957).

The spectra of these solutions, including 0.1mM tyrosine and tryptophan were recorded in the Unicam SP 500 between 270 m μ and 320 m μ .

1.5 mg of coat protein, dissolved in 4 ml of 0.1N NaOH, was adjusted to pH 13.0 using 1N NaOH (Radiometer Glass Electrode G202B) and the spectra recorded in the same range as the free amino acid mixtures.

(f) Tryptic digestion of the coat protein

Tryptic digestions were carried out using the method of Guidotti, Hill and Konigsberg (1962).

Trypsin, 10 mg/ml in 0.001N HCl was kept at 4°C for 4 hours to reduce the chymotryptic activity. 10 - 50 mg of coat protein was denatured on a boiling water bath

for 2 hours, dispersed with an MSE soniprobe for 10 minutes and then adjusted to pH 8.0 with 0.05N NaOH.

An aliquot of trypsin was added and after 2 hours a further aliquot was added to give a final enzyme : substrate ratio of 1:100.

The digestion was carried out at 37°C in a Radiometer TT₁ pH Stat (pH 8.0), the pH of the digestion mixture being maintained with either 0.01N or 0.05N NaOH. Digestions were carried out for periods of 4 - 24 hours, and were terminated by adjusting the pH of the solution to 4.0 with glacial acetic acid, followed by lyophilization.

The soluble peptides from the digestion were extracted with water, lyophilized and taken up in a small amount of water. This procedure was repeated twice, the peptides were finally lyophilized and stored at -20°C.

(g) Separation of tryptic peptides.

Tryptic peptides were separated by column chromatography and analytical or preparative paper electrophoresis.

Paper electrophoresis of tryptic peptides was carried out on a Pherograph high voltage electrophoresis apparatus according to Wieland and Pfleiderer (1955, using pyridine : acetic acid : water buffer (10:0.4:90) pH 6.4, (Michl, 1951).

Whatman No. 1 paper was either prewashed with the electrophoresis buffer or 75% formic acid and then washed with water.

Loadings of approximately 0.014 M/cm (30 μ l) of peptide material were applied over a 25 cm wide strip. A marker solution containing glycine and lysine was also applied on either side of the band of peptide material.

Electrophoresis was conducted for 1 - 3 hours at 30 - 50 V/cm with the temperature held below 5 $^{\circ}$ C. The paper was then placed in a fume cupboard and dried. A marker strip was removed and briefly immersed in cadmium acetate - ninhydrin reagent (Dreyer and Bynum, 1963), and heated at 100 $^{\circ}$ C for 10 minutes.

From these marker strips the peptides were identified and eluted using descending chromatography, the eluant being 5% acetic acid or 75% formic acid. In some cases a marker strip was cut from each edge of the paper, the peptides being identified, and eluted while the paper was still wet.

Each peptide isolated was tested for homogeneity by electrophoresis of a small aliquot at pH 6.4, lyophilized and then stored at -20 $^{\circ}$ C.

The method of Schroeder (1967) was used to separate the tryptic peptides by ion exchange chromatography. Approximately 2 μ M of tryptic peptides were dissolved in 0.5 ml of 0.2M pyridine acetate buffer pH 3.1 and applied to a AG 50W-X2 column (1 x 25 cm).

The column was eluted with 100 ml of 0.2M pyridine acetate buffer pH 3.1 and 100 ml of 1.0M pyridine acetate buffer pH 5.0 arranged in a simple exponential gradient (Alm, Williams and Tiselius, 1952). A further 100 ml of 2.0M pyridine acetate buffer pH 5.0 was added to elute the highly basic peptides.

The flow rate was 30 ml/hour and 2 ml fractions were collected using an LKB fraction collector. 10% of each fraction was subjected to alkaline hydrolysis followed by ninhydrin analysis (Crestfield, Stein and Moore, 1963). An aliquot of each fraction was evaporated at 110°C in a polypropylene test tube, 0.15 ml of 13.5N NaOH was then placed in each tube which was autoclaved by heating with steam to 15 psi for a period of 20 minutes. After the pressure was released and the tubes cooled, the alkali was neutralised by the addition of 0.25 ml glacial acetic acid. Each tube received 0.50 ml of ninhydrin reagent which was then heated for 15 minutes in a covered boiling water bath. After cooling the tubes for 10 minutes in a cold water bath, 2.5 ml of 50% ethanol was added to each tube, the tubes then being vigorously shaken for 45 seconds. The absorbance of each tube was measured at 570 m μ in a Hitachi 101 Spectrophotometer.

Fractions within each peak were pooled and lyophilized, and tested for homogeneity by electrophoresis.

The tryptic peptides obtained from digestion of the protein for 24 hours were also separated as DNS-peptides by a procedure employing two dimensional thin layer chromatography (Atherton and Thompson, 1969).

(h) Separation of core and protein resulting from tryptic digestion of the f1 coat protein.

The insoluble material resulting from tryptic digestion of the protein was dissolved in as small an amount of 90% formic acid as possible.

This was placed on a column (100 x 1.7 cm) of Sephadex G-50, fine (20 - 80 μ) previously equilibrated with 1M formic acid.

The flow rate of the column was maintained at 20 ml/hour, and 3.5 ml fractions were collected which were monitored at 280 m μ with a Hitachi 101 spectrophotometer.

The effluent from the tubes under each peak were combined and lyophilized. Amino acid analysis of each peak was carried out using the Beckman 120C amino acid analyser.

(i) Amino acid composition of the tryptic peptides.

The method for acid hydrolysis of the peptides was the same as for the hydrolysis of the protein. Approximately 0.04 - 0.2 μ M of peptide was hydrolysed in 1 ml of 5.7 N HCl for 24 and 48 hours.

Each sample was analysed on the Beckman 120C amino

acid analyser.

(j) Edman degradation of the f1 coat protein.

4.5 mg of coat protein was denatured by the procedure used for tryptic digestion and then subjected to Edman degradation using the procedure of Eriksson and Sjoquist (1960).

The protein was suspended in 1 ml of distilled water and 2 ml of pyridine : triethylamine : PITC (100 : 3 : 1) were added and the suspension incubated for 90 minutes at 40°C.

The suspension was extracted 5 times with 2 mls of benzene : ethylene chloride (3:1) (previously washed with 0.1N NaOH), followed by centrifugation. The upper organic layer was discarded and any remaining benzene was evaporated at 40°C under a stream of nitrogen.

Cleavage of the N-terminal amino acid derivative from the PTC-protein was achieved by adding 1 ml of water and 2 ml of glacial acetic acid, saturated with HCl gas, and incubating at 40°C for 2 hours.

After lyophilization the residue was suspended in 2ml of ethyl acetate : methyl ethyl ketone (2:1) (saturated with water), and the PTH-amino acid was extracted 3 times from the water phase with the above mentioned solvent. The protein was then recycled and the combined organic extracts containing the N-terminal derivative was

evaporated to dryness at 40°C and the residue dissolved in 100 μ l of acetone.

The N-terminal amino acid was identified by thin layer chromatography on 20 x 20 cm silica gel GF₂₅₄ plates. Aliquots of the unknown PTH-amino acid were chromatographed along with 1 μ l aliquots of a solution containing 2 mg/ml of the PTH-derivatives of Asp, Ala, Pro, Gly, and Glu.

Solvent systems used were:- Solvent E (see chromatography section), 2 and 3. After identification of the derivative, the plate was further dried, and the PTH-derivative of the N-terminal amino acid together with a blank of the same area was removed from the plate and eluted with absolute methanol for 1 hour at 40°C.

The spectra of this derivative was recorded between 230 - 300 m μ using a Hitachi 101 spectrophotometer.

(k) Carboxypeptidase digestion of the coat protein.

25 μ l of carboxypeptidase A suspension, (50 mg/ml) was prepared according to the modified procedure adapted from Fraenkel-Conrat et al. (1952) as outlined by Ambler (1967).

The carboxypeptidase A was treated with diisopropylfluorophosphate (Fraenkel-Conrat et al., 1952) and made up to a final concentration of 1 mg/ml with 0.2M N-ethyl morpholine acetate buffer, pH 8.5.

4.65 mg of coat protein was denatured as for tryptic digestion, lyophilized and suspended in 1 ml of N-ethyl morpholine buffer.

Carboxypeptidase A was added to the protein to give a final enzyme : substrate ratio of 1:50 (mole/mole) and incubated at 25°C for 11 hours. Aliquots were removed at intervals, and adjusted to pH 2 - 3 with Dowex 50-X2 (20 - 50 mesh), the beads being shaken for 1 hour.

The beads were washed twice with 2 resin volumes of distilled water, and the amino acids were eluted with 3 washings of 2 bed volumes of 5M NH₃, these washings were pooled and lyophilized.

Each sample was analysed on the Beckman 120C amino acid analyser. A blank was also digested under the same conditions, containing everything but coat protein to detect any liberation of free amino acids by autodigestion of carboxypeptidase A.

(1) Edman degradation of the tryptic peptides

The method of Schroeder, Shelton, Balog-Shelton and Cormick (1963) in which the degradation is carried out on paper strips was used.

Tryptic peptides (T₁, T₂, T₃, T₄, T_β, and T₅) separated by paper electrophoresis on Whatman No. 1 were identified, dried and then cut into strips 1 x 7 cm which were used as a support for the degradation reaction.

Coupling of the peptides with PITC was facilitated by saturating the strips with 20% PITC in dioxane and suspending them (two strips per jar) in jars containing 15 ml of pyridine : dioxane : water (1:1:1) which were sealed with aluminium foil and incubated at 40°C for 3 hours.

Each strip was extracted twice with benzene (approximately 10 ml) for 1½ hours, then finally overnight. The strips were dried for 1 hour, and the PTH-amino acid was cleaved by suspending the strips on a glass rack for 7 hours at 10 mm Hg pressure in a dessicator containing separate beakers of glacial acetic acid and 5.7N HCl.

After drying overnight the strips were extracted with acetone, the extracts for each peptide were pooled and lyophilized at 40°C. The PTH-amino acids were identified by thin layer chromatography in solvent 4, solvent E, and solvent 1. Peptides T₂ and T₅ were also degraded using the method of Sjoquist, Bloomback and Wallen, (1961).

CHAPTER 3

RESULTS

(a) Characterization of the f1 phage coat protein

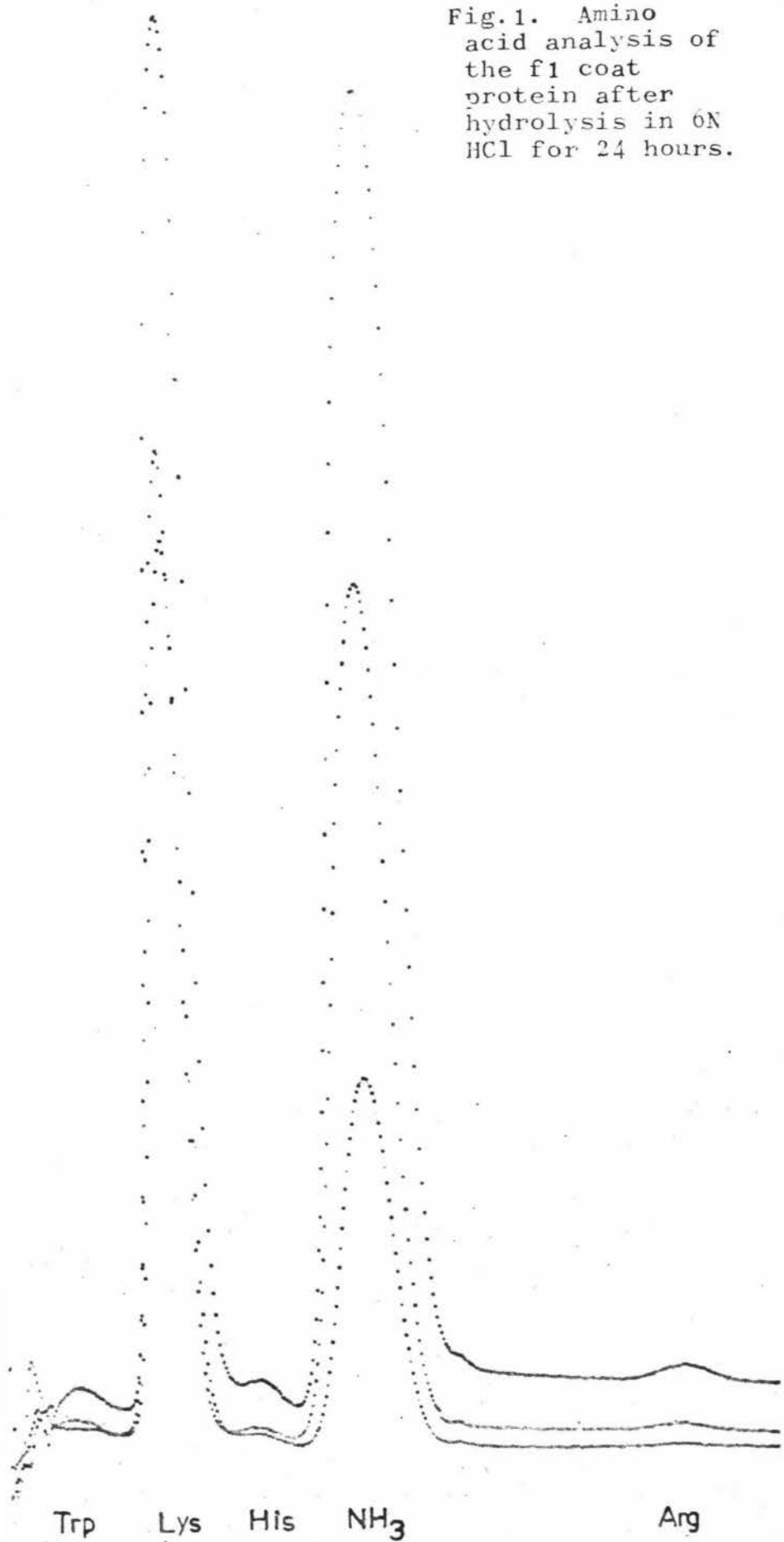
It was found that low concentrations (15%) of acetic acid were required to maintain the purified coat protein in solution, while at concentrations above 50% the protein aggregated into clear amorphous particles which rapidly precipitated. In addition to pH, prolonged storage at -20°C induced aggregation of coat protein solutions, at concentrations greater than about 5 mg/ml.

The $E_{280/260}$ ratio for the purified coat protein was 1.33 indicating a contamination with nucleic acid of approximately 1.1%. The concentration of the protein as determined by the Folin-Lowry method was 6.0 mg/ml.

(b) Amino acid analysis of the coat protein

Duplicate samples of 1.5 mg of protein were hydrolysed with acid for periods of 24, 48, 72 and 96 hours. The concentration of each amino acid present in the hydrolysate was calculated from chromatograms obtained from the Beckman amino acid analyser, by the height/width method using norleucine as an internal standard (fig. 1). The values obtained for each amino acid (μM) were converted to molar ratios by assuming that the molecule contained three aspartic acid residues. The molar ratios

Fig. 1. Amino acid analysis of the f1 coat protein after hydrolysis in 6N HCl for 24 hours.



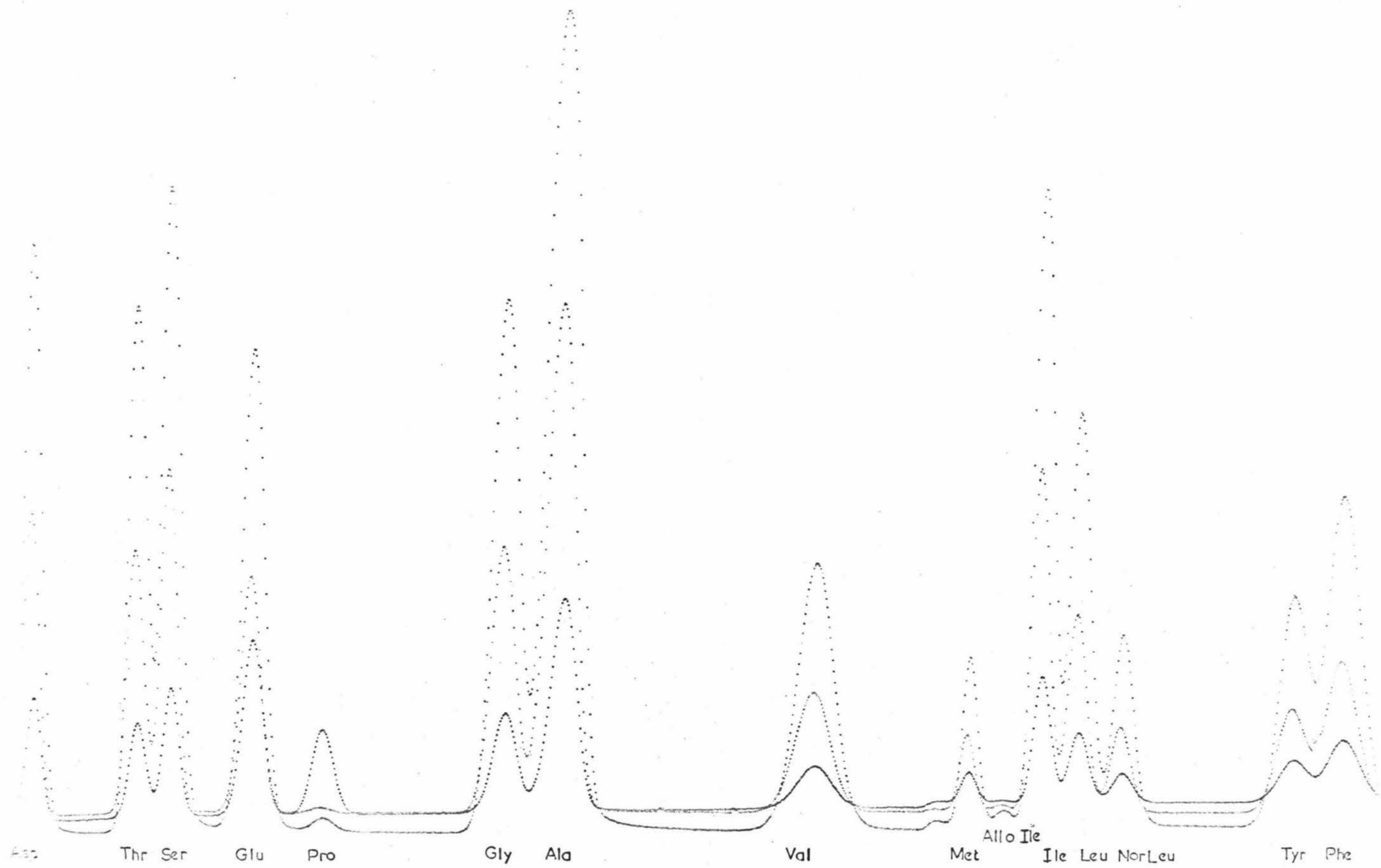


TABLE 1Amino acid composition¹ of the phage coat protein

Hydrolysis Time (Hrs)	24	24	48	48	72	72	96	96	Averaged Value	Nearest integral value.
Amino Acids										
Lys	4.4	5.0	5.4	4.3	4.6	5.2	4.1	4.1	4.6	5
Asp	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3
Thr	2.5	2.6	2.1	2.5	2.5	2.3	2.3	2.0	2.7 ²	3
Ser	3.4	2.7	3.0	3.0	2.2	2.5	2.2	2.5	3.3 ²	3
Glu	2.7	2.6	3.1	3.1	2.4	3.2	2.5	2.5	2.8	3
Pro	0.9	1.0	1.2	1.0	0.8	0.6	0.5	0.5	0.8	1
Gly	3.4	4.0	4.1	4.0	3.1	4.2	4.2	3.8	3.8	4
Ala	8.9	10.6	10.1	9.4	10.1	11.0	9.9	9.7	10.0	10
Val	1.8	2.6	3.7	3.7	3.5	4.2	4.2	4.3	4.2 ³	4
Met	0.5	0.4	0.7	0.7	0.8	0.8	0.8	0.8	0.7	1
Ile	2.9	3.2	3.8	3.5	2.9	3.7	3.7	3.5	3.6 ³	4
Leu	1.7	1.9	1.6	2.0	2.1	2.1	1.6	2.2	1.9 ³	2
Tyr	1.4	1.7	1.6	2.0	1.6	2.0	1.5	2.0	1.7	2
Phe	2.7	2.5	2.8	2.6	2.8	2.6	2.8	2.6	2.7	3
Trp									1.0 ⁴	1
									Total	49

NOTE:-

1. Values for the amino acid composition are expressed as moles of amino acid per mole of protein based on the assumption of 3 moles of aspartic acid per mole of protein.

2. Threonine and serine have been corrected for decomposition based on a standard decomposition of 5% and 10% respectively after 24 hours hydrolysis (Moore and Stein, 1963).

TABLE 1 (cont'd)

3. Values for valine, isoleucine and leucine were extrapolated to infinite time. (Moore and Stein, 1963).

4. Tryptophan was determined using the spectrophotometric method of Bencze and Schmid (1957).

TABLE 2

Values of S^1 Calculated from the Spectra of Known Mixtures of

Tyrosine and Tryptophan

S^1	R^2
-2.9	1.0
0.8	1.5
6.4	2.0
9.2	2.5

Coat protein

4.1 1.85³.

1. The value S is derived from the slope of the line tangential to the two peaks in the absorption curve, divided by A_{\max} and multiplied by 1000.

2. R is the molar ratio of tyrosine and tryptophan.

3. The molar ratio for the f_1 coat protein is 1.85, giving a value of 1 tryptophan residue per mole of protein.

of each amino acid are presented in Table 1. These values were corrected for decomposition of the labile amino acids threonine and serine, and for the slow release of such amino acids as leucine, isoleucine and valine, from peptide bonds resistant to hydrolysis. The results which are corrected to the nearest integral value are also presented in Table 1.

(c) Tryptophan analysis

The determination of tryptophan was carried out according to the method of Bencze and Schmid (1957). The absorption curve for 1.5 mg of coat protein dissolved in 0.1 N NaOH adjusted to pH 13.0 with 1.0N NaOH is shown in fig. 2a. The curve shows two absorption maxima at 285 $m\mu$ and 295 $m\mu$ due to tryptophan and ionized tyrosine respectively.

The tryptophan content of the protein was determined from the absorption curve as a tyrosine to tryptophan ratio by drawing a line tangential to the two absorption maxima. The slope a/b of this line which was divided by A_{\max} , the absorption maxima of the tyrosine peak, and multiplied by 1000 for convenience, was designated as S. The value of S for the coat protein (fig. 2a) was calculated from its absorption curve. Correlation of this value with the standard curve obtained from a plot of S versus the molar ratio of tyrosine to tryptophan for

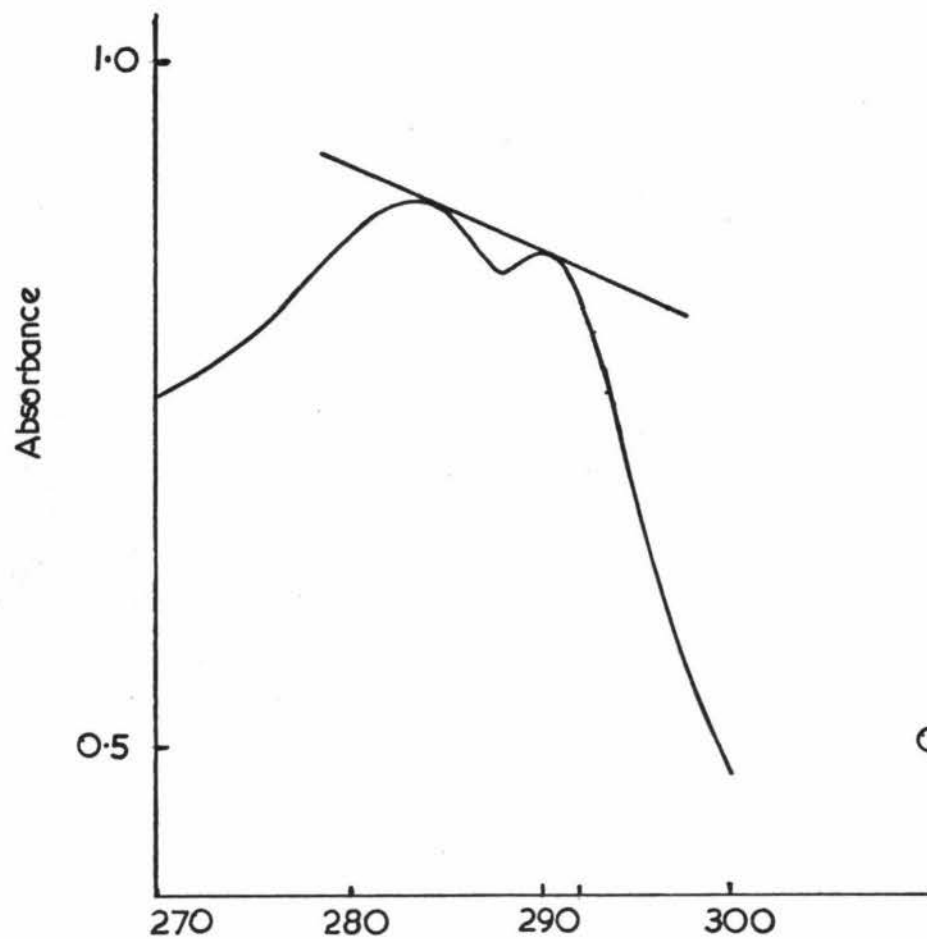


fig. 2a
 Ultra-violet spectra of the f_1 coat
 protein obtained at pH 13.0

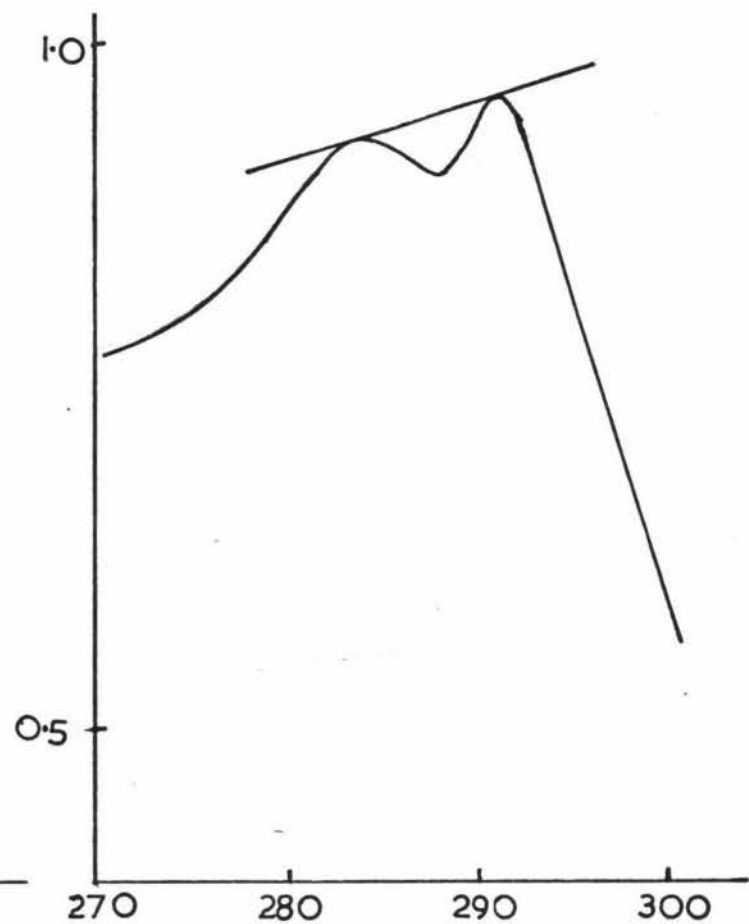


fig. 2b
 Ultra-violet spectra of the f_1 coat
 protein obtained at pH 12.5

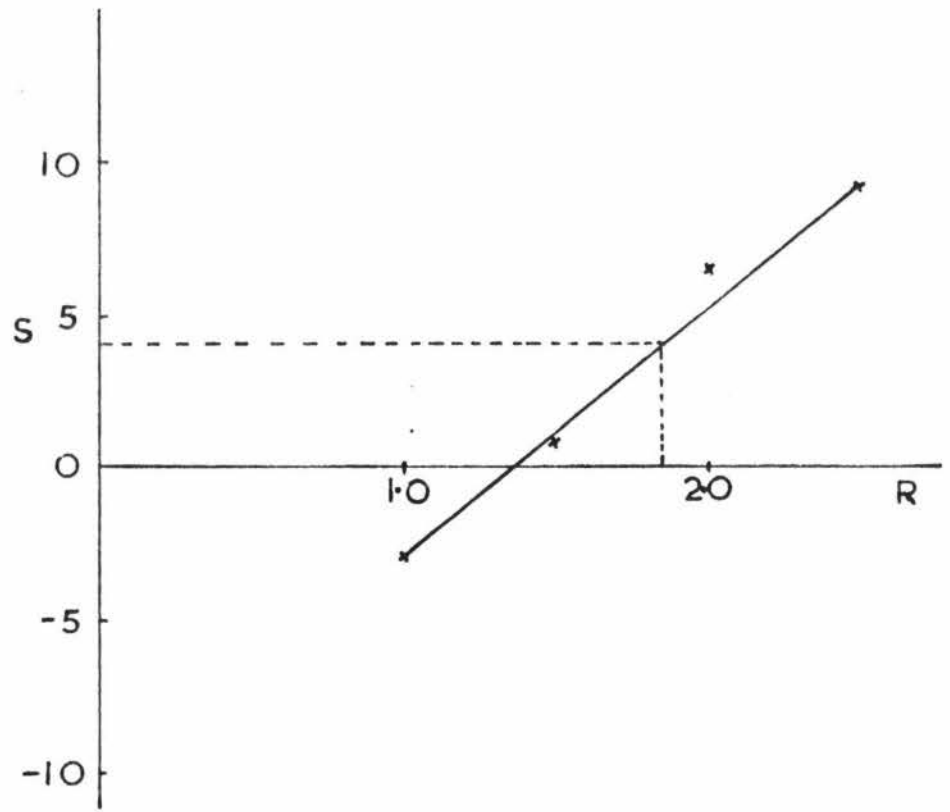


fig. 3 Value of S as a function of R for the Standard amino acid mixtures used for estimating the tryptophan content of the protein.

The value S is obtained from the slope of the line ($\frac{a}{b}$) tangential to the two tyrosine and tryptophan peaks divided by A_{\max} and multiplied by 1000.

R is the molar ratio of tyrosine to tryptophan.

standard amino acid solutions (fig. 3 and Table 2) revealed the molar ratio of tyrosine to tryptophan for the coat protein to be 1.85. This gives a value of 1 tryptophan residue per mole of protein.

Initially a variable molar ratio of tyrosine to tryptophan for the protein of 0.7-1.3 was obtained. This was attributed to the pH of the protein in 0.1 N NaOH being 12.5 rather than 13.0. Titration of the solution from pH 12.5 to pH 13.0 with 1.0 N NaOH changed the absorption spectra from one with a negative tangential slope for S (fig. 2b) at pH 12.5 to one with a positive tangential slope for S (fig. 2a) at pH 13.0. The phenolic hydroxyl group of tyrosine is not completely ionized until pH 13.0 and this caused the marked change in the absorption spectra.

The amount of extraneous absorption in the protein, measured by the absorption at 320 $m\mu$, compared with that for the mixture of tyrosine and tryptophan having the same molar ratio was negligible. In cases where extraneous absorption does occur the correction of Beaven and Holiday (1952) could be applied.

(d) Tryptic digestion of the coat protein

Digestion of denatured coat protein as described under methods for periods of 3-5 hours was found to give the optimal yield of peptides, without the appearance

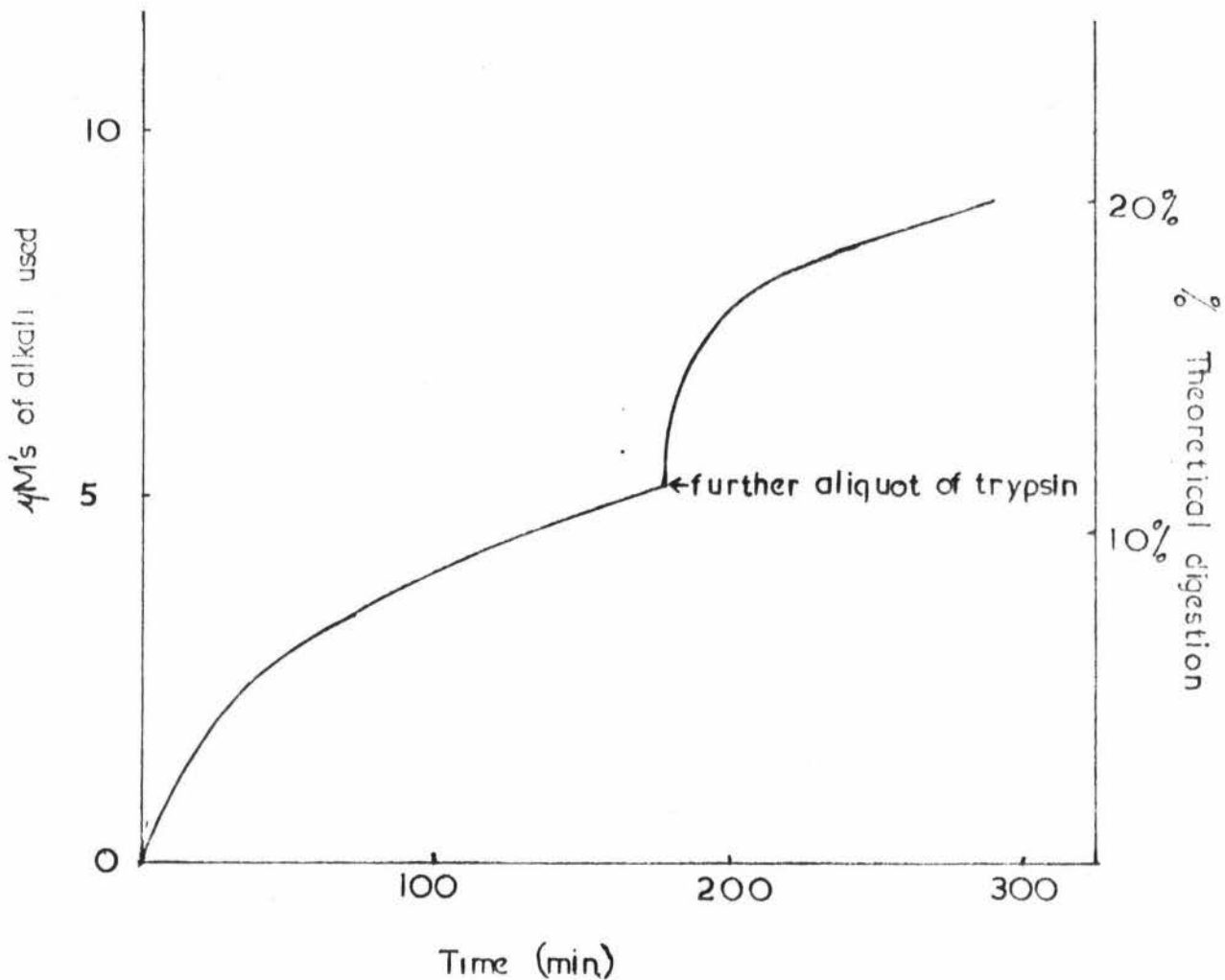


fig. 4 Tryptic digestion of the f_1 coat protein. Coat protein solution was denatured and an aliquot of trypsin added to give an initial enzyme substrate ratio 1:200. The protein was digested at 37°C for 3 hrs and then a further aliquot of trypsin was added to give a final enzyme substrate ratio of 1:100. The digestion was stopped after 5 hrs by adjusting the pH of the solution to pH 4.

of additional peptides due to non-specific cleavage.

Mild conditions of denaturation resulted in a low yield of peptides, whereas periods of hydrolysis longer than 5 hours resulted in a large number of peptides resulting from non-specific cleavage of the protein. The degree of digestion, calculated from the uptake of alkali and based on the complete cleavage of all the lysine C-terminal peptide bonds varied from 20-70%. A typical titration curve for the tryptic digestion of the coat protein is shown in fig. 4.

After digestion the peptides were lyophilized and taken up in water, the insoluble core material which remained could not be dissolved in any buffers or detergents. This core material is a large tryptic peptide which because of its hydrophobic nature is insoluble. This type of tryptic peptide has also been reported during tryptic digestion of the coat protein of the R17 bacteriophage (Weber, 1967).

(e) Separation of the tryptic peptides

After electrophoresis and staining of a marker strip with cadmium acetate-ninhydrin, five major peptides were found. These were called T_1 , T_2 , T_3 , T_4 , and T_5 in order of their migration, at pH 6.4 (fig. 5).

T_1 is an acidic peptide, T_2 is neutral, having the same mobility as the glycine marker. T_5 has an

electrophoretic mobility equal to that of the free lysine marker. During latter experiments T_2 another peptide was identified which had a mobility intermediate between that of T_4 and T_5 . Tryptic digestion for 24 hours frequently resulted in other peptides which gave a much fainter ninhydrin colour after electrophoresis than the tryptic peptides (fig. 5). These were due to non-specific cleavage of the protein by chymotryptic activity of the trypsin preparation and were disregarded for the purpose of this investigation.

The electrophoretic mobility of each peptide was expressed relative to the migration of the lysine marker (Table 3). The colour developed by the peptides T_1 to T_4 with the cadmium acetate-ninhydrin reagent was the characteristic pink colour developed by peptides, whereas T_5 had a deep purple colour similar to that of free lysine. The order of colour development of the peptides with cadmium acetate-ninhydrin was T_2 T_1 , T_5 T_3 , T_4 , T_5 . T_2 always gave a very intense colour reaction compared to the other peptides, and the electrophoresis of purified peptides (T_1 to T_5) invariably showed contamination by T_2 .

Treatment of a tryptic digestion (after digestion for 24 hours) with dansyl chloride and separation of the DNS-peptides formed by a 2-dimensional thin layer map (Atherton and Thompson, 1969) identified a total of

TABLE 3

The Electrophoretic Mobility (E_m) of the Tryptic Peptides at pH 6.4 in Pyridine Acetate Buffer

<u>Tryptic Peptide</u> ¹	<u>E_m</u> ²
T ₁	-0.38
T ₂	0.11
T ₃	0.51
T ₄	0.68
T	0.79
T ₅	1.0

1. Other peptides identified had electrophoretic mobilities relative to lysine of -0.69, -0.61, -0.51, -0.15, -0.06, 0.55, and 1.1.

2. The electrophoretic mobility of each peptide was measured relative to the free lysine marker.

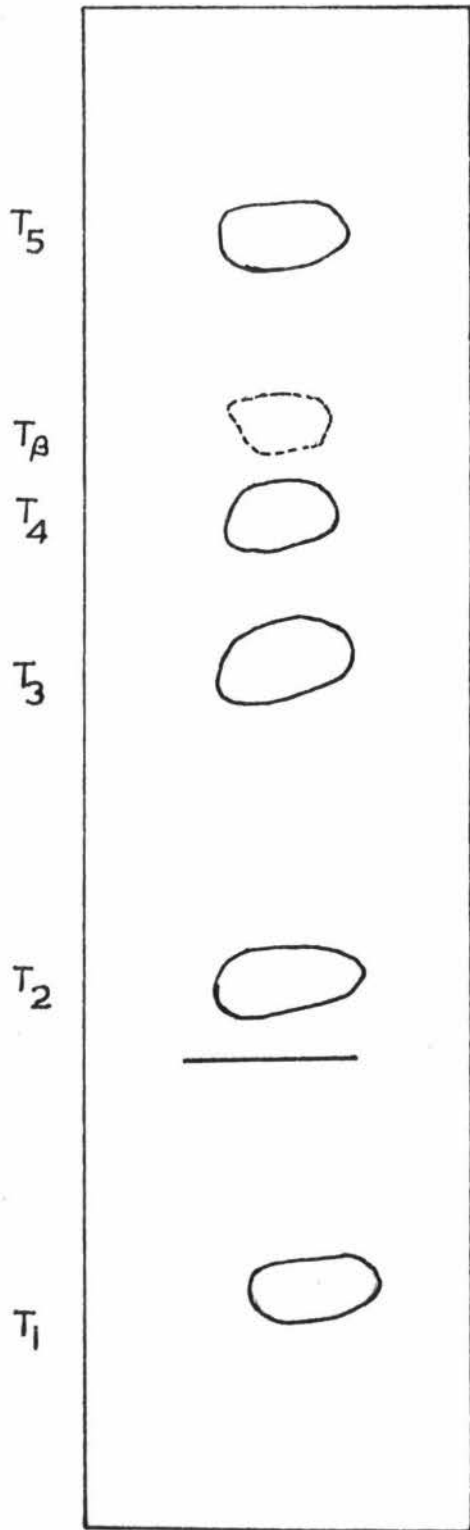


fig. 5 High voltage paper electrophoresis of the tryptic peptides, using pyridine : acetic acid : water buffer, pH 6.4.

After electrophoresis the paper was dried and the peptides identified using the cadmium acetate-ninhydrin reagent.

25 peptides (fig. 6). Most of these peptides were due to non-specific enzymatic cleavage of the coat protein.

Purification of large quantities of peptide by paper electrophoresis was necessary as the yield of peptide recovered was only 10-20% (Konigsberg and Hill, 1962). Because of the low yields of peptide from paper, further purification of peptides by a second stage of electrophoresis was not possible. In view of the low yields of purified peptide obtained from high voltage electrophoresis, attempts were made to purify the peptides from tryptic digests using AG 50W-X2 ion exchange resin. This method, although it did not give a measurably better purification of peptides (fig. 7) compared to electrophoresis, permitted larger quantities of peptide to be purified.

Paper electrophoresis of the pooled and lyophilized fractions under each peak showed the peptides to be contaminated with T_2 , similar to those peptides separated by paper electrophoresis. Since ion exchange chromatography was not significantly better than paper electrophoresis, the latter technique was used for isolating all the tryptic peptides characterized in this investigation.

This would suggest that there is a large amount of T_2 in each of the tryptic digests compared to the other peptides.

Chloroform: ethyl alcohol: acetic acid (38:4:3)

Methylacetate: isopropanol: ammonia (9:7:4)

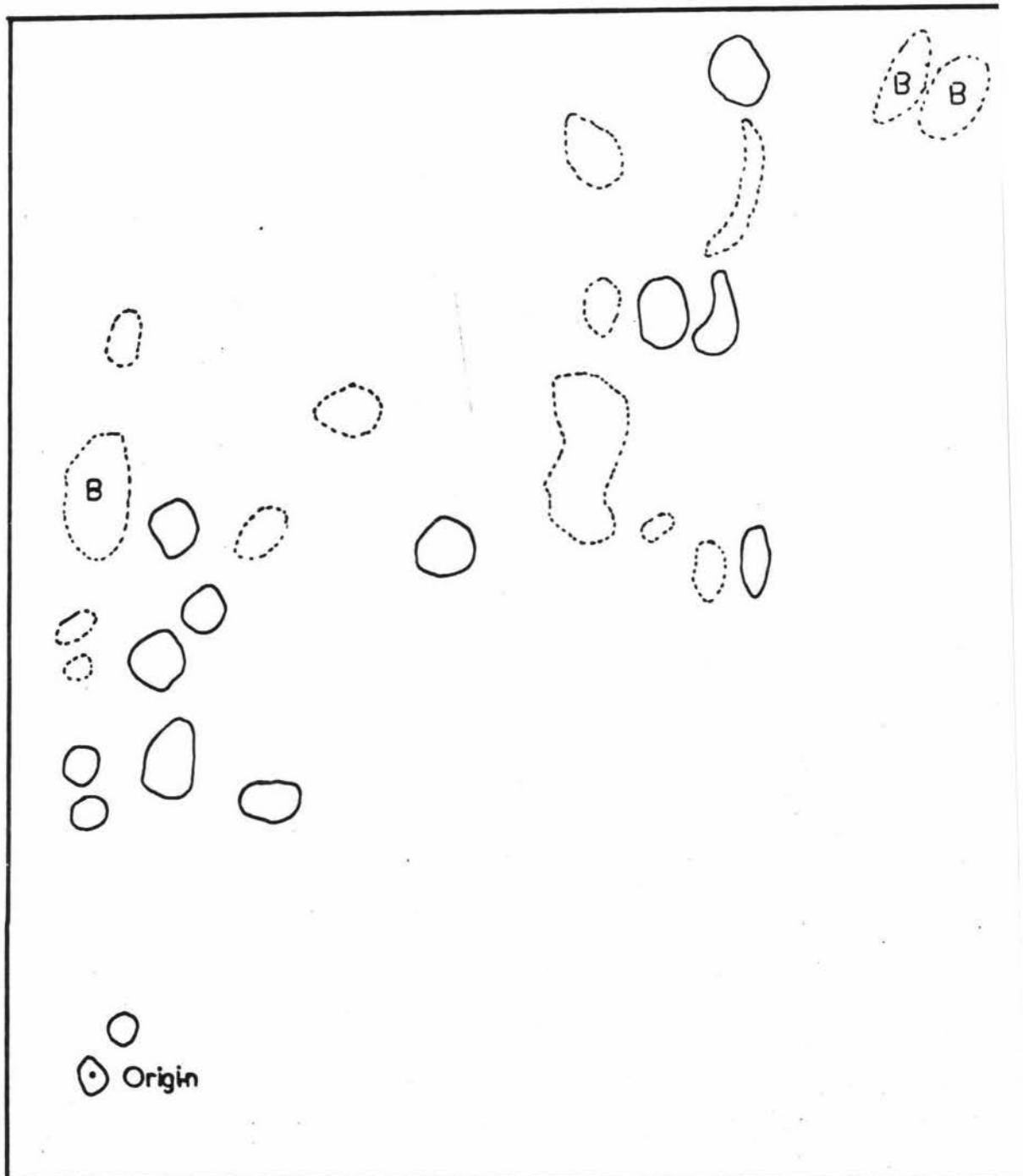


fig. 6 Thin layer finger print map of a tryptic digest after hydrolysis for 24 hours. The peptides are present as DNS-derivatives. B refers to the blue by-products of the coupling reaction.

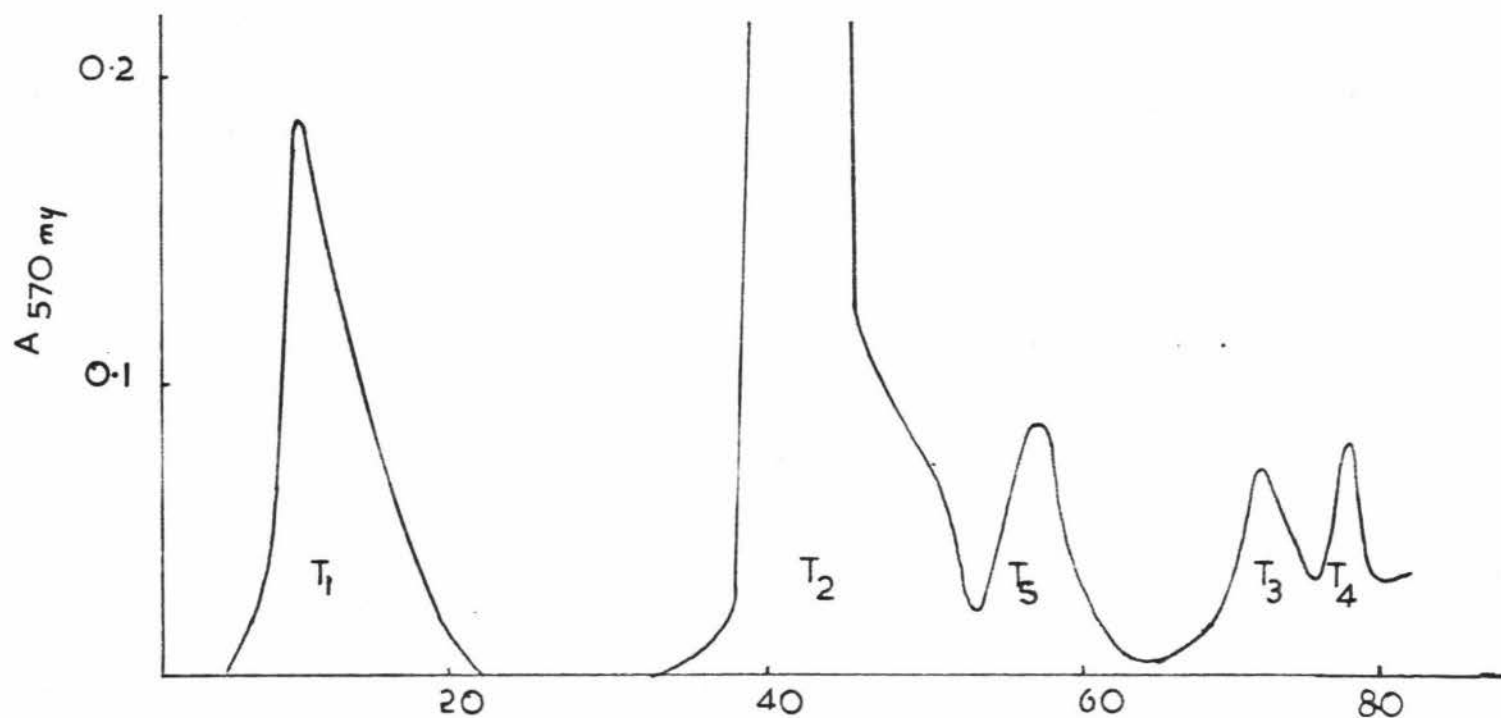


fig. 7 Fraction number

Separation of the tryptic peptides of the f1 coat protein on a AG 50W-X2 column (1 x 25cm). Approximately 4mg of peptide material was applied to the column and eluted with an exponential gradient from 0.2M pyridine acetate to 2.0M pyridine acetate buffer. The flow rate was maintained at 30ml/hr, 2ml fractions being collected. 10% of each fraction was subjected to alkaline hydrolysis followed by ninhydrin analysis.

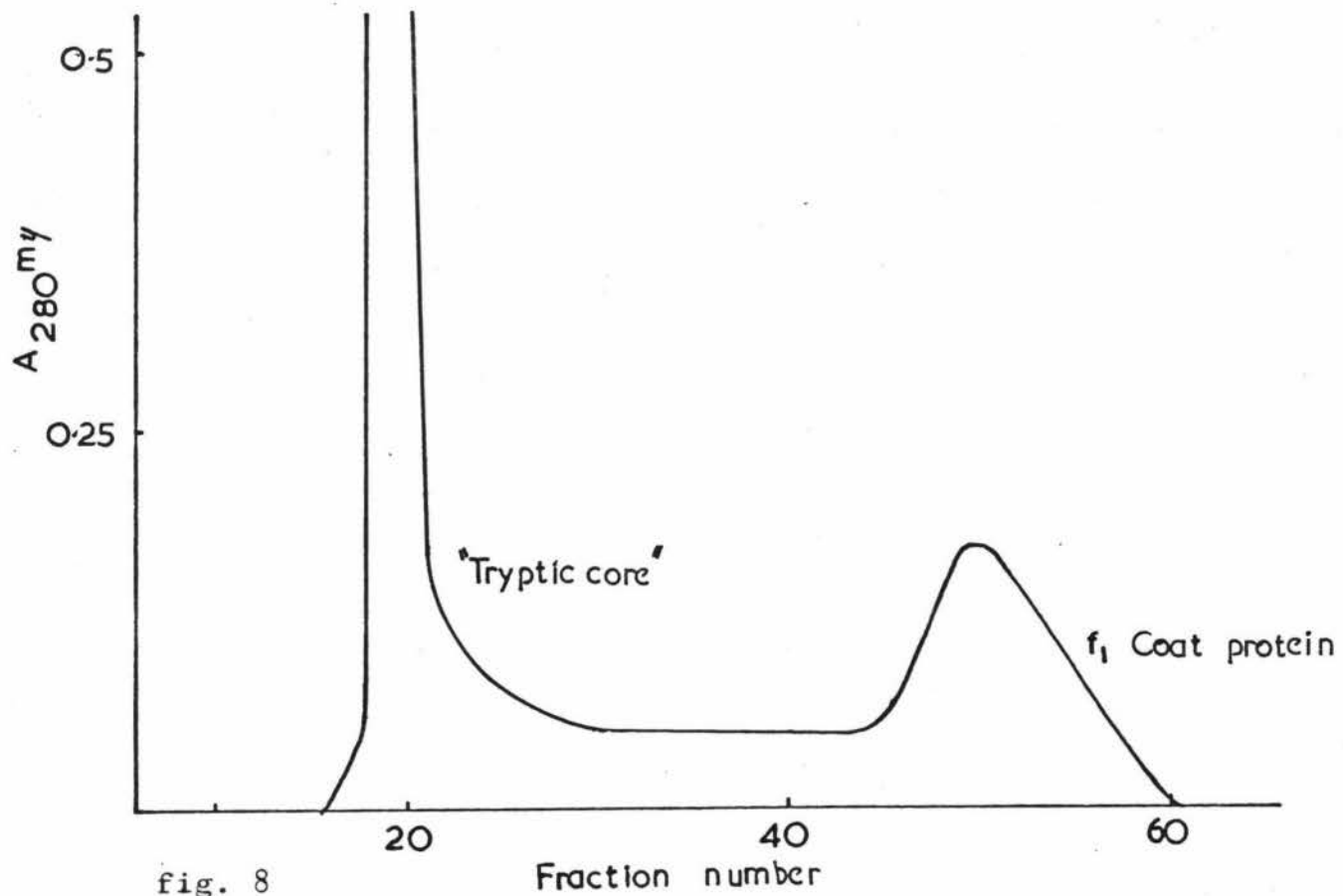


fig. 8

Separation of the core and coat protein by chromatography on Sephadex G-50.

(f) Separation of the tryptic core and f1 coat protein

Gel exclusion chromatography on Sephadex G-50 was used to separate the "insoluble" tryptic core material from the undigested protein. Initial experiments using Biogel P-2 and Sephadex G-25 resulted in the appearance of only one peak. Fig. 8 shows a typical elution pattern obtained after the chromatography of approximately 20 mg of the water insoluble material dissolved in 90% formic acid and eluted from a 1.7 x 100 cm column of Sephadex G-50 with 1 M formic acid.

Amino acid analysis of the material under the two peaks showed the first to be tryptic core material, as no proline was present. This peak corresponds to material with a molecular weight of greater than or equal to 30,000. The second peak which was very broad and diffuse was shown by amino acid analysis to represent the undigested protein. The broad peak obtained indicates the sample is heterogeneous and has a molecular weight of approximately 5,000.

(g) Amino acid composition of the tryptic peptides

The tryptic peptides isolated by paper electrophoresis were contaminated with small quantities of other peptides, mainly T₂. The composition of each peptide was determined after acid hydrolysis by amino acid analysis. The results

TABLE 4

Amino Acid Composition of the Tryptic Peptides of the f1
Phage Coat Protein

	<u>T₁</u>		<u>T₂</u>		<u>T₃</u>		<u>T₄</u>		<u>T₅</u>	
	μM	M.R.	μM	M.R.	μM	M.R.	μM	M.R.	μM	M.R.
Lys	.040	1.0			.030	1.0	.035	1.0	.01	1.0
Asp	.071	1.77								
Thr					.044	1.46				
Ser			.056	0.89	.045	1.48				
Glu	.046	1.15								
Pro	.047	1.17								
Gly	.046	1.15								
Ala	.082	2.04	.063	1.0						
Leu							.025	0.71		
Phe					.035	1.15	.032	0.91		

Results are expressed in micromoles (μM) of each amino acid. These results are converted to molar ratios (M.R.) by assuming there is one lysine residue per mole of tryptic peptide, except in the case of T₂. With T₂ it was assumed there was one alanine residue per mole of peptide.

(Table 4) are based on the assumption that except for T_2 each peptide will carry a lysine residue at the carboxyl end since trypsin specifically cleaves peptide bonds involving the carboxyl terminal lysine or arginine and the adjacent amino acid. Since amino acid analysis showed there was no arginine in the protein then all the tryptic peptides except the peptide derived from the C-terminal of the protein, must contain lysine. The composition of each peptide has therefore been calculated by assuming one molecule of lysine per peptide molecule except for T_2 .

The composition of each peptide based on the results in Table 4 is:-

T_1	Lys ₁ , Asp ₂ , Glu ₁ , Pro ₁ , Gly ₁ , Ala ₂ .
T_2	Ala ₁ , Ser ₁ . (The carboxyl terminal peptide)
T_3	Lys ₁ , Thr ₁ , Ser ₁ , Phe ₁ .
T_4	Lys ₁ , Leu ₁ , Phe ₁ .
T_5	Lys ₁ .

The degree of contamination with other peptide material varied from approximately 5% to 25% and undoubtedly this has made a significant contribution to the values for the amino acid composition, particularly in the case of T_3 .

If the lysine value for T_3 is too low compared to the other amino acids, then when the results are based

on lysine, this will give a higher set of figures than is actually the case. This is feasible since lysine is analysed on the short basics column and the other amino acids on the neutrals and acidics column of the amino acid analyser.

(h) N-Terminal sequence of coat protein

4.5 mg of denatured protein dissolved in 1 ml of water was subjected to 5 cycles of the Edman degradation, the PTH-amino acids being identified by thin layer chromatography. At the end of the fifth cycle no further N-terminal amino acids could be identified except for the by-products of the reaction, MPTU and DPTU.

After each cycle the chromatograms became obscured by increasing amounts of ultraviolet quenching material that remained at the origin of the plate, irrespective of the solvent system used.

The results of each degradation cycle are presented in Table 5 and fig. 9.

The N-terminal amino acid derivative (PTH-Ala) was scraped off the thin layer plate and eluted with absolute methanol at 40°C for one hour. The ultraviolet spectra of this amino acid derivative was determined between 230-300 m μ (fig. 10). The amount of PTH-Ala recovered, determined from its spectra (Fraenkel-Conrat, Harris and Levy, 1952) was 0.46 μ M, which represented a

TABLE 5Edman Degradation¹ of the f1 Coat Protein

<u>Stage</u>	<u>PTH - amino acid²</u>	<u>Solvents used³</u>
Cycle 1	Ala	E, and 2
Cycle 2	Glu	E, and 3
Cycle 3	Gly	E, and 2
Cycle 4	Asp	E
Cycle 5	Asp	E

1. The PTH-amino acid (PTH-amino acid) at each stage was identified by thin layer chromatography.

2. Therefore the N-terminal sequence of the f1 coat protein is:- NH₂-Ala-Glu-Gly-Asp-Asp-

3. For details of the solvent systems used see the methods section.

TABLE 6

Amino Acids Released During the Carboxypeptidase A Digestion of the f1 Coat Protein

Amino Acid Released	Digestion Time		
	1 hour	2 hours	11 hours
serine	0.09	0.10	0.70
alanine	0.05	traces	0.58
lysine	-	-	0.09

Values for amino acids are expressed as moles of amino acid released per mole of protein.

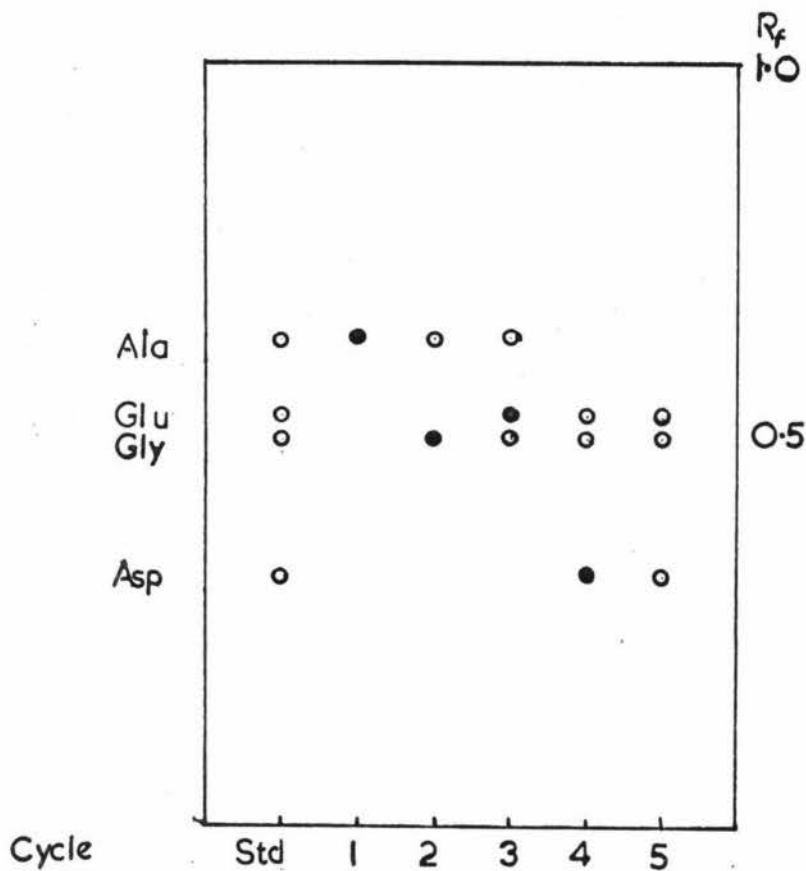


fig. 9 Edman degradation of the f1 coat protein and identification of the PTH-amino acids formed using solvent E.

● Major PTH-amino acid visible

○ Minor PTH-amino acid visible

These PTH-amino acid derivatives, except for cycle 5 were checked by chromatography in another solvent system.

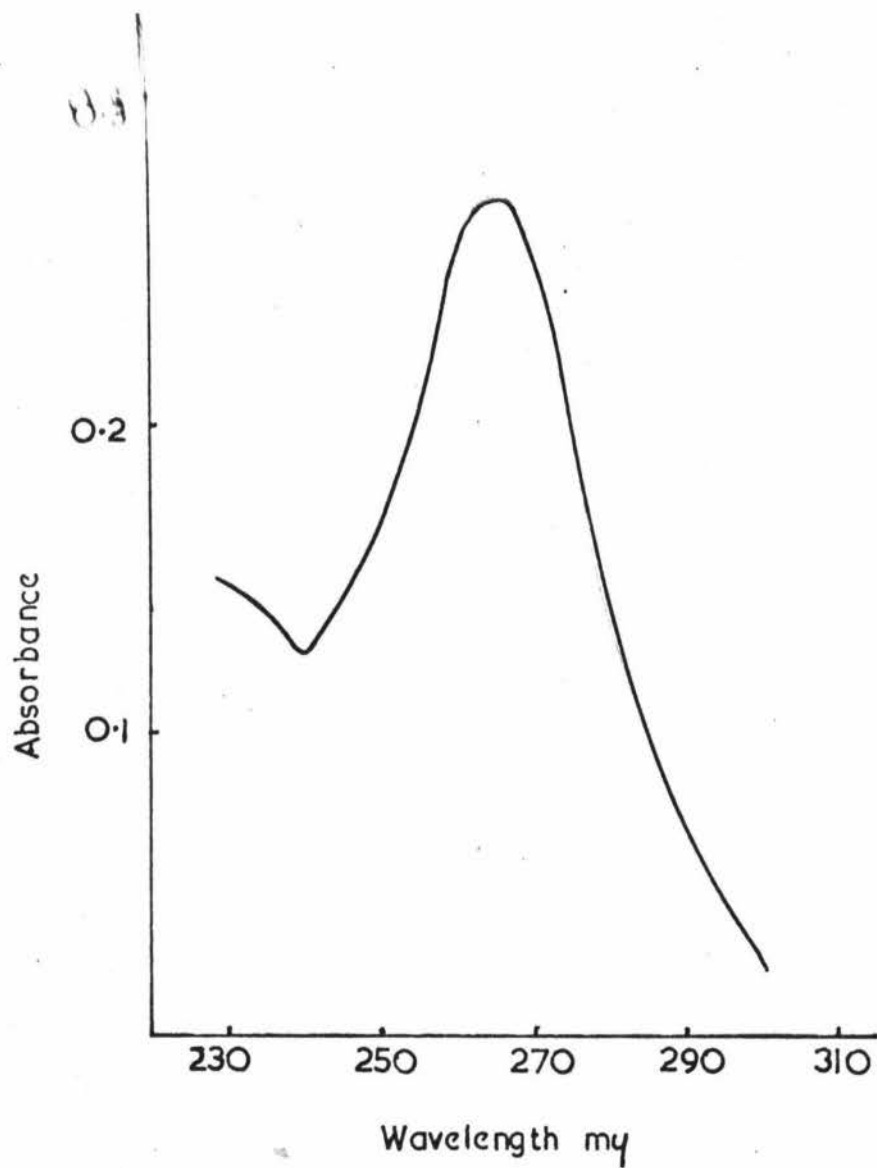


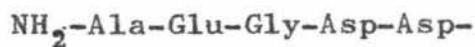
fig. 10 Ultra-violet spectra of the N-terminal PTH-Ala of the f1 coat protein in absolute methanol.

yield of 50% for the first degradation step.

The ratios $E_{260:265}$ (Fraenkel-Conrat *et. al.*, 1952) and $E_{245:265}$ (Sjoquist, 1957) were 0.89 and 0.46 respectively. These ratios give an indication of the purity of the PTH-amino acid. Both values indicate a small amount of contamination with non-cyclised material (probably PTC-Ala) since the spectra of pure PTH-Ala has a ratio $E_{260:265}$ of 0.85 and a ratio $E_{245:265}$ of 0.40. Non-cyclised PTC-Ala is identified by the fact that spectra of PTC-amino acids have a maximum at approximately 240 $m\mu$.

After the first cycle of the degradation contaminating PTH-amino acids from the previous step could be identified. Attempts to remove these contaminants by coupling twice with phenylisothiocyanate as suggested by Sjoquist (1959) did not improve the yield of the new PTH-derivative appreciably, or lower the degree of contamination from the previous cycles.

Since no new N-terminal PTH-amino acid could be identified from cycle 5 there is the possibility of having a -Asp-Asp sequence. The results from the Edman degradation of the coat protein show the sequence of the N-terminal is:-



(i) Digestion of the coat protein with carboxypeptidase A

The sequence of the carboxyl end of the coat protein was determined using the enzyme carboxypeptidase A. This enzyme sequentially removes amino acids from the carboxyl end of the protein (Fraenkel-Conrat *et. al.*, 1952). The contribution of amino acids released by autodigestion of the carboxypeptidase itself, determined by using a parallel blank solution containing all the components of the system except coat protein was negligible. Treatment of the carboxypeptidase A with diisopropylfluorophosphate to remove contaminating trypsin and chymotrypsin was found to be necessary, as the chromatograms showed evidence of the presence of small peptides and free amino acids. The results for digestion of the protein with carboxypeptidase A are presented in Table 6, aliquots being removed from the digestion mixture after 1, 2, and 11 hours digestion.

From these results the C-terminal sequence was deduced to be:-

-Lys-Ala-Ser-COOH

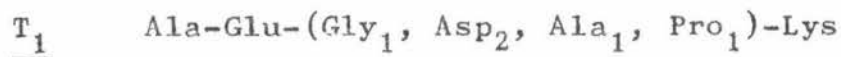
(j) The Edman degradation of the tryptic peptides

The N-terminal of each peptide previously purified by paper electrophoresis was determined by the Edman

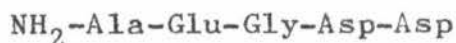
technique. Only the PTH-amino acids from the first two cycles could be positively identified, further successive cycles were obscured by an accumulation of by-products, MPTU and DPTU together with a large amount of ultraviolet quenching material which remained at the origin.

The results from each sequence are presented in fig. 11 and Table 7, together with the appropriate solvent systems used to identify the PTH-amino acids.

The results of the Edman degradation showed the sequence of the tryptic peptides to be:-



Since the Edman degradation of the protein (Chapter 3, section h) showed the N-terminal sequence to be:-



Peptide T_1 was probably derived from the N-terminal section of the protein. Hence the sequence of T_1 is:

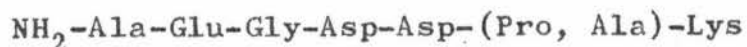


TABLE 7

Edman Degradation of the Tryptic Peptides

<u>Peptide</u>	<u>Cycle 1</u>	<u>Cycle 2</u>
T ₁	PTH-Ala	PTH-Glu
T ₂	PTH-Ala	PTH-Ser
T ₃	PTH-Phe	PTH-Thr
T ₄	PTH-Leu	PTH-Phe
T ₅	PTH-Lys	-
T _β	PTH-Pro	-

The N-terminal derivatives were identified at each cycle by co-chromatography on silica gel GF₂₅₄ thin layer plates.

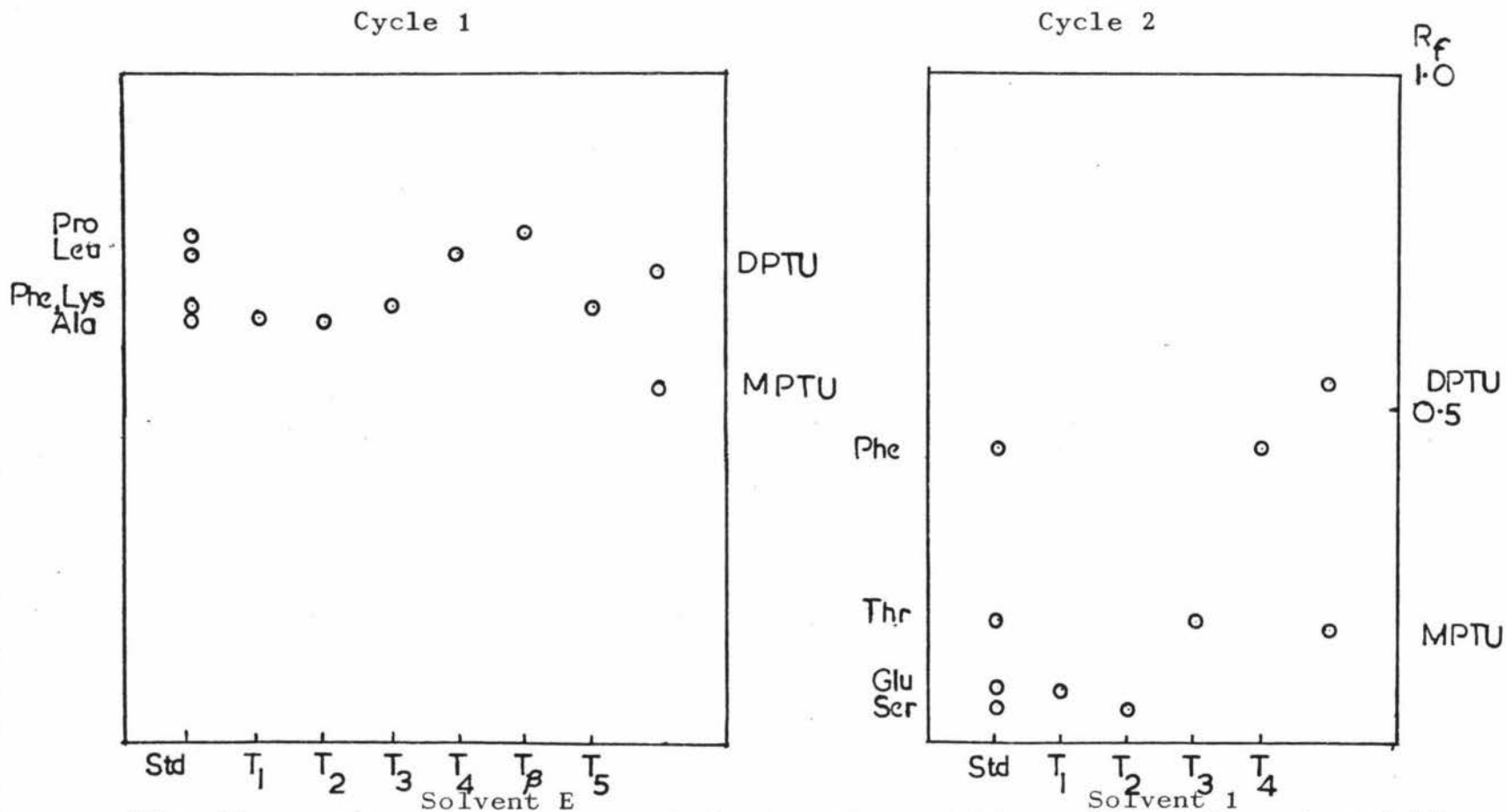


fig. 11 Edman degradation of the tryptic peptides. After Edman degradation of the peptides, aliquots of each derivative were co-chromatographed on Silica Gel GF₂₅₄ plates with known standard PTH-amino acid derivatives. The identity of the unknown PTH-amino acid derivatives was verified by co-chromatography in another solvent system.

T₂ Ala-Ser

This is the tryptic peptide derived from the C-terminal of the protein, as shown by carboxypeptidase A digestion of the protein (Chapter 3, section i), and from the known specificity of trypsin (Chapter 3, section g) since it contains no lysine.

T₃ Phe-Thr-(Ser, Lys)

Since Lys must be the C-terminal of the peptide, then the sequence of T₃ is:-

Phe-Thr-Ser-Lys

T₄ Leu-Phe-Lys

The Edman degradation gives the sequence Leu-Phe and the lysine must be the C-terminal residue.

T_β Pro

No further derivatives could be identified. Since there is only one proline residue per mole of protein, this probably was derived from T₁. The significance of

this will be discussed later.

T₅ Lys

T₅ was identified as a single lysine residue.

CHAPTER 4DISCUSSION(a) Purification of the f1 coat protein

Most of the studies to date on the group of small DNA bacteriophages have been on the coat protein of the fd bacteriophage. Since these bacteriophages belong to a group which are morphologically and serologically related, with coat proteins of similar amino acid composition (Asbeck et. al., 1969), it seems likely that the coat proteins of this group would also exhibit a similar behaviour in solution. What evidence that there is available would certainly indicate this.

The f1 coat protein was found to aggregate very easily in aqueous solutions. This behaviour is similar to that of the fd coat protein which Hoffmann-Berling, Kaerner and Knippers (1966) reported as existing as a uniformly sized oligomer (probably a heptamer) in a number of solvents and which polymerised in neutral solutions to give rod shaped aggregates. These workers found that high concentrations of detergents were necessary to prevent the formation of protein aggregates.

The existence of a second kind of protein subunit in the coat proteins of the f1 type viruses has been suggested on the basis of conditional lethal mutants

(Pratt, Tzagoloff, Eidah and Henry, 1967). Rossomando and Zinder (1968) have suggested that the disassembly of the f1 phage during infection is controlled by a small A protein which constitutes only a few percent of the total coat protein. Evidence for this is derived from studies with C¹⁴-histidine and C¹⁴-arginine which they show to be incorporated into the protein of the f1 bacteriophage. Since the analyses of the f1 coat protein indicate that these two amino acids are lacking, then it is assumed that they could only have originated from another protein. Possibly this could also explain the very small amount of arginine and histidine detected on the chromatograms during the amino acid analysis of the f1 coat protein (fig. 1) carried out in the present investigations. This would be consistent with the evidence of Rossomando and Zinder (1968), as the purification procedure used for the coat protein in the present investigation would also contain any other proteins that were present in the phage.

(b) Amino acid analysis of the f1 coat protein

The amino acid composition of a protein or peptide may be determined to within $\pm 3\%$ by using the Beckman 120C Amino Acid Analyser. This accuracy depends on a number of factors of which the principal ones are the experimental conditions, destruction of amino acids, racemization of amino acids and incomplete liberation of amino acids during hydrolysis.

During hydrolysis of the sample the oven must be maintained at $110^{\circ}\text{C} \pm 1^{\circ}\text{C}$ as there is chromatographic evidence of peptides still being present in the hydrolysate after 20 hours at 105°C (Moore and Stein, 1963). It is also necessary for the tubes to have been sealed under a high vacuum (0.01 mm Hg) to prevent losses of methionine and tyrosine due to oxidation (Moore and Stein, 1963).

During the present investigation it was assumed that the amount of threonine and serine destroyed by acid hydrolysis after 24 hours was 5% and 10% respectively (Moore and Stein, 1963). Desnuelle and Casal (1948) found that these amino acids are completely liberated during that time. The decomposition of serine and threonine apparently follows first order kinetics (Hirs, Stein and Moore, 1954) although some uncertainty does exist about the kinetics of the reaction (Kassell and Laskowski, 1961).

Tryptophan is completely destroyed during acid hydrolysis (Moore and Stein, 1963) presumably through cleavage of the indole ring, although after 24 hours hydrolysis there is generally some chromatographic evidence for the presence of tryptophan (fig. 1). Tryptophan was analysed spectrophotometrically during the present investigation using the method of Bencze and Schmid (1957). Asparagine and glutamine are also completely destroyed during acid hydrolysis, being converted to aspartic acid and glutamic acid respectively (Moore and Stein, 1963).

During prolonged acid hydrolysis racemization of isoleucine to allo-isoleucine occurs (fig. 1) (Piez, 1954). During the present investigation very little evidence of the presence of allo-isoleucine was found presumably because acid hydrolysis was only continued for periods up to 96 hours. Asbeck et. al. (1969) found appreciable quantities of allo-isoleucine in acid hydrolysates of the fd coat protein. However these workers used prolonged periods of hydrolysis for the fd coat protein of up to 200 hours.

The peptide bonds of leucine, isoleucine and valine are particularly resistant to acid hydrolysis, being only slowly released (Moore and Stein, 1963). In the present investigation leucine and isoleucine were nearly

completely liberated after 24 hours, although the isoleucine value was still low. A striking feature of the results is the slow release of valine (Table 1). Only 2.2 residues of valine per mole of protein were released after 24 hours hydrolysis, this increasing to 3.7 residues of valine per mole of protein after 48 hours. The Ile-Val peptide bond is known to be particularly resistant to acid hydrolysis (Harfenist, 1953), as is the Val-Val peptide bond (Christensen, 1943). Asbeck et. al. (1969) report a Val-Val-Val-Ile-Val sequence (residues 28-32) for the fd coat protein. A region such as this would be extremely resistant to acid hydrolysis and since it is likely f1 contains this same sequence, this could explain the slow release of valine.

The amino acid composition of the coat proteins of f1, fd, M13 and ZJ-2 (Asbeck et. al., 1969) are shown in Table 8. The lack of any histidine, arginine and cysteine in the coat proteins is a dominant feature of this group of proteins. In addition these proteins have a high content of aliphatic amino acids, especially alanine.

The difference between the amino acid composition of f1 (the present investigation) and that reported by Asbeck et. al. (1969) is 4 serine residues and 8-9 alanine residues in Asbeck's results compared to 3 serine and 10 alanine residues for the analysis in this

TABLE 8

Amino Acid Composition¹ of the f1, fd, M13 and ZJ-2 Coat Proteins

Amino Acid	f1 ³	f1 ²	fd ²	M13 ²	ZJ-2 ²
Lys	5	5	5	5	5
Asp	3	3	3	3	3
Thr	3	3	3	3	2
Ser	3	4	4	4	4
Glu	3	3-4	3	3	3
Pro	1	1	1	1	1
Gly	4	4	4	4	4
Ala	10	8-9	9	9	10
Val	4	4	4	4	4
Met	1	1	1	1	1
Ile	4	4	4	4	4
Leu	2	2	2	2	2
Tyr	2	2	2	2	2
Phe	3	3	3	3	3
Trp	1	1	1	1	1
Total	49	48-50	49	49	49

The only amino acids different in the composition of any of the proteins are Thr, Ser, Glu and Ala.

1. The results are expressed as moles of amino acid per mole of protein.
2. These results are taken from Asbeck et. al. (1969).
3. These results are obtained from the present investigation.

investigation. The amino acid replacement serine to alanine is acceptable since it only involves a change of U → G in the triplet codon. The amino acid composition suggested by Asbeck et. al. (1969) for the f1 coat protein indicates uncertainty over the number of glutamic acid and alanine residues. During the present investigation we have been unable to account for the differences between the results obtained in the present study and those reported by Asbeck and his co-workers for the composition of the f1 coat protein. In some investigations (Ikawa and Snell, 1961), evidence has been found of the formation of an ester of serine and glutamic acid which is formed during evaporation of acid hydrolysates, leading to a lower value for these two amino acids. However as there is no evidence on the amino acid analysis chromatograms of this ester being formed, this has been discounted.

There is a difference in the host used for isolation of the phage coat proteins in the two investigations. Asbeck and his co-workers used E. coli K12 as the host for growth of the phage, whereas Professor G.B. Petersen, from whom our phage preparation was obtained, used E. coli K13. However there does not seem to be any reason why this should change the composition of the f1 coat protein. Asbeck et. al. (1969) suggests f1 has one more glutamine residue and one less alanine residue than the fd coat protein. If it is assumed that this

change in composition depicts replacement of one alanine residue in fd for a glutamine residue in the f1 coat protein such a replacement would involve the change of two consecutive nucleotide bases which is unlikely.

If we now compare the differences in the amino acid composition of the f1 coat protein (the present investigation) and the results obtained by Asbeck et. al. (1969) for fd, M13 and ZJ-2 we find a replacement of alanine to serine for fd (Table 8). This can be explained in terms of a single nucleotide base change in the triplet codons of G → U (Dayhoff, 1969).

Ala:- GCA, GCG, GCC, GCU →

Ser:- UCA, UCG, UCC, UCU, AGC, AGU

The amino acid composition of fd and M13 is identical. The difference between f1 and ZJ-2 is a replacement of threonine to serine. This can also be explained by the triplet code in terms of a change of A - U.

Thr:- ACA, ACG, ACC, ACU →

Ser:- UCA, UCG, UCC, UCU, AGC, AGU

Single nucleotide base changes in the genetic code

were also found to explain the differences in the amino acid composition of the coat proteins of R17, MS2 and f2. The amino acid analysis of the coat protein of the fr phage has indicated this is an entirely different protein (Weber, 1967; Enger and Kaesburg, 1965; Lin, Tsung and Fraenkel-Conrat, 1967; and Wittmann-Liebold, 1966).

The molecular weight of the f1 coat protein determined from the amino acid analysis is 5012. This excludes the contribution due to any amide groups, which may be present. The molecular weight of the fd coat protein is 5169, and this is reputed to be the smallest protein having a known structural function (Braunitzer et. al., 1967).

(c) Tryptophan analysis

The coat proteins of f1, fd, M13 and ZJ-2 (Asbeck et. al., 1969) were found to contain one tryptophan residue per mole of protein.

Because spectra of the f1 coat protein in alkali showed a bathochromic shift in the absorption maxima of up to 4 μ (fig. 2), the method of Goodwin and Morton (1946) which employs absorbance measurements at a fixed wavelength was not considered to be sufficiently accurate.

The Bencze and Schmid technique can only be used for proteins having no extraneous absorption. Two assumptions that had to be made were that the incorporation

of tyrosine and tryptophan into protein did not change the extinction coefficients of these amino acids (Beaven and Holiday, 1952), and that the tyrosine residues in the protein were in the fully ionized state.

Initially the coat protein was dissolved in 0.1 N NaOH and the spectra recorded (fig. 2b). This was subsequently shown to give spectra with a tyrosine to tryptophan ratio of 0.7 to 1.3. Measurement of the pH of the solution showed it to be 12.5. Titration of the coat protein (dissolved in 0.1 N NaOH) to pH 13.0 with 1 N NaOH, immediately gave an ultraviolet spectra with a tyrosine to tryptophan ratio of 1.85 (fig. 2a). Incomplete ionization of tyrosine at pH 12.5 caused the low tyrosine to tryptophan ratio earlier noted.

Crammer and Neuberger (1943) found the tyrosine residues of ovalalbumin behaved abnormally and adjustment to pH 13 was necessary for maximum ionization.

(d) Tryptic digestion of coat protein and separation of the tryptic peptides

Trypsin possesses the highest fundamental specificity of all the known proteolytic enzymes, being specific for the peptide bond between the carboxyl group of lysine or arginine and the amino group of the adjacent amino acid (Bergmann, Fruton and Pollock, 1939). Cleavage of the peptide bond is slower when the basic residue is

adjacent to an acidic residue and does not occur at all when the residue is followed by proline (Tsugita, Gish, Young, Fraenkel-Conrat, Knight and Stanley, 1960). Transpeptidation, the limitation of enzymatic hydrolysis has not been shown to occur to any appreciable extent during tryptic digestion (Katchalski and Lela, 1959).

Rigorous heat denaturation of the f1 coat protein followed by tryptic digestion for a period of 4 hours and separation of the soluble peptides by paper electrophoresis indicated only low yields of hydrolysis products. Asbeck et. al. (1969) report aggregation of the fd coat protein after heat denaturation, leading to poor enzymatic digestions. Presumably the aggregation of the protein being due to the inability to denature the hydrophobic core region.

Longer periods of digestion produced a large number of peptides due to non-specific cleavage (fig. 6).

These peptides could be due to:

1. the hydrolysis being conducted at 37°C instead of 25°C. This is known to lead to the hydrolysis of less susceptible peptide bonds (Smyth and Elliott, 1964).
2. commercial preparations of trypsin have been found to contain approximately 0.1-0.5% of chymotryptic activity (Young and Carpenter, 1961). Trypsin preparations which were used during this investigation

were kept in 0.001 N HCl at 4°C for 4 hours. These conditions are known to reduce the chymotryptic activity (Guidotti et. al., 1962).

Cleavage during tryptic hydrolysis of peptide bonds normally cleaved by chymotrypsin has been demonstrated in the case of polypeptides and synthetic ester substrates such as acetyl-L-tyrosine ethyl ester (Inagami and Sturtevant, 1960). Cleavage of this ester occurs at a rate 1/12th of that during chymotryptic hydrolysis (this hydrolysis cannot be accounted for by chymotrypsin contamination). During a 50 hour tryptic hydrolysis of glucagon (25°C, pH 7.8), cleavage occurred at lysine and arginine and also the C-terminal of phenylalanine and tryptophan (Bromer, Staub, Sinn and Behrens, 1957b). Digestion for 2¼ hours resulted in cleavage only at the lysine and arginine bonds. There is evidence of chymotryptic activity being an inherent part of the trypsin molecule as chromatography of trypsin will not remove the chymotryptic activity (Cole and Kinkade, 1961).

Problems were encountered with the separation of tryptic peptides by paper electrophoresis. To obtain good separation of the peptides, after soaking the paper in buffer it had to be dried with blotting paper until it was only slightly damp. The origin line where

the peptide mixture was applied had to be dried further still with a thin strip of blotting paper to stop excessive diffusion of the sample, before the current is applied. If the paper was too wet, then the peptides diffused excessively during electrophoresis causing the peptide bands identified to be diffuse and poorly separated.

The temperature setting of the apparatus influenced separation, the margin between the paper overheating and freezing being small. During electrophoresis the temperature was monitored at approximately 4°C.

The main problem encountered in separating peptides by electrophoresis was the poor recovery from the paper after the peptide band was identified, yet this method has the advantage of being quick and simple. Eluants for washing the identified peptide band off the paper ranged from distilled water, and 5% acetic acid to 75% formic acid. This problem was also reported by Konigsberg and Hill (1962), who obtained recoveries of peptides after electrophoresis between 10-20% depending on the peptide.

The two classes of tryptic peptides resulting from the digestion of f1 protein could be classified as "soluble" and "insoluble". The "soluble" tryptic peptides, soluble in water or pyridine acetate buffer,

could be separated into five major bands by high voltage electrophoresis, whereas the "insoluble" tryptic core could only be isolated on Sephadex G-50 using 50% acetic acid as the solvent.

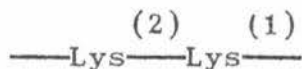
The tryptic peptide T_2 was isolated in larger quantities than any of the other peptides, suggesting that the peptide bond Lys-Ala was cleaved much more readily than any other peptide bonds.

Since the N-terminal group in T_β was proline then the peptide must have originated from T_1 as there is only one proline residue per mole of protein. This is conceivable since, removal of a peptide having the composition Pro, Ala and Lys from T_1 would result in a new acidic peptide having the composition Ala₁, Glu₁, Gly₁, Asp₂. This could explain one of a large number of acidic peptides observed in minor amounts on various occasions. There is evidence during the studies on the TMV coat protein of the lability of the peptide bond, Asp-Pro (Anderer, 1959) in weak acids.

T_5 which was shown to be free lysine must have been obtained from a region in the coat protein with the sequence -Lys-Lys-. Matsubara and Smith (1963) found free lysine from tryptic digests of cytochrome c, as well as tryptic peptides containing lysine as the N-terminal residue, which probably resulted from a

Lys-Lys sequence. During this investigation a basic peptide was identified which had an electrophoretic mobility slightly greater than free lysine. Since only the major peptides were isolated during the present investigation, this peptide was not further characterized.

The presence of tryptic peptides with lysine as the N-terminal residue is due to the inability of trypsin to cleave a peptide bond adjacent to free amino group.



If peptide bond (1) is cleaved first, then peptide bond (2) may also be cleaved, resulting in the liberation of free lysine. If peptide bond (2) is cleaved first, then peptide bond (1) cannot be cleaved, resulting in a peptide containing an N-terminal lysine.

The "insoluble" tryptic core peptide was highly aggregated during elution on Sephadex G-50 using 50% acetic acid as the eluant since the position at which it was eluted represented a molecular weight of greater than or equal to 30,000. This can be deduced from the known characteristics of Sephadex G-50 which is capable of separating polypeptides with a molecular weight of up to 30,000. The "insoluble" core material was distinguished from the intact f1 protein by amino acid

analysis. The "insoluble" peptide was heterogeneous and could not be purified.

Asbeck et. al. (1969) also experienced problems with an insoluble peptide from fd which was shown to have a molecular weight of greater than 200,000. These workers proposed that the peptide which resulted from an extremely hydrophobic region had aggregated to give a polypeptide 80 times its normal size. Even the broad peak obtained for the f1 coat protein in the present investigation indicated that it was heterogeneous.

The f1 coat protein has a molecular weight of slightly greater than 5000 as demonstrated by amino acid analysis. Hence removal of the regions susceptible to trypsin should, if one compares by homology the f1 and fd coat proteins leave a core with a molecular weight of 3,000. Hence for the core material to be eluted first from a column of Sephadex G-50, it must have been extensively aggregated. This mixture of core and coat protein should have been capable of being separated on Sephadex G-25, on which the mixture gave only one peak.

(e) Edman degradation of the f1 coat protein and the tryptic peptides

The number of steps which the Edman degradation can undergo is theoretically unlimited. The problem arises

after several steps of degradation when the presence of excessive amounts of by-products gradually obscures the analysis. Edman and Begg (1967) using the automated protein sequenator report yields of 98.5% at each step. All attempts to obtain a higher yield at each step have failed.

If we consider a degradation reaction which has a yield of 98.5% at each stage, then we can calculate the theoretical yield of product.

Step 1	98.5%
Step 2	97.0%
Step 3	95.5%
Step 4	94.0%
Step 5	92.5% etc.

The N-terminal sequence studies on the f1 coat protein indicated a yield of only 50% at the first step. Initially this low yield was attributed to incompetent handling of the technique. However subsequent studies on synthetic peptides indicated that this was not the problem, rather it was the characteristics of the protein that caused poor yields of product. This can be verified by the studies of Asbeck et. al. (1969) who were unable to obtain a yield of the N-terminal PTH-Ala in excess of 40-50%. On this basis if one considers the theoretical yield at each step we have:

Step 1	50.0%
Step 2	25.0%
Step 3	12.5%
Step 4	6.25%

Beyond this stage there is too much contamination to positively identify the N-terminal derivative. PTH-Asp could be identified at both step 4 and 5 but no new derivatives could be seen at step 5. Consequently placing aspartic acid at position 5 can only be tentatively inferred by homology with the fd coat protein. The most likely reason for a poor yield of the N-terminal derivative of the protein is the effect which aggregation of the protein has on the degree of coupling of the N-terminal derivative of the protein with PITC. If the protein even in the denatured state is extensively aggregated, then we could expect steric reasons alone to account for a poor coupling reaction.

Sequencing of the tryptic peptides using the method of Schroeder et. al. (1963) was only successful, if after paper electrophoresis, the peptides were identified by a marker band and the peptides coupled with phenylisothiocyanate while using the chromatography paper as support. By using this method two Edman degradation steps were completed. Again this was limited because of the presence of by-products in the

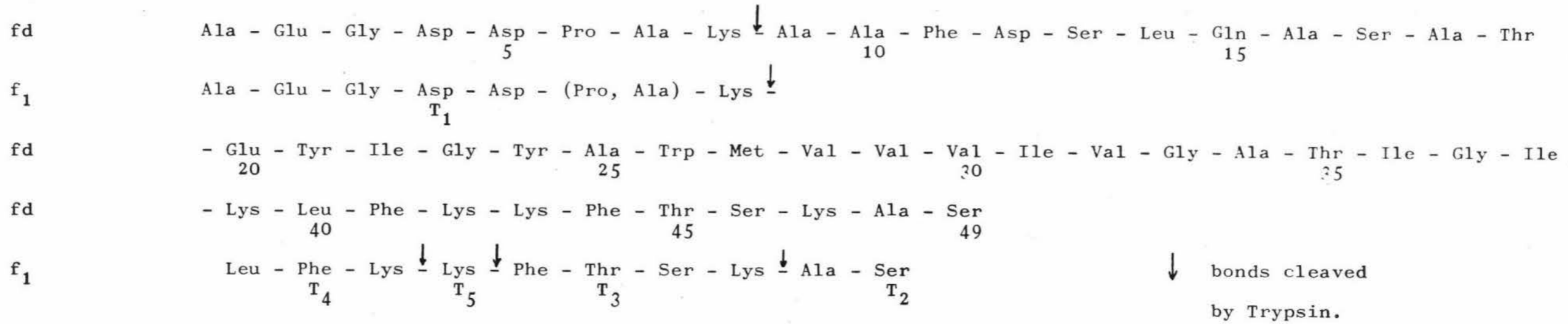
third degradation step. However these two sequencing steps were enough to completely determine the sequence of all the tryptic peptides except T_1 . Further Edman degradation steps would have been possible if a higher yield of peptides was obtained from paper electrophoresis. To have washed the peptides off the paper prior to Edman degradation would have resulted in a very poor recovery of peptide, certainly not enough for a sequence determination.

The sequence of the tryptic peptides of f1 determined in this investigation is the same as the tryptic peptides of the fd coat protein. This means that any differences in the composition between f1 and fd exist in the insoluble tryptic core peptide.

The sequence of the tryptic peptides of the f1 coat protein can thus be aligned by homology with those of the fd coat protein (fig. 12). Due to the low yields of each peptide obtained, there was certainly insufficient material at each stage for an amino acid analysis and a sequence analysis from the same sample. The only peptide with which no problems were experienced was T_2 as there was approximately 10 times more than the other peptides.

The Edman degradation reaction itself has presented many problems in previous investigations (Smyth et. al., 1963). Difficulty has been experienced during the cyclization reaction in strong acid (100°C for 1 hour in

fig. 12 Comparison of the amino acid sequence of the f1 coat and fd coat proteins.



The sequence of the fd coat protein is obtained from Asbeck *et. al.*, (1969). The sequences of the tryptic peptides from the f₁ coat protein were obtained during this investigation.

acetic acid:HCl) with peptides containing a glutamine residue at the amino terminal of a peptide since this rapidly cyclises to a pyrrolidone carboxylic acid residue, which cannot be released from the peptide. Internal serine and threonine residues become O-acetylated and as soon as this residue becomes the amino-terminal, an O → N shift of the acetyl group occurs. Peptides which contain aspartic acid or asparagine can undergo imide formation and then hydrolysis of the imide to form a β-peptide can occur, thus preventing further degradation.

These side reactions even occur when longer reaction times and lower temperatures are used, as the half lives of the first two reactions of the Edman degradation are one hour at 50°C. This means that the conditions used in the present investigation (40°C for 2 hours) could cause problems. None of these problems were found to occur during the Edman degradation of the intact f1 coat protein. This would also indicate that the glutamic acid residue at position 2 is not glutamine as further degradation steps were obtained in good yield. If there is a change of glutamic acid in the fd coat protein for glutamine in the f1 coat protein, then it is possibly one of the residues in the core region.

The paper strip modification of Schroeder et. al. (1963) was used during the present investigation because Schroeder claims that none of the above problems occur

when the degradation is carried out on paper. In fact we found no evidence of these problems occurring during the degradation of the peptides using Schroeder's method, even though the peptides contained amino acids that were readily susceptible to the modification reactions.

(f) C-Terminal Sequence of the Coat Protein

The prerequisite for carboxypeptidase A digestion of a protein is the presence of a free α -COOH group (Fraenkel-Conrat et. al., 1952). With this enzyme some amino acids are cleaved rapidly from the C-terminal while others are either released slowly or not at all.

Long periods of digestion (11 hours) were necessary to determine the C-terminal sequence of the f1 coat protein. After 11 hours there was a 74% release of serine, a 61% release of alanine and a 9% release of lysine (calculated on the basis of the number of micromoles of protein used). Digestion for 2 hours only released 10% of the C-terminal serine. Serine and lysine are two of the amino acids which are only slowly released by carboxypeptidase A, whereas alanine is rapidly released (Ambler, 1967). The poor initial yield of amino acids from the C-terminal end using carboxypeptidase A can be partly attributed to the C-terminal serine.

These results indicated that the f1 coat protein was extremely resistant to digestion with carboxypeptidase A, presumably for steric reasons.

Asbeck et. al. (1969) found the tendency for the fd coat protein to aggregate made it very difficult to digest the coat protein with carboxypeptidase A. They found carboxypeptidase A digestion of the fd coat protein without prior denaturation did not yield any C-terminal amino acids, whereas after denaturation these amino acids were released, although in low yield (10%) (Braunitzer et. al., 1967; Asbeck et. al., 1969). Braunitzer et. al., (1967) proposed that the C-terminal of the protein which is hydrophilic is responsible for the exchange mechanism with the DNA. The similar C-terminal sequences of the f1, fd, TMV vulgare and TMV dahlemense coat proteins provides support for this hypothesis (Table 9).

TABLE 9

Comparison of the C-Terminal Sequences of the f1, fd,
TMV vulgare and TMV dahlemense Coat Proteins.

<u>Virus</u>	<u>C-Terminal Sequence</u>
Phage f1	Lys - Ala-Ser-OH ⁴⁹
Phage fd ¹	144 Phe-Thr-Ser-Lys - Ala-Ser-OH ⁴⁹
TMV vulgare ²	152 Trp-Thr-Ser-Gly-Pro-Ala-Thr-OH ¹⁵⁸
TMV dahlemense ³	152 Trp-Thr-Ser-Ala-Pro-Ala-Thr-OH ¹⁵⁸

1. Asbeck et. al. (1969)
2. Nii and Fraenkel-Conrat (1955)
3. Wittmann (1965)

(g) General Summary

If the known sequence of fd and f1 is compared the differences in the composition of each protein must exist in the sequence of the insoluble core peptide. If it is assumed that the sequence of f1 and fd is identical except for the changes in composition then we obtain the comparative sequence as shown in fig. 12. This can be tentatively inferred as the coat proteins of the RNA phages R17, M12, MS2 and f2 have all been shown to be similar except for single amino acid replacements. The exception to this is the fr coat protein which differs by 20 amino acids (Enger and Kæsborg, 1965).

The model of the fd protein (Braunitzer et. al., 1967) suggests:

Constitution of side chains of fd and their probable function

Chain character	hydrophilic (acid)	hydrophobic (acid)	hydrophilic (basic)
Part of chain	residues 1-20	residues 21-34	residues 35-49
Function	Stabilization of virus in aqueous media	Polymerization of protein (quaternary structure)	Exchange mechanism with the DNA

This model also fits the f1 coat protein since

1. Only a 50% yield of the N-terminal derivative of the f1 coat protein may be obtained by Edman degradation, similar to that for the fd coat protein.
2. Poor yields of all the tryptic peptides except T₂, the C-terminal peptide, were obtained.
3. Extensive digestion with carboxypeptidase A was necessary to release the amino acids from the C-terminal of the protein, similar to that for the fd coat protein.
4. The extensive aggregation of the core peptide resulting from tryptic digestion on Sephadex G-50 chromatography in 50% acetic acid. Asbeck et. al., (1969) also found evidence of aggregation of the insoluble peptide from the fd coat protein.
5. The protein has a large number of hydrophobic residues.
6. Braunitzer et. al. (1967) found that denaturation of the fd coat protein was necessary to release the C-terminal amino acids. Even after denaturation of the protein a poor yield of the C-terminal amino acids was obtained (10%).

The evidence obtained in this investigation and also that of Braunitzer et. al. (1967) and Asbeck et. al. (1969) suggested that fd and f1 coat proteins could be described as preferring a highly aggregated state, which

can be induced under very mild conditions, which would prevent the action of proteolytic enzymes. Hence it would seem that steric hindrance prevents hydrolysis of the coat protein by enzymes. Because the yield of the C-terminal peptide is much higher than that of any of the other peptides, this suggests the peptide bonds which represent these other tryptic peptides must be very close to the hydrophobic region. This is feasible since Asbeck et. al. (1969) believe the hydrophobic region is aggregated under all conditions. If one accepts the f1 and fd coat proteins as having 3 main regions when denatured, that is, the N-terminal region (residues 1-20) which is free to interact with the environment, residues 21-34 which exist in a highly coiled state, and residues 35-49 which are hydrophilic (basic), then it would seem that peptide bonds 35-43 are too close to the central core to be cleaved easily with trypsin.

This protein is indeed remarkable since studies on other proteins have generally been successful after the protein has been denatured. Insoluble tryptic peptides have been encountered with other proteins but these only represent a small fraction of the molecule.

Asbeck et. al. (1969) commented on the extreme difficulty they encountered in trying to sequence the fd

coat protein by conventional methods. For this reason these workers used a different approach but they still encountered considerable problems.

It is necessary for the complete sequence of the f1 coat protein to be determined as well as the sequences of the other proteins of this class, M13 and ZJ-2, to be understood before the final conclusions can be made about the structure and behaviour of the proteins belonging to this group.

A better approach to the sequencing of this protein would be to disrupt the molecule at the methionine residue (residue 27 in fd) and to separate the fragments. This gives a low yield of cyanogen bromide peptides (as found from the present work) so it would be necessary to start with several hundred milligrams of coat protein. This would cleave the centre of the extremely difficult core region and so hopefully avoid the problem of low recoveries of peptide.

At this stage the two peptides produced should be more amenable to standard techniques of protein sequencing.

REFERENCES

- Alm, R.S., Williams, R.J.P., and Tiselius, A. (1952).
Acta chem. scand. 6, 826.
- Ambler, R.P. (1967). Methods in Enzymology, Vol. 11.
Academic Press, New York, N.Y.
- Anderer, F.A. (1959). Z. Naturf. 14b, 363.
- Arbor, W. (1966). J. molec. Biol. 20, 483.
- Asbeck, F., Beyreuther, K., Kohler, H., von Wettstein, G.,
and Braunitzer, G. (1969). Hoppe-Zeyler's Z. Physiol.
Chem. 350, 1047.
- Atherton, R.S. and Thompson, A.R. (1969). Biochem. J.
111, 797.
- Barton-Wright, E.C. (1946). Analyst, London. 71, 267.
- Barton-Wright, E.C. and Curtis, N.S. (1948). Analyst,
London. 73, 330.
- Beaven, G.H. and Holiday, E.R. (1952). Adv. Prot. Chem.
7, 319.
- Bencze, W.L. and Schmid, K. (1957). Analyt. Chem.
29, 1193.
- Bergmann, M., Fruton, I.S. and Pollock, H. (1939).
J. biol. Chem. 127, 643.
- Blackburn, S. and Lowther, A.B. (1951). Biochem. J.
48, 126.
- Boulton, A.A. and Bush, I.E. (1964). Biochem, J. 92,
11p.
- Bradley, D.E. (1964). J. gen. Microbiol. 35, 471.

- Braunitzer, G. (1964). Ber. Bunsengesellschaft 68, 133.
- Braunitzer, G., Asbeck, F., Beyreuther, C., Kohler, H.
and von Wettstein, G. (1967). Hoppe-Zeyler's Z.
Physiol. Chem. 348, 725.
- Brenner, M., Niederwieser, A. and Pataki, G. (1961).
Experienta 17, 145.
- Bromer, W.W., Staub, A., Sinn, L.G. and Behren, O.K.
(1957a). J. Am. chem. Soc. 79, 2801
- Bromer, W.W., Staub, A., Sinn, L.G. and Behren, O.K.
(1957b). J. Am. chem. Soc. 79, 2807.
- Christensen, H.N. (1943). J. biol. Chem. 151, 319.
- Cole, R.D. and Kincade, J.M. (1961). J. biol. Chem.
236, 2443.
- Consdan, R., Gordon, A.H. and Martin, A.J.P. (1944).
Biochem. J. 38, 224.
- Crammer, J.L., Neuberger, A. (1943). Biochem. J.
37, 302.
- Crestfield, A.M., Stein, W.H., and Moore, S. (1963).
J. biol. Chem. 238, 618.
- Day, L.A. (1966). J. molec. Biol. 15, 395.
- Day, L.A. (1969). J. molec. Biol. 39, 265.
- Dayhoff, M. (1969). Atlas of Protein Structure, National
Biomedical Research Foundation, Maryland, U.S.A.
- Dent, C.E. (1948). Biochem. J. 43, 169.
- Desnuelle, P. and Casal, A. (1948). Biochem. biophys.
Acta 2, 64.

- Dreyer, W.J. and Bynum, E. (1963). *Methods in Enzymology*, Vol. 11 Academic Press, New York, N.Y.
- Edman, P. (1950). *Acta chem. scand.* 4, 277, 283.
- Edman, P. (1956). *Acta chem. scand.* 10, 761.
- Edman, P. (1957). *Proc. Roy. Aust. Chem. Inst.* 24, 434.
- Edman, P. (1960). *Ann. N.Y. Acad. Sci.* 85, 602.
- Edman, P. and Begg, G. (1967). *Eur. J. Biochem.* 1, 80.
- Edman, P. and Sjoquist, J. (1956). *Acta chem. scand.* 10, 1507.
- Enger, M.D. and Kaesburg, P. (1965). *J. molec. Biol.* 13, 260.
- Enger, M.D., Stubbs, E.A., Mitra, S. and Kaesburg, P. (1963). *Proc. natn. Acad. Sci. U.S.A.* 49, 857.
- Eriksson, S. and Sjoquist, J. (1960). *Biochem. biophys. Acta* 45, 290.
- Fischer, E. (1906). *Ber.* 39, 530.
- Folin, O. and Denis, W. (1912). *J. biol. Chem.* 12, 245.
- Fraenkel-Conrat, H., Harris, J.I. and Levy, A.L. (1952). *Methods of Biochemical Analysis*, Vol. 2, Interscience, New York, N.Y.
- Goodwin, T.W. and Morton, R.A. (1946). *Biochem. J.* 40, 628.
- Gray, W.R. and Hartley, B.S. (1963). *Biochem. J.* 89, 59p, 379.
- Guidotti, G., Hill, R.J. and Konigsberg, W. (1962). *J. biol. Chem.* 237, 2184.

- Hanes, C.S. (1961). *Can. J. Biochem. and Physiol.* 39, 119.
- Hanes, C.S., Harris, C.K., Moscarello, M.A. and Tigane, E. (1961). *Can. J. Biochem. and Physiol.* 39, 163.
- Hanke, M.T. and Koessler, K.K. (1920). *J. biol. Chem.* 43, 527.
- Harfenist, E.J. (1953). *J. Am. chem. Soc.* 75, 5528.
- Harris, C.K., Tigane, E. and Hanes, C.S. (1961). *Can. J. Biochem. and Physiol.* 39, 439.
- d'Herelle, F. (1917). *Acad. Sci. Paris* 165, 373.
- Hir, C.H.W., Moore, S. and Stein, W.H. (1960). *J. biol. Chem.* 235, 633.
- Hirs, C.H.W., Stein, W.H. and Moore, S. (1954). *J. biol. Chem.* 211, 941.
- Hoffmann-Berling, H., Kaerner, C. and Knippers, R. (1966). *Adv. Virus Res.* 12, 329.
- Hoffmann-Berling, H., Marvin, D.A. and Durwald, H. (1963). *Z. Naturf.* 18b, 876.
- Hofschneider, P.H. (1963). *Z. Naturf.* 18b, 203.
- Ikawa, M. and Snell, E.E. (1961). *J. biol. Chem.* 236, 1955.
- Inagami, T. and Sturtevant, J.M. (1960). *J. biol. Chem.* 235, 1019.
- Jeppsson, J. and Sjoquist, J. (1967). *Analyt. Biochem.* 18, 264.

- Kassel, B. and Laskowski, M. (1961). J. biol. Chem. 236, 941.
- Katchalski, E. and Lela, M. (1959). Adv. Prot. Chem. 14, 392.
- Konigsberg, W. and Hill, R.J. (1962). J. biol. Chem. 237, 2547.
- Kossel, A. and Kutscher, F. (1900). Z. Physiol. Chem. 31, 165.
- Lin, J.Y., Tsung, C.M. and Fraenkel-Conrat, H. (1967). J. molec. Biol. 24, 1.
- Loebb, T. and Zinder, N. (1961). Proc. natn. Acad. Sci. U.S.A. 47, 282, 742.
- Lowry, O.H., Rosebrough, N.J., Carr, A.I. and Randall, R.J. (1951). J. biol. Chem. 193, 265.
- McCoy, R.H., Meyer, C.E. and Rose, W.C. (1935). J. biol. Chem. 112, 283.
- Marvin, D.A. (1966). J. molec. Biol. 15, 8.
- Marvin, D.A. and Hoffmann-Berling, H. (1963a). Z. Naturf. 18b, 884.
- Marvin, D.A. and Hoffmann-Berling, H. (1963b). Nature 197, 517.
- Matheson, A.T., Tigane, E. and Hanes, C.S. (1961). Can. J. Biochem. Physiol. 39, 417.
- Matsubara, H. and Smith, E.L. (1963). J. biol. Chem. 238, 2732.

- Michl, H. (1951). *Montash Chem.* 82, 489.
- Morgan, F.J. and Henschen, A. (1969). *Biochim. biophys. Acta* 181, 93.
- Moore, S. and Stein, W.H. (1963). *Methods in Enzymology*, Vol. 6, Academic Press, New York, N.Y.
- Nathans, D. (1965). *J. molec. Biol.* 13, 521.
- Nonoyama, M., Yuki, A. and Ikeda, Y. (1963). *J. gen. applied Microbiol.* 9, 299.
- Notani, G.W., Engelhardt, D.L., Konigsberg, W. and Zinder, N.D. (1965). *J. molec. Biol.* 12, 439.
- Nui, C.Z. and Fraenkel-Conrat, H. (1955). *Arch. biochem. Biophys.* 59, 388.
- Petersen, G.B. and Reeves, J. (1969). *Biochim. biophys. Acta* 179, 510.
- Piez, K.A. (1954). *J. biol. Chem.* 207, 77.
- Pratt, D., Tzagoloff, H., Eidah, W.S. and Henry, T.J. (1967). *Symp. on the Molecular Biology of Viruses*, Academic Press, New York, N.Y.
- Ray, D.S. (1968). *Molecular Basis of Virology* ACS Monograph 164, Reinhold.
- Rossomando, E.F. and Zinder, N.D. (1968). *J. molec. Biol.* 36, 387.
- Rubenstein, I., Thomas, C.A. and Hershey, A.D. (1961). *Proc. natn. Acad. Sci. U.S.A.* 47, 1113.
- Sakaguchi, G. (1925). *J. Biochem. (Tokyo)* 5, 13, 25, 133.

- Salivar, W.D., Tzagoloff, H. and Pratt, D. (1964).
Virology 24, 359.
- Sanger, F. (1945). Biochem. J. 39, 507.
- Sanger, F. (1949). Biochem. J. 45, 563.
- Sanger, F. and Thompson, E.O.P. (1953). Biochem. J.
53, 353, 366.
- Sanger, F. and Tuppy, H. (1951). Biochem. J. 49, 463,
481.
- Schlesinger, M. (1936). Nature 138, 50.
- Schroeder, W.A. (1967). Methods in Enzymology, Vol.
11, Academic Press, New York, N.Y.
- Schroeder, W.A., Shelton, J.R., Balog-Shelton, J. and
Cormick, J. (1963). Biochemistry 2, 992.
- Scott, D.W. (1965). Virology 26, 85.
- Senn, M. and McLafferty, F.W. (1966). Biochem. biophys.
Res. Comm. 23, 381.
- Sinsheimer, R.L. (1959a). J. molec. Biol. 1, 37.
- Sinsheimer, R.L. (1959b). J. molec. Biol. 1, 43.
- Sjoholm, I. (1964). Acta chem. scand. 18, 889.
- Sjoquist, J. (1953). Acta chem. scand. 7, 447.
- Sjoquist, J. (1955). Biochim. biophys. Acta 16, 283.
- Sjoquist, J. (1957). Ark. Kemi 11, 151.
- Sjoquist, J. (1959). Ark. Kemi 14, 291.
- Sjoquist, J. (1960). Biochim. biophys. Acta 41, 20.
- Sjoquist, J., Bloombach, B. and Wallen, P. (1961).
Ark. Kemi 16, 425.

- Smyth, D.G. and Elliot, D.F. (1964). *Analyst* 89, 81.
- Smyth, D.G., Stein, W.H. and Moore, S. (1963). *J. biol. Chem.* 238, 227.
- Spackman, D.H., Stein, W.H. and Moore, S. (1958). *Analyt. Chem.* 30, 1190.
- Spackman, D.H., Stein, W.H. and Moore, S. (1960). *J. biol. Chem.* 235, 648.
- Stanley, W.M. (1935). *Science* 81, 644.
- Strauss, J.H. and Sinsheimer, R.L. (1963). *J. molec. Biol.* 7, 43.
- Tessman, E.S. and Shleser, R. (1963). *Virology* 19, 239.
- Tigane, E., Wade, E.H.M., Wong J.T. and Hanes, C.S. (1961). *Can. J. Biochem. and Physiol.* 39, 427.
- Tsugita, A., Gish, D.T., Young, J., Fraenkel-Conrat, H., Knight, C.A. and Stanley, W.M. (1960). *Proc. natn. Acad. Sci. U.S.A.* 46, 1463.
- Wade, E.H.M., Matheson, A.T. and Hanes, C.S. (1961). *Can. J. Biochem. and Physiol.* 39, 141.
- Warburg, O. and Christian, W. (1941). *Biochem. Z.* 310, 384.
- Weber, K. (1967). *Biochemistry* 6, 3144.
- Weber, K. and Konigsberg, W. (1967). *J. biol. Chem.* 242, 356.
- Wieland, T. (1948). *Angew. Chemie* 60, 313.

- Wieland, T. and Gebert, U. (1963). *Analyt. Biochem.*
6, 201.
- Wieland, T. and Pfleiderer, G. (1955). *Angew. Chemie*
67, 257.
- Wittmann, H.G. (1965). *Z. Naturf.* 20b, 1213.
- Wittmann-Liebold, B. (1966). *Z. Naturf.* 21b, 1249.
- Young, J.D. and Carpenter, F.H. (1961). *J. biol. Chem.*
236, 743.
- Zahler, S.A. (1958). *J. Bact.* 75, 310.
- Zinder, N.D. (1965). *Ann. Rev. Micro.* 19, 455.
- Zinder, N.D., Valentine, R.C., Roger, M., Stoeckenius,
W. (1963). *Virology* 20, 638.