THE STRUCTURE OF DOWN FEATHER KERATIN

A thesis submitted by

IAN DOUGLAS WALKER, B.Sc. (HONS.),

to the University of Adelaide,

South Australia,

for the degree of

Doctor of Philosophy.

DEPARTMENT OF BIOCHEMISTRY,
UNIVERSITY OF ADELAIDE,
SOUTH AUSTRALIA.

FEBRUARY, 1974.

SUMMARY

The work of this thesis is concerned with the structural basis for the heterogeneity of chick down feather keratin chains. The number of keratin chains was investigated by examining discrete fractions of the S-carboxymethylated feather protein by polyacrylamide gel electrophoresis at two pH values and at least nineteen prominent proteins were found to comprise the down feather. Each of these proteins was remarkably similar to the others with respect to molecular weight and amino acid composition.

Sequence analysis of the $[^{14}C]$-SCM-tryptic peptides of down feather keratin indicated that the electrophoretic variants of feather keratin differed only at a few sites in the primary structure from one another. The nature of these amino acid substitutions were compatible only with the existence of multiple genes, each encoding a specific keratin variant. However, few of these substitutions, if any, could be the result of alleles and their existence is due to the presence in the chick genome of mutationally altered replicates of an ancestral keratin gene. Extensive homology exists between the amino acid sequences of down feather keratin and that of a purified protein of emu feather rachis, and this homology allowed the deduction of a partial amino acid sequence for down feather keratin.

Amino-terminal sequence analysis of the keratin chains of chick down feather, adult barbs and scales established that whereas the two feather tissues may have possessed common keratin chains, all the keratin chains of scales are unique.

The origin and the possible arrangement of keratin genes within the chick genome is discussed.

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge, this thesis contains no work previously performed by any other person except where due reference is made in the text.

Signed:

IAN D. WALKER

ACKNOWLEDGEMENTS

I wish to thank Professor W.H. Elliott for permission to work in this Department and for his advice and encouragement.

I sincerely thank my supervisor, Dr. G.E. Rogers for his wise and patient support throughout the course of this work.

Thanis are due especially to Drs. I.J. O'Donnell, Wool Research Laboratories, Department of Protein Chemistry, Melbourne and H.J. Harding of this Department for useful advice, and to Dr. Harding and Mr. M. Cichorz for amino acid analyses.

I am indebted to Dr. D.J. Kemp for his collaboration with me in the experiments of Chapter 8 and to Mr. P.M. Gibbs for his collaboration in the experiments of Chapter 6.

For many helpful discussions, thanks are due to Dr. D.J. Kemp and Mr. G.A. Partington whose work in parallel with my own prompted many of the experiments of this thesis.

For technical assistance, thanks are due to Miss L. Bissel.

This project was carried out while in receipt of a Commonwealth Postgraduate Award. The work was supported by grants from the Wool Board and the Australian Research Grants Committee (to Dr. G.E. Rogers).

ABBREVIATIONS

dansyl	1-dimethylaminonapthalene-5-sulphonyl chloride
HVPE	high voltage paper electrophoresis
MW	molecular weight
NEMA	N-ethylmorpholine/acetate buffer
PTH	phenylthiohydantoin
PMA	pyridine/N-ethyl morpholine/acetate buffer
SCM-	S-carboxymethyl
TCA	trichloroacetic acid
TPCK	L-(1-tosylamido-2-phenyl) ethyl chloromethyl ketone.

CONTENTS

CHAPTER 1

INTRODUCTION

INTRODUCTION

A. The Structure of Feathers

Feathers are composed of dead and dehydrated cells which are filled with the intracellular fibrous protein, keratin. Morphologically, they are complicated structures and the shape and size of a feather is determined during growth by the number and nature of mitotic divisions that the epidermal cells, from which feathers develop, undergo. The morphological parts of a feather each consist of cell types which have differentiated during feather morphogenesis but which are collectively termed keratinocytes. Keratin synthesis by such cells is the terminal process in that differentiation.

The cellular structure of feather tissue and the physico-chemical properties of the keratin molecules present in its cells will now be discussed.

(i) The Morphology of Feathers

The chick down feather consists of a short base, the calamus, to which is attached ten to fifteen barbs (Fig. 1.1a). Each barb consists of a cylinder of flattened cortical cells surrounding a core of medullary cells (Fig. 1.1b). Two rows of smaller structures, the barbules, are arranged diametrically opposite one another on each barb: each barbule, a linear arrangement of ten to twenty single cells, tapers towards the tip. (Watterson, 1942.)

Feathers of adult birds differ slightly in structure depending upon their position on the bird but in general they

Figure 1.1

Structure of the embryonic chick feather (down feather)

Figure 1.1a

Feather from newly hatched chick after discarding sheath

Figure 1.1b

Diagram, illustrating details of the structure of barbs and barbules

B: barbs

B.BL: barbules

CAL: calamus

C: cortex

M: medulla

(From Watterson, 1942)

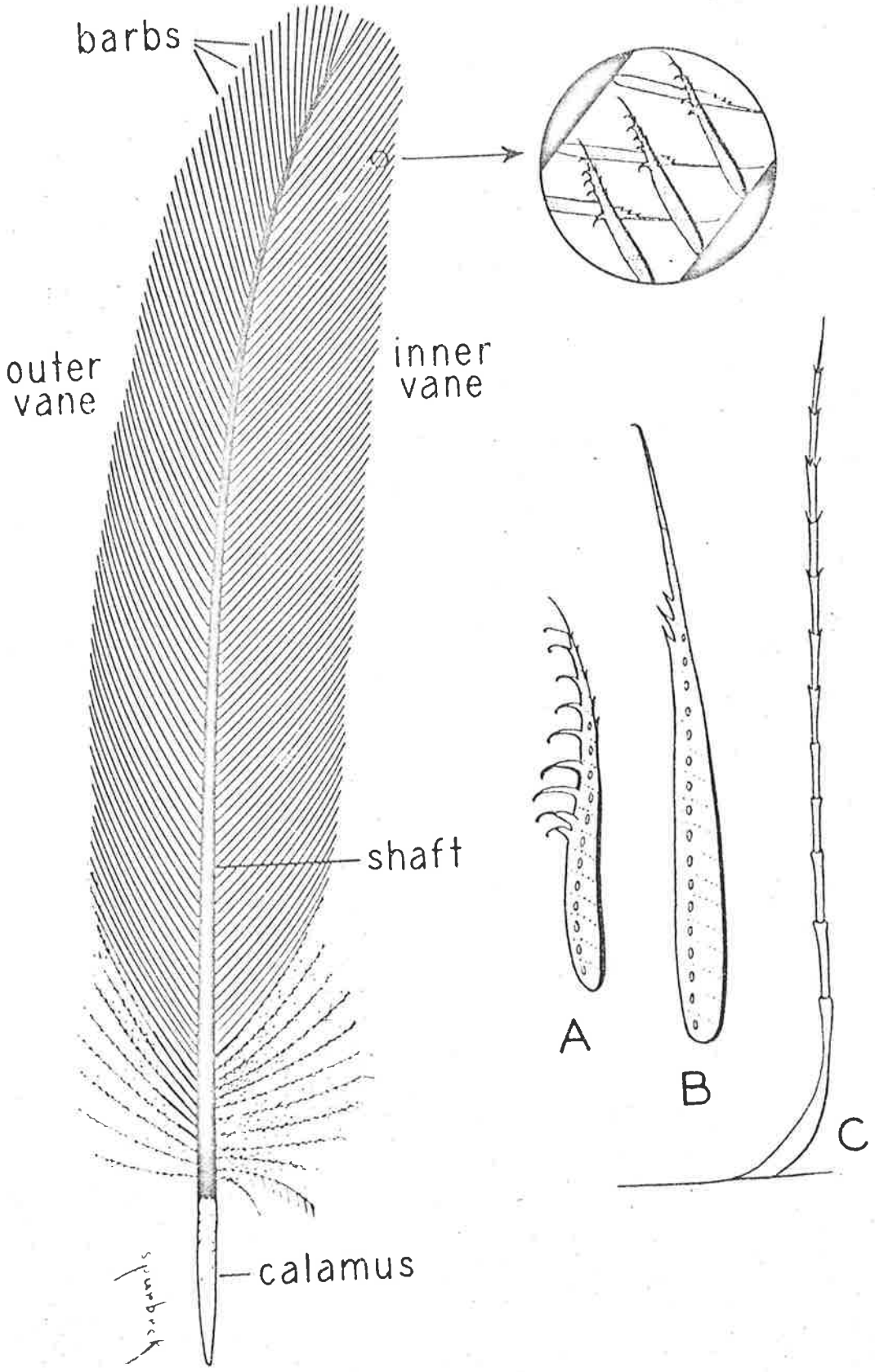

barbs
outer vane
inner vane
shaft
calamus
A
B
C

Figure 1.2

Structure of the adult feather

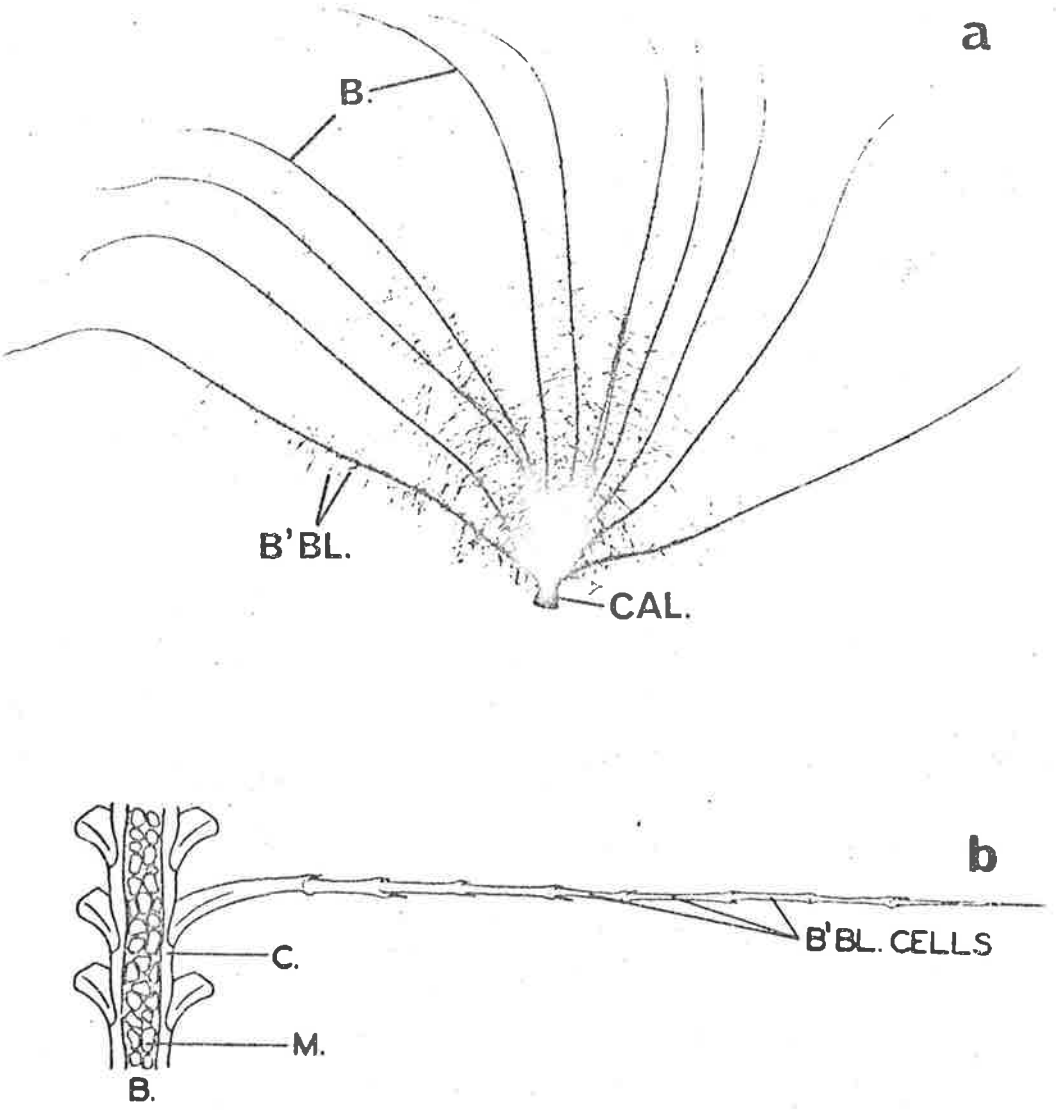
INSET Detail showing barbules branching off the barbs

A. Anterior barbule bearing hooklets (hamuli)

B. Posterior barbule

C. Downy barbule, from the lower, fluffy region of the feather

(From Rawles, 1965)


a
B.
B'BL.
CAL.
b
B'BL. CELLS
C.
M.
B.

consist of an approximately cylindrical shaft, the upper part or rachis bearing numerous barbs and the lower part, the calamus lacking barbs and fixed in the feather follicle.(Voitkevich (1966), (Fig. 1.2). The shaft encloses a pithy medulla. Two rows of barbules are attached to each barb which in contrast to the down feather are of three different types. The barbules closest to the base of the feather are morphologically similar to those of the down feather whereas the barbs of the upper portions have microscopic hooklets or barbicells which enable the barbules to be linked together. The barbules borne on the dorsal (upper) side of the barb possess more pronounced and exagerated barbicells than those on the ventral (lower) side - the two types are quite distinct in morphology (Rawles, 1965).

The cells of the mature feather are dead, dehydrated and filled with keratin. The toughness and resistance properties of feathers depend upon the physico-chemical properties of the keratin molecules which exist in feather cells as regular arrays of extended protein chains.

(ii) The Molecular Structure of Feather Keratin

Feather keratins possess a filmaentous fine structure and exist within cells as long cylindrical rods, approximately 30Å in diameter and closely packed together parallel to the feather axis (Filshie and Rogers, 1962; Rogers and Filshie, 1963). The arrangement of individual polypeptide chains within the feather filaments is not known but several salient features are clear regarding their state in native

feathers. Approximately 30% of each polypeptide exists in an antiparallel β-pleated sheet conformation as determined by infrared spectroscopy (Fraser et al., 1971). Four such pleated sheet regions, each of 8 amino acid residues, are believed to constitute a repeating unit, many of which are arranged into a left-hand helix of pitch 95Å. Two such left-hand helicies are entwined in an antiparallel fashion along the longitudenal axis of the feather filament which they compose (Fraser et al., 1972).

The x-ray diffraction pattern, on which the model for feather keratin is based is specific for feather keratin and clearly distinct from that of α- keratin. (Fraser et al., 1972).

B. The Properties of Feather Keratin

The understanding of keratin structure requires a thorough knowledge of the properties of the constituent protein molecules. Keratin molecules are insoluble and difficult to examine in their native state owing to the high proportion of intra-and intermolecular disulphide crosslinks. Normally they are extracted with a reducing agent which converts cystine to cysteine residues and they may be prevented from reoxidising by their conversion to a stable derivative. The S- carboxymethyl (SCM-) derivatives have proved satisfactory for the preparation of soluble and stable keratin chains derived from both hair (Goddard and Michaelis, 1935, O'Donnell

and Thompson, 1962) and feather (Harrap and Woods, 1964a, Kemp and Rogers, 1972, O'Donnell, 1973a).

The properties of SCM-keratin extracted from mammalian hair and avian feathers or scales respectively are clearly different in many respects and a comparative summary of the salient features of both is undertaken below.

(i) The Molecular Weight of Keratin Molecules

Two major classes of protein molecules, the high-sulphur fraction and the low-sulphur fraction, exist in mammalian hair (Crewther et al., 1965) and the major differences in chemical properties between the two classes are believed to reflect the distinct structural role of each (Fraser et al., 1972). The two prominent fractions which constitute the low-sulphur proteins, components 7 and 8 have molecular weights of 51,400 and 45,000 respectively (Thompson and O'Donnell, 1965, Jeffrey, 1969, Jeffrey, 1970). The high-sulphur proteins of wool are also heterogeneous with respect to molecular weight and range from 10,800 to 28,000 (Gillespie, 1965, Haylett et al., 1971).

In contrast, all feather keratins studied are uniform with respect to molecular weight. Woodin (1954) determined the molecular weight of chick feather keratin by both sedimentation and osmotic pressure measurements and found that a high proportion of feather protein was homogeneous in molecular weight and obtained a value of approximately 10,000. Harrap and Woods (1964b) determined the molecular weight of

both reduced and of SCM- chick rachis keratin and obtained a value of 10,400 for the reduced monomer. On the basis of the amino acid sequence of a purified component of SCM- emu rachis keratin O'Donnell (1973b) determined its molecular weight as 10,459.

(ii) Amino Acid Composition

The amino acid compositions of the high-sulphur and low-sulphur proteins of wool differ significantly in the amounts of many amino acids. Half-cystine, notably, is enriched in the high-sulphur fraction which contains approximately 18 residues per 100 residues (Fraser et al., 1972). The low sulphur fraction is relatively poor in cystine but contains relatively high amounts of glutamic and aspartic acids. Methionine is absent from both fractions and the low-sulphur proteins possess significant amounts of lysine which is present only in trace amounts in the high-sulphur proteins. (Fraser et al., 1972) Attempts to isolate high- and low-sulphur fractions from feather keratin did not prove successful (Harrap and Woods, 1964a), and samples of chick feather keratin whether obtained from different morphological parts of adult feathers, from down feathers (Kemp and Rogers, 1972) or from purified fractions of adult feather contain similar amounts of half-cystine (8 residues per 100 residues). Typically, feather keratins are rich in glycine, serine and aliphatic amino acids. They contain no methionine and little if any histidine or lysine. O'Donnell (1973a) found only minor differences in amino acid composition between four

purified components of the keratin of emu feather rachis. These studies suggested that feather keratin molecules were much more alike with respect to amino acid composition and presumably primary structure than the high-sulphur proteins were to one another or to the low-sulphur proteins.

(iii) The Heterogeneity of Keratin

The keratin proteins of both high-sulphur and low-sulphur groups of mammalian hair are extremely heterogeneous and no reliable upper limit for the heterogeneity has been established. The high-sulphur proteins are composed of at least forty components (Darskus, 1972).

Studies of the SCM-derivatives of feather keratin have not indicated such a high degree of heterogeneity as for α-keratin. Harrap and Woods (1964a) identified four components of chick rachis and calamus by ascending boundary electrophoresis. Woods (1971) isolated discrete fractions of SCM-feather rachis by DEAE-cellulose chromatography and their examination by polyacrylamide gel electrophoresis indicated the presence of ten constituents. Typically, SCM-feather keratin from different morphogical parts of adult feathers when examined on polyacrylamide gels is resolved into four to six major bands (Kemp and Rogers, 1972: O'Donnell 1973a) SCM-keratin from down feathers was resolved into seven major bands and six major bands when examined by polyacrylamide gel electrophoresis at pH 9.5 and pH 2.7 respectively. (Kemp and Rogers, 1972)

(iv) Primary Structure Studies on Keratins

Sequence analysis of purified high-sulphur proteins (Fraser et al., 1972) has indicated that all high sulphur proteins so far studied bear a close similarity to one another in amino acid sequence. Many high-sulphur proteins possess N-acetyl-alanine as amino-terminal residue (Haylett et al., 1971, Elleman and Dopheide, 1972) but others possess unblocked amino-termini (Swart, 1973). Cystine (or cysteine) has been reported as the carboxyl-terminal residue of all high-sulphur proteins so far examined (Fraser et al., 1972, Elleman and Dopheide, 1972, Swart, 1973). The presence of a repeating unit of ten residues has been reported (Elleman and Dopheide, 1972) in the sequence of certain high-sulphur proteins and Swart (1973) re-interpreted the data to indicate a repeating unit of five residues. On the basis of the amino acid sequences of a number of high-sulphur proteins, Swart (1973) was able to propose the nucleotide sequence of an ancestral keratin gene from which the high-sulphur protein family presumably arose by gene duplication and mutational divergence.

A partial sequence has been determined for the amino-terminal portion of a low-sulphur protein (component 8) and the presence of amino acid substitutions at a number of sites was detected indicating microheterogeneity in protein composition (Fraser et al., 1972). The low-sulphur proteins therefore are also a family of closely-related protein chains. No obvious sequence similarities are apparent between the families of high-sulphur and low-sulphur proteins respectively.

The sequence analysis of feather keratin molecules has not been as extensive as that of α-keratins. Schroeder et al., (1957) determined the sequences of some short peptides derived from partial acid hydrolysates of feather calamus. O'Donnell (1971) reported the isolation of the amino-terminal peptide N-acetyl-ser-cys-tyr in molar amounts, from SCM- calamus of the goose, indicating that all the keratin constituents of goose calamus possess a common amino-terminal sequence.

The complete amino acid sequence of a purified constituent of emu rachis keratin has been determined (O'Donnell, 1973b). The amino-terminal sequence of the protein was N-acetyl-ser-cys-tyr, as in goose calamus. The carboxyl-terminal residue was half -cystine. A feature of interest is that the half-cystine residues were found to be confined to the amino-terminal and carboxyl-terminal regions of the sequence whereas the central portion of the molecule contained no half-cystine but a preponderance of the aliphatic amino acids, especially isoleucine and valine. No repeating sequences such as were found in the high-sulphur proteins, were obvious. The high content of aliphatic amino acids in the central portion of the sequence suggested that this portion was involved in the formation of the β-pleated sheet part of the molecular structure proposed by Fraser et al., (1972). Indeed, Suzuki (1973) on the basis of infrared spectral measurements, produced evidence in support of this view.

(v) Species Specific Differences Between Keratins

The low-sulphur proteins from different species have been examined by gel electrophoresis (Shechter et al., 1969). Keratins fromclosely-related species, for instance monkey and human, were very similar and no major qualitative differences were apparent. The low-sulphur proteins of rabbit or guinea pig hair were also alike but were clearly different to those of human or monkey hair. The elution profiles of high-sulphur protein fractions from pig, cow and goat from DEAE- cellulose chromatography were compared and each was found to be quite distinctive (Darskus and Gillespie, 1971). The high degree of heterogeneity of both high- and low-sulphur proteins however would not reveal differences between individual proteins which differred between closely related species.

When the keratin of the calami of feathers from six different avian species, namely cockatoo, duck, emu, domestic fowl, silver gull and turkey was examined by gel electrophoresis, major qualitative differences were evident in all cases (O'Donnell, 1973a). Peptide maps of components purified from the SCM- calami of emu and silver gull, indicated that whereas most or all emu components possessed the amino-terminal sequence N-acetyl-ser-cys-tyr, no silver gull component possessed this sequence. Thus the differences between feather keratins from different avian species are more distinctive than the differences in hair keratins between species.

C. The Development of Feathers and the Synthesis of Keratin

The purpose of this section is to review aspects of feather development and keratin synthesis which are pertinent to the nature of the protein products of avian keratinocytes. However, little mention is made of the initial events which commit epidermal cells in early embryos to feather morphogenesis and keratin synthesis.

(i) The Onset of Keratin Synthesis in Embryonic Feathers

Feathers develop from populations of epidermal cells which proliferate and differentiate in response to their interaction with dermis (Rawles, 1965, Wessells, 1965). By day 11 of embryonic life, extensive morphodifferentiation has taken place and although keratin synthesis has not yet begun, barbule, barb and calamus cells are clearly developed. (Matulionis, 1970). By day 12 the presence of small amounts of keratin can be detected and by day 14, keratin synthesis accounts for over 50% of the protein synthesis occurring in feather cells (Kemp et al., 1974). All the electrophoretic variants of reduced and S-carboxymethylated keratin found in the mature down feather are synthesised simultaneously in embryonic down feathers (Kemp et al., 1974). The presence in embryonic feather cells of mRNA which encodes the synthesis of most or all of these keratin chains has been demonstrated by Partington et al., (1973) who reported the translation of keratin mRNA in reticulocyte lysate. Recently, Kemp et al., (unpublished results) have isolated a purified sample of RNA from developing down feather which migrated on formamide-

acrylamide gels as a single band, and which was capable of programming the synthesis of most or all feather keratin chains in either a heterologous cell free system (or when injected into an amphibian oocyte).

These studies indicated that the synthesis of most or all keratin chains was encoded by mRNA molecules all very similar in size and which were present in the cytoplasm of at least some feather cells by day 12 of embryonic life.

(ii) The Tissue Specificity of Keratin Synthesis

In an examination of the keratin chains present in different morphological parts of the adult feather, in the down feather and in adult and embryonic scales, Kemp and Rogers (1972) found that the SCM- keratin from each tissue was distinguishable from the others when examined by gel electrophoresis both at pH 9.5 and at pH 2.7. Thus, adult barb keratin possessed some components which were not present in rachis keratin from the same feather or in down feather keratin. All feather tissues, however, showed electrophoretic bands in common with other feather tissues at both pH 9.5 and at pH 2.7. Scale tissues on the other hand possessed no prominent constituents in common with any feather tissue. It was concluded that in each tissue only some of a potentially much larger set of keratin genes is expressed.

D. Aims of the Project

The project is concerned mainly with the structure and the properties of the keratin proteins of the down feather. It was undertaken in parallel with other studies on the molecular biology of feather growth and keratin synthesis and is primarily a study of gene expression in feather cells. From the properties and in particular the primary structure of the keratins of the newly-hatched chick down feather it was hoped to infer some of the salient features of the keratin genes. Operationally the aims of the project were threefold.

(i) Heterogeneity of Down Feather Keratin

A central problem in the study of keratins concerns the number of different proteins present in a keratinised tissue. Whereas the α-keratins of mammalian hair were exceeded only by immunoglobulins in chemical complexity, feather keratins appeared from published work to be rather less heterogeneous. It was therefore decided to critically examine the number of keratin proteins that constituted the chick down feather.

(ii) The Relationships between the Keratin Variants of a Single Tissue

Previous studies had suggested that feather keratins closely resembled each other in primary structure and that the similarities between them were even more pronounced that those between individual high sulphur proteins from mammalian hair. A study of the primary structure of down feather keratin was

undertaken to investigate the nature and the origin of the differences between the keratin variants of a single tissue.

(iii) Differences in Keratin Chains between Tissues

The appearance in each epidermal tissue of different electrophoretic variants of keratin suggested the differential expression of specific keratin genes. A study was undertaken to confirm this hypothesis and to investigate the differences in primary structure between the keratin chains of three tissues.

CHAPTER 2

MATERIALS AND GENERAL METHODS

A. Materials

(i) Tissue

Feathers and scales were from white leghorn fowls, strain Para 3, a pure bred strain of the domestic fowl Gallus gallus. Animals were obtained from the Parafield Poultry Station, of The Department of Agriculture, Parafield, South Australia.

(ii) Enzymes and Proteins

Trypsin (TPCK-treated), chymotrypsin and carboxypeptidase B(DFP-treated) were obtained from Worthington Biochemical Corporation, Freehold, N.J., U.S.A. Thermolysin was obtained from Daiwa Kasei K.K., Osaka, Japan. Carboxypeptidase A, Bovine serum albumin, ovalbumin, sperm whale myoglobin, bovine ribonuclease, egg-white lysozyme and horse heart cytochrome c were obtained from Sigma Chemical Company, St. Louis, U.S.A. and were the purest samples available. SCM- rabbit globin was a gift from Dr. D.J. Kemp. Plastocyanin,, and SCM- insulin B chain were gifts from Drs. P.R. Milne and P.M. Stienert respectively.

(iii) Chemicals

Acrylamide, TEMED, N-ethyl morpholine,phenylisothiocyanate and trifluoroacetic acid were obtained from Eastman Organic Chemicals, New York, U.S.A. Tris, sodium dodecyl sulphate, iodoacetic acid, β-mercapto-ethanol and dithiothreitol were obtained from Sigma Chemical Company,

St. Louis U.S.A. NN' methylenebisacrylamide and dansyl-chloride were obtained from BDH Chemicals, Poole, England. 1,4-bis-(5-phenyloxazolyl)-benzene, and 2,5-diphenyloxazole were obtained from Packard Instruments Co. Inc., La Grange, U.S.A. Coomassie blue R-250 was obtained from ICI Ltd., England. Ninhydrin and methyl cellosolve were obtained from Pierce Chemicals.

Dansyl-amino acids were a gift from Dr. P.R. Milne.

Batches of other chemicals were obtained from two or more different sources but in all cases the results obtained were independent of source.

The following chemicals were routinely purified by redistillation: glacial acetic acid, trifluoroacetic acid, pyridine, N-ethyl morpholine and phenylisothiocyanate.

(iv) Resins for Column Chromatography

DEAE- cellulose (DE-11) and phosphocellulose (P-11) were Whatman Chromedia. (W. and R. Balston. Ltd. England). Sephadex G-100 (fine) and Sephadex G-50 (fine) were obtained from Pharmacia Fine Chemicals, Uppsala, Sweden. Dowex-1 (AG 1-X5 minus 400 mesh) and Dowex-50 (AG 50W-X2, 100-200 mesh) were obtained from Bio-Rad Laboratories, Richmond, Calif., U.S.A. Dowex-50 (Chromo-Beads, type P) was obtained from Technicon Chemical Co. Inc., New York, U.S.A.

(v) Radioactive Compounds

2- $[^{14}C]$ -iodoacetic acid of specific activity 30 mCi/m mole was obtained from The Radiochemical Centre, Amersham, Buckinghamshire, England.

B. General Methods

(i) Amino Acid Analysis

Samples of peptides or proteins were hydrolysed with 1.0 ml of 6N HCl (redistilled-constant boiling point) to which thiodiglycol (10 µl) was added. Hydrolysis was performed at 110°C in glass tubes (K-89685, Kontes Glass Co., N.J.) essentially according to the procedure of Moore and Stein (1963). Hydrolysates were dried by rotary evaporation and analysed by the procedure of Piez and Morris (1960) using a Beckman 120C analyser modified as described by Harding (1971).

(ii) Determination of Radioactivity

Radioactivity present in volatile buffers was determined by drying samples (50-200 µl) on glass fibre discs (Whatman GF/C) at 110°C which were then counted in toluene based scintillation fluid containing 0.35% w/v 2,5-diphenyl-oxazole and 0.035% w/v 1,4-bis-(5-phenyloxazolyl)-benzene. Radioactivity present in buffers containing high concentrations of urea was determined initially by TCA precipitation of radioactive protein and filtration onto glass fibre discs which were then washed with ether, dried and counted. Latterly however radioactivity in urea buffers was determined as follows. The sample (50 - 100 µL) was dissolved 1.0ml of triton-scintillation fluid (triton X-100 : scintillation fluid 3:5 (v/v)) with the dropwise addition of 10M acetic acid if required to clarify the solution (Wells, personal

communication). A Packard Tricarb Liquid Scintillation Spectrometer was used for radioactivity determinations.

CHAPTER 3

THE SPECIFICITY OF THE S-CARBOXYMETHYLATION REACTION

INTRODUCTION

The work presented in this thesis is concerned almost exclusively with the structure and properties of keratin from feathers subjected to reduction and S-carboxymethylation. Many of the conclusions in subsequent chapters require that the S-carboxymethylation of feather keratin is entirely specific and complete. It was crucial therefore, to establish that the amino acid residue which was modified by reduction of feathers and their subsequent reaction with iodoacetic acid was exclusively cysteine. It was important, in addition, to establish that the reaction of cysteine with iodoacetic acid was complete - that every cysteine reside in the keratin molecules was converted to its S-carboxymethyl-(SCM-) derivative. The SCM- derivatives of feather keratin were chosen for study for several reasons. The alkylation of cysteine residues prevents the reformation of disulphide bonds in protein molecules after the removal of reducing agent. SCM-cysteine in peptide linkage is chemically stable and S-carboxymethylated proteins are amenable to study under conditions involving high and low pH without the loss or rearrangement of the carboxymethyl group. They are therefore particularly suitable for study by polyacrylamide gel electrophoresis at a variety of pH values and are amenable to fractionation by ion exchange chromatography. Furthermore, these derivatives are susceptible to digestion by proteolytic enzymes and the resultant peptides which contain SCM-cysteine

can be readily purified by established techniques. Primary sequence analysis of such peptides is straightforward and SCM-cysteine can be easily identified either as its phenylthiohydantoin (PTH) or its dansyl derivative. Reduced proteins can be alkylated with radioactively labelled iodoacetic acid and very high specific activities can be obtained. This facilitates the very sensitive detection of such "labelled" proteins and peptide fragments derived from them.

For the above reasons and in view of the high content of cystine in feather keratin, the SCM-derivatives were chosen for study.

Previous studies on SCM-keratin from both mammalian hair (Fraser et al., 1972) and avian feathers (Harrap and Woods, 1964a, Kemp and Rogers 1972, O'Donnell 1973a) have established the existence of multiple constituents of keratin. Chapter 4 of this thesis is devoted to an examination of the heterogeneity of the SCM-derivatives of keratin of the newly hatched chick. The main purpose of the present chapter is to investigate whether any of the heterogeneity observed arose as a result of the preparation of the SCM-derivatives of feather keratin.

Amino acid analysis of feather keratin before and after S-carboxymethylation showed no evidence of the modification of any residue other than cysteine by the procedure. Furthermore, gel electrophoresis of two protein samples before and after S-carboxymethylation could detect no additional bands due to artefactual or incomplete carboxymethylation.

METHODS

(i) Reduction and carboxymethylation of Feathers

Feathers were plucked from newly hatched chicks and reduction and carboxymethylation was carried out by the method of Kemp et al. (in press). 100 mg of feathers were incubated with stirring in 10 ml of a solution containing 8 M urea, 0.1 M β-mercaptoethanol, 0.5 M ethanolamine, pH 10.5, (HCl), at 37°C for 1 hr. The mixture was then homogenised in a Potter-Elvehjem homogeniser and the resultant suspension was incubated for a further 2 hr at 37°C. 6 ml of 3.0 M tris- HCl buffer pH 8.0, was added. 3g of iodoacetic acid was dissolved by the addition firstly of 1 ml of 3.0M tris- HCl buffer pH8 and the subsequent dropwise addition of 5M NaOH. The pH was adjusted to 8.2 and the volume of the iodoacetate made up to 10.0ml. 2ml of this iodoacetate solution was added to the reduced suspension of feather keratin and the alkylation of thiol groups allowed to proceed for 15 min. After this time the solution was always negative to the nitroprusside test indicating the absence of any thiol groups. 0.30ml of β-mercaptoethanol were added and incubation at 37°C was carried out for a further 1 hr. 2ml of iodoacetate solution, pH 8, were added and incubation carried out for 15 min. Finally 0.4ml of β-mercaptoethanol was added to remove any unreacted iodoacetate. The suspension was centrifuged at 18000 rpm for 30 min in an MSE centrifuge to remove cell debris and the clear supernatant dialysed exhaustively against four changes of bidistilled H_2O. After

dialysis, the protein solution contained a precipitate but no steps were taken to remove this and the total dialysate was freeze dried.

When proteins other than keratin were subjected to reduction and carboxymethylation, the procedure followed was the same as that described above, except that the homogenisation step was omitted.

(ii) S-carboxymethylation of Feathers with $[^{14}C]$-iodoacetic Acid

100 mg of feathers were incubated in 10ml of a solution of 8M urea, (0.5M ethanolamine, pH 10.5, containing 12.6mg (81.8 μmoles) of dithiothreitol (DTT) for 1 hr at 37°C under a nitrogen atmosphere. Homogenisation and re-incubation was carried out as described previously and 6ml of 3.0M tris- HCl buffer, pH 8.0, added. 50μCi of $[^{14}C]$ iodoacetic acid (S A : 30 m Ci/ m mole) was added to the suspension which was then incubated for 15 min to allow all the $[^{14}C]$ - iodoacetate to react. 2.0ml of unlabelled iodoacetate (30% w/v) was added and a further cycle of reduction and carboxymethylation carried out as previously described. The protein suspension was centrifuged and the supernatant dialysed and freeze-dried. The specific activity of the $[^{14}C]$-SCM-keratin was determined by withdrawing two small samples of protein before freeze-drying, one for the determination of radioactivity by liquid scintillation counting

and one for the determination of protein by the method of Lowry et al., (1951). The specific activity of samples of keratin alkylated in this manner with $[^{14}C]$- iodoacetic acid was $2.7 \times 10^5 \pm 10\%$ cpm/mg.

(iii) Gel Electrophoresis at pH 9.5

10% Polyacrylamide gels, 5M in urea and running at pH 9.5 were prepared in 6mm x 10cm tubes as described by Canal Industrial Corp., Bethesda, Maryland. Routinely, known amounts of protein samples for electrophoresis were freeze-dried and dissolved in 8M urea at a concentration of 1mg/ml. 50μg (50μl) was loaded per gel and electrophoresis was carried out at 2 mA./tube using bromophenol blue as tracker dye. After electrophoresis the protein samples were fixed in the gel with 10% TCA and stained with 0.05-0.10% coomassie blue in 10% TCA (Chrambach et al., 1967).

(iv) Gel Electrophoresis at pH 2.7

15% polyacrylamide gels 2.5M in urea and running at pH 2.7 were prepared as described by Panyim and Chalkley (1969). Electrophoresis was carried out for 6 hr or shorter times as stated in the text. Fixing and staining was carried out as described for pH 9.5 gels.

(v) Gel Autoradiography

Polyacrylamide gels of protein samples alkylated with $[^{14}C]$- iodoacetic acid were prepared for autoradiography essentially by the method of Fairbanks et al., (1965). A

longtitudinal slice, 2mm in width, was cut from the gel after fixing and staining and immersed in 3 changes (100ml each) of 10% acetic acid (v/v) to remove TCA. The gel slice was placed on a sintered PVC disc ("Corvic". Chloride Batteries of Aust. Ltd.) and the PVC disc overlaid with a cellophane film which was brought into tight contact with the gel slice by applying suction from a water pump. The gel slice was dried, intact, onto the upper cellophane sheet by suction on the pump overnight. On the gel slice, firmly affixed to the cellophane sheet could clearly be seen the protein bands which could be adequately photographed. The autoradiogram was prepared by placing the dried gel slice in contact with x-ray film (Ilford Red Seal) and exposing for 7 days. The exposed x-ray film was developed as recommended by the manufacturers specifications.

RESULTS

(i) Polyacrylamide gel Electrophoresis of SCM- Feather Keratin

Samples of SCM- feather keratin were subjected to gel electrophoresis at pH 9.5 and at pH 2.7 (Fig. 3.1a). The pH 9.5 gel shows the presence of 7 prominent bands which are labelled according to the nomenclature of Kemp and Rogers (1972). On the pH 2.7 gels 6 bands (A-H) were apparent and an additional band, Fp, could be detected when shorter electrophoresis times were used. Figure 3.1b shows a pH 9.5 gel of $\{^{14}C\}$- SCM- feather keratin, prepared as described

FIGURE 3.1 Polyacrylamide gel electrophoresis of SCM-feather keratin at pH 9.5 and pH 2.7

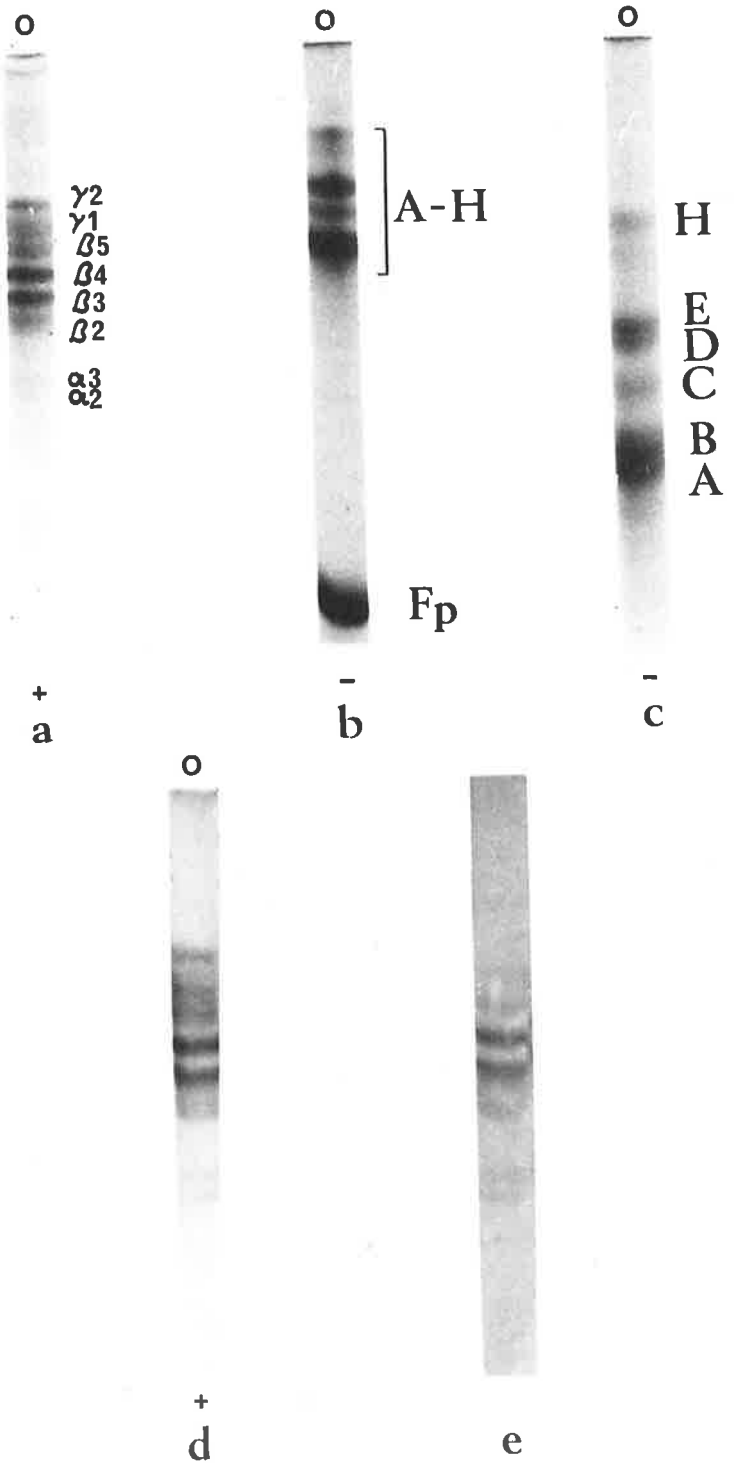
Samples of SCM- feather keratin, (50 μg in gel a and 100 μg in gels b and c) were subjected to polyacrylamide gel electrophoresis and subsequently fixed and stained as described in Methods.

gel	a	pH 9.5
gel	b	pH 2.7, 3 hr
gel	c	pH 2.7, 6 hr

50 μg of $[^{14}C]$- SCM- feather keratin (1.3×10^4 cpm) was subjected to electrophoresis and the gel stained and subjected to autoradiography

(d) 2mm gel slice photographed after drying

(e) Autoradiogram of the same gel slice.


o
γ2
γ1
β5
β4
β3
β2
α3
α2
+
a
o
A-H
Fp
-
b
o
H
E
D
C
B
A
-
c
o
+
d
e

in Methods, and an autoradiogram of the same gel; the staining and radioactivity patterns are indistinguishable.

These gel profiles are entirely reproducible and the band patterns of Figure 3.1 have been repeated over 100 times on various different preparations of SCM- feather keratin and have never been shown to differ detectably between different batches of down feathers. Neither has quantitative denitometry of gels of such preparations revealed significant differences in the proportions of the various bands. To establish that the multiple bands were not the result of incomplete S-carboxymethylation, a sample of SCM- feather keratin containing no radioactivity was subjected to a further round of alkylation with $[^{14}C]$-iodoacetic acid (Methods). Normally, 50% of the radioactive iodoacetic acid reacts with reduced feather keratin molecules but less than 0.01% of the radioactive iodoacetic acid reacted with SCM- feather keratin. The efficiency of the S-carboxymethylation was at least 99.98% on this basis.

(ii) The Amino Acid Composition of Feather Keratin

Samples of feather keratin, reduced but not S-carboxymethylated, and SCM- feather keratin were hydrolysed and subjected to amino acid analysis (Table 3.1). The two analyses differ significantly only in that SCM- feather keratin contains 7.7 mole % (residues / 100 residues) of SCM- cysteine and no detectable amounts of cysteine or cystine. In neither preparation is there detectable amounts of methionine. The low values for lysine and histidine (0.3 and 0.9 respectively in reduced feather keratin and 0.5 and 0.7 respectively in SCM- feather keratin) do not differ significantly between the

TABLE 3.1 Amino Acid Compositions of reduced feather keratin and of SCM- feather keratin

Two 5mg samples of reduced feather keratin (SH-FK) and of SCM- feather keratin (SCM-FK) were hydrolysed and amino acid analysis carried out on 400μg portions of each. The values are expressed as mole % of each residue (residues /100 residues).

TABLE 3.1

	SCM-FK	SH-FK
SCM-cys	7.8	0.0
asn+asp	5.6	5.2
thr	4.3	4.4
ser	12.6	13.1
gln+glu	7.4	7.6
pro	11.5	11.9
gly	12.3	14.1
ala	4.5	4.2
$\frac{1}{2}$cys	0.0	5.7
val	8.3	7.5
met	0.0	0.00
ile	4.5	4.3
leu	7.4	7.8
tyr	2.6	2.7
phe	4.1	4.5
lys	0.5	0.3
his	0.7	0.9
arg	5.6	5.7

two preparations. It is likely from the molecular weight estimate of SCM- feather keratin (Chapter 4) that few if any feather keratin chains possess even a single residue of either lysine or histidine.

(iii) S-carboxymethylation of Defined Proteins

French bean plastocyanin was examined by pH 9.5 gels before and after S-carboxymethylation and a major band accounting for over 80% of the protein material of each gel was detected. Small amounts of three other bands were noticed in both preparations and presumably were minor impurities.

Harlow (unpublished results) examined the histones of red blood cells before and after S-carboxymethylation carried out by the procedure outlined in Methods. The only difference between the SCM- histone preparation and its intact counterpart when examined by pH 2.7 gels was the absence of a slow band previously identified as a disulphide linked dimer of histone F3 (Fambrough et al., 1968) in the SCM- histone preparation. The absence of significant amounts of bands other than those characteristic of intact histones (before S-carboxymethylation) suggests that no partial modification of residues other than cysteine has taken place.

DISCUSSION

The purpose of the experiments reported in this Chapter was to show that the extraction and alkylation of keratin from chick feathers as described in sections (i) and (ii) of Methods produced a fully S-carboxymethylated but otherwise intact protein product. The importance of this conclusion in evaluating the results of subsequent experiments is fundamental.

Iodoacetic acid is widely used for chemical modification of proteins and its usefulness is in part due to the fact that it modifies different amino acid residues to varying degrees depending upon the pH value. The carboxymethylation of thiol groups proceeds very rapidly at pH values above 7.0 (Gurd 1967) but alkylation of methionine residues at pH values greater than 2 (Vithayathil and Richards, 1960), of histidine residues above pH 5 (Crestfield et al., 1963) and of lysine residues above pH 8.5 (Heinrikson, 1966) have been reported. The resolution of SCM- feather keratin into multiple bands by electrophoresis at both pH 9.5 and at pH 2.7 could concievably have been due to incomplete modification of residues other than cysteine and to test this possibility, amino acid analysis of reduced and of SCM- feather keratin was carried out. The complete absence of methionine from both protein samples precludes the possibility that its partial modification in SCM- feather keratin molecules could account for any of the observed bands. In both protein

samples, the values for histidine and lysine were also very low and in any case not significantly altered by S-carboxymethylation.

S-carboxymethylation of plastocyanin which contains 1 cysteine, 2 methionine and 4 lysine residues (Milne and Wells, 1970) did not produce multiple bands as determined by pH 9.5 gels - nor did S-carboxymethylation of avian red blood cell histones produce extra bands on pH 2.7 gels which could have been due to carboxymethylation of lysine or histidine residues.

The gel of $[^{14}C]$-SCM- feather keratin and its autoradiogram indicate that all the pH 9.5 bands (Fig. 3.1b) represent cysteine rich proteins alkylated by iodoacetic acid and the unreactivity of SCM- feather keratin to further reaction with $[^{14}C]$-iodoacetic acid establishes that each band of the gels of Figure 3.1 contains no unreacted cysteine residues.

In summary, of the amino acid residues potentially susceptible to carboxymethylation, feather keratin contains prominent amounts of only cysteine and there is no evidence that the small amounts of histidine and lysine present are susceptible to carboxymethylation under the conditions used. In accord with this, S-carboxymethylation of plastocyanin or red blood cell histones introduced no additional heterogeneity as determined by gel electrophoresis at pH 9.5 or at pH 2.7. The multiple bands detected by gel electrophoresis at pH 9.5

and at pH 2.7 of SCM- feather keratin therefore reflected the existence in feather cells of a number (at least 7) of different keratin chains.

CHAPTER 4

THE NUMBER OF PROTEINS IN THE DOWN FEATHER

INTRODUCTION

This chapter is concerned with the number of proteins which comprise the down feather of the chick and their relation to one another. The importance of this question must be considered from two standpoints. Firstly, studies on feather keratin from different morphological parts of the same feather (Harrap and Woods, 1964a; Kemp and Rogers, 1972) and of feather keratin from different avian species (O'Donnell, 1973a) hitherto have not revealed the high degree of heterogeneity in feather keratin which is associated with the mammalian keratins of hair and wool (Fraser et al., 1972).

Typically, S-carboxymethylated samples of feathers produce 7 or fewer major bands when examined by gel electrophoresis at one pH whereas extensive fractionation of wool proteins has revealed the presence of at least 60 different protein constituents. The apparent simplicity of protein composition of avian feathers compared to that of mammalian hair seemed worthy of further investigation.

Kemp and Rogers (1972) proposed that for each type of keratin forming tissue, be it scale, claw or morphologically distinct parts of the feather, an individual set of structural genes was activated. The simultaneous synthesis of all the electrophoretic variants of SCM- feather keratin reflected the activation of one such set of genes. It was important to establish how many genes were involved and to determine this, the number of proteins in the embryonic

chick feather was investigated.

METHODS

(i) DEAE -Cellulose Chromatography

Whatman DE - 11 was freed of fines by repeated decantation of DE - 11 suspensions, in water. It was then washed successively with 0.5 M NaOH, H_2O, 0.5 M HCl and H_2O until it was completely free of brown pigment. The resin was then suspended in 8.0 M urea, 0.01 M Tris, 0.001 M EDTA, pH 7.0 (HCl), packed into a glass column (90 x 1.8 cms) and equilibrated with that buffer overnight. SCM- feather keratin was dissolved in the above buffer (100 mg/ml) and dialysed against a 100-fold excess of fresh buffer for 3 hr. The protein solution was then applied to the DEAE-cellulose column and eluted with a linear salt gradient from 0 - 0.2 M KCl made up in the same buffer. The eluate was continuously monitored at 280 nm with an LKB Uvicord and fractions collected with an LKB fraction collector. The conductivity of selected fractions was determined with a Radiometer conductivity meter. Appropriate fractions (see Results) were pooled and dialysed against four changes of double distilled water, freeze dried and redissolved in water with the dropwise addition of 0.01 M NH_4OH if required. Protein concentrations were estimated by the method of Lowry et al., (1951), and known amounts of protein lyophilised for electrophoresis.

(ii) Gel Filtration on Sephadex G-100

Sephadex G-100 (fine) powder was suspended in elution buffer (either 8 M urea, 0.10 M Tris, pH 7.0, or 2.5 M urea, 5.4% acetic acid), allowed to swell overnight and packed into a glass column (1.0 x 90 cm). At no stage during the operation of the column was the hydrostatic head allowed to rise above 20 cm and was maintained at a constant level during packing and elution of samples using a Marriot flask as buffer reservoir. Protein samples (10 - 100 mg) were loaded without altering the hydrostatic head and their elution was detected by reading the absorbance of fractions in 2mm quartz cells at 280 nm in a Shimadzu Spectrophotometer (Model QV50). The elution of $[^{14}C]$ - SCM- feather keratin was detected by precipitation of aliquots from eluted fractions with 10% TCA and subsequent determination of radioactivity was carried out as described in Chapter 2. The protein samples used to calibrate the Sephadex G-100 column were reduced and carboxymethylated before use as described in Chapter 3.

(iii) SDS Gel Electrophoresis

Electrophoresis of protein - SDS complexes on polyacrylamide gels was based on the method of Laemli (1970). The running gel was prepared by mixing equal volumes of two solutions A and B with solid ammonium persulphate added to a concentration of 0.038% (w/v). The compositions of A and B were -

A: 0.75 M Tris, 0.2% SDS (w/v) and 0.05% TEMED (v/v) pH 8.8

B: 30% acrylamide (w/v), 0.8% NN' methylenebisacrylamide (w/v).

The stacking gel was prepared by mixing two solutions C and D in equal volumes and adding ammonium persulphate to 0.076% (w/v). The compositions of C and D respectively were -

C: 0.25 M Tris, 0.2% SDS (w/v) 0.05% TEMED (v/v) pH 6.8 (HCl)

D: 6% acrylamide (w/v) 0.8% NN' methylenebisacrylamide (w/v).

The gel buffer was that described by Laemli (1970). Proteins were dissolved at a concentration of 10mg/ml in a solution of 1% SDS, 4 M urea, 0.1 M β-mercapto-ethanol, 0.005% bromophenol blue and 10% sucrose (Weber and Osborn, 1969). Electrophoresis was carried out at 3-4 mA per gel until the tracker dye had migrated 90% of the distance of the running gel. The position of the tracker dye was recorded by the lateral insertion of a piece of steel wire and staining and destaining carried out by the method of Fairbanks et al., (1971).

(iv) Elution of protein bands from Polyacrylamide Gels

After electrophoresis of a protein sample, the bands were localised in the gel by its immersion in 10% TCA. The bands of interest were cut out and eluted from the gel

slice by allowing it to stand in several changes of 8 M urea for 3 one hour intervals. The extracts were pooled, dialysed and lyophilised.

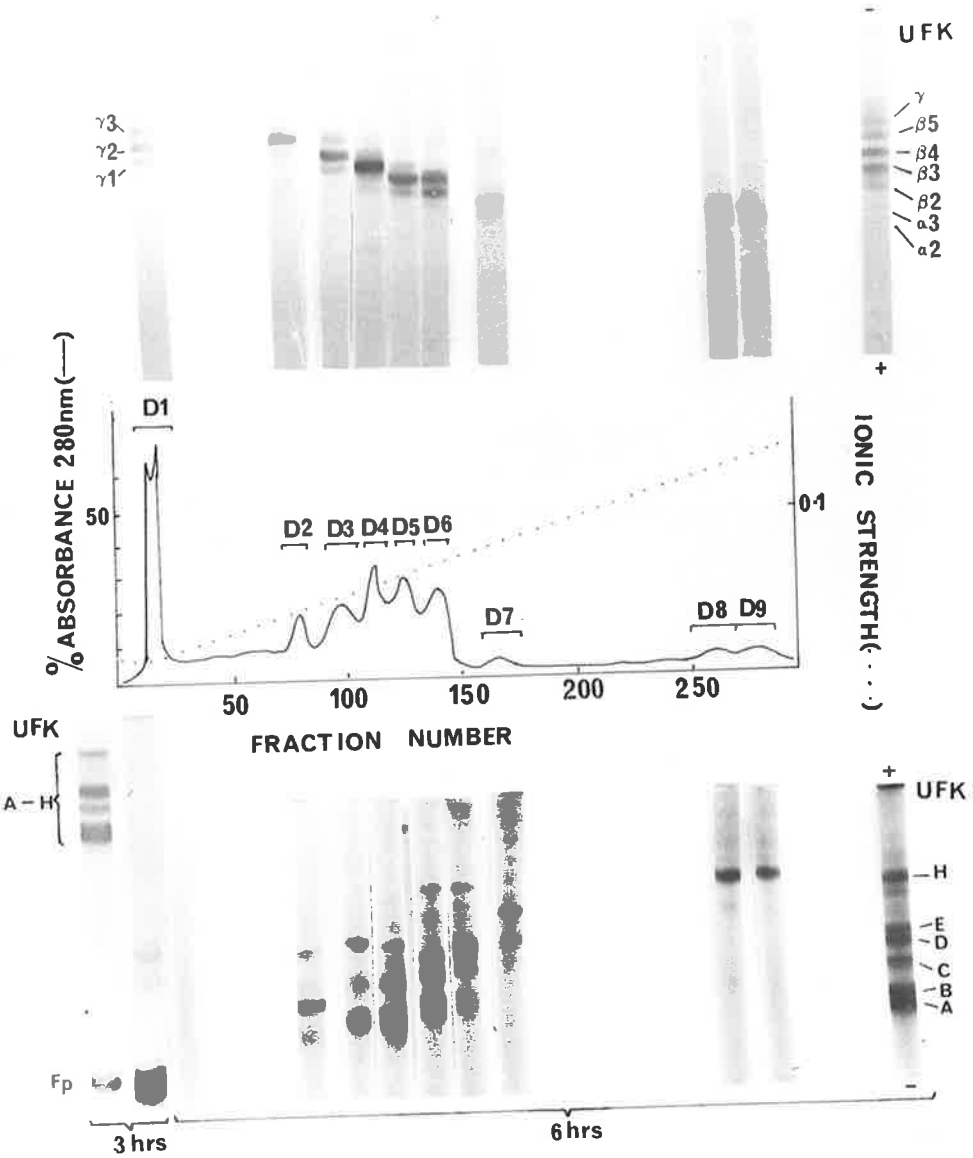
RESULTS

In order to investigate the number of discrete proteins in chick down feathers, their SCM- derivatives were separated by DEAE- cellulose chromatography into protein fractions which contained essentially single bands when examined by electrophoresis at pH 9.5. Examination of these fractions by electrophoresis at pH 2.7 allowed an estimate of the minimal number of proteins comprising the chick down feather.

(i) DEAE - Cellulose Chromatography

The results of DEAE- cellulose chromatography of SCM- feather keratin are shown in Figure 4.1. Polyacrylamide gel electrophoresis (pH 9.5) of the proteins in the pooled fractions from the peaks are shown in the upper portion of Figure 4.1 and demonstrated that each peak eluted from DEAE-cellulose contained essentially one major band. The proteins comprising the separate bands, eluted from the column in order of increasing electrophoretic mobility on pH 9.5 gels. The proteins of band β_3 eluted as two prominent peaks, D5 and D6, and this is consistent with the fact that this band is sometimes observed to resolve into a doublet on pH 9.5 gels. In order to establish the number of proteins present in each

Figure 4.1: DEAE-cellulose chromatography of SCM-feather keratin. 500 mg of SCM-feather keratin was applied to the DEAE-cellulose column, fractions were pooled as shown and samples prepared for polyacrylamide gel electrophoresis as described in Methods. The upper set of gels are of protein samples from the column run at pH 9.5. The lower set are an identical set of protein samples run at pH 2.7. In both cases 50 μg of protein was loaded on to each gel. The two gels at the bottom left are pH 2.7 gels of unfractionated SCM-feather keratin and fraction D1 run for $2\frac{1}{2}$ hours in order to retain the Fp component. The gels labelled "UFK" are of unfractionated feather keratin.


UFK
γ3
γ2
γ1
γ
β5
β4
β3
β2
α3
α2
+
D1
D2
D3
D4
D5
D6
D7
D8
D9
50
0·1
% ABSORBANCE 280nm(—)
IONIC STRENGTH(· · ·)
50
100
150
200
250
FRACTION NUMBER
UFK
A – H
Fp
+
UFK
H
E
D
C
B
A
–
3 hrs
6 hrs

of the major bands observed by gel electrophoresis of the DEAE- cellulose fractions, the proteins of fractions D1 - D9 were examined by electrophoresis on pH 2.7 gels. It is clear from Figure 4.1 that the only fractions which were homogeneous on both pH 9.5 and pH 2.7 gels were D8 and D9. The major bands β_2 - β_5 are all composed of multiple constituents as determined by pH 2.7 gels. The bands observed on pH 2.7 gels from fractions D1 - D9 can be related to the bands in unfractionated feather keratin and this argues against any modification during the analytical procedures. The component Fp was detected only in fraction D1.

Some of the protein fractions of Figure 4.1 when examined by electrophoresis at pH 9.5, in addition to the major band, contained minor contaminating amounts of adjacent bands - hence D5 contained predominently band β_3 (80%) but traces of β_4 (8%) and β_2(12%) were also present. In some fractions, as in the case of D5, these contaminants were present in a sufficiently high proportion to account for one or more of the multiple bands observed on the corresponding pH 2.7 gel. Thus the level of contamination of β_2 in fraction D5 (12%) could conceivably account for the presence of band H in the pH 2.7 gel of D5 (10%) although not for bands D (20%), C (20%) or A+B (50%). Table 4.1 presents the compositions of each fraction as determined by densitometry of the corresponding pH 9.5 and pH 2.7 gels and those values in parentheses refer to minor pH 2.7 bands whose proportion

Table 4.1. Compositions of protein fractions D1-D9

The pH 9.5 and pH 2.7 gels of Figure 4.1 were subjected to destitometry and the proportion of each band expressed as % of total absorbance on the gel. Values in parenthesis refer to bands present in lower amount on the pH 2.7 gels than contaminants on the corresponding pH 9.5 gel.

[1] D7 contained 2, pH 2.7 bands one of which was (D), the other of which was not a prominent constituent of unfractionated SCM-feather keratin

TABLE 4.1

Compositions of Protein fractions D1 to D9

Protein Fraction	Composition (% of total) pH 9.5									pH 2.7					
	γ_2	γ_1	β_5	β_4	β_3	β_2	α_3	α_2	H	E	D	C	(A+B)	Fp.	No. of Components
D1	40	30	30											100	
D2		90	10							(10)		60	30		3
D3		20	70	10						25		(15)	70		3
D4				>90						20		20	60		4
D5				8	80	12			(10)		20	20	50		4
D6					75	25			(5)		50	(15)	30		1
D7						80					40[1]				2
D8							85	15	100						1
D9							10	90	100						1

Total = 19

in the protein fractions can be accounted for by high pH bands contaminating the main band of each fraction. When the degree of purity of the pH 9.5 bands is taken into account, a total of nineteen prominent electrophoretic variants can be deduced (Table 4.1).

(ii) Amino Acid Compositions of Protein Fractions D1-D9

Table 4.2 shows the amino acid compositions of fractions D1-D9 from the DEAE- cellulose experiment. The values for SCM- cysteine were found in two independent sets of hydrolysates to be variable and where no reasonable correspondence between the two experiments was found the values have been omitted. The only fraction markedly distinct from the others in amino acid composition is D1 whose content of the amino acids histidine, phenylalanine and tyrosine is remarkably higher than all the others. The value of SCM-cysteine was reproducibly much lower in this fraction than in the others. A high value for lysine in fraction D2 places it apart from the other fractions and reproducible variations in the aromatic amino acids occur between the other fractions (D3 - D8+9). The similarity in amino acid composition between fractions D2-D9 is however the most obvious feature of Table 4.2 and their amino acid compositions bear close similarities to SCM- feather keratin prepared from other species and to that of adult chick. Feather keratins are typically rich in serine glycine, aliphatic amino acids and cysteine. Lysine,

Table 4.2 Amino Acid Compositions of Protein Fractions D1-D9

400 μg samples of the fractions D1-D9 were hydrolysed for 20 hr with 6 N HCl and analysed as described by Piez and Morris (1960). The amino acid composition of each residue is expressed as residues per 100 residues.

a. SCM- cysteine was subject to variable destruction during acid hydrolysis and where two independent analyses did not yield values within 20% of one another they have been omitted from the table (n.d.).

b. The protein sample D8+9 was obtained from a previous experiment and was a mixture of components α2 and α3.

TABLE 4.2 Amino Acid Compositions of Protein Fractions D1-D9

Amino Acid	Protein Sample							
	D1	D2	D3	D4	D5	D6	D7	D8+9[b]
SCMC[a]	00.11	n.d.	n.d.	8.06	5.61	7.01	7.50	7.35
ASP	5.01	6.19	7.12	5.49	6.39	5.75	5.33	4.96
THR	1.14	4.66	4.93	4.45	4.42	5.19	5.49	5.34
SER	8.63	16.57	16.43	14.68	16.22	15.64	16.18	15.68
GLU	5.21	8.74	8.06	7.27	7.96	8.49	7.76	7.31
PRO	3.86	8.38	12.05	11.12	11.65	11.68	12.28	12.10
GLY	17.80	17.39	12.53	12.05	12.36	12.44	11.56	11.39
ALA	0.97	5.72	5.22	4.78	4.76	4.61	4.50	4.46
VAL	0.61	6.15	8.64	7.41	7.42	6.96	8.15	8.09
ILE	0.62	4.29	5.38	4.68	4.62	4.19	3.99	4.24
LEU	7.52	6.83	7.28	7.46	7.78	7.47	7.27	6.64
PHE	8.89	2.27	3.44	4.45	3.90	3.57	2.47	3.51
TYR	12.80	3.85	2.11	1.12	1.15	1.92	1.13	2.19
HIS	12.12	1.03	0.00	0.67	0.00	0.00	0.72	0.96
LYS	1.09	2.03	0.00	0.43	0.00	0.00	0.51	0.71
ARG	7.62	5.90	6.02	5.79	5.80	5.05	5.11	5.00

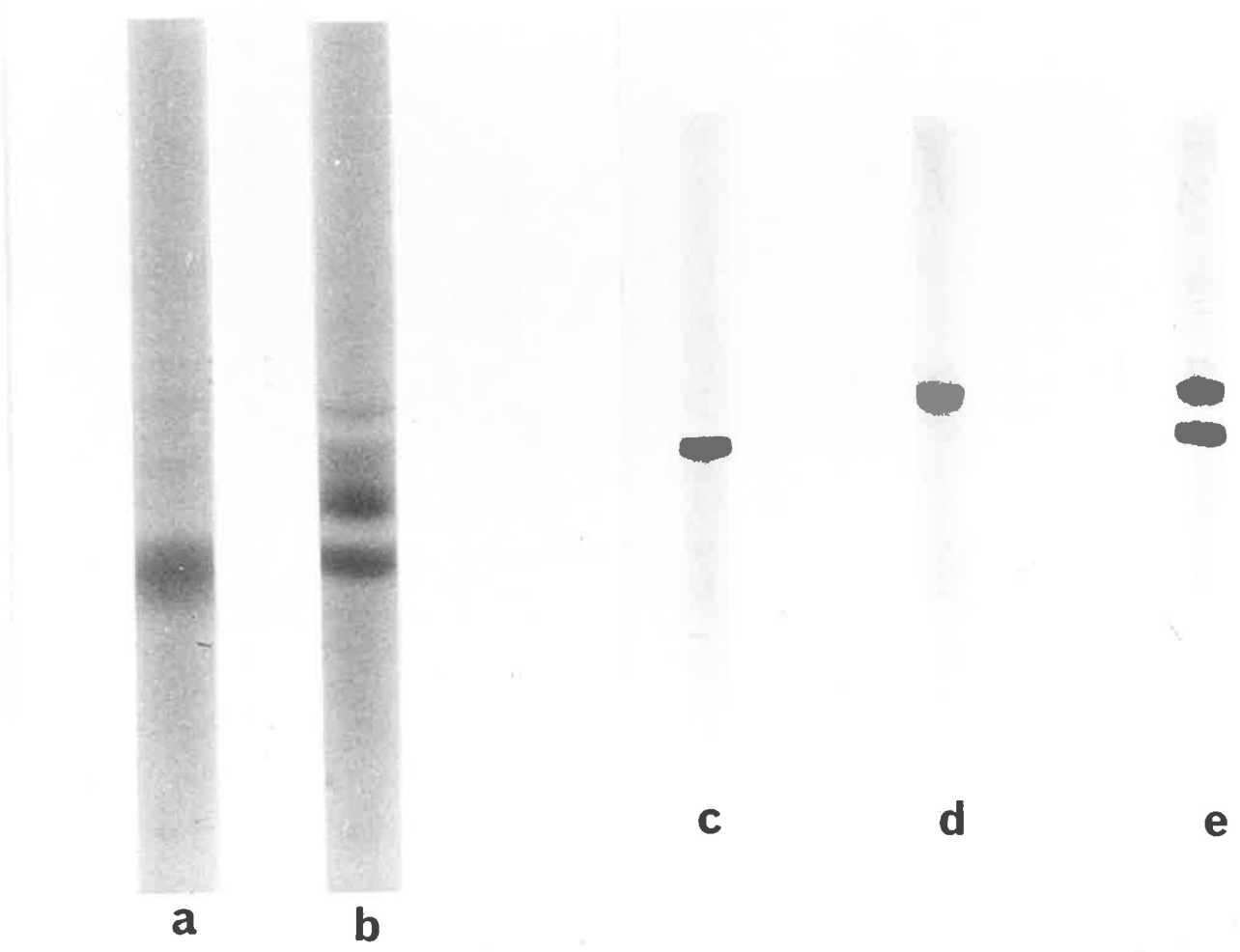
histidine and methionine are present only in trace amounts if at all (Harrap and Woods, 1964a, Woods, 1971. Kemp and Rogers, 1972. O'Donnell, 1973a, 1973b).

(iii) Determination of the Molecular Weight of Feather Keratin by SDS Gels

The determination of the molecular weight of feather keratin by SDS gel electrophoresis was carried out both on a sample of keratin which had been reduced but not S-carboxymethylated and on a sample of SCM- feather keratin. Figure 4.2 shows the results of SDS gel electrophoresis of reduced feather keratin (a) and SCM- feather keratin (b). The band in gel (a) was suggestive of molecular weight heterogeneity in the proteins of reduced feather keratin and densitometry of this gel clearly demonstrated asymmetry of the peak (Fig. 4.3a). This densitometer profile was interpreted as the composite of two sets of protein molecules, one with a molecular weight of 11,000 and the other set having a somewhat higher mobility on SDS gels and therefore lower molecular weight. SCM- feather keratin was resolved into two distinct major species by SDS gel electrophoresis (Fig. 4.2, gel (b)). The faster of the two bands aligned with the asymmetric shoulder on the leading edge of the reduced feather keratin peak as can be seen by a comparison of the densitometer profiles in Figure 4.3. The slower of the two bands of SCM-feather keratin (Fig. 4.2, gel b) has an apparent molecular

Figure 4.2. SDS polyacrylamide gels of reduced and of S-carboxymethylated feather keratin and RNase

50 μg aliquots of reduced feather keratin (gel a), SCM-feather keratin (gel b), reduced RNase (gel c), SCM- RNase (gel d) and reduced RNase + SCM- RNase gel (e) were subjected to SDS electrophoreses on 15% polyacrylamide gels

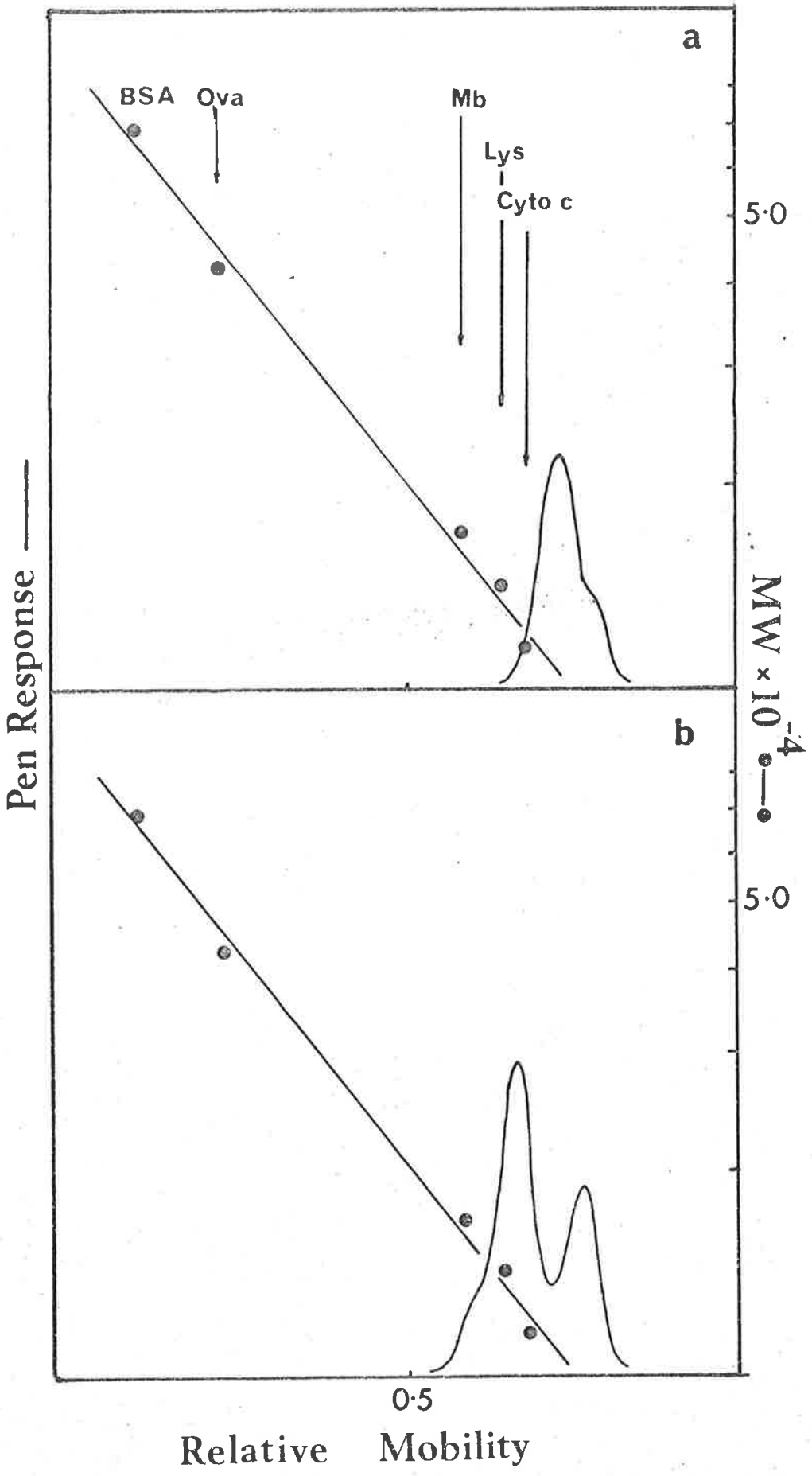

a
b
c
d
e

weight of 13500 and was considerably slower on SDS gels than either component of reduced feather keratin. To investigate this anomalous reduction in mobility and apparent increase in molecular weight caused by S-carboxymethylation of feather keratin, RNase was subjected to gel electrophoresis as its reduced form and its SCM- derivative. Figure 4.2 demonstrates that the mobility of SCM-RNase (gel d) is considerably lower than that of reduced RNase (gel c). Gel (e) containing both reduced and SCM-RNase allows a quantitative comparison between the two mobilities. S-carboxymethylation of RNase lowers its mobility by 11% whereas a reduction in mobility of only 1% is expected due to the molecular weight increase caused by addition of 8 carboxymethyl groups per molecule.

Presumably the reduction in mobility on SDS gels of one band of SCM- feather keratin compared to its reduced counterpart reflects the S-carboxymethylation of a set of feather proteins rich in cysteine. Conversly, that the mobility of the faster band of reduced feather keratin (Fig. 4.2, gel a) is not affected by S-carboxymethylation indicates that the proteins of this band contain few if any cysteine residues. The gel of reduced feather keratin allows the molecular weight of the major, slower peak to be estimated at approximately 11,000 (Fig. 4.3 a). The validity of this determination is considered in Discussion.

Figure 4.3. Densitometry of SDS gels of reduced feather keratin and SCM- feather keratin

The gels of Figure 4.2 were subjected to densitometry on a Photovolt Spectrophotometer. Protein markers (20 μg of each) BSA (Bovine Serum Albumin), Ova (Ovalbumin) Mb (Myoglobin) Lys (lysozyme) and Cyto c (cytochrome c) were also subjected to electrophoresis, staining and densitometry. The mobilities of each protein standard were measured and expressed relative to that of the bromophenol blue tracker dye.


a
BSA
Ova
Mb
Lys
Cyto c
5·0
b
5·0
Pen Response
MW × 10⁻⁴
0·5
Relative Mobility

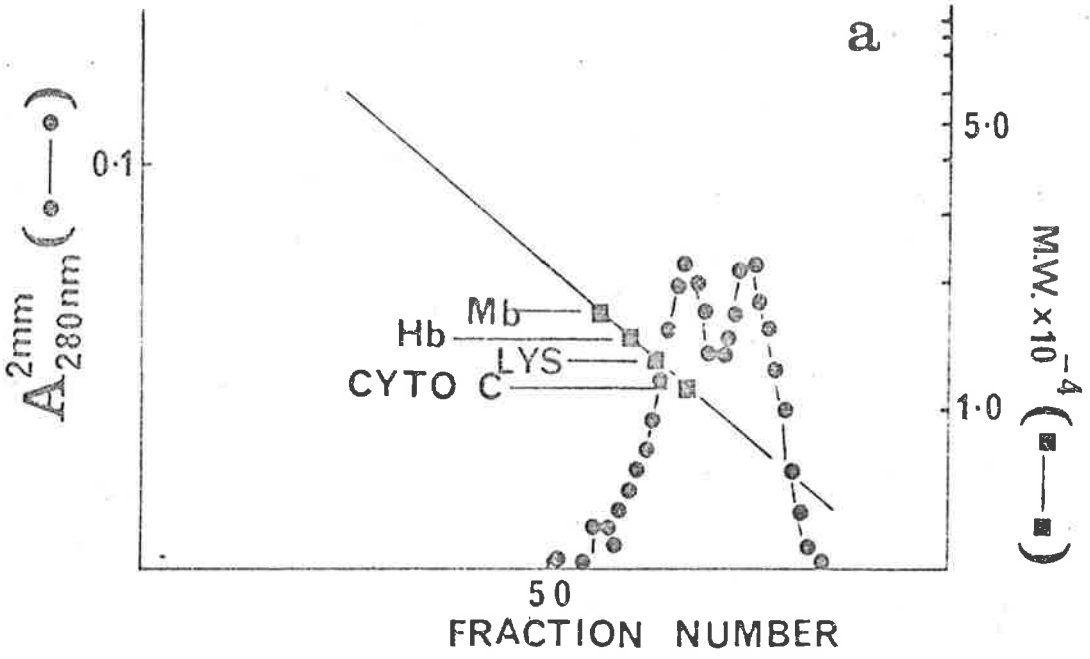
(iv) The Determination of the Molecular Weight of Feather Keratin by Gel Filtration on Sephadex G-100

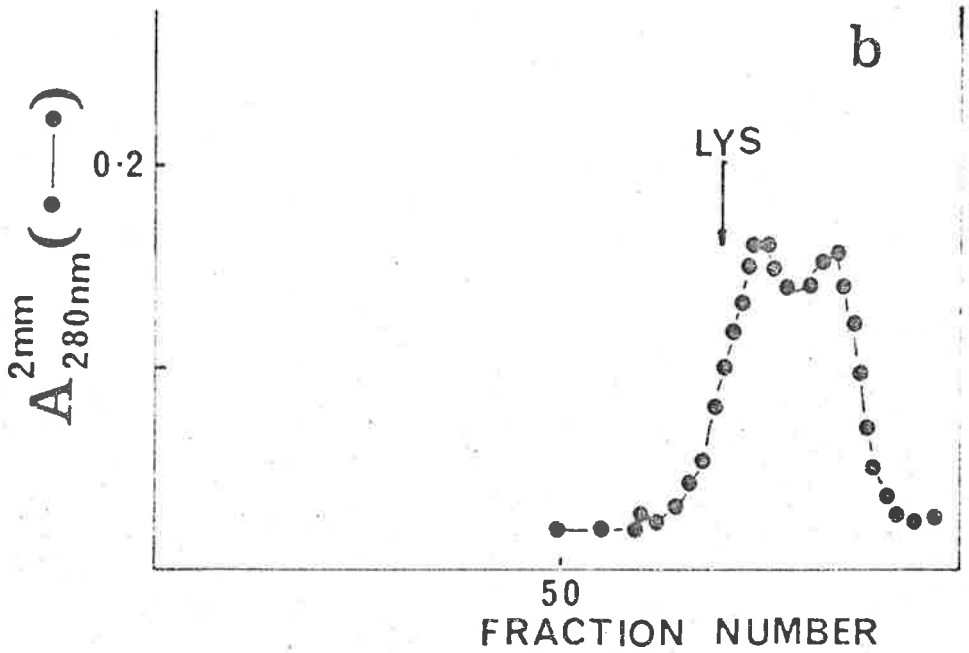
To obtain a more accurate molecular weight for the two sets of proteins distinguished by SDS gels of feather keratin and to prepare samples of each set for further examination, samples of SCM- feather keratin were subjected to gel filtration on Sephadex G-100.

a. Sephadex G-100 chromatography at pH 7.0 and pH 2.7

SCM- feather keratin was intially subjected to chromatography on Sephadex G-100 columns run under two sets of conditions. One column was run in 8 M urea, 0.1 M Tris pH 7.0 and was calibrated using the S-carboxymethylated protein markers myoglobin, globin, lysozyme, and cytochrome C. The chromatography of SCM- feather keratin on this column produced two major peaks indicating the elution of two sets of protein molecules with molecular weights of 11,700 and 8,400 respectively (Fig. 4.4a). The other column was run in 2.5 M urea, 5.4% acetic acid at pH 2.7 and chromatography of SCM- feather keratin also produced two peaks whose elution volume in relation to SCM- lysozyme (Fig. 4.4b) confirmed their identity to their counterparts from Figure 4.5a. The detection of two classes of proteins by chromatography of SCM- feather keratin on Sephadex G-100 is in accord with the results obtained by SDS gel electrophoresis (Fig. 4.2) and the molecular weight calculated for the peak of higher molecular weight from Sephadex chromatography is in excellent agreement with that obtained for the slower component of reduced feather keratin on SDS gels (Fig. 4.3). The results of gel filtration of

Figure 4.4. Chromatography of SCM- feather keratin on Sephadex G-100. The Sephadex G-100 column in (a) was run in 8 M urea, 0.1 M Tris, 0.001 M EDTA, pH 7.0 and that in (b) was run in 5.4% acetic acid, 2.5 M urea. Protein samples were loaded onto both columns at a concentration of 10 mg/ml. The absorbance of the fractions was determined at 280 nm in 2.0 mm quartz cells. The standards, myoglobin (Mb), haemoglobin (Hb) lysozyme (Lys) and cytochrome C (C_{yto} c) were reduced and carboxymethylated before chromatography.


a
0·1
5·0
1·0
Mb
Hb
LYS
CYTO C
50
FRACTION NUMBER


b
0·2
LYS
50
FRACTION NUMBER

SCM- feather keratin are consistent with the existence of monomers and not dimers or oligomers of the protein molecules at the concentrations of denaturant ionic strength and at the pH used for both sets of chromatography. The importance of this conclusion is considered in Discussion.

b. Sephadex Chromatography of $[^{14}C]$- SCM- feather keratin

To examine the properties of the two sets of protein molecules of feather keratin resolvable by gel filtration 20 mg (5.28×10^5 cpm) of $[^{14}C]$- SCM- feather keratin was applied to a column of Sephadex G-100 and aliquots of each fraction were precipitated with 10% TCA for measurement of radioactivity. Figure 4.5a shows the absorbance and radioactivity profiles of this experiment. The $[^{14}C]$ radioactivity eluted in a manner coincident with only the peak of higher molecular weight: no radioactivity was associated with the lower molecular weight peak. The tubes from this experiment were pooled as shown in Figure 4.5a and the three resulting protein fractions were exhaustively dialysed and freeze dried. The amount of protein in each fraction was determined (Lowry et al., 1951) and 50μg aliquots run on pH 9.5 and 100μg aliquots on pH 2.7 polyacrylamide gels. Figure 4.5b shows pH 9.5 polyacrylamide gels of the three fractions. Qualitatively all three samples share the major bands designated β_2, β_3, β_4, β_5 and γ_1 though in different proportions. Gel I of Figure 4.5b shows an enrichment of γ_1 and closely associated components when compared to unfractionated SCM- feather keratin. Gel II, that of protein

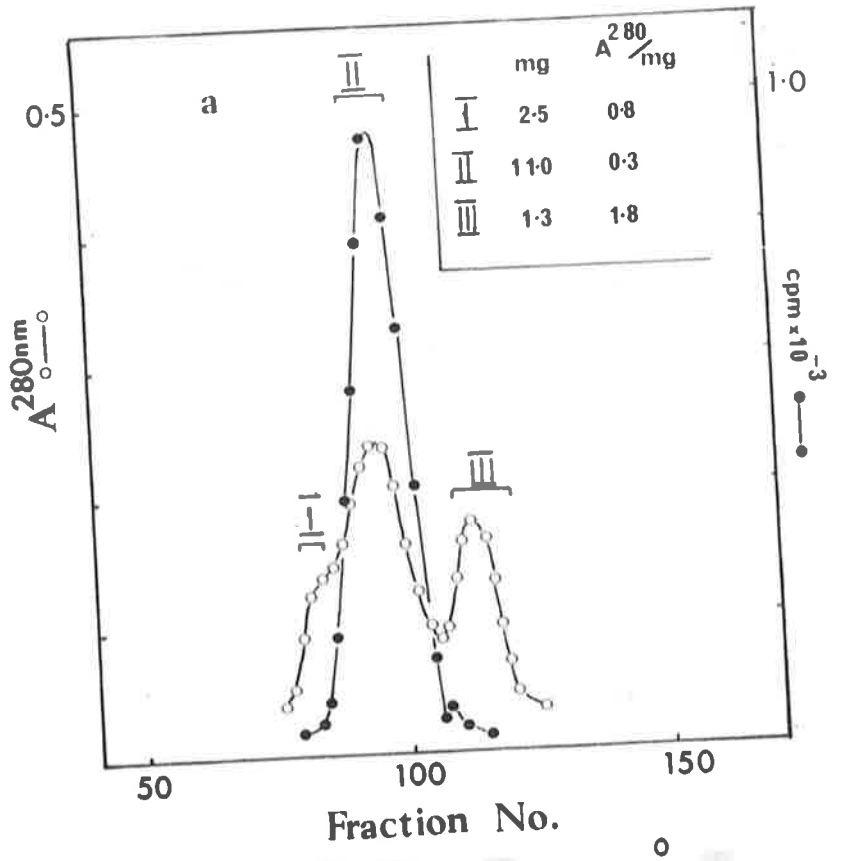
Figure 4.5. Chromatography of $[^{14}C]$ - SCM- feather keratin on Sephadex G-100

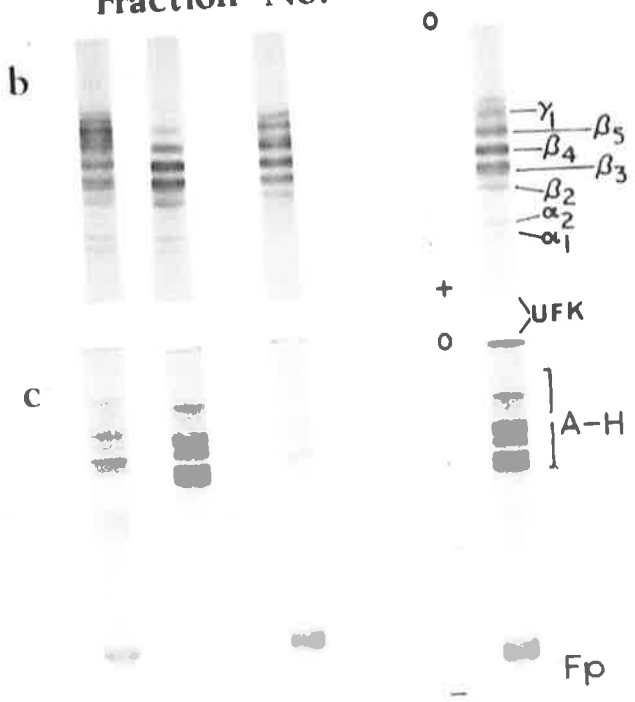
(a) Absorbance and radioactivity profiles of 20 mg of $[^{14}C]$ - SCM- feather keratin from Sephadex G-100. The column was run in 8 M urea, 0.10 M tris pH 7.0, and 50 μL aliquots sampled from each fraction for the determination of radioactivity.

(b) pH 9.5 polyacrylamide gels of dialysed and lyophilised fractions pooled as shown in (a) 50 μg of each protein fraction was loaded onto the gels.

(c) pH 2.7 polyacrylamide gels of fractions from (a). 100 μg of each protein sample was loaded onto the gels

The gels labelled "UFK" are of unfractionated SCM-feather keratin


a
mg
A^{280}/mg
I 2·5 0·8
II 11·0 0·3
III 1·3 1·8
0·5
1·0
A^{280nm}
cpm x 10^{-3}
50
100
150
Fraction No.


b
c
o
γ_1
β_5
β_4
β_3
β_2
α_2
α_1
+
UFK
o
A–H
Fp
–

pooled from the major peak of the profile (Fig. 4.3a), shows bands $\alpha 2$, $\alpha 3$, $\beta 2$, $\beta 3$, $\beta 4$, $\beta 5$ and $\gamma 1$ in the same proportions as in unfractionated feather keratin and gel III, of protein pooled from the lower molecular weight peak, displays no detectable amounts of $\alpha 2$, and $\alpha 3$ and significant enrichment of the bands $\beta 5$, $\gamma 1$ and $\gamma 2$. The differences between the protein fractions are more pronounced when examined by polyacrylamide gels run at pH 2.7 (Fig. 4.5c). Fractions I and III alone contain the component designated Fp - gel II contains no significant amount of Fp but contains the components A-H in the same proportions as they occur in unfractionated SCM- feather keratin. Gel III contains only Fp.

The constituent proteins of Fp eluted from the Sephadex G-100 column before and after the radioactive protein and are therefore heterogeneous with respect to molecular weight. The sharpness and symmetry of elution of the single peak of radioactivity in Figure 4.5a suggested that most or all the protein molecules of feathers susceptible to S-carboxy-methylation with iodoacetic acid were uniform with respect to molecular weight and that the differences between the proteins in bands A-H on pH 2.7 gels must result from charge and not from size. Protein fraction I (Fig. 4.5a) contains low amounts of bands A-H in addition to a prominent amount of Fp (Fig. 4.5c). The occurence of these small amounts of bands A-H in this fraction presumably reflected the incomplete resolution of bands A-H, eluting as a discrete radioactive

peak from the Sephadex G-100 column, from the higher molecular weight fraction of Fp which appears as a shoulder of optical density but not of radioactivity in Figure 4.5a.

(v) Elution of Fp from pH 2.7 Gels

On each of ten pH 2.7 gels was run 200μg of SCM- feather keratin. The bands were localised by immersion of the gels in 10% TCA - after 10 minutes the bands were clearly visible as white precipitates. The region of gel containing the Fp band was excised and protein eluted as described in Methods. After dialysis and freeze drying samples of Fp protein were subjected to pH 9.5 electrophoresis. Four major bands β_3, β_4, β_5 and γ_1, and two minor ones β_2 and γ_2 were present. In another experiment [^{14}C] - SCM- feather keratin (1.15 x 10^6 cpm) was subjected to pH 2.7 electrophoresis and the region of gel containing band Fp and bands A-H cut out separately and the amount of radioactivity in each fraction determined. Bands A-H contained 3.78 x 10^5 cpm and band Fp contained 2.30 x 10^2 cpm.

DISCUSSION

(i) The two Sets of Proteins in the Chick Down Feather

Despite the high degree of heterogeneity for the proteins of chick down feather which has been established and described in this Chapter, there appears to be two major sets

of protein molecules whose properties when examined by a number of analytical criteria were distinctive enough to justify their separate classification. It is the purpose of this section to define the properties of the two sets of protein molecules and to relate these properties to the general structure of each set. To facilitate this Discussion, the two sets of proteins have been considered individually and each set is examined on the basis of molecular weight, amino acid composition, chromatographic properties and behaviour on polyacrylamide gels.

(ii) The Properties of the Proteins of Set F

The evidence presented in Results and the subsequent list of properties favour the interpretation that set F proteins are the sole constituents of the band designated Fp on pH 2.7 acrylamide gels. They are heterogeneous with respect to size and charge and have a distinct amino acid composition. It is unlikely that they can be strictly classified as keratin molecules.

a. Molecular Weight

The results of both SDS gels (Fig. 4.2) and Sephadex G-100 chromatography (Figs. 4.4 and 4.5) indicate that there are two major classes of protein molecule in chick feather. The SDS gel (a) in Figure 4.2 is of reduced feather keratin and shows two barely resolved bands of protein the slower of which has a molecular weight of approximately 11000. The mobility of a protein SDS complex is in inverse proportion to the

logarithim of its molecular weight (Shapiro et al., 1967). However this relationship is valid only for protein molecules whose molecular weights are greater than about 11,000 (Bryce and Crighton, 1971; Inoyue, 1971; Williams and Gratzer, 1971). Smaller protein molecules migrate on SDS gels more slowly than the empirical relationship mobility$\propto$ -log mW dictates. The incomplete resolution of the two major bands in Figure 4.2 (gel a) presumably results from the inability of SDS gels to adequately separate small protein molecules although the actual differences in their size may be considerable.

S-carboxymethylation of RNase was found to significantly lower the mobility relative to that of the reduced form on SDS gels (Fig. 4.2, gels c and d). Similar observations have been made with other derivatives of protein molecules which are more acidic than the intact molecule. Tung and Knight (1972) showed that the mobility on SDS gels of maleylated proteins relative to their intact counterparts was reduced by more than could be accounted for by the molecular weight increase due to modification.

S-carboxymethylation reduced the mobility of only one of the reduced feather keratin bands on SDS gels. Gel (b) in Figure 4.2 is of SCM- feather keratin and the two major bands are completely resolved. The faster of the two bands in gel (b) migrates at the same rate as its counterpart in gel (a) - S-carboxymethylation apparently has not changed its mobility on SDS gels. The content of SCM- cysteine in the

fast band must therefore be very much lower than that of the slower band whose mobility is reduced considerably by S-carboxymethylation. These experiments establish the presence of two sets of protein molecules one with a molecular weight of 11,000 and with a high content of cysteine and the other with a molecular weight lower than 11,000 and containing little if any cysteine. The properties of the fast band are consistent with those proposed for constituents of protein set F.

Chromatography of SCM- feather keratin on Sephadex G-100

Sephadex G-100 chromatography (Figs. 4.4 and 4.5) separated the constituent proteins into two well resolved peaks of protein whose molecular weights were 11,700 and 8,400 respectively. $\left[^{14}C\right]$ SCM- feather keratin produced only one radioactive peak on Sephadex chromatography (Fig. 4.5a) which aligned precisely with the absorbance peak of 11,700. The low molecular weight peak was devoid of radioactivity and thus also of SCM- cysteine; it was composed entirely of proteins which imigrated on pH 2.7 polyacrylamide gels as the Fp band (Fig. 4.5C, peak III). These properties strongly imply that the low molecular weight protein (peak III) of Figure 4.5c and the fast band of the SDS gels of Figure 4.2 are one and the same and that it is a prominent constituent of set F proteins.

The shoulder of absorbance on the leading (high molecular weight) edge of peak II in Figure 4.5a when examined by gel electrophoresis at pH 2.7 was shown to

contain the band Fp in addition to contaminating amounts of bands A-H derived from the main peak. The protein constituents of set F are therefore heterogeneous with respect to size.

b. Chromatographic properties on DEAE- cellulose

DEAE- cellulose chromatography of SCM- feather keratin (Fig. 4.1) resolved the constituent proteins into 9 prominent peaks and examination of fractions D1-D9 revealed that only D1 contained protein corresponding to the Fp band on pH 2.7 gels. The pH 2.7 gel of D1 in Figure 4.1 contained only the Fp band and this fraction as a consequence consisted solely of set F protein constituents. This does not imply that all the set F protein constituents eluted in fraction D1 and other members of this set may have eluted in the region of the chromatogram between D1 and D2 (Fig. 4.1). However, the absence of band Fp in the pH 2.7 polyacrylamide gels of fractions other than D1, eliminated the possibility of contamination of set F protein constituents in fractions D2-D9.

c. Amino Acid Composition

The amino acid composition of protein fraction D1 (Fig. 4.1) is remarkably different from the other fractions (D2-D9) examined. Its high content of the basic amino acids histidine and arginine (Table 4.2) accounts for the high electrophoretic mobility observed on pH 2.7 polyacrylamide gels. The absence of SCM- cysteine in fraction D1 is consistent

with the absence of radioactivity observed in the peaks of set F proteins eluted from Sephadex G-100 (Fig. 4.5a). Furthermore the fact that the Fp band contained less than 0.1% of the radioactivity present in bands A-H from the pH 2.7 separation of a sample of $[^{14}C]$-SCM- feather keratin points to the absence of SCM- cysteine in <u>any</u> prominent member of the set F proteins.

d. <u>Electrophoretic properties</u>

Two samples of set F protein, those of peak III of the Sephadex G-100 chromatography in Figure 4.5a and those of protein fraction D1 from DEAE- cellulose chromatography in Figure 4.1 were examined by electrophoresis at pH 2.7 and at pH 9.5. Both samples were composed entirely of the Fp band on pH 2.7 gels but both gave rise to multiple bands on pH 9.5 gels. The low molecular weight peak from Sephadex G-100 chromatography (peak III Fig. 4.5 a) gave rise to prominent amounts of bands β_3 β_4 β_5 and γ_1 on pH 9.5 gels (Fig. 4.5 b). Fractions D1 (Fig. 4.1) however gave rise to β_5 γ_1 and γ_2 when analysed on pH 9.5 gels.

When the Fp band was eluted from pH 2.7 gels of SCM- feather keratin and the proteins examined by pH 9.5 electrophoresis, bands corresponding to β_3, β_4, β_5 γ_1 and γ_2 were prominent. This preparation of Fp represents the <u>total</u> set F proteins and contains all the bands present in both fraction D1 from DEAE- cellulose (Fig. 4.1) and peak

III from Sephadex G-100 (Fig. 4.5a). All 3 samples of set F protein each differring in protein composition according to their method of preparation were heterogeneous when examined by pH 9.5 gels.

One other distinguishing feature of set F proteins is their high relative absorbance. The insert of Figure 4.5a shows the yields of protein in fractions I, II and III and the absorbance at 280 nm per milligram of protein as determined by the Lowry estimation. The ratio A^{280}_{1cm}/mg of protein is five fold higher for peak III than peak II. The high relative absorbance presumably reflects the high content of tyrosine, phenylalanine and possibly tryptophan of set F protein constituents. One consequence of this is that methods which determine protein concentration by absorbance at 280 nm would detect set F proteins with fivefold greater sensitivity than the other major set (set K) of proteins.

In summary, the proteins of set F

1. Exclusively constitute the band Fp on pH 2.7 polyacrylamide gels
2. Are heterogeneous with respect to molecular weight and also charge
3. Are devoid of cysteine or cystine and most if not all set F proteins are rich in the aromatic amino acids and histidine
4. Constitute only a minor amount (approximately 10%) of the proteins of SCM- feather keratin.

None of these properties is characteristic of keratin molecules examined from other avian sources or of the other major set (K) of proteins of embryonic chick feather. Set F proteins do not fulfil any of the chemical requirements such as amino acid composition which would justify their designation as keratin and it is therefore of interest that in embryonic feathers of increasing age the band Fp becomes detectable at least 4 days after bands A-H have appeared at day 13 of embryonic age (D. Kemp., PhD Thesis). The localisation of set F protein molecules within feather tissue, whether intracellular or extracellular may provide an insight as totheirbiological function.

(iii) The Properties of Protein Set K

The proteins comprising set K are the main constituents of the embryonic chick feather. They are heterogeneous with respect to charge and while having similar amino acid compositions, all members of the set are the same size. All their properties are consistent with those reported for keratin molecules from different avian sources and this class of protein molecule is the subject of all subsequent chapters of this thesis.

a. Molecular Weight

Sephadex G-100 chromatography of $[^{14}C]$SCM- feather keratin (Fig 4.5a) results in the elution of only onesymmetrical peak of radioactive protein. This peak aligns precisely with

the most prominent absorbance peak corresponding to a molecular weight of 11,700. Tubes pooled from the centre of this peak (Fraction II) when examined by electrophoresis at pH 9.5 (Fig. 4.5b) and pH 2.7 (Fig. 4.5c) respectively were shown to contain every band characteristic of unfractionated SCM- feather keratin with the exception of band Fp on pH 2.7 gels which was totally absent. The bands α_2 - γ_1 of Fraction II (Fig. 4.5b, gel II) were present on pH 9.5 gels in the same proportions as they occurred in unfractionated SCM- feather keratin and likewise the bands, A-H, of Fraction II were qualitatively and quantitatively indistinguishable from their counterparts in unfractionated SCM- feather keratin. Despite the obvious heterogeneity of the proteins of fraction II on both pH 9.5 and pH 2.7 gels, there is no evidence of any heterogeneity with respect to molecular weight.

SDS gels of SCM- feather keratin exhibited two major well resolved bands (Fig. 4.2b). The slower of these major bands had an apparent molecular weight of 13,500 and showed no evidence of asymmetry which would be indicative of molecular weight heterogeneity. The value of 13,500 is certainly an overestimate due to the effect of S-carboxymethylation on the mobility of protein molecules rich in cysteine residues on SDS gels. The molecular weight of 11,000 deduced for the major band of reduced feather keratin (Fig. 4.3) is in close accord with that of 10,400 reported for the reduced proteins of chick feather rachis determined by ultracentrifugation

studies (Harrap and Woods, 1964b) and the value of 10452 for a prominent constituent of emu rachis whose amino acid sequence was determined (O'Donnell, 1973b).

b. Chromatographic properties on DEAE- cellulose

DEAE- cellulose chromatography of SCM- feather keratin allowed the isolation of nine discrete protein fractions D1-D9 (Fig. 4.1) and only D1 contained detectable amounts of the Fp band when analysed on pH 2.7 polyacrylamide gels. The protein samples D2-D9 are fractions of protein set K. Each of the fractions D2-D9 contains predominantly a single band as determined by pH 9.5 polyacrylamide gels and the order of elution of peaks containing the proteins of these bands is that of increasing electrophoretic mobility at pH 9.5. The concentration of urea during DEAE- cellulose chromatography was 8 M, sufficient to ensure the dissociation of set K molecules to monomers as evidenced by the elution of a single peak of radioactive (set K) protein of molecular weight 11700 from Sephadex G-100 run in 8 M urea. The resolution of fractions D2-D7 into multiple components by electrophoresis at pH 2.7 reflects the existence of multiple monomeric variants in each fraction and cannot be accounted for by association of a single monomer into dimers oroligomers before or during electrophoresis at pH 2.7 - Sephadex G-100 chromatography of SCM- feather keratin run in 5.4% acetic acid and 2.5 M urea, the same concentrations as are used in pH 2.7 gels (Panyim and Chalkley, 1969) results in the elution

of 2 peaks whose elution volumes are consistent with monomeric set K and set F proteins but not with dimers or oligomers (Fig. 4.3b). The resolution of protein samples D2-D7, each of which migrates as a single monomeric band on pH 9.5 gels into multiple monomeric sub species by electrophoresis at pH 2.7 must reflect the existence within each fraction of primary structure variants of set K proteins for the following reasons

1. Both pH 9.5 and pH 2.7 gels were run under conditions sufficient to ensure the existence of monomers
2. The specificity and efficiency of the S-carboxymethylation preclude the existence of partially alkylated species (Chapter 3)
3. No heterogeneity with respect to molecular weight could be detected in the set K proteins by Sephadex G-100 chromatography.

Figure 4.1 and Table 4.1 indicate that at least 19 primary structure variants constitute the set K proteins of SCM- feather keratin. This estimate does not take into account the presence of the minor bands, α_1, α_4 and β_1, which are commonly resolved on pH 9.5 gels (Kemp and Rogers, 1972). Nor does it take into account the possibility that electrophoresis at pH 2.7 will not completely resolve <u>all</u> protein constituents of each of the fractions D2-D9 (Fig. 4.1). The above estimate of 19 variants of SCM- feather

keratin is therefore likely to be an underestimate.

c. Electrophoretic properties

All the proteins of set K are electrophoretically distinct from those of set F when examined by pH 2.7 gels (Fig. 4.5c), the K set comprising bands A-H, the F set comprising band Fp. Many of the K set proteins (β_3- γ_1) co-migrate with F set proteins on pH 9.5 polyacrylamide gels (Fig. 4.5b) but are present in SCM- feather keratin in much higher amounts and for most purposes the contribution of set F proteins to these bands in unfractionated SCM- feather keratin may be neglected.

d. Amino Acid Composition

The amino acid compositions of fractions D2-D9 are remarkably similar (Table 2) and strongly suggest that most or all of the set K protein molecules bear close resemblance to one another in terms of primary structure.

In summary, the proteins of set K

1. Are rich in SCM- cysteine
2. Are homogeneous with respect to molecular weight (11700)
3. Consist of at least 19 different electrophoretic variants whose differences reside in their primary structure
4. Are very similar to one another with respect to amino acid composition.

Kemp et al., (1973) have shown that SCM- feather keratin prepared from a single chick embryo is indistinguishable by pH 2.7 and pH 9.5 gel electrophoresis from that

prepared from 10 or 20 newly hatched chicks as were the samples of SCM- keratin used in the experiments of this Chapter. The existence of at least 19 similar variants of the set K proteins of SCM- feather keratin is not due to the existence of genetic polymorphism within a large population of animals but is an intrinsic property of the proteins of the feathers of individual chicks and in all probability of individual feathers.

CHAPTER 5

THE TRYPTIC PEPTIDES OF SCM-DOWN FEATHER KERATIN

INTRODUCTION

The existence of a large family of keratin molecules (set K) in feathers of the newly hatched chick, all of which were rich in cysteine and of the same or very similar molecular weight was established in Chapter 4. There are two major possibilities which could account for the existence of such a family of proteins. One or a few genes could be active in producing protein molecules which are subsequently modified in a manner which alters their electrophoretic properties. Acetylation of specific lysine residues of some histone F2a molecules produces electrophoretic variants differing in positive change depending upon the extent of acetylation (Gorovsky, 1973). Similarly the two different chains of rabbit muscle aldolase are electrophoretically distinguishable and differ only in that a specific asparagine residue in one is aspartic acid in the other (Lai et al., 1970). Deamidation of the specific asparagine residue after the synthesis of aldolase is responsible for this change (Midelfort and Mehler, 1972). Both of these events in their respective tissues produce electrophoretic variants of single gene products and the results of Chapter 3 are compatible with the existence of such mechanisms in feather cells. On the other hand, extensive primary structure analysis of hair and wool keratins (Fraser et al., 1972) have established that most or all of the multiple protein molecules of these tissues are each the product of a different gene. Immunoglobulin chains which

probably consist of even larger sets of protein molecules than those of hair or wool keratin are each the product (in part at least) of a different gene (Gally and Edelman, 1972).

In order to investigate the origins of the heterogeneity of the protein molecules constituting feather keratin extensive primarystructure analysis of SCM- feather keratin molecules was carried out. In the present Chapter a study of the properties of the tryptic peptides of $\{^{14}C\}$ - SCM- feather keratin is reported. The results showed that tryptic digestion of SCM- feather keratin produced twelve major peptides and that the positions of tryptic cleavage in the primary structure of most or all of the set K proteins (Chapter 4) were very similar.

METHODS

(i) Tryptic Digestion

Digestion of substrates with trypsin was carried out in 0.2M N-ethyl morpholine acetate at pH 8.30 and 37°C. The enzyme and substrate concentrations were 1mg/ml and 100 mg/ml respectively. The digestion was allowed to proceed for 3 hr and terminated by rotary evaporation, freeze drying or in some cases, direct application to Sephadex G-50. The two substrates used in the experiments reported in this chapter were $[^{14}C]$- SCM- feather keratin (see Chapter 3) and SCM- insulin B-chain which was a gift from Dr. P. Stienert.

(ii) High Voltage Paper Electrophoresis (HVPE)

Peptide samples were spotted onto sheets of Whatman 3MM paper at a loading of not more than 2mg per inch. HVPE at pH 2.2 was carried out on a water-cooled flat place apparatus. The electrode buffer was 1.0 M formic acid (Kotaki 1962). HVPE at pH 6.5 was carried out in a perspex tank with the paper immersed in an organic solvent (Shellsol T) essentailly according to Michl (1951). The composition of the electrode buffer was pyridine, glacial acetic acid, water (25:1:244, v/v). The power supply unit (Paton Industries Pty. Ltd., Stepney, S.A.) was operated at constant voltage. After electrophoresis, papers were dried in a forced-draught oven at 55-60°C for 60 minutes. Visualisation of peptides was carried out after drying by spraying the electrophoretogram with 1% ninhydrin in acetone and allowing the spots to develop at 55-60°C for 5-60 minutes.

(iii) Amino-Terminal Analysis

Peptides (0.5-5 n moles) were dansylated and hydrolysed as described by Gray (1972). After hydrolysis the 6N HCl was removed in a vacuum dessicator maintained at 60-80°C over NaoH pellets and P_2O_5 powder as dessicants. The residue was extracted with H_2O-saturated ethyl acetate (2 x 50 μl) and the extract dried. Both ethyl acetate soluble and insoluble dansyl-amino acids were dissolved in pyridine (approximately 2 μ l) and separately spotted onto polyamide layers (5 x 5 cms) as described by Woods and Wang (1967).

Development of the chromatograms was carried out in the first dimension with the solvent formic acid: H_2O (3:200, v/v) and after careful drying, in the perpendicular direction using benzene:glacial acetic acid (9:1, v/v) as solvent. All dansyl-amino acids except the pairs, dansyl-aspartic acid and dansyl-glutamic acid, dansyl-serine and dansyl-threonine and dansyl-alamine and dansyl-amide could be unambiguously assigned at this stage. Development in the second direction with a third solvent, ethyl acetate : acetic acid : methanol (20:1:1, v/v) allowed resolution of these 3 pairs of dansyl derivatives (Hartley, 1970). Inclusion of SCM- cysteine (0.2 moles/ml) in the 6N HCl used for hydrolysis of dansyl-peptides prevented the variable destruction of dansyl-SCM-cysteine and the appearance of fluorescent breakdown products.

(iv) Gel Filtration on Sephadex G-50

Seaphadex G-50 (fine) was thoroughly equilibrated with a buffer (PMA) whose composition was pyridine : N-ethyl morpholine : H_2O (30:50:3920 v/v), pH 9.2 (acetic acid). After degassing, it was packed into a glass column (1.8 x 70 cm). The elution volumes of standards of known molecular weight were used to obtain a molecular weight calibration curve for the column (Figure 5.1a). The standards used were SCM-insulin B chain, SCM- insulin A chain, SCM- plastocyanin and a cyanogen bromide fragment of SCM- plastocyanin (CNBr I, Milne and Wells, 1970). The former two peptides were gifts

from Dr. P. Steinert and the latter two were gifts from Dr. P.R. Milne, Fractions were collected from the column as described for the Sephadex G-100 chromatography of Chapter 4.

(v) Detection of Peptides from Gel Filtration

Aliquots (50 - 200 µl) of fractions were subjected to alkaline hydrolysis and subsequent reaction with ninhydrin as described by Hirs et al., (1956).

RESULTS

(i) The specificity of Tryptic Cleavage

In order to ensure that digestion of SCM- feather keratin with trypsin produced peptides due only to cleavage of the polypeptide chains at sites carboxyl-terminal to basic amino acids, the specificity and efficiency of tryptic cleavage of SCM- insulin B chain were examined.

SCM-insulin B chain was digested as described in Methods and 20 µl (approx. 40 n moles) were withdrawn for dansylation and 200 µl subjected to HVPE at pH 2.2.

Examination of the tryptic digestion by dansylation revealed the presence of prominent amounts of the dansyl-derivatives of phenylalanine, glycine and alanine which derive from the amino-terminus of intact SCM- insulin B chain, and cleavage of the bonds ^{22}arg-^{23}gly and ^{29}lys-^{30}ala respectively. No traces of the dansyl-derivatives of valine, leucine or

tyrosine whose occurrence would indicate the presence of chymotryptic activity could be detected. As a consequence of this experiment, the preparation of trypsin used was judged to contain a negligibly low amount of chymotryptic activity and tryptic digestions of other substrates (SCM- feather keratin) was likely to be specific.

HVPE of the tryptic digest of SCM-insulin B chain at pH 2.2 produced 3 bands detectable with ninhydrin. These corresponded to the SCM- insulin B chain fragments ^{30}ala., ^{23}gly ... ^{29}lys. and ^{1}phe ... ^{22}arg. No other bands which could indicate the presence of peptides arising from the incomplete cleavage of SCM- insulin B chain were detected. By this criterion, the tryptic digestion was judged to have gone to completion.

(ii) <u>The Separation of Tryptic Peptides of SCM- feather Keratin by Sephadex G-50 Chromatography</u>

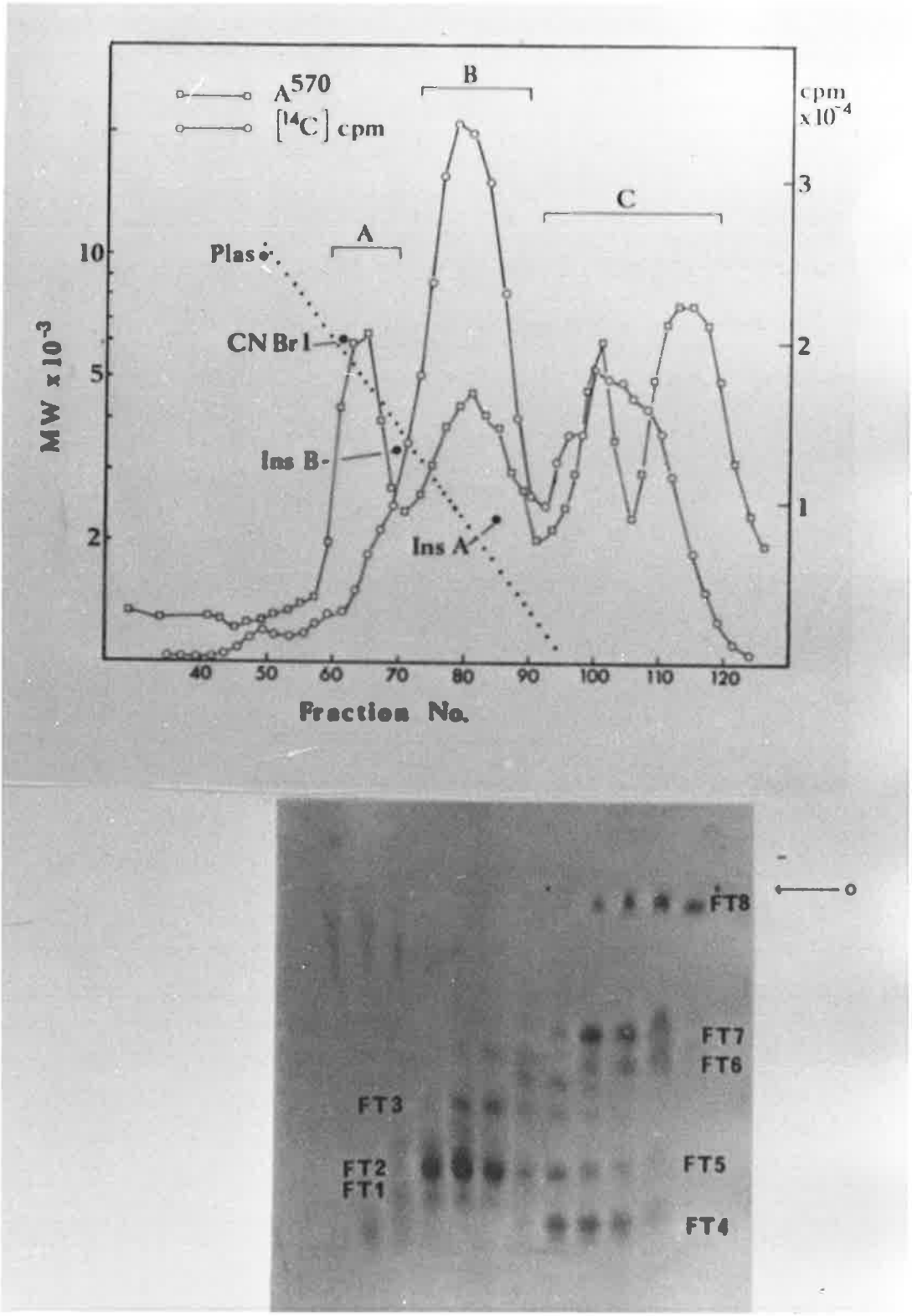
A tryptic digest of 100 mg (2.67×10^7 cpm) of $[^{14}C]$SCM- feather keratin was subjected to chromatography on Sephadex G-50 and fractions analysed for their content of radioactivity and eluted peptide material. Figure 5.1a shows the results of this experiment and the elution of three major peaks of peptide material indicated three fairly discrete size classes of peptides A, B and C. Previous calibration of the Sephadex G-50 column allowed the approximate molecular weights of the pooled fractions A, B and C to be estimated at 5000

FIGURE 5.1 Sephadex G-50 chromatography and HVPE of tryptic peptides of $[^{14}C]$- SCM- feather keratin

a. $[^{14}C]$- SCM- feather keratin (100 mg) was digested with trypsin (1 mg) and subjected to chromatography on Sephadex G-50. Two 50 μL aliquots were withdrawn from each 3.0 ml fraction for determination of radioactivity and peptide material. The elution volumes of SCM- plastocyanin (Plas) a cyanogen bromide fragment of plastocyanin (CNBr I), SCM- insulinBchain (Ins B) and SCM- insulin A chain (Ins A) are shown.

b. Aliquots (200 μL) from selected tubes were subjected to HVPE at pH 6.5 (2000 V, $1\frac{1}{2}$ h) and peptides detected by autoradiography. The Sephadex G-50 elution profile and the autoradiogram have been aligned with respect to tube number.

Fractions (A, B and C) were pooled as shown.


A570
[14C] cpm
cpm x10-4
MW x10-3
A
B
C
Plas
CN Br1
Ins B
Ins A
10
5
2
3
2
1
40
50
60
70
80
90
100
110
120
Fraction No.
FT8
FT7
FT6
FT5
FT4
FT3
FT2
FT1

2000 and less than 1000. Peak A which contained peptides of high molecular weight contained little radioactivity and the amount detected (9.56×10^5 cpm) could not account for the presence of even a single $\left[^{14}C\right]$- SCM- cysteine residue in every constituent peptide. This is based on the assumption that there are eight SCM- cysteine residues in every keratin molecule (Chapter 3). Peak B contained peptides rich in SCM- cysteine as judged by the preponderance of $\left[^{14}C\right]$ radioactivity in the peak. The properties of the constituent peptides of the three peaks were examined by high voltage paper electrophoresis (HVPE), amino acid analysis and gel electrophoresis.

(iii) Size-charge Peptide Maps

Samples from selected tubes of the Sephadex G-50 profile (Fig. 5.1a) were subjected to HVPE at pH 6.5. The $\left[^{14}C\right]$ peptides were detected by autoradiography. Figure 5.1b shows the autoradiogram and it is clear that peaks B and C both contain a number of radioactive peptides. Peak B contained three radioactive peptides (TF1, TF2 and TF3 in Figure 5.1b) all of which eluted at the same volume from the Sephadex G-50 column as judged from their simultaneous intensification and subsequent dissappearance from the autoradiogram (Fig. 5.1b). Peak B (Fig. 5.1a) contained five major radioactive peptides (TF4-TF8) each of which eluted from the Sephadex G-50 column at discrete volumes. That some peptide

spots (TF2, TF4 and TF7) in Figure 5.1b are relatively intense compared to the others reflects in part their high content of $[^{14}C]$- SCM- cysteine and in part their relative abundance in the tryptic digest of $[^{14}C]$- SCM- feather keratin.

(iv) Amino Acid Compositions of Peaks A, B and C

Table 5.1 shows the amino acid compositions of the three pooled peptide fractions A, B and C from Figure 5.1a. Significant differences exist between the compositions of all three fractions. In particular it can be seen that fraction A contains very large amounts of isoleucine and valine and very little SCM- cysteine compared to fractions B and C.

(v) Polyacrylamide Gel Electrophoresis of Fraction A

The high molecular weight of the peptide material of peak A (Fig. 5.1a), its distinctive amino acid composition, and its high colour yield after alkaline hydrolysis and reaction with ninhydrin suggested that it derived from homologous regions of the primary structure of most of the proteins of SCM- feather keratin (see Discussion). To examine whether peak A contained a homogeneous peptide corresponding to a region of the primary sequence common to most or all of the constituents of SCM- feather keratin, a sample of peak A material (equivalent to 100 μg of SCM- feather keratin) was subjected to gel electrophoresis at pH 9.5. Figure 5.2 shows the results of this experiment in which four prominent bands are apparent. The gel

TABLE 5.1 Amino acid compositions of the three pooled peptide fractions from Figure 5.1a

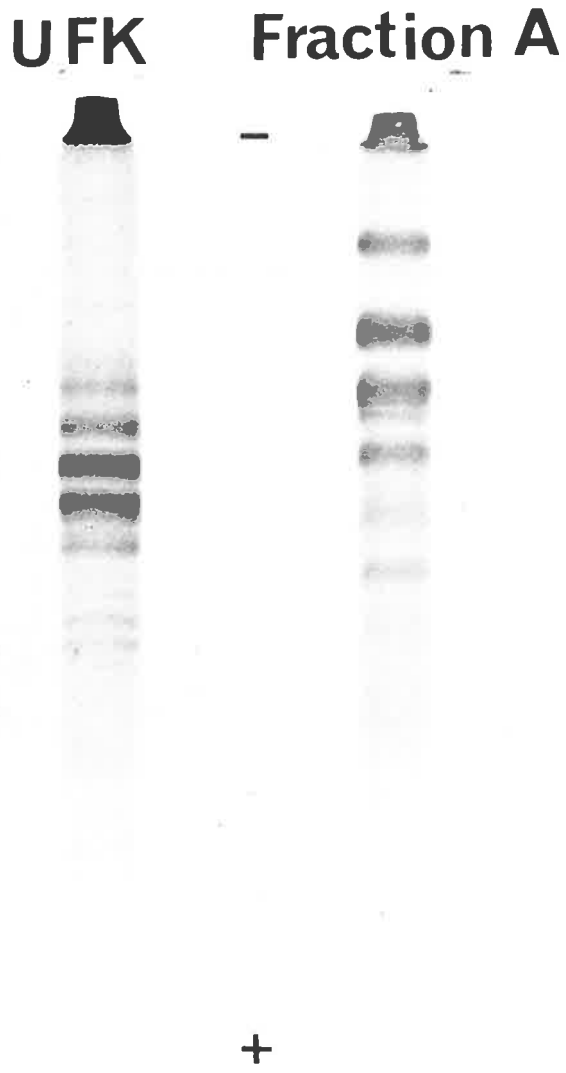
The values are expressed as residues per 100 residues of each amino acid.

TABLE 5.1

	A	B	C
SCM-cys	0.29	20.43	18.83
asp	3.85	8.83	4.53
thr	5.73	1.97	1.02
ser	18.37	10.40	10.95
gln	8.05	13.49	3.21
pro	10.78	9.54	7.30
gly	14.89	5.68	15.18
ala	5.42	1.18	0.88
val	10.17	1.81	0.73
met	0.00	0.00	0.00
ile	6.80	1.41	0.73
leu	7.17	6.62	7.30
tyr	0.31	3.00	7.88
phe	4.26	4.42	7.74
lys	0.28	0.79	0.73
his	0.33	1.89	0.50
arg	2.19	7.89	14.60

FIGURE 5.2 Gel electrophoresis at pH 9.5 of fraction A

A sample of the peptide material of peak A (Fig. 5.1a) equivalent to 100 μg of undigested SCM- feather keratin was freeze dried and subjected to polyacrylamide electrophoresis at pH 9.5. For comparison a gel (50 μg) of intact SCM- feather keratin (UFK) run simultaneously is included.


UFK
Fraction A
–
+

contained 5 M urea and it is therefore unlikely that the four bands of Figure 5.2 were aggregates of smaller peptides.

DISCUSSION

(i) The specificity and efficiency of tryptic cleavage

Trypsin is the most widely used of all proteolytic enzymes for the preparation of peptide fragments for sequence analysis. This is due mainly to its high specificity - tryptic cleavage of polypeptide chains occurs at the carboxyl-terminal side of lysine, arginine and certain other synthetic derivatives of amino acids which possess a side chain primary amino group (for example S-amino ethyl cysteine, Rafferty and Cole, 1963). The usefulness of trypsin as a means of specifically cleaving polypeptide chains depends upon its purity since the presence of low amounts of chymotrypsin in trypsin preparations can lead to the production of more peptides than those due solely to cleavage at lysine and arginine residues. In order to test for the presence of chymotryptic activity in the trypsin preparation used, a defined substrate, SCM-insulin B chain, was chosen which possessed several peptide bonds susceptible to chymotrypsin, ^{14}tyr - ^{15}leu, ^{25}phe-^{26}tyr and ^{26}tyr - ^{27}thr (Sanger and Tuppy, 1951). Digestion of SCM- insulin B chain with trypsin produced no trace of peptides with leucine, tyrosine or threonine at amino-termini as judged by dansylation of a sample of the digest. The only new residues exposed by tryptic cleavage were alanine and glycine which were

carboxyl-terminal to ^{22}arg and ^{29}lys respectively. The limit of detection of dansyl amino acids on polyamide layers is less than 10^{-1} n moles (Bruton and Hartley, 1970) and the presence of 1% chymotryptic cleavage of SCM- insulin B chain would have been detectable by the techniques used.

Analysis of tryptic peptides of SCM- insulin B chain by HVPE at pH 2.2 indicated that only three peptides were present and no traces of peptides due to incomplete digestion were evident. The fact that complete cleavage of SCM- insulin B chain had taken place under the conditions did not necessarily ensure that all potential sites of tryptic cleavage of SCM- feather keratin would be as prone to digestion as those of SCM- insulin B chain; the presence of an acidic residue adjacent to the trypsin-sensitive basic residue slows the rate of tryptic cleavage of that peptide bond considerably (Hirs et al., 1956).

It is therefore impossible to eliminate the possibility of incomplete cleavage of some peptide bonds in SCM-feather keratin by trypsin without more detailed structural analysis of the resultant peptides.

(ii) The Number of Tryptic Peptides of SCM- feather Keratin

The autoradiogram of Figure 5.1b is in effect a two dimensional peptide map of the $[^{14}C]$ - SCM- cysteine containing peptides. Peptides have separated on Sephadex G-50 in accordance with diminishing size on the horizontal axis and by HVPE at pH 6.5 on the vertical axis, which separates peptides on

the basis of both size and charge. At least eight peptides which contain SCM- cysteine were clearly resolved by this technique. Moreover, analysis of the high molecular weight group of tryptic peptides (peak A, Fig. 5.1a) by polyacrylamide gel electrophoresis at pH 9.5 indicated the presence of an additional four major peptides. Therefore, there are a total of at least twelve tryptic peptides of SCM- feather keratin. This number of peptides can readily be accounted for by the existence of nineteen electrophoretic variants of SCM- feather keratin (Chapter 4, Figs. 5.1b and 5.2) which contain similar amounts of SCM-cysteine (approximately 8 mole %) and arginine (4-5 mole %). The fact that only 12 tryptic peptides could be identified indicates the likelihood that some of them at least do not derive from only one species of the nineteen electrophoretic variants but that many variants possess identical regions of primary structure. This deduction depends upon the ability of the procedures used (namely gel electrophoresis, Sephadex G-50 chromatography and HVPE at pH 6.5) to completely resolve most or all the tryptic peptides produced.

The amino acid composition data of the three size classes of tryptic peptides, A, B and C (Table 5.1) and the gel of fraction A (Fig. 5.2) however support the idea that some of the nineteen electrophoretic variants of SCM- feather keratin share identical regions of primary structure. The contents of valine, isoleucine and threonine in the material of fraction A are very much higher than in fractions B and C - in fact,

fraction A contains isoleucine almost exclusively. This, and the fact that all major fractions (D2-D9) of SCM- feather keratin contained similar amounts of isoleucine (4-5 mole%, Table 4.2) suggests that the isoleucine rich portion of most or all SCM- feather keratin chains is represented solely by the peptides of peak A and that isoleucine residues are confined to the same general area and possibly to precisely the same positions within the sequence of all SCM-feather keratin molecules. Only four prominent bands were detected when these large isoleucine rich peptides were examined by gel electrophoresis at pH 9.5 and each band presumably represents the isoleucine rich tryptic peptide of a number of SCM- feather keratin molecules.

Their sharpness and symmetry of elution from Sephadex G-50 suggest the uniformity of the peptides of peaks A and B (Fig. 5.1a) with respect to size and in consequence that the location of tryptic cleavage sites (arginine residues) within the primary structure of SCM- feather keratin has been conserved at least in most of the different molecules. This observation prompted a detailed study of the primary structure of the tryptic peptides of SCM- feather keratin - the fundamental conclusion arrived at from the results of the present Chapter was that such a study should reveal the nature of the structural differences between the nineteen variants of SCM-feather keratin.

CHAPTER 6

THE PARTIAL SEQUENCE OF SCM-FEATHER KERATIN

INTRODUCTION

In this Chapter the results of sequence analysis of the tryptic peptides of SCM- feather keratin are presented. The results of the previous Chapter established the existence of twelve major peptides in tryptic digests of SCM- feather keratin, eight of which contained SCM- cysteine and which were small enough to permit their complete sequence determination by the dansyl-Edman procedure. This study impinges on three aspects of the structure of SCM- feather keratin molecules.

1. Whether the differences in primary structure of the multiple constituents demonstrated in Chapters 4 and 5 arise from post synthetic modification of one or a small number of primary gene products or from the expression of multiple genes each determining the synthesis of a specific keratin chain.

2. Whether the differences in the primary structure of the electrophoretic variants of SCM- feather keratin were confined to a particular region of the amino acid sequence as has been demonstrated for the immunoglobulin molecules. IgG molecules have been shown to consist of heavy and light polypeptide chains both of which possess a variable amino-terminal region and a carboxyl- terminal constant region. The variable regions of different IgG molecules are suually remarkably divergent with respect to primary structure whereas few differences are evident in the constant regions between molecules of the same class (Porter, 1970, Gally and Edelman 1972).

3. The possible evolutionary relationships between the different keratin chains.

Primary structure studies of homologous proteins from different species have enabled the construction of phylogenetic trees relating existing proteins to evolutionary ancestors (Fitch and Margoliash, 1966). Studies on purified α-keratin chains of wool have established that within a tissue of a single organism, a large family of related proteins derived from a common ancestral protein is synthesised (see Introduction).

Studies on the primary structure of SCM- feather keratin chains were expected to permit a critical examination of these 3 aspects of protein structure in relation to the heterogeneity established in Chapter 4. The sequence studies of this Chapter were carried out on the three tryptic peptide fractions of SCM- feather keratin, A, B and C whose isolation was discussed in the previous Chapter. A separate section has been devoted to the purification and structure of the constituent peptides of each fraction and their proposed arrangement within the primary structure is discussed.

METHODS

(i) Enzyme Digestions

Peptides were digested at a concentration of 1 μmole/ml with 0.01% chymotrypsin or thermolysin overnight at 37°C in 0.2 M NEMA, pH 8.3. 1mm $CaCl_2$ was present for thermolytic digestions. Digests were either dried by rotary evaporation or loaded directly onto Whatman 3MM paper for HVPE.

Carboxypeptidase A solutions (1mg/ml in 0.2M NEMA)

were prepared as described by Ambler (1967) and digestions were performed as described for chymotrypsin except that digestion times were 3h. Carboxypeptidase B digestions were performed as described for carboxypeptidase A without prior purification of the enzyme. Digests were directly applied to Whatman 3MM paper for HVPE. To detect the release of free amino acids with carboxypeptidases, samples containing 1-10 n moles of substrate were withdrawn after digestion, solvent removed in a heated vacuum dessicator and dansylation carried out as described in Chapter 5. Dansyl-amino acids were extracted without prior hydrolysis and identification of dansyl-amino acids carried out as described in Chapter 5.

(ii) Chromatography on Dowex-50

Separation of peptide mixtures was carried out using the Dowex-50 resin, chromo-beads Type P (Technicon), in a jacketed column (0.80 x 90 cm) operated at 42-45°C. The gradient used for elution of peptides was based on that of Schroeder et al., (1962) and was prepared in a Techicon varigrad (9 chambers). The three solutions used to prepare the gradient were (i) 0.2M pyridine pH 3.1 (acetic acid), (ii) 2.0M pyridine, pH 5.0 (acetic acid), (iii) 2M pyridine (acetic acid) pH 6.5 (Schroeder, 1962). Consecutive chambers of the varigrad contained solutions (i), (ii) and (iii) in the following proportions

Chamber	Contents
1	(i) 60ml
2	(i) 60ml
3	(i) 60ml
4	(i) 40ml, (ii) 20ml
5	(i) 20ml, (ii) 40ml
6	(ii) 60ml
7	(ii) 30ml, (iii) 30ml
8	(iii) 60ml
9	(iii) 60ml

The flow rate was maintained at 30ml/h with a Beckman Accu-Flow pump and 2.5 ml fractions were collected. Aliquots were removed from each fraction to detect the elution of peptide material and of radioactivity as described in Chapter 5.

(iii) Purification of Peptides by HVPE

HVPE at pH 6.5 was carried out as described in the previous Chapter. HVPE at pH 2.7 was carried out in a Michl tank using a buffer, 10% glacial acetic acid adjusted to pH 2.7 with pyridine. Peptides were first subjected to HVPE at pH 6.5, localised by autoradiography and cut out on paper strips which were sewn onto a new piece of Whatman 3MM paper (Harris, 1967). HVPE was then carried out at pH 2.7, peptides localised by autoradiography and eluted from paper strips with 1% pyridine.

(iv) Edman Degradation

Sequence analysis was carried out using the Edman degradation procedure described for peptides by Gray (1972).

Samples of peptide material (1-5 n moles) were withdrawn after each cycle for dansylation (Chapter 5).

(v) Purification of Reagents

N-ethyl morpholine and pyridine were redistilled from ninhydrin (1g /l). Glacial acetic acid, ethyl acetate and phenylisothiocyanate were redistilled. N-ethyl morpholine and phenylisothiocyanate were stored in brown bottles with tight-fitting lids after removal of oxygen by flushing with nitrogen. The 6N HCl used for acid hydrolysis was redistilled constant boiling point HCl (Vogel, 1961).

RESULTS

(i) The Partial Structure of T3

At least four major peptides all of molecular weight 5000 compose peak A (Fig. 5.1a). When the amino acid composition of this material (Table 5.1) is expressed assuming one mole of arginine per mole of peptide a molecular weight of 4850 can be calculated (Table 6.1), and this is in accord with the value (5000) calculated by gel filtration. It was then likely that most or all of the peptides constituting peak A possessed only one residue of arginine which was carboxyl-terminal in each case. The amino-terminal and carboxyl-terminal sequences of peak A peptides indicated that they composed a family of homologous tryptic peptides and they were collectively designated T3. The nomenclature of these and all

TABLE 6.1 The properties of the tryptic peptides T3

Molecular Weight: 5000 (Sephadex G-50). [1]4850 (Amino Acid) Analysis

Amino Acid Composition

Amino Acid	[2]Residues (arg=1)
SCM-cys	0.13
asp	1.76
thr	2.61
ser	8.39
glu	3.68
pro	4.92
gly	6.80
ala	2.47
val	4.64
met	0.00
ile	3.11
leu	3.27
tyr	0.14
phe	1.95
lys	0.13
his	0.15
arg	1.00

Amino Terminal Sequence : val-val-ile-glx-pro-ser-pro-.....

[1]The molecular weight was calculated by summation of the total number of residues (relative to arginine) and assuming an average residue molecular weight of 110

[2]Residue ratio calculated from Fig. 5.1 assuming 1 residue of arginine

other peptides examined accorded with the order in which the peptides occur in the intact SCM- feather keratin molecule (see Discussion).

(a) The amino-terminal Sequence of T3

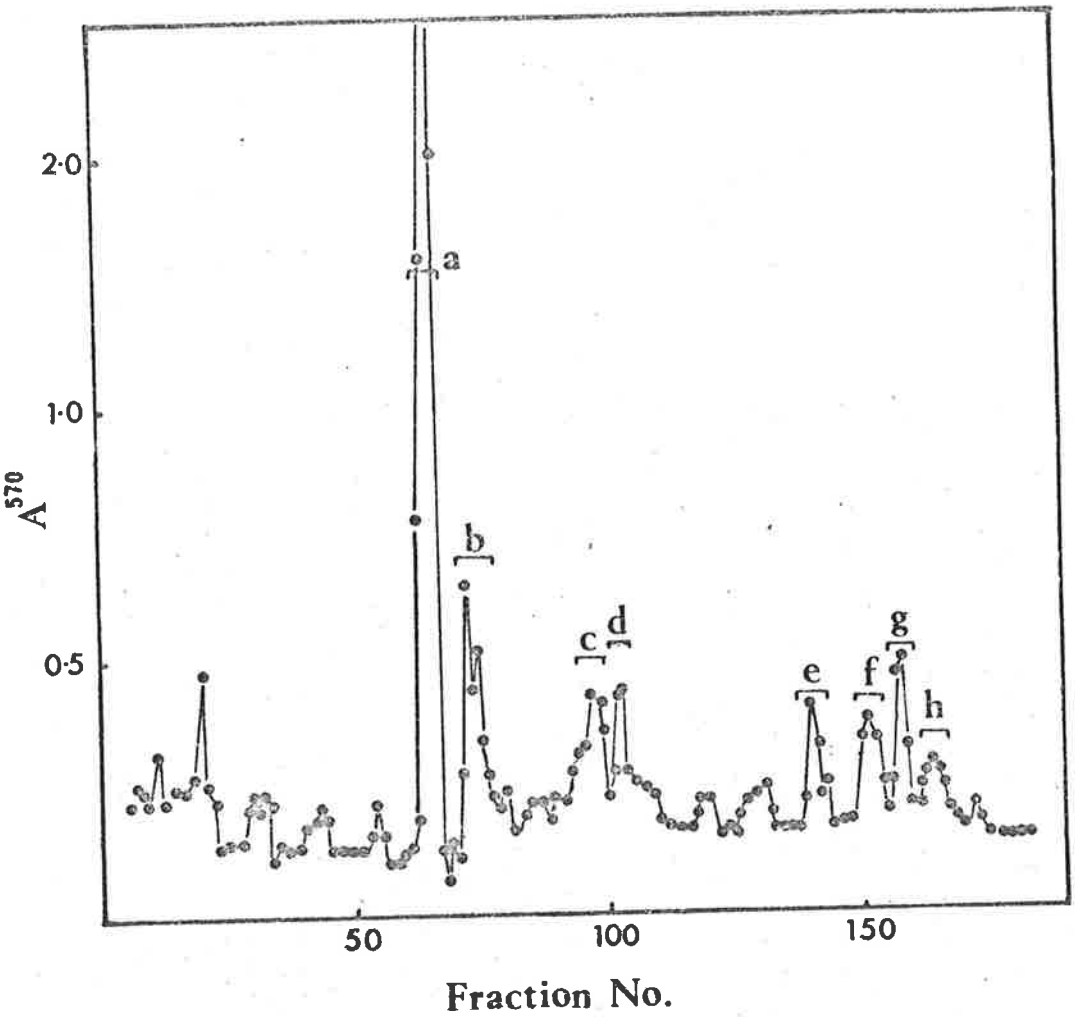
Approximately 20 n moles (calculated from amino acid analyses) of T3 was subected to 6 cycles of the dansyl-Edman procedure (Methods). Each cycle of the Edman reaction exposed only one amino-terminal residue on subsequent dansylation and the amino-terminal sequence of T3 was found to be val-val-ile-glx-pro-ser-pro ...

(b) The Carboxyl-terminal Sequences of T3

The amino acid composition (Tables 5.1 and 6.1) indicated that the peptides T3 were rich in hydrophobic amino acids, serine and glycine and that they contained only one mole of arginine presumably at the carboxyl-terminus of the peptides. To obtain peptides derived from the carboxyl-terminus, a thermolytic digest of T3 was prepared and subjected to chromatography on Dowex-50 as described in Methods. Such basic peptides were expected to bind strongly to the anionic resin and to elute from it only at high ionic strengths thus enabling their selective purification. Fractions from the column were analysed for the presence of peptide material as described in Methods and the elution profile from Dowex-50 of the thermolytic peptides is shown in Figure 6.1. To determine which of the peaks contained carboxyl-terminal

FIGURE 6.1 Dowex-50 separation of thermolytic peptides of T3

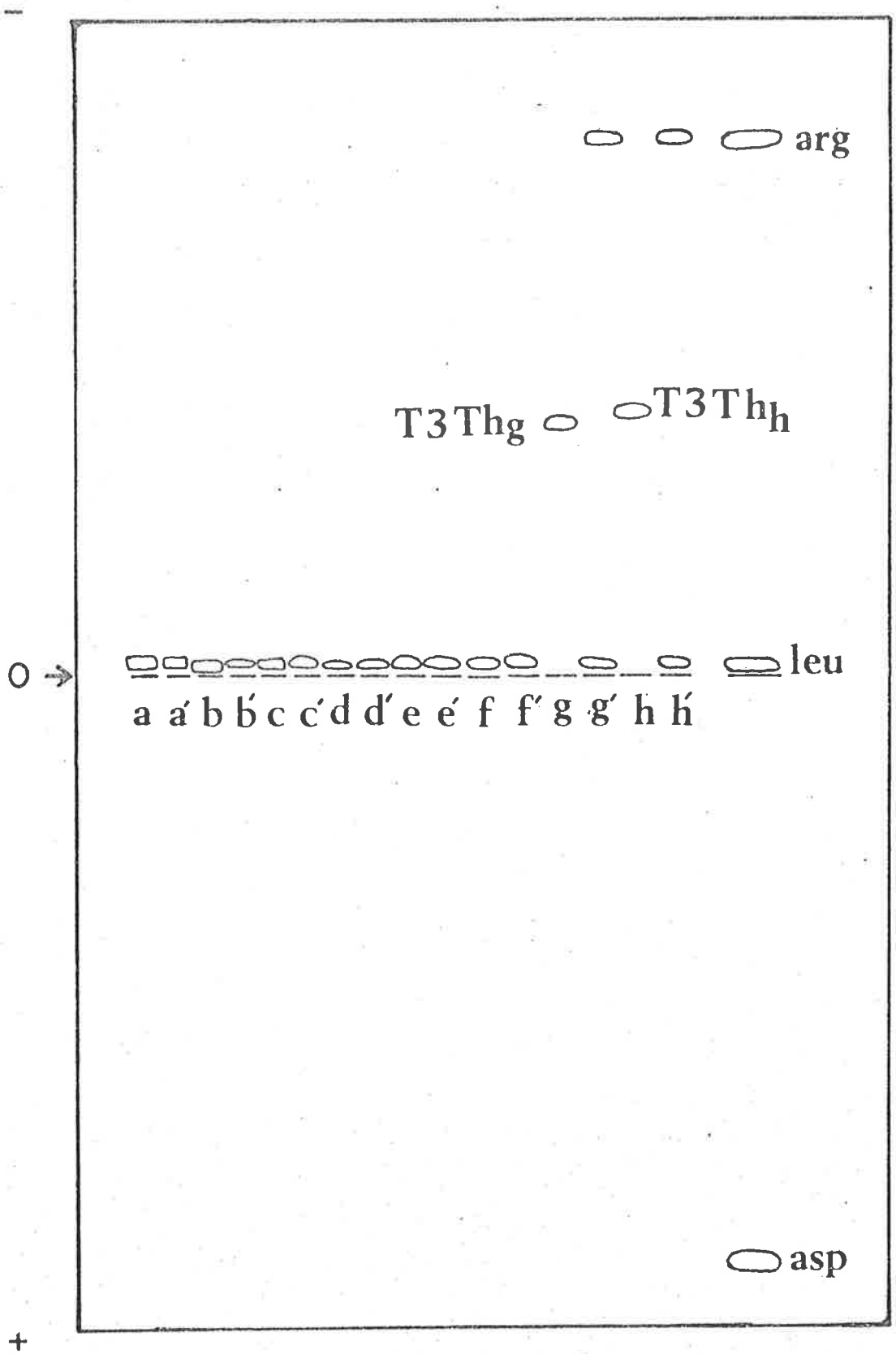
A sample of T3 (equivalent to 50mg of SCM- feather keratin) was digested with thermolysin and Dowex-50 chromatography carried out as described in Methods. 50 μL was withdrawn from each fraction for determination of peptide material. Fractions (a-h) were pooled as shown.


2·0
1·0
0·5
A^{570}
a
b
c
d
e
f
g
h
50
100
150
Fraction No.

peptides two aliquots each representing 10% of each pooled fraction (Fig. 6.1) were analysed in the following manner. One aliquot was digested with carboxypeptidase B (20 μg in 200 μl of NEMA) and applied to Whatman 3MM paper for HVPE at pH 6.5. The other aliquot of each sample was applied undigested, to electrophoresis paper alongside its degested counterpart. After electrophoresis and drying, the electrophoretogram was sprayed with 1% ninhydrin and the results are presented in Fig. 6.2. Only fractions g and h released free arginine on digestion with carboxypeptidase B and were therefore two carboxyl terminal peptides of T3. The mobilities of the undigested peptides T3.Th.$_g$ and T3.Th.$_h$ allowed their molecular weights to be estimated at 420 and 440 respectively (Offord, 1966). Amino acid analysis of aliquots (10%) of peptides T3.Th.$_g$ and T3.Th.$_h$ indicated the compositions ser(1), gly(1), leu(1), arg(1) and gly(2), Leu(1), arg(1), respectively (Table 6.2). The molecular weights of 431 and 401 for T3.Th.$_g$ and T3.Th$_h$ respectively deduced from their respective amino acid compositions are in excellent accord with those deduced from mobility measurements. The sequences of the two peptides were established by subjecting each to three cycles of dansyl-Edman procedure which allowed the assignment of the first residues, leu-gly-ser. and leu-gly-gly. in the sequences of T3.Th$_g$ and T3.Th.$_h$ respectively. The remaining arginine residue in both sequences was therefore carboxyl terminal as predicted from the susceptibility of both peptides to carboxypeptidase B (Fig. 6.2). The yields of both peptides in μmoles were calcul-

FIGURE 6.2 Identification of the carboxyl-terminal thermolytic fragments of T3

10% of fractions a-h pooled from the Doxex-50 chromatogram (Fig. 6.1) were digested with carboxy-peptidase B and the digests applied separately to electrophoresis paper, (a', b', c' etc). Adjacent to each digest was loaded an undigested aliquot, (a, b, c etc). HVPE at pH 6.5 was carried out for 1.0 hr at 2000V. The electrophoretogram was dried and sprayed with ninhydrin.


−
arg
T3Thg
T3Thh
0 →
leu
a a′ b b′ c c′ d d′ e e′ f f′ g g′ h h′
asp
+

ated from their amino acid analyses relative to the amount of T3 used as substrate for themolytic digestion and are expressed as % yield in Table 6.2. Their combined yield of 62% may reflect the loss incurred by the selective pooling of fractions (Fig. 6.1) or alternatively may be due to the existence of other carboxyl-terminal peptides present in much lower amounts than T3.Th.$_g$ and T3.Th.$_h$. The relation of the 2 sequences to one another is considered in Discussion.

(ii) The Structure of The Peptides of Peak B

The amino acid composition and the large amount of radioactivity associated with peak B (Fig. 5.1a) indicated that its constituent peptides were rich in $[^{14}C]$- SCM- cysteine. The use of HVPE and autoradiography facilitated the isolation of four of these peptides in markedly different amounts. The high content of $[^{14}C]$SCMcysteine allowed their purification and the detection and characterisation of cleavage products obtained by enzyme digestion of the peptides with other enzymes. Primary structure analyses of two such tryptic peptides allowed the elucidation of their respective amino acid sequences. No primary structure analysis of the other two peptides was possible because of low yields but their properties and those of their chymotryptic peptides established their relation to the two prominent peptides whose sequences were determined. All four peptides derive from the amino - terminal protion of different SCM- feather keratin molecules and are labelled Tl(a), Tl(b) etc.

TABLE 6.2 The carboxyl-terminal peptides of T3

[1] The electrophoretic mobilities (M rel) of the peptides are expressed relative to aspartic acid and their molecular weights (MW) calculated by the method of Offord (1966)

[2] The amino acid compositions are expressed in absolute μmoles and represent 1/10th the total sample

[3] The amino acid compositions are expressed assuming 1 mole of arg/mole of peptide

[4] The yield is based on the total number of μmoles of each peptide recovered from Dowex-50 compared to the number of μmoles of arginine present in the T3 sample used for thermolytic digestion and subsequent chromatography.

TABLE 6.2 The carboxyl-terminal peptides of T3

	T3 Thq	T3 Th_{hr}
[1]M(rel)	0.48	0.50
[1]MW	440	420

Amino Acid Composition

	[2]μmoles	[3]Res.(arg=1)	[2]μmoles	[3]Res.(arg=1)
SCM-cys	-	-	-	-
asp	-	-	-	-
thr	-	-	-	-
ser	0.126	1.00(1)		
glu	-	-	-	-
pro	-	-	-	-
gly	0.132	1.04(1)	.114	2.02(2)
ala	-	-	-	-
val	-	-	-	-
met	-	-	-	-
ile	-	-	-	-
leu	0.134	1.05(1)	0.054	0.95(1)
tyr	-	-	-	-
phe	-	-	-	-
lys	-	-	-	-
his	-	-	-	-
arg	0.126	1.00(1)	0.056	1.00(1)

[4]Yield (moles arg/mole T3) : 43% 19%

Sequence: $\overrightarrow{\text{leu-gly-ser}}$-arg $\overrightarrow{\text{leu-gly-gly}}$-arg

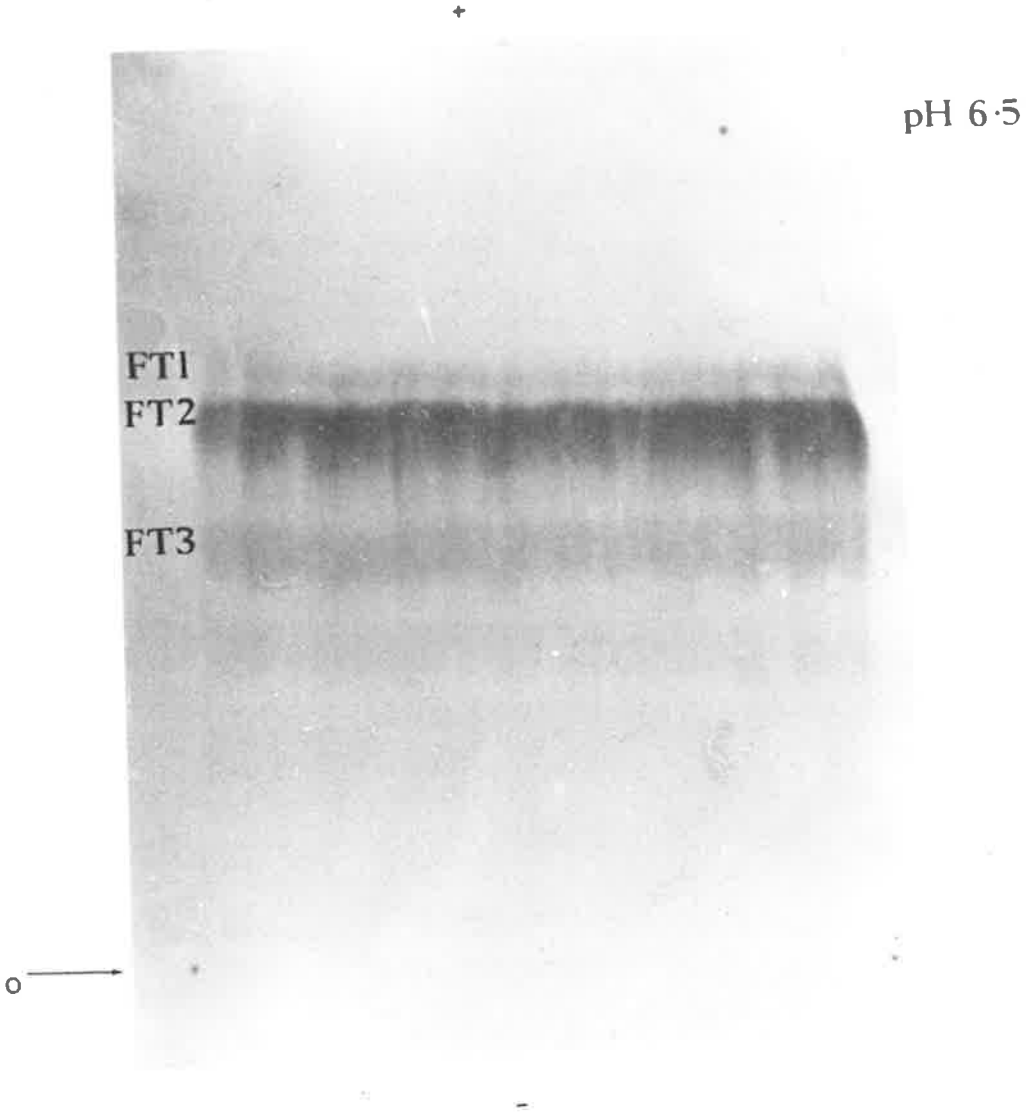
The isolation of the four peptides is dealt with in part (i) of this section and each subsequent part (ii-v) deals with the structure of the peptides separately.

(a) The Isolation of Tryptic Peptides T1a, T1b, T1c and T1d

Figure 5.1b shows the presence of three radioactive peptides (FT1, FT2, FT3) in peak B of the Sephadex G-50 profile, separated by HVPE at pH 6.5. For purification on a preparative scale, the fractions of peak B were pooled and applied to electrophoresis paper as a 10" band and subjected to HVPE at pH 6.5. The electrophoretogram was used to expose the autoradiogram of Figure 6.3. Strips of paper containing peptides FT1, FT2 and FT3 were cut out and the three peptide bands separately subjected to HVPE at pH 2.7 as described in Methods. Figure 6.4 shows the results of HVPE at pH 2.7 and autoradiography revealed four major peptides (T1a - T1d). Guide strips cut from the side of all 3 autoradiograms were sprayed with 1% ninhydrin but no detectable band was observed on any of the 3 electrophoretograms. However chlorination followed by the starch-iodide spray (Rydon and Smith, 1952) revealed patterns identical to the autoradiograms. Therefore the peptides constituting peak B (Fig. 5.1a) were N-blocked and this suggested that they derived from amino-termini of SCM-feather keratin chains. To confirm that peak B peptides, T1a-T1d, were amino-terminal and to examine the relationship between them, each peptide band was cut out and eluted from

FIGURE 6.3 Preparative HVPE at pH 6.5 of Peak B Peptides

Fractions constituting peak B (Fig. 5.1a) were pooled and solvent removed by rotary evaporation. The peptides were redissolved in a small amount (500 μL) of 50% pyndine:H_2O and applied to a 10" origin (O). HVPE at pH 6.5 was carried out for 1.0 hr at 3000 V as described (Methods). After drying, the electrophoretogram was allowed to expose a strip of x-ray film which was subsequently developed (Methods). The peptide bands FT1, FT2 and FT3 are labelled in accordance with Fig. 5.1b. The three intense black spots are due to the use of radioactive $\left[^{14}C\right]$ ink on the electrophoretogram which allowed its precise allignment with the autoradiogram.


+
pH 6·5
FT1
FT2
FT3
o
-

electrophoresis paper as described in Methods.

The specific activity of $[^{14}C]$ - SCM - cysteine was determined by relating the radioactivity in a sample of undigested substrate to its content of SCM- cysteine. This specific activity was useful in calculating the amount (in μmoles) of SCM- cysteine in a particular peptide derived from a given preparation of $[^{14}C]$ - SCM- feather keratin. The specific activity of the batch of radioactive substrate from which peptides T1a-T1d were derived was 1.07×10^5 cpm/ μmole (of SCM- cysteine).

(b) Peptide T1a

The eluted peptide T1a (Fig. 6.4) contained 2.74×10^4 cpm, equivalent to 0.25 μmoles of SCM- cysteine. Other experiments (see peptide T1b) indicated that all T1 peptides of peak B (Fig. 5.1a) probably contained 5 moles SCM- cys/ mole of peptide and therefore a yield for T1a of only 0.05 μmoles was estimated. Complete structural analysis was not attempted on such a small amount of material and the experiments carried out were designed with the aid of foreknowledge gained by initial sequence analysis of peptides T1b and T1c.

T1a was digested with chymotrypsin as described in the legend to Figure 6.5 and the digest subjected to HVPE at pH 6.5 (Fig. 6.5a). 3 radioactive fragments T1a C1, T1a C2 and T1a C3 were obtained. T1a C1 had previously been shown to be pure by HVPE at pH 2.7 followed by paper chromato-

FIGURE 6.4 Isolation of Peptides T1a-d by HVPE at pH 2.7

The strips of paper containing radioactive peptide bands FT1, FT2 and FT3 were cut out and subjected to HVPE at pH 2.7. Electrophoresis was carried out for 3 h at 3000 V. After drying the electrophoretograms autoradiography was carried out as described (Methods).

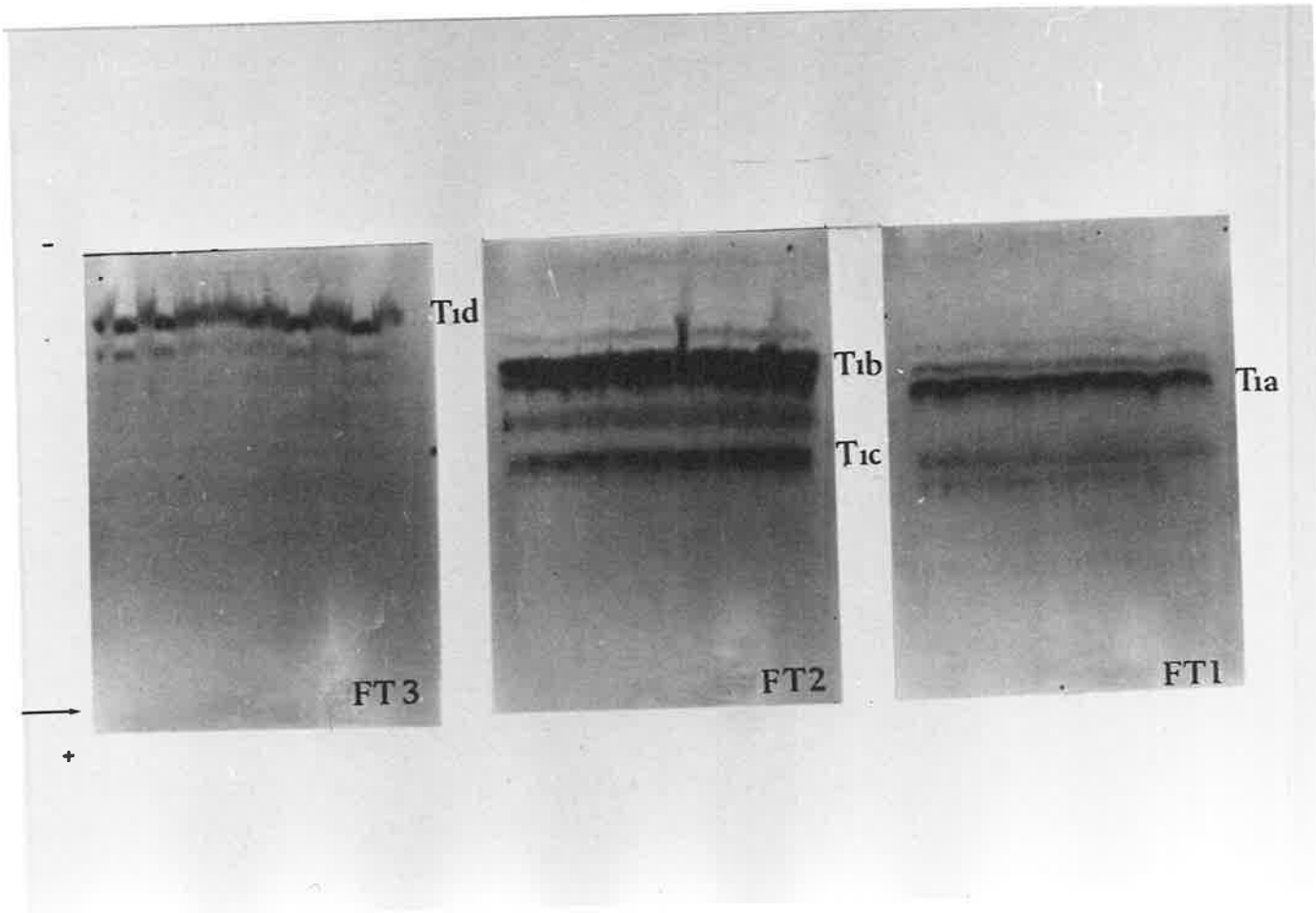
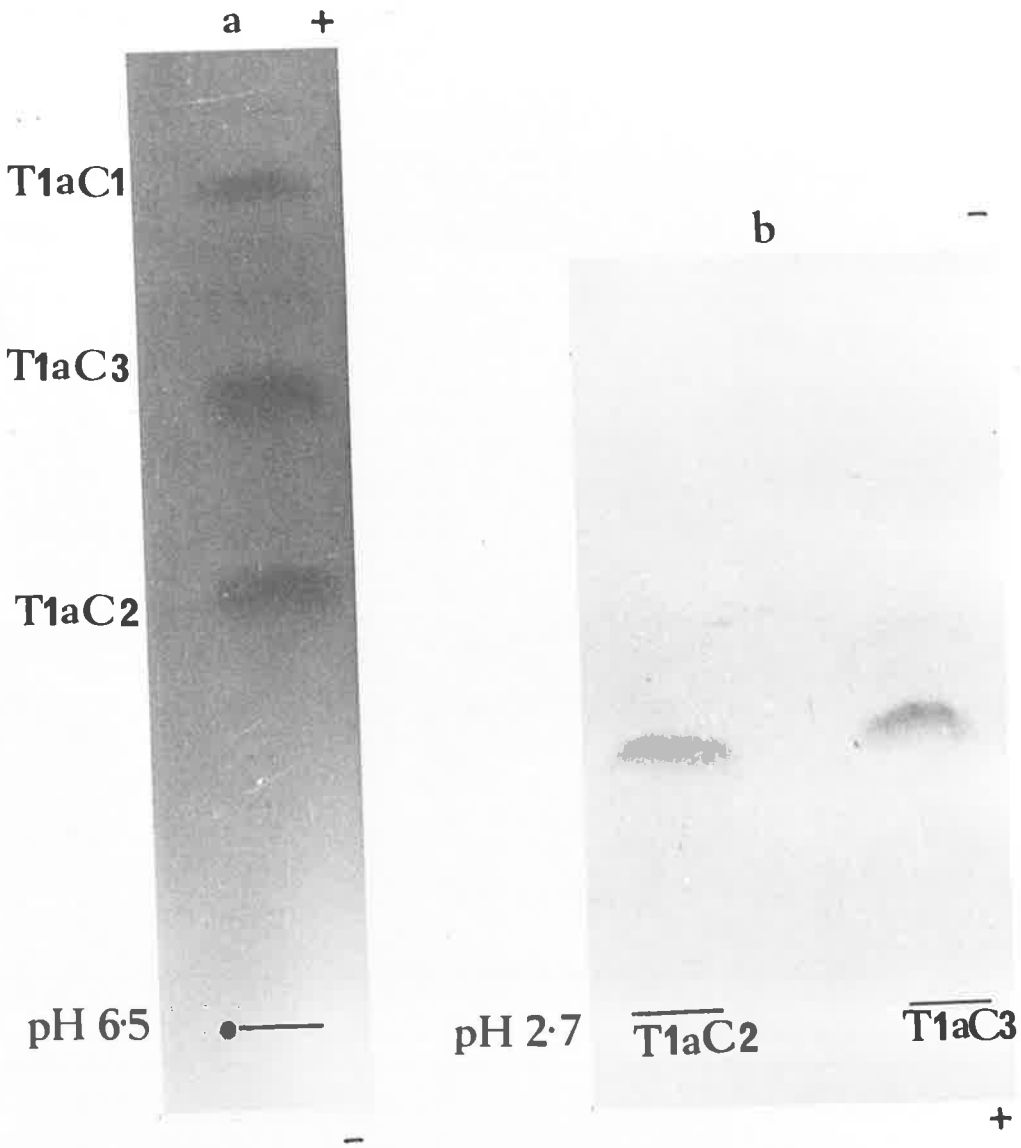

-
+
T1d
T1b
T1c
T1a
FT3
FT2
FT1

FIGURE 6.5 The purification of the chymotryptic peptides of Tla

Tla (0.05 μmole) was digested with chymotrypsin (8 μg) in N-EMA, pH 8.3 (80 μL) and subjected to HVPE at pH 6.5 (a) for 1 h at 3000 V and autoradiography carried out as described in Methods. The bands labelled Tla C2 and Tla C3 in (a) were subjected to HVPE at pH 2.7 at 3000 V for 1.5 h (b) and autoradiography.


a
+
T1aC1
T1aC3
T1aC2
pH 6·5
–
b
–
pH 2·7
T1aC2
T1aC3
+

graphy carried out as described by O'Donnell (1973b) and in the present experiment was eluted from paper without further purification. Peptides Tla C2 and Tla C3 were subjected to HVPE at pH 2.7 to establish their purity. Figure 6.5b shows the latter autoradiogram demonstrating that both peptides were homogeneous by HVPE at both pH 6.5 and pH 2.7. Tla C2 and Tla C3 were in turn eluted from paper and the three pure peptides were hydrolysed and subjected to amino acid analysis. The results of this experiment (Table 6.3) indicated that Tla C1 was a tripeptide (containing ser, SCM- cys, phe) of molecular weight 455. Its relative mobility (0.91) indicated that it possessed a net charge of -2 at pH 6.5 and it was therefore N- blocked and must represent the amino-terminal sequence of peptide Tla. Experiments with chymotryptic peptides derived from Tlb and Tlc (Tlb C1, Tlc C1) whose properties were identical to those of Tla C1 indicated that its sequence was N-acetyl-ser-SCM-cys-phe.

The amino acid composition of Tla C2 (Table 6.3) indicated that it contained eleven residues two of which were SCM- cysteine and its relative mobility indicated a net charge of -2.

The presence of arginine in this peptide cannot be deduced from the net negative charge since Tla C2 contains one mole of asx which, if present as asp, would mask the effect of arginine on the mobility of the peptide. Its relative mobility indicated that Tla C3 possessed a charge of

TABLE 6.3 The properties of the chymotryptic peptides of Tla

The total amount of the peptides Tla C1, Tla C2 and Tla C3 were hydrolysed and analysed on the long column of the Beckman amino acid analyser and as a consequence the contents of basic amino acids were not determined directly but have been inferred (see Text). The scale expander was used to facilitate integration of the three chromatograms.

(a) Absolute μmoles

(b) Residues per mole (relative to a convenient standard)

1 The exact molecular weight of Tla C1 was calculated from its amino acid composition taking into account one acetyl group per mole. The approximate molecular weights of Tla C2 and Tla C3 were calculated assumming eleven and nine residues respectively and an average residue weight of 110.

2 The net charge carried by each peptide was calculated from its relative mobility (Mrel) and its molecular weight by the method of (Offord, 1966).

TABLE 6.3.

	Tla C1		Tla C2		Tla C3	
	a	b	a	b	a	b
SCM-cys	0.026	0.78(1)	0.026	1.86(2)	0.025	1.78(2)
asp			0.016	1.14(1)	0.027	1.89(2)
thr			0.014	1.00(1)	-	
ser	0.048	1.43(1)			0.014	1.00(1)
glu					0.016	1.11(1)
pro			0.052	3.43(3)	0.017	1.22(1)
gly	0.011	0.33(0)	0.016	1.14(1)	0.005	0.03(0)
ala					0.014	1.00(1)
val						
met						
ile						
leu			0.028	2.00(2)		
tyr						
phe	0.033	1.00(1)				
lys						
his						
arg	0	0	?	(1)	?	(1)

	$^{1}_{455}$	$^{1}_{1200}$	$^{1}_{1100}$
MW			
Mrel	0.91	0.46	0.69
^{2}Net Charge	-2	-2	-3

-3 using the molecular weight calculated from its amino acid composition (Table 6.3). Comparison of its amino acid composition with that of the corresponding fragments of T1b and T1c (see below) made it likely that T1a C3 contained one residue of carboxyl-terminal arginine. Therefore to maintain a net charge of -3, two of the threeacidicamino acid residues in the sequence of T1a C3 must be present as their acid and not amide forms. The likely structure of T1a and the arrangement within the sequence of its three chymotryptic cleavage products (T1a C1, C2, C3) is considered in Discussion.

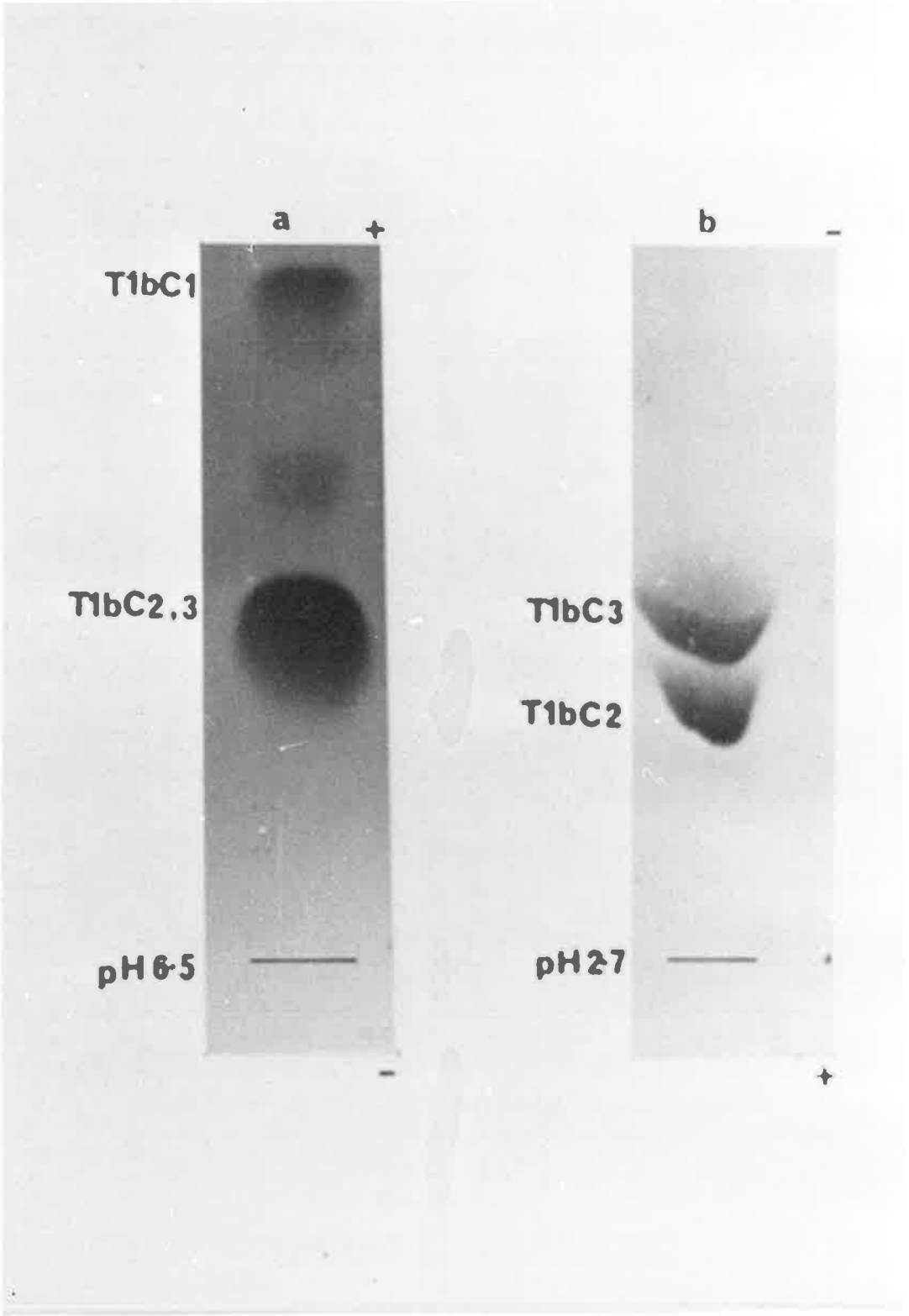
(c) Peptide T1b

Peptide T1, isolated by preparative HVPE (Figs. 6.3, 6.4), was eluted from paper and found to contain 4.47 x 10^5 cpm equivalent to 4.41 μmoles of SCM- cysteine. This amount was adequate for the complete determination of its amino acid sequence.

0.10 μmoles of T1b was withdrawn for amino acid analysis (Table 6.4) and the remainder was digested with chymotrypsin and subjected to HVPE at pH 6.5 followed by autoradiography of the dried electrophoretogram (Fig. 6.6). Two prominent radioactive bands were apparent, the faster of which (T1b.C1) had a relative mobility of 0.89. The slower band (T1b.C2 + T1b.C3), of relative mobility 0.48 was cut out and subjected to HVPE at pH 2.7 as described in Methods and the dried electrophoretogram subjected to autoradiography

FIGURE 6.6 The purification of the chymotryptic peptides of T1b

T1b (approx. 0.7 μmoles) was digested with 66 μg of chymotrypsin in 0.66 ml of NEMA, pH 8.3 and the digest was subjected to HVPE at pH 6.5 (2000 V, 1 h) and the electrophoretogram autoradiographed (a). The slower radioactive band (T1b C2, T1b C3) was cut out and subjected to pH 2.7 HVPE (3000 V, 1.5 h) and autoradiography (b).


a
+
T1bC1
T1bC2,3
pH 6·5
-
b
-
T1bC3
T1bC2
pH 2·7
+

(Fig. 6.6b). HVPE at pH 2.7 separated the two peptides Tlb C2 and Tlb C3 completely. The three peptides Tlb C1, C2 and C3 were cut out and eluted from paper with 1% pyridine as described in Methods. Previous experiments (not shown) had established that both Tlb C2 and Tlb C3 had unblocked amino-termini as judged by their reaction with ninhydrin whereas Tlb C1 was blocked and was therefore the amino-terminal peptide of Tlb. Amino acid analyses were carried out of samples containing approximately 0.1 μmole of each peptide (Table 6.4) and the primary structure of each of the three peptides was determined.

Peptide Tlb C1

Tlb C1 (approx. 10 n moles) was digested with carboxypeptidase A and the released amino acids detected by dansylation (Methods). Only phenylalanine was released by carboxypeptidase A confirming its expected placement at the carboxyl-terminus of the tripeptide. O'Donnell (1973 a,b) and Kemp (PhD Thesis) have established that the amino-terminal sequences of feather keratins from a number of avian sources is N-acetyl-ser-SCM-cys-^{tyr}phe and the structure of Tlb C1 was assumed to be N-acetyl ser-SCM-cys-phe.

Peptide Tlb C2

The amino acid composition of Tlb C2 (Table 6.4) was consitent with a peptide eleven residues in length and carrying a charge of -2 at pH 6.5. The asx present in Tlb C2 is therefore present as aspartic acid and not asparagine.

50 n moles of Tlb C2 were subjected to 10 cycles of Edman degradation and aliquots were removed for amino-terminal analysis by dansylation after each successive step. The sequence was established as : asp-leu-SCM-cys-arg-pro-SCM-cys-gly-pro-thr-pro-leu.

Measurement of the amount of radioactivity retained in the shortened peptide after the removal of the third residue ($\left[^{14}C\right]$ -SCM-cysteine) by the Edman procedure allowed the overall efficiency of three steps to be estimated at 87% and thus for individual steps at 96%. This is in good accord with reported efficiencies (Gray, 1972).

Peptide Tlb C3

The amino acid composition and the relative mobility of Tlb C3 (Table 6.4) indicated a peptide of ten residues and with a charge of -2 at pH 6.5 - therefore two of the three acidic amino acids were present in the amide form although which of the three could not be determined. Seven cycles of the Edman procedure established the sequence: ala-asx-ser-SCM cys-asx-glx-pro-SCM cys (val,arg)

Incubation of Tlb C3 (10 n moles) with carboxypeptidase A alone released no free amino acids detectable by dansylation but addition of carboxypeptidase B released free arginine and valine which established the carboxyl-terminal sequence of this peptide as -val-arg, completing the sequence and establishing that Tlb C3 was the carboxyl-terminal peptide

TABLE 6.4 The Amino Acid Compositions of T1b and its Chymotryptic Fragments

Approximately 0.10 μmoles of each peptide was hydrolysed and analysed as described in Methods.

(a) Values are expressed as absolute μmoles

(b) Values are expressed as ratios relative to a convenient standard: numbers adjacent in brackets refer to the number of residues determined by dansyl-Edman procedures.

1 Molecular weights (MW) of T1b, T1b.C2 and T1b.C3 were calculated from the nearest integralvalue of each residue in their respective amino acid compositions assuming the average residue weight of 110. The molecular weight of T1b C1 is that of the tripeptide Nacetyl-ser SCM-cys-phe.

2 Relative mobilities (Mrel) were measured and the net charge of each peptide determined by the method of Offord (1966).

3 The proline content of T1b.C2 was higher in this particular analysis than in others of the same peptide and is likely to be an overestimate. Furthermore, only three proline residues were detected in the course of primary sequence determination.

TABLE 6.4

	Tlb		Tlb C1		Tlb C2		Tlb C3	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
SCM-cys	0.229	4.50(5)	0.117	0.84(1)	0.114	1.72(2)	0.161	1.56(2)
asp	0.149	2.87(3)			0.066	1.00(1)	0.188	1.83(2)
thr	0.056	1.13(1)			0.069	1.04(1)	0.017	0.17(0)
ser	0.105	2.13(2)	0.139	1.00(1)		0.00(0)	0.105	1.03(1)
glu	0.062	1.25(1)				0.00(0)	0.100	0.97(1)
pro	0.187	3.77(4)			0.290[3]	4.38(3)	0.109	1.06(1)
gly	0.060	1.21(1)			0.079	1.19(1)	0.025	0.24(0)
ala	0.056	1.13(1)			-	0.00(0)	0.103	1.00(1)
val	0.050	1.00(1)			-	0.00(0)	0.109	1.06(1)
met	0.00	0 (0)			-	0.00(0)		
ile	0.005	0.10(0)			-	0.00(0)	0.005	0.05(0)
leu	0.105	2.12(2)			0.161	2.42(2)	0.013	0.13(0)
tyr	0.008	0.15(0)						
phe	0.042	0.85(1)	0.130	0.93(1)			0.007	0.06(0)
lys	0.003	0.06(0)						
his	0.007	0.13(0)						
arg	0.094	1.90(2)			0.072	1.09(1)	0.077	1.00(1)

[1]MW	2640	455	1210	1100
[2]Mrel	-	0.89	0.48	0.48
[2]Net Charge	-	-2	-2	-2

of the tryptic peptide T1b.

The complete sequence of T1b is

```
          1      2    3   4   5   6       7   8   9
* N-acetyl-ser-SCM cys-phe-asp-leu-SCM cys-arg-pro-SCM cys-

10  11  12  13  14  15  16  17   18      19  20  21   22
gly-pro-thr-pro-leu-ala-asx-ser-SCM cys-asx-glx-pro-SCM cys-

23  24
val-arg
```

The resistance to tryptic digestion of the bond arg(7)-pro(8) is in accord with that established for arg-pro bonds in other proteins (Smyth, 1967). The sequence of T1b is in good agreement with the amino acid compositions of the intact peptide and with its chymotryptic peptides. The proline content of T1b C2 (4.38) from amino acid analysis (Table 6.4) indicates that at least one more residue of proline than can be accounted for by the sequence data (three) but the unambiguous detection of successive amino-terminal residues exposed after each cycle of the Edman reaction suggests that the high proline value given in Table 6.4 was an overestimate.

(d) Peptide T1c

Peptide T1c was isolated by preparative HVPE at pH 6.5 and pH 2.7 (Figs. 6.3, 6.4) - the amount recovered by

* The residues are numbered in the order in which they occur in the sequence of T1b. In Fig. 6.12 they are aligned with an homologous sequence of emu rachis keratin and the numbering system of O'Donnell (1973b) has been adopted for comparative purposes.

elution from paper was estimated as 0.18μmoles. A sample was taken for amino acid analysis (Table 6.5) and the remainder subjected to chymotryptic digestion. The chymotryptic digest was separated by HVPE at pH 6.5 into three radioactive peptides Tlc C1, Tlc C2 and Tlc C3 (Fig. 6.7a)and the barely-resolved peptides Tlc C2 and Tlc C3 separated completely by HVPE at pH 2.7 (Fig. 6.8b). The three radioactive peptides were eluted form paper as described in Methods and samples from each withdrawn for amino acid analysis (Table 6.5).

Tlc C1

The amino acid composition and relative mobility of this peptide (Table 6.5) were identical to that of peptide Tlb C1 (Table 6.4) and the sequences were assumed to be identical (N acetyl-ser-SCM cys-phe).

Tlc C2

The amino acid composition of this peptide was very similar to that of Tlb C2. There were two differences - Tlb C2 contained no serine residues, one arginine residue and three proline residues whereas Tlc C2 contained two serine residues, no arginine residues and two proline residues. Both peptides contained the same number of residues and both carried a charge of -2 at pH 6.5 (Tables 6.4 and 6.5). The asx of Tlc C2 was asn on this basis (Offord, 1966) and from seven cycles of the dansyl-Edman procedure performed on Tlc C2 (0.03 μmoles) the partial sequence asn-leu-SCM cys-ser-pro-SCM cys-gly-pro (thr, ser, leu) was established - lack of

FIGURE 6.7 The purification of the chymotryptic peptides of T1C

T1c (approx. 0.1 μmoles) was digested with 13 μg of chymotrypsin in 0.13 ml of NEMA, pH 8.3 and the digest was subjected to HVPE at pH 6.5 (2000 V, 1 h) and the electrophoretogram autoradiographed (a). The peptides T1c C2 and T1c C3 were cut out together and separated by electrophoresis at pH 2.7 (3000 V, 1.5 h) and the electrophoretogram autoradiographed.

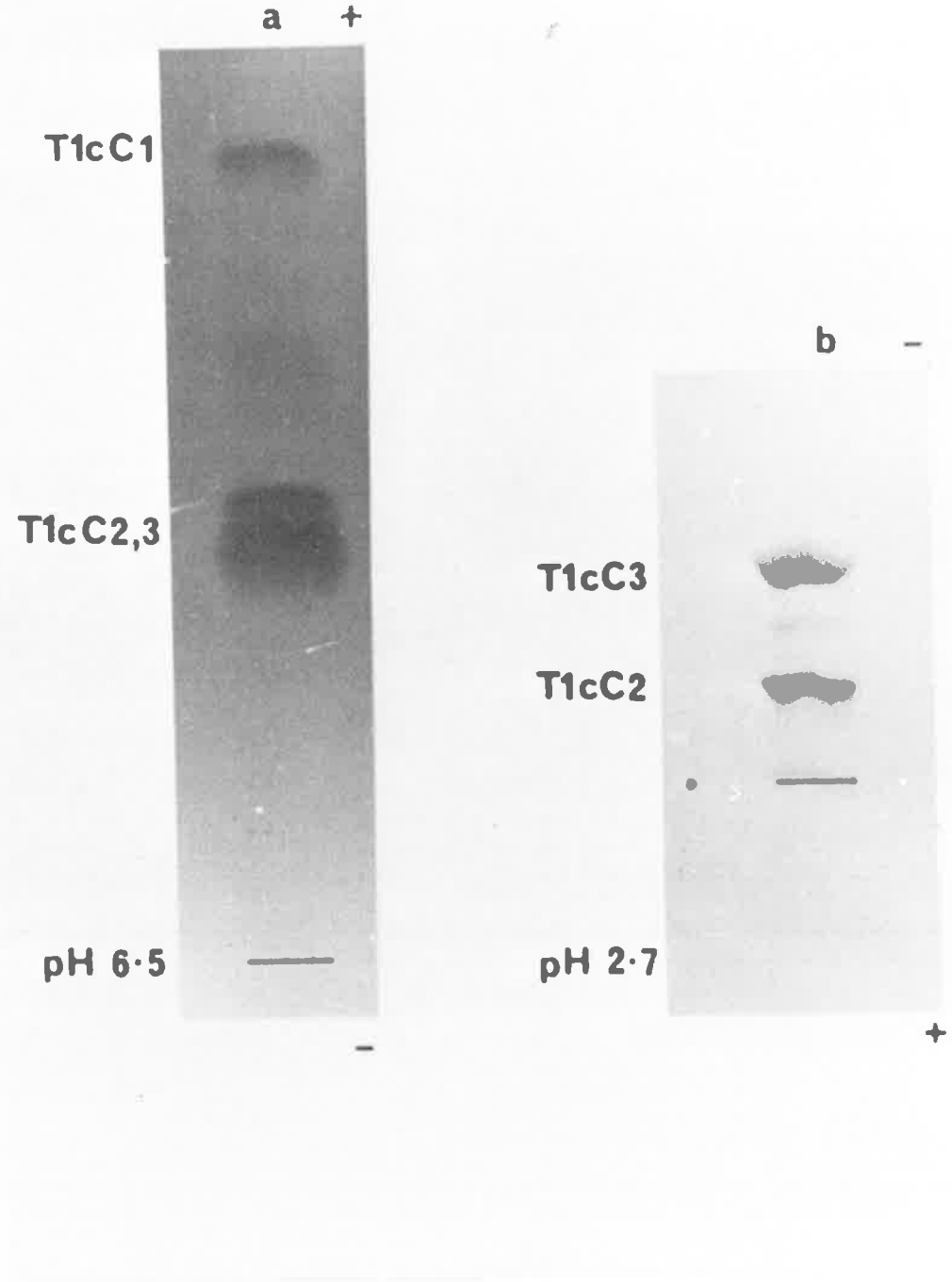

a
+
T1cC1
T1cC2,3
pH 6·5
-
b
-
T1cC3
T1cC2
pH 2·7
+

TABLE 6.5 The amino acid compositions of peptide T1c and its chymotryptic fragments

T1c and T1c C1 (approx 0.05 μmoles) were hydrolysed and analysed as described. Approximately 0.01 μmoles of T1c C2 and T1c C3 were hydrolysed and analysed using the scale expander.

(a) Values are expressed as absolute μmoles

(b) Values are expressed as ratios relative to a convenient standard. Numbers adjacent in brackets are the estimated number of residues present in the peptide as determined by sequence analysis.

[1] Values likely to be underestimates (see Text).

TABLE 6.5. The amino acid compositions of peptide Tlc and its chymotryptic fragments

	Tlc		Tlc C1		Tl cC2		Tlc C3	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
SCM-cys	0.197	5.18(5)	0.079	0.91(1)	0.024	1.84(2)	0.022	1.69(2)
asp	0.063	[1]1.66(3)			0.012	0.92(1)	0.022	1.69(2)
thr	0.055	1.45(1)			0.013	1.00(1)	0.007	0.54(0)
ser	0.125	[1]3.29(4)	0.095	1.11(1)	0.024	1.85(2)	0.017	1.31(1)
glu	0.043	1.13(1)			-	(0)	0.013	1.00(1)
pro	0.144	3.78(3)			0.026	2.00(2)	0.009	0.69(1)
gly	0.038	1.00(1)			0.013	1.00(1)	0.006	0.46(0)
ala	0.040	1.05(1)			-		0.015	1.15(1)
val	0.028	0.74(1)			-		0.011	0.85(1)
met	-	(0)			-			
ile	-	(0)			-			
leu	0.076	2.00(2)			0.023	1.77(2)		
tyr	-	(0)						
phe	0.033	0.87(1)	0.086	1.00(1)				
lys	-	(0)						
his	-	(0)						
arg	0.032	0.84(1)			-	0	0.011	0.85(1)

MW	2640	455	1210	1100
Mrel	-	0.90	0.50	0.46
Net Charge	-	-2	-2	-2

peptide material prevented completion of the sequence.

Tlc C3

The relative mobility of this peptide at pH 6.5 (Table 6.5) and at pH 2.7 were identical to that of Tlb C3 : its amino acid composition (Table 6.5) showed no significant differences to that of Tlb C3 (Table 6.4). Seven cycles of dansyl-Edman procedure established the sequence ala-asx-ser-SCM cys-asx-glx-pro-SCM cys (val, arg) and thus, the two peptides Tlb C3 and Tlc C3 on this basis and on that of their mobilities at pH 6.5 and pH 2.7 were identical.

The amino acid composition of Tlc (Table 6.5) revealed less aspartic acid (1.66) and serine (3.29) than the total amounts of aspartic acid (3) and serine (4) in the three chymotryptic fragments, but since all three aspartic acid residues and two of the four serine residues were directly identified by dansylation in the course of Edman degradation of peptides Tlc C2 and Tlc C3 the low values for aspartic acid and serine in Tlc (Table 6.5) must be underestimates.

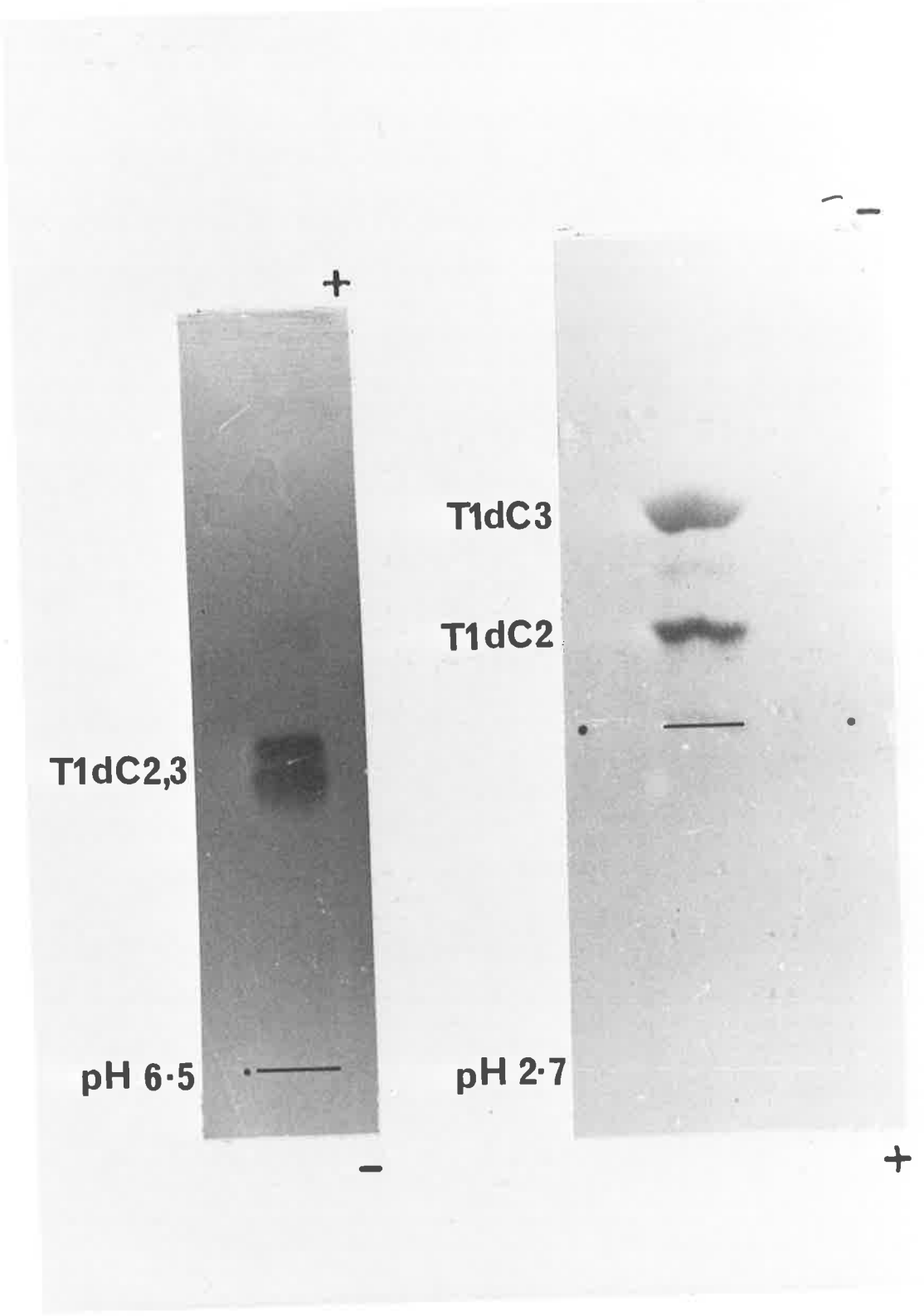
The sequence of Tlc was established as N-acetyl-ser-SCM cys-phe-asn-leu SCM cys-ser-pro SCM cys-gly-pro-(thr, ser,leu)-ala-asx-ser-SCM cys-asx-glx-pro-SCM cys-(val,arg).

(e) Peptide Tld

Peptide Tld was isolated by HVPE at pH 6.5 and pH 2.7 (Figs. 6.3, 6.4) and 0.06 μ moles was recovered by elution of this peptide from paper. Chymotryptic digestion was carried out and the digest subjected to HVPE at pH 6.5 (Fig. 6.8a). Two

FIGURE 6.8 The purification of the chymotryptic peptides of T1d

T1d (approx) 0.06 μmoles was digested with 8 μg of chymotrypsin in 80 μL of NEMA, pH 8.3 and the digest subjected to HVPE at pH 6.5 (2000 V, 1 h) and the electrophoretogram autoradiographed (a). The peptides T1d C2 and T1d C3 were cut out and electrophoresed together at pH 2.7 (3000 V, 1.5 h) and the electrophoretogram autoradiographed (b).


+
T1dC2,3
pH 6·5
−
−
T1dC3
T1dC2
pH 2·7
+

peptides T1d C2, T1d C3 of relative mobilities (0.46 and 0.50 respectively) were barely resolved and were cut out together and completely separated by HVPE at pH 2.7 (Fig. 6.8). The peptides T1d C2 and T1d C3 after elution from paper (0.026 and 0.023 μ moles respectively) were hydrolysed for amino acid analysis (Table 6.6).

T1d C2

The amino acid composition of T1d C2 was identical to that of T1b C2 except that glu replaced asp. The relative mobility of T1d C2 (Table 6.6) indicated that it possessed a net charge of -2 and the glx in its amino acid composition was therefore glu in the intact peptide (Offord, 1966). SCM- cysteine in the amino acid composition of T1d C2 (Table 6.6) has been underestimated (1.29 residues) since determination of radioactivity of a sample of T1d C2 before hydrolysis indicated the presence of 0.026 μ moles of SCM- cys (1.86 residues).

T1d C3

The amino acid composition and the relative mobility of T1d C3 (Table 6.6) were not significantly different to those properties of peptides T1b C3 (Table 6.4) or T1c C3 (Table 6.5). The absence of the radioactive tripeptide N-acetyl-ser-SCM-cys-phe in chymotryptic digests of T1d is considered in Discussion.

(iii) The Structure of the Peptides of Peak C

At least five major peptides containing SCM-

TABLE 6.6 The amino acid compositions of the chymotryptic fragments of peptide T1d

T1d.C2 (approx 0.02 μmoles) and T1d C3 (approx 0.023 μ moles) were hydrolysed and amino acid analysis carried out with the use of the scale expander.

(a) Values expressed as absolute μmoles

(b) Values expressed relative to a convenient standard.

TABLE 6.6 The amino acid compositions of the chymotryptic fragments of peptide T1d

	T1d C2		T1d C3	
	(a)	(b)	(a)	(b)
SCM-cys	0.018	1.29(2)	0.028	1.56(2)
asp	0.004	0.29(0)	0.030	1.67(2)
thr	0.014	1.00(1)		
ser	0.006	0.43(0)	0.020	1.11(1)
glu	0.012	0.86(1)	0.020	1.11(1)
pro	0.042	3.00(3)	0.022	1.22(1)
gly	0.020	1.43(1)	0.008	0.44(0)
ala			0.018	1.00(1)
val			0.016	0.89(1)
met				
ile				
leu	0.028	2.00(2)		
tyr				
phe				
lys				
his				
arg	0.016	1.14(1)	0.018	1.00(1)

MW	1210	1100
Mrel	0.46	0.50
Net Charge	-2	-2

cysteine compose peak C of the Sephadex G-50 profile (Figs. 5.1a, b). The molecular weights of all the constituent peptides of peak C were less than 1000 indicating that all such peptides had ten or less residues. To obtain pure samples of each peptide for sequence analysis peak C material (Fig. 5.1a) was subjected to chromatography on Dowex-50 and the peaks of peptide material obtained were further purified by HVPE. Samples of pure peptides, thus obtained, were subjected to amino acid analysis and primary structure determination. Part (i) of this section deals with the isolation of ten prominent peptides of peak C and part (ii) with the elucidation of their respective sequences.

(a) The Purification of Peak C Peptides

Peak C peptides, derived from 100 mg of $[^{14}C]$ - SCM- feather keratin were subjected to chromatography on Dowex-50 as described in Methods and aliquots of fractions withdrawn for the determination of peptide material by the ninhydrin method and of radioactivity. Seven major peaks of peptide material, five of which contained radioactivity eluted from the Dowex-50 column (Fig. 6.9) and these were labelled a-g and tubes pooled as shown. Purification of the constituent peptides of each peak was carried out as follows.

Peak (a)

Peak (a) peptides were subjected to HVPE at pH 6.5 and the electrophoretogram autoradiographed (Fig. 6.10a). The two prominent radioactive peptide bands T5a and T1e,

FIGURE 6.9 Separation of peak C tryptic peptides by Dowex-50 chromatography

Peak C peptides (equivalent to 100 mg of [^{14}C] - SCM-feather keratin, Fig. 5.1a) were applied to Dowex-50 and eluted with the gradient described in Methods. 1.65 ml fractions were collected and 50μL aliquots withdrawn for both the determination of radioactivity and the estimation of peptide material as described in Methods.

Tubes were pooled as shown.

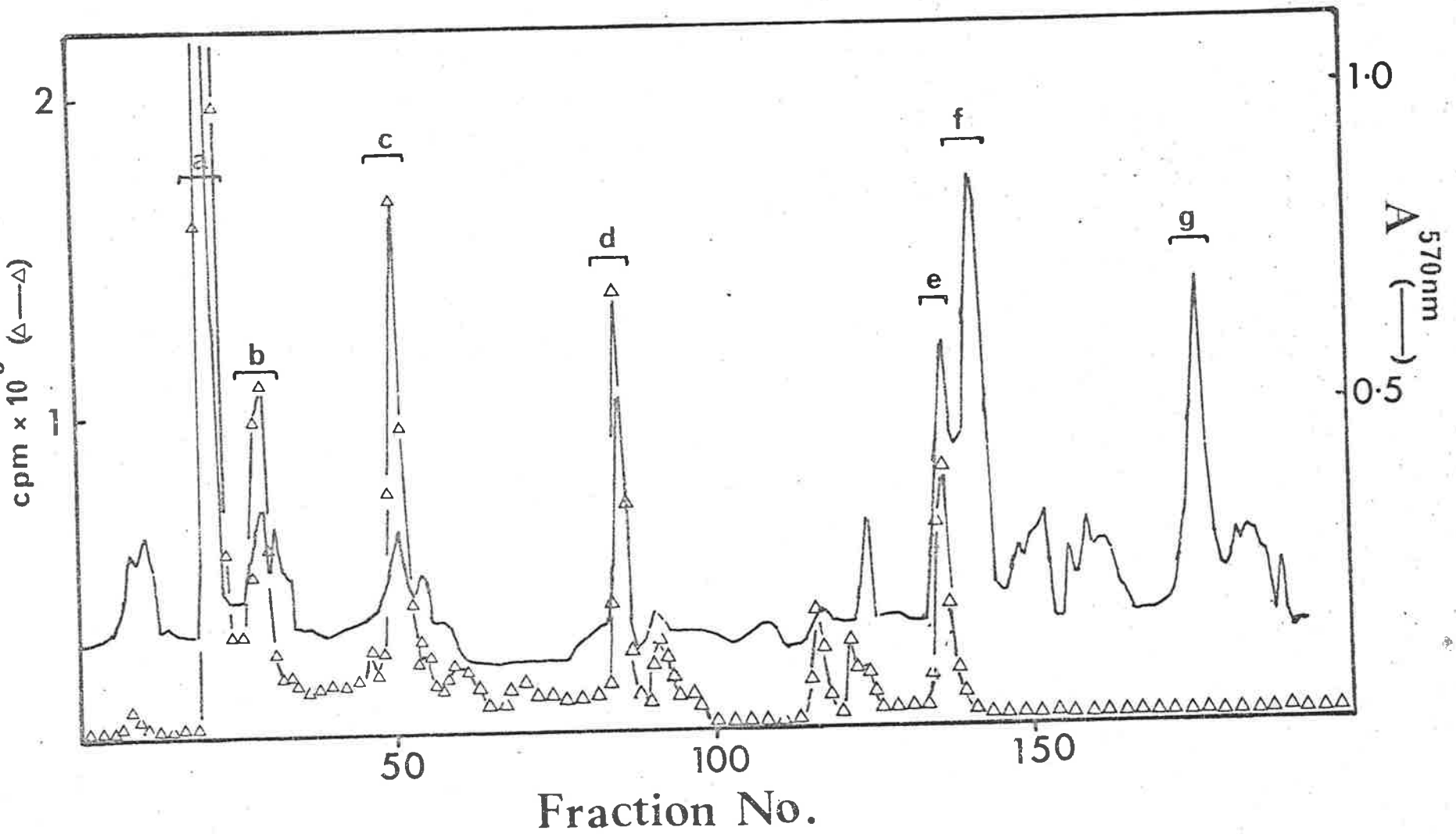
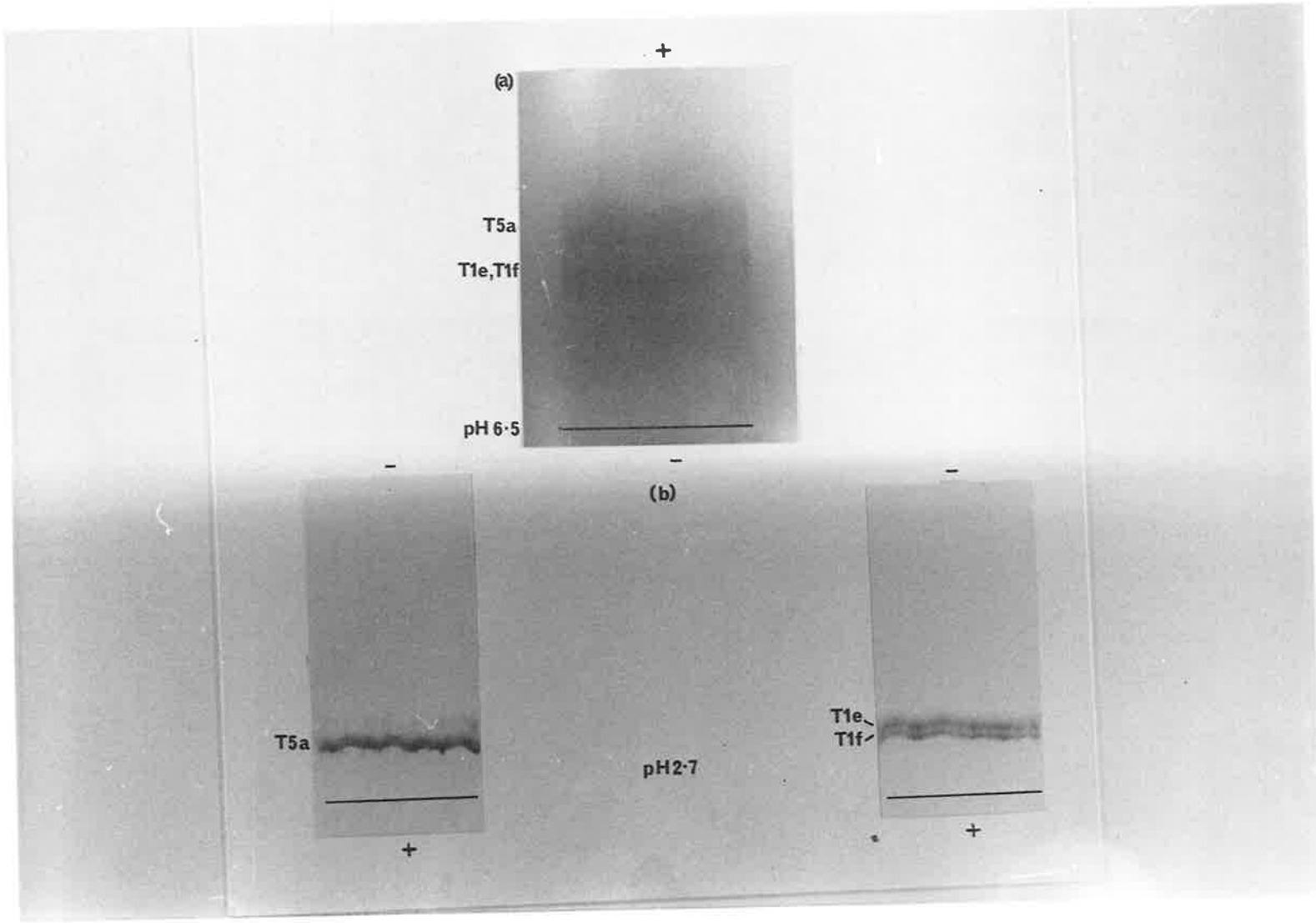

cpm × 10⁻³ (△—△)
A 570nm (——)
2
1
1·0
0·5
a
b
c
d
e
f
g
50
100
150
Fraction No.

FIGURE 6.10 Purification of peptides T5a, T1e and T1f

The peptides of peak a of the Dowex-50 chromatogram (Fig. 6.10) were subjected to HVPE at pH 6.5 (2000 V, 1.5 h) and the electrophoretogram autoradiographed (a). The two bands of (a) (T5a and T1e, T1f) were subjected separately to HVPE at pH 2.7 (3000 V, 1.5 h) and the electrophoretogram autoradiographed (b).


+
(a)
T5a
T1e,T1f
pH 6·5
–
–
–
(b)
T5a
pH 2·7
T1e
T1f
+
+

Tlf were cut out and further purified by HVPE at pH 2.7 (Fig. 6.10b). All three peptides were eluted from the electrophoretograms and samples removed for amino acid analysis.

Peaks b, c, d and e

HVPE at pH 6.5 of the peptides of peaks b, c, d and e was carried out and the autoradiogram (Fig. 6.11a) indicated that each peak contained only one radioactive peptide except peak c which contained minor amounts of a radioactive peptide whose properties were not further investigated. Samples of peaks f and g (Fig. 6.10) were subjected to HVPE at pH 6.5 and the electrophoretogram sprayed with ninhydrin (Fig. 6.11b). Both peaks contained only 1 peptide band and the peptides (T4b and Tx) were not purified further.

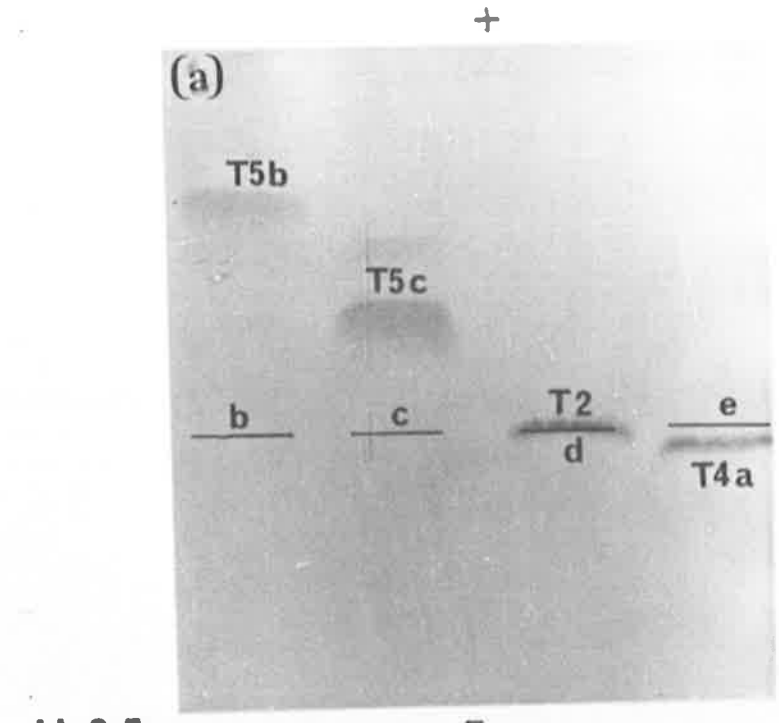
(b) The Amino Acid Sequences of Peak C Peptides Tle and Tlf

The amino acid compositions of Tle and Tlf (Table 6.7) were very similar, the former having no tyrosine and two glutamic acid residues the latter having one tyrosine and one glutamic acid residue but otherwise the two peptides were identical in amino acid composition. Dansylation of both peptides indicated that they possessed no free amino-terminal residue and therefore may be amino-terminal peptides. Their relative mobilities indicated a charge of -3 (Table 6.7) for both peptides and thus two of the acidicamino acid residues in each peptide were present as carboxylate and not amide forms.

FIGURE 6.11 Purification of peptides T5a, T5b, T2, T4a T4b and Tx

(a) HVPE at pH 6.5 (2000 V, 1 h) of the peptides of peaks b, c, d, e (Fig. 6.10). Autoradiography was performed as previously described.

(b) Aliquots (10%) of peaks f and g (Fig. 6.10) were subjected to HVPE at pH 6.5 (2000 V, 1 h) and the peptides localised by spraying with 1% ninhydrin (Methods).



pH 6·5

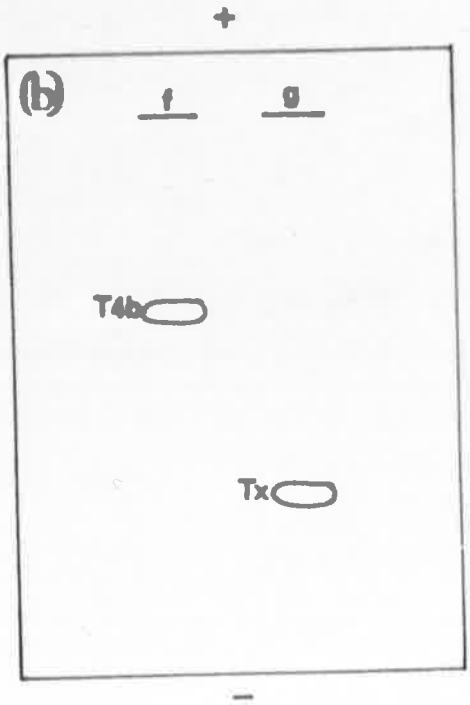


TABLE 6.7 The amino acid compositions of peak C peptides

A known amount (0.1-0.2 μmoles) of each of the peptides of peaks (a-g) (Fig. 6.10) was subjected to amino acid analysis after further purification (Figs. 6.11, 6.12).

(a) Values expressed as absolute μmoles

(b) Values expressed relative to a convenient standard.

[1] Molecular weights (MW) of all peptides are based on integral numbers of residues (column b). The molecular weights of Tle and Tlf were calculated assuming an approximate residue weight of 110 whereas those of all other residues are exact values.

TABLE 6.7

Peak (D50)	a		a		a		b	
Peptide	T1e		T1f		T5a		T5b	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
SCM-cys	0.258	1.79(2)	0.176	2.51(2)	0.189	2.07(2)	0.186	2.27(2)
asp	0.184	1.23(1)	0.086	1.23(1)				
thr								
ser	0.150	1.04(1)	0.070	1.00(1)				
glu	0.252	1.75(2)	0.066	0.94(1)				
pro	0.150	1.04(1)	0.080	1.14(1)	0.091	1.00(1)	0.095	1.16(1)
gly								
ala								
val								
met								
ile								
leu	0.144	1.00(1)	0.072	1.03(1)	0.089	0.98(1)		
tyr			0.048	0.71(1)				
phe							0.082	1.00(1)
lys								
his								
arg	0.104	0.72(1)	0.070	1.00(1)				

^{1}MW	1000	1000	550	584
Mrel	0.62	0.62	0.72	0.78
Net Charge	-3	-3	-2	-2

TABLE 6.7 (Contd.)

Peak(D50)	c		d		e		f	
Peptide	T5c		T2		T4a		4b	
SCM-cys	0.135	0.89(1)	0.114	0.97(1)	0.088	1.00(1)		
asp	0.213	1.40(1)	0.118	1.00(1)				
thr								
ser			0.138	1.17(1)			0.446	1.00(1)
glu			0.290	2.46(2)				
pro								
gly					0.122	1.39(1)	0.466	1.04(1)
ala								
val								
met								
ile								
leu	0.152	1.00(1)						
tyr	0.150	0.98(1)			0.088	1.00(1)		
phe							0.412	0.92(1)
lys								
his								
arg			0.104	0.88(1)	0.086	0.98(1)	0.416	0.93(1)

MW	552	795	555	455
M rel	0.43	0	0	-0.48
Net Charge	-1	0	0	+1

TABLE 6.7 (Contd.)

Peak D50		g
Peptide		?
	(a)	(b)
SCM-cys		
asp		
thr		
ser		
glu		
pro		
gly		
ala		
val		
met		
ile		
leu		
tyr		
phe		
lys		
his		
arg	0.128	1.00(1)

MW	174
Mrel	0.90
Net Charge	+1

No further sequence investigations were carried out on these peptides but their possible structures are considered in Discussion.

T2

The amino acid composition and the relative mobility of T2 (Table 6.7) indicated a neutral peptide of six residues. Therefore since T2 contains one residue of arginine, all three acidic amino acids must be present as their amide forms. Dansylation of T2 however revealed no amino-terminal residue suggesting the presence of amino-terminal pyrrolidone carboxylic acid (PCA). Examination of carboxyl-terminal residues released from T2 by a mixture of carboxypeptidases A and B revealed serine and arginine. The structure of T2 is probably gln-SCM-cys-glu-asn-ser-arg. (Discussion).

T4a, T4b, T5a, T5b, T5c

Four successive cycles of the dansyl-Edman procedure were carried out on 50 nmoles each of the peptides T4a, T4b, T5a, T5b and T5c and the following sequences were established.

T4a tyr-SCM-cys-gly-arg

T4b phe-ser-gly-arg

T5a SCM-cys-leu-pro-SCM-cys

T5b SCM-cys-phe-pro-SCM-cys

T5c tyr-leu-pro-SCM-cys

These amino acid sequences are in accord with the amino acid compositions and the relative mobilities of

each of the peptides (Table 6.7).

Tx

The amino acid composition and the relative mobility of peptide Tx (Table 6.7) established that it was free arginine.

DISCUSSION

The purpose of this discussion is to consider the possible arrangement of the tryptic peptides whose sequences have been determined, within the primary structure of SCM- feather keratin and to evaluate the significance of the multiple substitutions within the sequence with respect to their genetic origin. The homology which exists among all SCM- feather keratin chains from the embryonic chick and between them and a purified protein of emu rachis keratin whose primary structure has been determined (O'Donnell, 1973a) has enabled the tentative placement of the tryptic peptides of SCM- feather keratin within the primary structure. To facilitate this interpretation section (i) is devoted to a brief consideration of the general topic of protein homology. Section (ii) is concerned with the application of homology considerations to the primary structure of SCM- feather keratin. The evolutionary events which can account for the generation of sequence diversity in SCM- feather keratin molecules are the subject of section (iii) and section (iv) deals with the relationship of the primary structure of feather keratin molecules to their function.

(i) Homology between Protein Molecules

The sequence of specific protein molecules derived from two species of organisms bear defined relationships to one another and the extent of similarity between two such homologous protein molecules depends upon how closely the two species are related with respect to their evolutionary origin (Fitch and Margoliash, 1967). Thus, the cytochrome c of man differs by only one substitution in the sequence from that of monkey and by six substitutions from the cytochrome c of pig. Differences in homologous protein molecules between species reflect changes in the genes and these genetic alterations can often be accounted for by changes at a single nucleotide within the gene which changes the coding specificity of the trinucleotide triplet to which it belongs. Thus the Erythrocebus Patas cytochrome c differs from human cytochrome c only by a substitution of $\overset{93}{\text{glu}}$ for $\overset{93}{\text{ala}}$ and this can be accounted for by alteration of a glu codon (GA$\overset{A}{G}$) to an ala codon (GC$\overset{A}{T}$) (Dayhoff, 1969). Commonly, two homologous proteins may differ in chain length in a manner which suggests the deletion of one or a number of trinucleotide codons from the gene of the shorter protein. As an example the cytochrome C of tuna lacks the first six residues in the primary structure when compared to the cytochrome C of other vertebrates (Fitch and Margoliash, 1967).

The mutation distance between two homologous genes is defined as the minimum number of nucleotides which must be altered in the base sequence of one gene for it to code for the other. This parameter is commonly used as a basis for

comparison of two homologous genes and the larger its value the more distantly related are the two genes (and by inference the two species) which it compares (Fitch and Margoliash, 1967).

Within a single organism there are examples of two or more distinct and non allelic genes which produce protein products whose primary structures are closely homologous. The β, γ and δ chains of human haemoglobin evidence this and it is believed that all three genes were derived from a common ancestral gene which duplicated twice and that each duplicate mutated independently leading to the present differences in primary structure (Dayhoff, 1969).

Sequence studies of the high sulphur and low sulphur proteins of wool indicate that within one tissue of a single organism there exists two or more large families of protein molecules, the individual members of which are closely homologous with respect to sequence (Fraser et al., 1972). By comparing the amino acid sequences of an homologous region in seven purified high sulphur proteins from wool, Swart (1973) was able to suggest a possible DNA sequence of the ancestral gene from which the seven high sulphur protein genes have evolved by gene replication and subsequent mutation.

(ii) The Primary Structure of SCM- Feather Keratin

(a) The amino-terminal sequence of SCM-feather keratin (T1)

The peptides T1a - T1c are all clearly amino-terminal tryptic peptides of SCM- feather keratin since each

FIGURE 6.12. The partial amino acid sequence of SCM- feather keratin

The amino acid sequences of the tryptic peptides of SCM- feather keratin have been arranged in order, in such a manner as to maximise homology with the sequence of a component of emu rachis (O'Donnell, 1973a). The SCM- chick feather keratin sequences have been numbered in relation to the emu rachis sequence.

Amino acid sequences in parenthesis have been inferred from amino acid compositions and homology considerations (see Text)

→ amino acid residue identified by dansylation in the course of sequential degradation of the peptide

← amino acids identified as dansyl derivatives after carboxypeptidase digestion of the peptide.

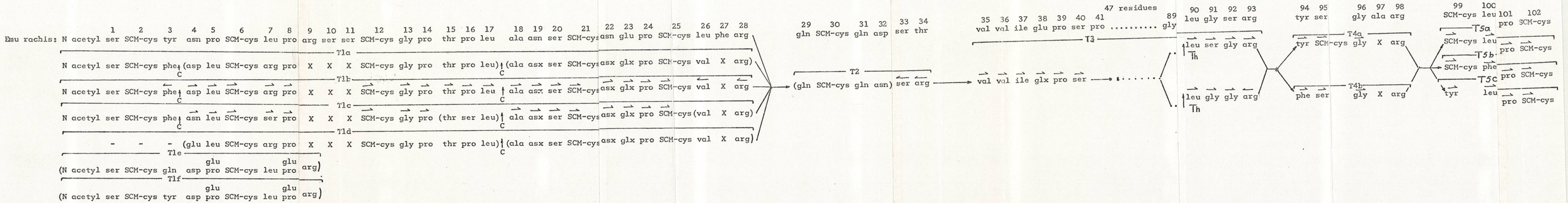
X amino acid deletion

– no assignment possible

↑ ↑

C and Th: sites of chymotryptic and thermolytic digestion.

FIGURE 6.12


Emu rachis: N acetyl ser SCM-cys tyr asn pro SCM-cys leu pro arg ser ser SCM-cys gly pro thr pro leu ala asn ser SCM-cys asn glu pro SCM-cys leu phe arg
T1a
T1b
T1c
T1d
T1e
T1f
T2
T3
T4a
T4b
T5a
T5b
T5c
Th
47 residues

yields the tripeptide N-acetyl-ser SCM cys-phe in approximately molaramounts on chymotryptic digestion. The complete sequence of T1b establishes that it is closely homologous to the amino-terminal region of the primary structure ofapurified emu rachis keratin (O'Donnell, 1973b). Figure 6.12 shows the alignment of T1b with the amino-terminal twenty-eight residues of emu rachis. The alignment has been made so as to maximise the homology between the two sequences and to permit the calculation of mutation distances (Fitch and Margoliash, 1967). The following differences are evident between the amino terminal sequences of emu rachis and of T1b. Emu rachis residues ^{3}tyr, ^{4}asn ^{5}pro and ^{7}leu in T1b are ^{3}phe, ^{4}asp, ^{5}leu and ^{7}arg: the emu rachis sequences ^{9}arg, ^{10}ser, ^{11}ser and ^{27}phe have been deleted in T1b resulting in it being four residues shorter than the homologous part of the emu rachis sequence. The charge (-2, Table 6.4) of peptide T1b C_3 indicates that one of the three acidic amino acids (^{19}asx ^{22}asx and ^{23}glx) in T1b is present as a carboxylate species and the other two are present as amides. This is consistent with the data for the same region of emu rachis (Fig. 6.12) in which only ^{23}glu is present as the carboxylate species. Which of the residues ^{19}asx ^{22}asx or ^{23}glx in T1b is present as the amide form cannot however be inferred with certainty on the basis of homology considerations alone.

T1a was not obtained in sufficient amount for direct primary sequence analysis. However the electrophoretic

mobilities and amino acid composition of the fragments Tla C1 and Tla C2 are not significantly different to those of peptides Tlb C1 and Tlb C2 (compare tables 6.3 and 6.4) and the region of sequence from ser^{1} - leu^{17} in peptides Tla and Tlb are probably identical. Tla C3 however differred from Tlb C3 in its possession of an extra net negative charge at pH 6.5, although its amino acid composition was identical to that of Tlb C3. The most likely explanation for this is that two of the three acidic amino acids in Tla C3 are present as carboxylate species conferring one extra negative charge as compared to Tlb C3. Only one of three acidic amino acids in Tla C3 is present as its amide form but it has not been identified and therefore residues 19, 22 and 23 in Tla are written in Figure 6.12 as asx, asx, and glx respectively. The sole difference between Tla and Tlb on the basis of electrophoretic mobilities and amino acid compositions of their respective chymotryptic fragments is the substitution of asp for asn at positions 19 or 22 or glu^{23} for gln^{23}.

Tlc differs from Tlb in having two extra serine residues and one residue less of proline and arginine (compare Tables 6.4 and 6.5). Substitution of arg^{7} for ser^{7} in Tlc C2 was detected in the course of primary structure analysis by the dansyl-Edman technique (see Results) and substitution of pro^{16} for ser^{16} has been inferred from the amino acid compositions of Tlb C2 and Tlc C2 (Tables 6.4 and 6.5). As shown in Figure 6.12 the sequence of Tlc is otherwise indistinguishable from

that of T1b.

T1d on digestion with chymotrypsin produced no amino-terminal tripeptide fragment detectable by autoradiography (Fig. 6.8,a) and therefore cys^{2} has been substituted by another amino acid. Otherwise however the similarity in amino acid composition of the two radioactive fragments of T1d (T1d C2 and T1d C3, Table 6.6) to T1b C2 and T1b C3 leave little doubt as to the identity of T1d as an amino-terminal tryptic peptide. The only other difference between the chymotryptic peptides of T1d and T1b is the substitution of asp^{4} (T1b C2) for glu (T1d C2) - presumably of asp^{4} (T1b) for glu (T1d).

The properties of peptides T1a - T1d and their chymotryptic fragments conclusively sstablish that they represent the amino terminal regions of distinct genetic variants of SCM- feather keratin. The mechanism by which such variants could have arisen in the course of evolution is discussed in section (iii).

It appears that T1e and T1f may also be amino terminal peptides derived from other genetic variants of SCM- feather keratin although the basis for this is tenuous compared to that for T1a - T1d. It rests mainly on the observations that T1e and T1f and blocked at their amino-termini and that the amino acid composition of each permits a sequence of amino acids closely homologous to each other and to the amino-terminal sequence of emu rachis (Fig. 6.12).

Peptide T1f may be related to emu rachis by permitting the substitutions ^{4}asp for ^{4}asn and ^{5}pro or ^{8}pro for glu - peptide T1e requires the additional substitution of ^{3}tyr by ^{3}gln. Detailed sequence analysis on T1e and T1f is necessary to test the above proposal.

(b) T2

The placement of T2 after T1 in the sequence of SCM-feather keratin depends upon its properties and amino acid composition. The fact that T2 was neutral at pH 6.5 (Table 6.7) established that its three acidic amino acids were present as amide and not as carboxylated species. The fact that T2 after storgge became resistant to amino-terminal analysis by dansylation supported the conclusion that there was an amino-terminal glutamine whose propensity to cyclise to pyrrdidone carboxylic acid is well established (Blomback, 1967). The carboxyl-terminal sequence ser-arg was that expected of a tryptic peptide of SCM- feather keratin homologous to residues 29-34 of the emu rachis sequence (Fig. 6.12) with the substitutions ^{32}asn for ^{32}asp and ^{34}arg for ^{34}thr. The other strong piece of evidence in support of the sequence and placement of T2 within the sequence is the amino-terminal sequence of T3 which is identical to that of emu rachis residues 35-41 (Fig. 6.12) suggesting the existence of a sequence bridging ^{28}arg and ^{35}val in the primary structure of SCM- feather keratin. T2 was the only tryptic peptide of SCM- feather keratin for which an homologous variant could not be found.

(c) T3

The amino-terminal sequence of T3 showed no evidence of sequence substitutions in the six amino-terminal

residues which were identical in sequence to residues 35 - 40 in emu rachis. However two carboxyl-terminal peptides ($T3Th_g$ and $T3Th_h$) indicated that T3 was not in fact homogeneous. Both carboxyl-terminal sequences (Fig. 6.12) were homologous to residues 90 - 93 of emu rachis, $T3Th_g$ differing by two substitutions. $(gly^{91} - ser^{91})$ and $(ser^{92} - gly^{92})$ and $T3Th_h$ differing by one substitution $(ser^{92} - gly^{92})$.

(d) T4

T4a and T4b were both homologous to the residues 94-98 in the emu rachis sequence. T4a differed from this region in emu rachis by the substitution of ser^{95} by cys and by the deletion of ala^{97}. T4b also lacked ala^{97} and tyr^{94} in emu rachis was substituted by phe^{94}.

(e) T5

The absence of arginine from peptides T5a, T5b and T5c established that they were carboxyl-terminal peptides of SCM- feather keratin. T5a was identical in sequence to residues 99-102 of emu rachis, T5b had phe^{100} substituted for leu^{100} and T5c had tyr^{99} substituted for cys^{99}. All three peptides were clearly homologous to residues 99-102 of emu rachis and to one another.

(f) Tx

The presence of free arginine in the tryptic digest of SCM - feather keratin (Fig. 6.9) establishes the existence of an arg-arg linkage within the primary structure of at least same species of SCM- feather keratin. The placement of the

arg-arg sequence within the primary structure cannot be inferred by homology considerations and must await the isolation of chymotryptic or thermolytic peptides of SCM-feather keratin which contain the arg-arg sequence.

(g) The necessity for the isolation of "overlap" peptides

The high degree of sequence similarity between the tryptic peptides of SCM- feather keratin and homologous regions of emu rachis establishes the overall order in which each of the tryptic peptides (T1-T5) occurs in the primary structures of SCM- feather keratin chains. However there are ten definite positions within the general sequence of SCM-feather keratin at which sequence substitutions occur (positions 2,3,4,7,16,91,94,95,99,100, Fig. 6.12). Presumably each of these substitutions is linked only to certain other substitutions and to investigate this linkage it is necessary to obtain a series of peptides which contain two or more sites at which substitutions occur. For example to test whether SCM- ^{95}cys (peptide T4a) is linked to ^{99}tyr (peptide T5c) (See Fig. 6.12) a chymotryptic digestion of SCM- feather keratin and the attempted isolation of a chymotryptic peptide containing both SCM-^{95}cys and ^{99}tyr (of presumptive sequence SCM-cys-gly-arg-tyr) would resolve the question. Alternatively homogeneous variants of SCM- feather keratin could be purified and the linkage between substitutions studied by sequence analysis of a number of such purified proteins. This topic is the subject of the subsequent chapter.

(iii) The Primary Structure and Evolution of Feather Keratin

The broad similarity in amino acid sequence of emu rachis and SCM- feather keratin makesit clear that the keratin proteins of the two species had a common ancestor from which they diverged by the accumulation of mutations during evolution. However, in the case of chick feather keratin, mutations were not accumulating in only one gene for feather keratin but in multiple genes which were, within the same species, diverging from one another.

In this section the evolution of emu keratin and chick feather keratin from a common ancestor is considered separately from the evolution and divergence of the primary structure variants of chick keratin.

(a) Evolutionary relationship between emu and chick feather keratin

Figure 6.13 shows the genetic relationships between all the positions in emu rachis at which different amino acids have been substituted in SCM- feather keratin. At least twenty genetic alterations must be invoked to account for the differences in the two sequences and with the exception of the proposed substitution tyr^{2} - gln, pro^{5} - glu and pro^{8} - glu each genetic alteration can be accounted for by a single event, for instance a single base change (tyr^{2}, UA py - phe UU py) or a deletion of one or more adjacent codons from the emu rachis keratin gene. However, many of those substitutions cannot occur within the same intact variant

FIGURE 6.13 Genetic interrelationships between the amino acid sequence of emu rachis and SCM- feather keratin

The numbered residues are those of the emu rachis sequence for which amino acid substitutions or deletions in homologous positions in the sequences of SCM- feather keratin have been found.

| represents a single genetic transformation required to interchange the amino acid residue in the emu rachis sequence to its counterpart(s) (shown above and below) in the SCM- feather keratin sequence.

| Two base changes invoked to explain the amino acid substitution.

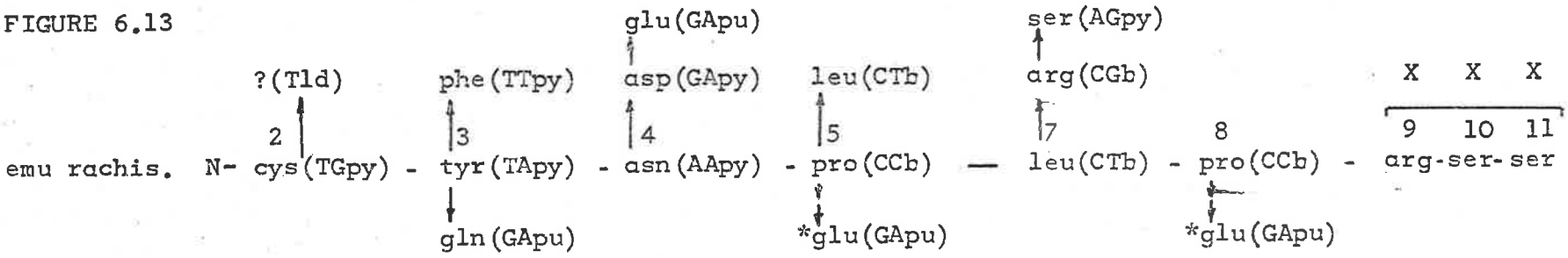
* Substitutions of glu at $\overset{5}{\text{pro}}$ or $\overset{8}{\text{pro}}$ are alternatives - glu substitutes either $\overset{5}{\text{pro}}$ or $\overset{8}{\text{pro}}$ but not both.

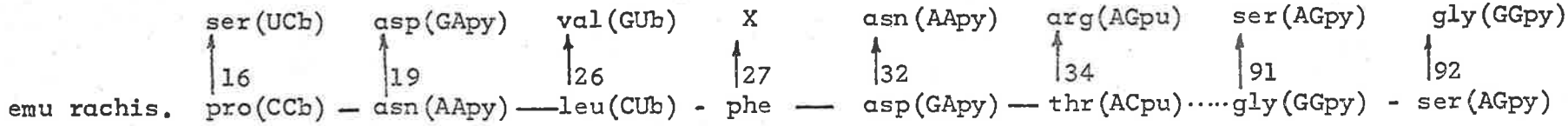
pu - purine nucleotide (A or G)

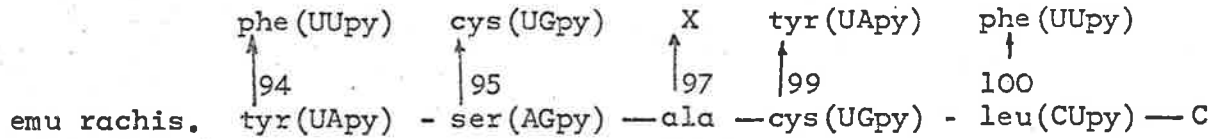
py - pyrimidine nucleotide (C or T)

b - (A G C or T)

FIGURE 6.13


glu(GApu)
ser(AGpy)
?(Tld)
phe(TTpy)
asp(GAPy)
leu(CTb)
arg(CGb)
X X X
2
3
4
5
7
8
9 10 11
emu rachis. N- cys(TGpy) - tyr(TAPy) - asn(AApy) - pro(CCb) — leu(CTb) - pro(CCb) - arg-ser-ser
gln(GApu)
*glu(GApu)
*glu(GApu)


ser(UCb)
asp(GAPy)
val(GUb)
X
asn(AApy)
arg(AGpu)
ser(AGpy)
gly(GGPy)
16
19
26
27
32
34
91
92
emu rachis. pro(CCb) — asn(AApy) —leu(CUb) - phe — asp(GApy) — thr(ACpu)····gly(GGPy) - ser(AGpy)


phe(UUpy)
cys(UGpy)
X
tyr(UApy)
phe(UUpy)
94
95
97
99
100
emu rachis. tyr(UApy) - ser(AGpy) —ala —cys(UGpy) - leu(CUpy) — C

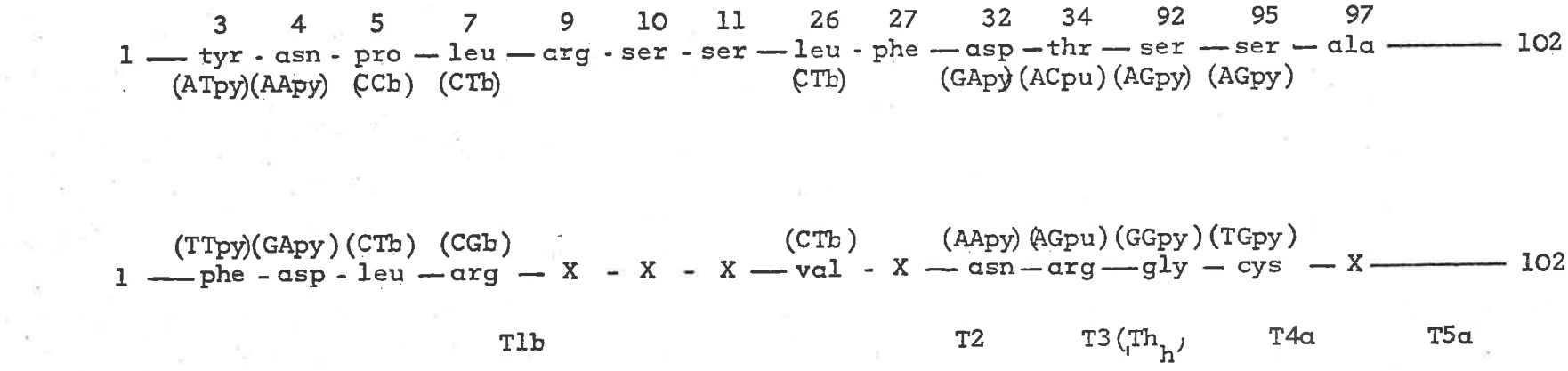
of SCM- feather keratin and are linked only to certain other substitutions. Thus, in T1c ^{7}leu and ^{16}pro have both been substituted by ser whereas ^{7}leu and ^{16}pro in T1b are ^{7}arg and ^{16}pro. To obtain a measure of the minimum mutation distance between emu rachis and SCM- feather keratin, emu rachis has been compared to a hypothetical sequence of SCM- feather keratin constructed by considering only the sequences of SCM- feather keratin tryptic peptides most similar to those of homologous emu rachis sequences (T1b - T2 - T3 (Th_h) - T4a - T5a). Figure 6.14 shows that at least twelve genetic interchanges are required to relate the two sequences (mutation distance = 12, Fitch and Margoliash, 1967). If a similar calculation is performed upon the two most unlike sequences possible within the SCM- feather keratin set by choosing homologous tryptic peptides so as maximise differences between the two sequences a mutation distance of eight can be calculated (Fig. 6.15). Only peptides on which direct primary structure data was collected were chosen for this comparison and peptides T1e and T1f whose primary structures are at best intelligent guesses based on very limited data have not been considered. It is clear from Figures 6.13 and 6.14 (omitting peptides T1e and T1f from the comparison) that all variants of SCM- feather keratin bear a closer evolutionary relationship to one another than they do to emu rachis and this is compatible with the mutational divergence of the multiple chick feather keratin genes from one another having taken place after the divergent speciation of chick and emu ancestral lines. However, much more data

FIGURE 6.14 Minimum *mutation distance between SCM-feather keratin and emu rachis

Sequences T1b, T2, T3(Th_h), T4a and T5a were chosen from Fig. 6.12 so as to minimise the number of genetic changes required to interconvert the sequence which they determined and that of emu rachis.
The residues differing between the two sequences and their possible DNA codons are shown.

* Mutation distance is calculated essentially according to Fitch and Margoliash (1967).
Mutation distance between two proteins is defined as the minimal number of nucleotides that would need to be altered in order for the gene of one protein to code for the other. In the present comparison, deletion of one or more adjacent amino acid residues from one sequence is assigned the same mutation value as a single nucleotide change.

FIGURE 6.14

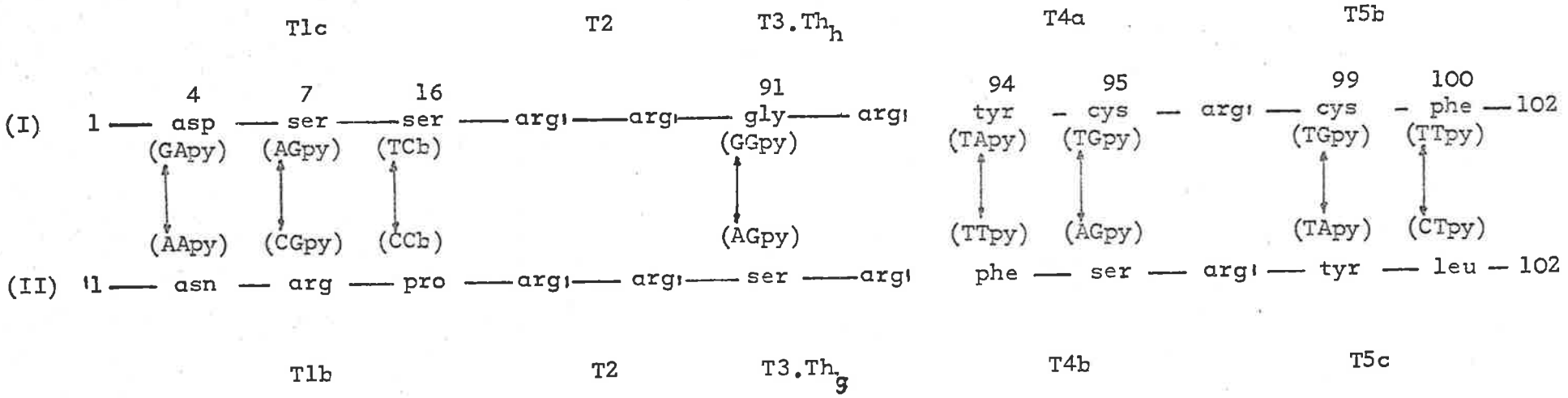


Mutation distance = 12

FIGURE 6.15 The maximum mutation distance between different variants of SCM- feather keratin

The sequences (I and II) represent the two hypothetical sequences most unlike each other with respect to primary structure that it is possible to derive from the data of Figure 6.12. DNA base changes necessary to interrelate the two amino acid sequences are connected by the symbol ↕

FIGURE 6.15



Sequence (I) : T1c T2 T3 (Thh) T4a T5b

Sequence (II) : T1b T2 T3 (Thg) T4b T5c

Mutation distance = 8

is required to substantiate this hypothesis.

(b) Evolution of multiple genetic variants of SCM- feather keratin

The existence of multiple amino acid substitutions in the primary structure of SCM- feather keratin reflects the existence of multiple feather keratin genes. Two alternatives exist to explain the occurence of multiple homologous genes.

(i) Multiple alleles

At a single genetic locus within a diploid organism no more than two alleles are possible. However within a population of organisms many alleles may exist at a single genetic locus, and such genetic polymorphisms are typified by the occurence of fifty-eight genetic variants of β-globin in human populations (Dayhoff, 1969). The InV variants (InV 1,2 and 3) of human immunoglobulin κ (light) chains (Terry et al., 1969) and the existence of multiple variants of human carbonic anhydrase C, at least two of which are allelic and co-dominant (Lin and Deutch, 1972) are other well defined examples of allelic polymorphism. In all three cases primary structure analysis of the allelic variants at each locus has correlated each allele with a single amino acid substitution, itself consistent with a single base change on the DNA of the allelic gene.

The sequence data of SCM- feather keratin (Fig. 6.12) cannot by itself absolutely exclude the existence of multiple alleles at a single locus. The feathers used to prepare the SCM- feather keratin were from a number of animals

which could concievably have been genetically non-uniform with respect to the genes for feather keratin. However, the pattern of sequence substitutions observed in SCM- feather keratin was unlike that observed in the allelic variants of the three examples mentioned above in that <u>multiple</u> substitutions existed in the <u>same</u> SCM- feather keratin chain. Thus peptides T1b and T1c differred not only at one but at two sites in the amino acid sequence (T1b: ^{7}arg, ^{16}pro; T1c: ^{7}ser, ^{16}ser). Similary peptides T4a and T4b differred at two sites (T4a: ^{94}tyr-^{95}cys; T4b ^{94}phe-^{95}ser).

Moreover, gel electrophoresis of SCM- feather keratin performed at pH 9.5 or separately at pH 2.7 could not distinguish the feather proteins of a single animal or from those obtained from a number of animals (Kemp, PhD Thesis, 1972)- electrophoretically the proteins of a single animal (at most expressing two alleles) would be expected to be considerably less heterogeneous than those obtained from many animals, that is, if genetic polymorphism was the factor responsible for the heterogeneity of SCM- feather keratin.

(ii) <u>Multiple non-allelic genes</u>

β, γ and δ genes of globin all exist within the one organism. Each of the three protein molecules bears a close resemblance to the other two with respect to primary structure and both γ and δ chains are believed to have originated by duplication of the β chain gene and subsequent mutational divergence (Dayhoff, 1969). γ and δ chains differ from β chains and from one another by multiple amino acid substitutions at

different sites. Similarly the two C genes for λ (light chain) immunoglobulin molecules occur at closely linked but distinct genetic loci and they determine the synthesis of protein products differing at only two sites within the amino acid sequence (Ein, 1968, Hess et al., 1971, Gally and Edelman, 1972). The high sulphur proteins (α-keratin) of mammalian hair and wool prepared from a single animal consist of an extremely heterogeneous family of homologous protein molecules differing from one another by amino acid interchanges at multiple sites within the sequence (Lindley and Elleman, 1972., Swart, 1973). The observed heterogeneity cannot possibly be due to polymorphism and must therefore represent the expression of multiple genes for α-keratin within a single diploid genome.

The existence of sequence heterogeneity in SCM-feather keratin is indicative of the existence of multiple non-allelic genes in two respects. The multiple amino acid substitutions that have been identified within the sequence of at least some genetic variants more closely resemble the pattern of substitutions within duplicated or reiterated and mutationaly divergent genes. Most conclusive however is the fact that the same electrophoretic complexity is obtained whether SCM- feather keratin is obtained from a single animal or a number of animals and this eliminates genetic polymorphism as the fundamental cause for the existence of electrophoretic and primary structure variants.

(iv) The Relationship of Structure to Function of Feather Keratin

(a) Salient features of the "general" primary structure

The amino acid sequences of SCM- feather keratin (Fig. 6.12) indicate that the cysteine residues of native keratin molecules are confined predominantly to amino-terminal and carboxyl-terminal regions of the "general" primary structure. Most or all of the arginine residues of feather keratin also exist in amino or carboxyl-terminal regions. On the other hand, isoleucine and valine are found almost exclusively in peptide T3 which constitutes a hydrophobic interior devoid of arginine and probably also of cysteine residues. Different arrangements of the tryptic peptides of SCM- feather keratin (Fig. 6.12) into a number of permutations of the primary structure indicate that at least six and at most nine SCM- cysteine residues occur in any one sequence but the primary structure data are still compatible with the existence of eight SCM- cysteine residues in every molecule. These general features of the sequence of SCM- feather keratin molecules are in accord with the findings of O'Donnell (1973b) who made similar observations on the primary structure of a purified constituent of SCM- emu rachis. Furthermore the presence of SCM- cysteine as the carboxyl-terminal residue of most or all variants of SCM- feather keratin agrees with the presence of SCM- cysteine residues at carboxyl-termini of all keratin molecules on which sequence data have been reported (Elleman and Dopheide, 1972., Fraser et al.,

1972., O'Donnell, 1973b, Swart, 1973). The presence of carboxyl-terminal SCM- cysteine (or cystine in native keratin chains) may play a fundamental role in the maintainence of the suprastructure of keratin fibrils in feather cells.

The amino acid sequences of the tryptic peptides of SCM- feather keratin although clearly homologous to corresponding regions in emu rachis show no obvious homology to the sequences of high-sulphur or low-sulphur proteins of mammalian hair presently available (Fraser et al., 1972., Swart, 1973). In particular no evidence of repeating units within the primary structure of SCM- feather keratin such as have been demonstrated in the high-sulphur proteins of wool (Elleman and Dopheide, 1972., Swart, 1973) was found.

(b) Conservation of primary structure

The amino-terminal regions of SCM- feather keratin molecules are shorter than the corresponding region of emu rachis keratin by four residues; there appears to be the same number of residues in the primary structures of all chick keratin variants and this is in agreement with the molecular weight homogeneity of all set K proteins (Chapter 4).

Fig. 6.12 demonstrates that no large region of the SCM- feather keratin sequence examined appears to be devoid of amino acid replacements. The complete sequence analysis of T3 was not attempted and this region of primary structure may be more stringently conserved than the regions represented by tryptic peptides T1, T2, T4 and T5. The gly, ser interchange

at position 92 indicates that T3 is at least not completely devoid of amino acid substitutions.

The substitution of ^{2}cys and ^{95}cys by other amino acids in peptides T1d and T4b respectively suggests that absolute conservation of cysteine residues within the sequence is not a fundamental prerequisite for the function of chick keratin. On the other hand, it is conceivable that the absence of cysteine at a particular position in the sequence may be compensated for by the addition of the same residue at another site and thus a constant number of cysteine residues (presumably eight) could be maintained in every chick keratin chain.

Two final remarks concerning the native of the sequence substitutions within the SCM- feather keratin sequence should be made.

Firstly, only one of the eleven amino acid interchanges which have been detected within the "general" sequence of SCM- feather keratin (Fig. 6.12) (excluding peptides T1e and T1f from consideration) could conceivably have arisen post synthetically. T1b and T1a can be interchanged by the diamidation of ^{19}asx, ^{22}asx or ^{23}glx in T1b and such an event could be responsible for the existence of T1a.

Secondly, the eleven amino acid interchanges detected almost certainly underestimate the true level of sequence heterogeneity since minor peptides whose amount seemed insignificant compared to that of their more prominent counterparts

were disregarded. Furthermore, the major portion of the T3 sequence, whose elucidation was not attempted may contain additional substitutions whose detection would increase the sequence heterogeneity even further.

CHAPTER 7

STUDIES ON PURIFIED FRACTIONS OF SCM-- FEATHER KERATIN

INTRODUCTION

It was established in Chapter 4 that SCM- feather keratin could be resolved into a large number of electrophoretic variants of very similar size and amino acid composition but differring in overall charge. Chapters 5 and 6 dealt with the amino acid sequences of tryptic peptides derived from unfractionated SCM- feather keratin and eleven sites were identified within the amino acid sequence at which amino acid substitutions occurred. The purpose of the experiments in the present chapter was to determine
(a) Whether individual electrophoretic variants represented homogeneous keratin chains encoded by only one gene or
(b) Whether individual electrophoretic variants were a mixture of two or more keratin chains whose sequences were too similar to allow their separation by gel electrophoresis. The question is an important one, since the estimate of the number of keratin chains in down feather (Chapter 4) depends upon the premise that polyacrylamide gel electrophoresis at pH 2.7 was capable of fully resolving a protein sample which was homogeneous at pH 9.5 into sub-components, each being a unique keratin chain representing a single amino acid sequence. Failure to completely separate all such sub-components would results in an underestimate of the overall complexity of SCM- feather keratin.

In Chapter 6 it was established that individual SCM- down feather keratin chains will produce no more than

four tryptic peptides containing SCM- cysteine and this is therefore a fundamental requirement ofapurified SCM- keratin sample composed entirely of a single amino acid sequence.

The final proof that each of the nineteen electrophoretic variants of SCM- feather keratin (Chapter 4) represents the expression of a unique gene(s) depends upon the demonstration that their respective amino acid sequences differ. The other important aspect of the work of this Chapter is an attempted isolation of peptides representing unique sequences from three purified electrophoretic variants.

METHODS

(i) HVPE at pH 2.7 and pH 9.5, autoradiography and amino acid analysis were carried out as described in Chapters 5 and 6.

(ii) Polyacrylamide gel electrophoresis at pH 9.5 and pH 2.7 were carried out as described in Chapter 3.

(iii) Phosphocellulose chromatography

Phosphocellulose chromatography was carried out essentially according to Darskus (1972). D4, (60mg) was dissolved in 0.2M citric acid, 5M urea at a concentration of 100mg/ml and loaded onto a phosphocellulose column (1.5x90cm) equilibrated in that buffer. Salt gradient elution was carried out with I.L of a linear gradient (500 ml of starting

buffer, 500 ml of starting buffer containing 0.2M NaCl). 5.2ml fractionswere collected and the absorbance of fractions analysed as described for DEAE-cellulose chromatography (Chapter 4). Appropriate tubes were pooled, dialysed and freeze dried for gel electrophoresis and tryptic digestion.

RESULTS

(i) Preparation and electrophoretic properties of three highly fractionated samples of SCM- feather keratin

A preparation of $[^{14}C]$- SCM- feather keratin (of Specific Activity, 4.74×10^5 cpm per mg) was used for the experiments of this Chapter. The high specific activity enabled sensitive detection of the small amounts of proteins and peptides used in this work.

DEAE-cellulose chromatography of a sample (approx. 500 mg) of $[^{14}C]$- SCM- feather keratin was carried out as described in Chapter 4 (Fig. 4.1) and fractions containing band β_4 (pH 9.5 gel) were pooled, dialysed and freeze dried as described (Chapter 4). Care was taken not to contaminate the preparation of band β_4 with minor amounts of the adjacent bands, β_3 or β_5. The pooled fraction from DEAE-cellulose (D4) was examined by gel electrophoresis at pH 9.5 (Fig. 7.1a, gel (i)) and at pH 2.7 (Fig. 7.1a gel (ii)). The former gel as expected showed only one band, β_4, with no traces of contaminants; the latter gel however showed prominent amounts of the three pH 2.7 bands A, B and C and a minor amount of

FIGURE 7.1 Polyacrylamide gel electrophoresis of D4 and sub fractions

(a) Polyacrylamide gel electrophoresis of D4 prepared by DEAE cellulose chromatography.

(i) pH 9.5, 50 μg

(ii) pH 2.7, 50 μg

(b) pH 9.5 gels of fractions D4P1(i), D4P2(ii) and D4P3(iii), (50 μg each).

(c) pH 2.7 gels of fractions D4P1(i), D4P2(ii) and D4P3(iii), (50 μg each).

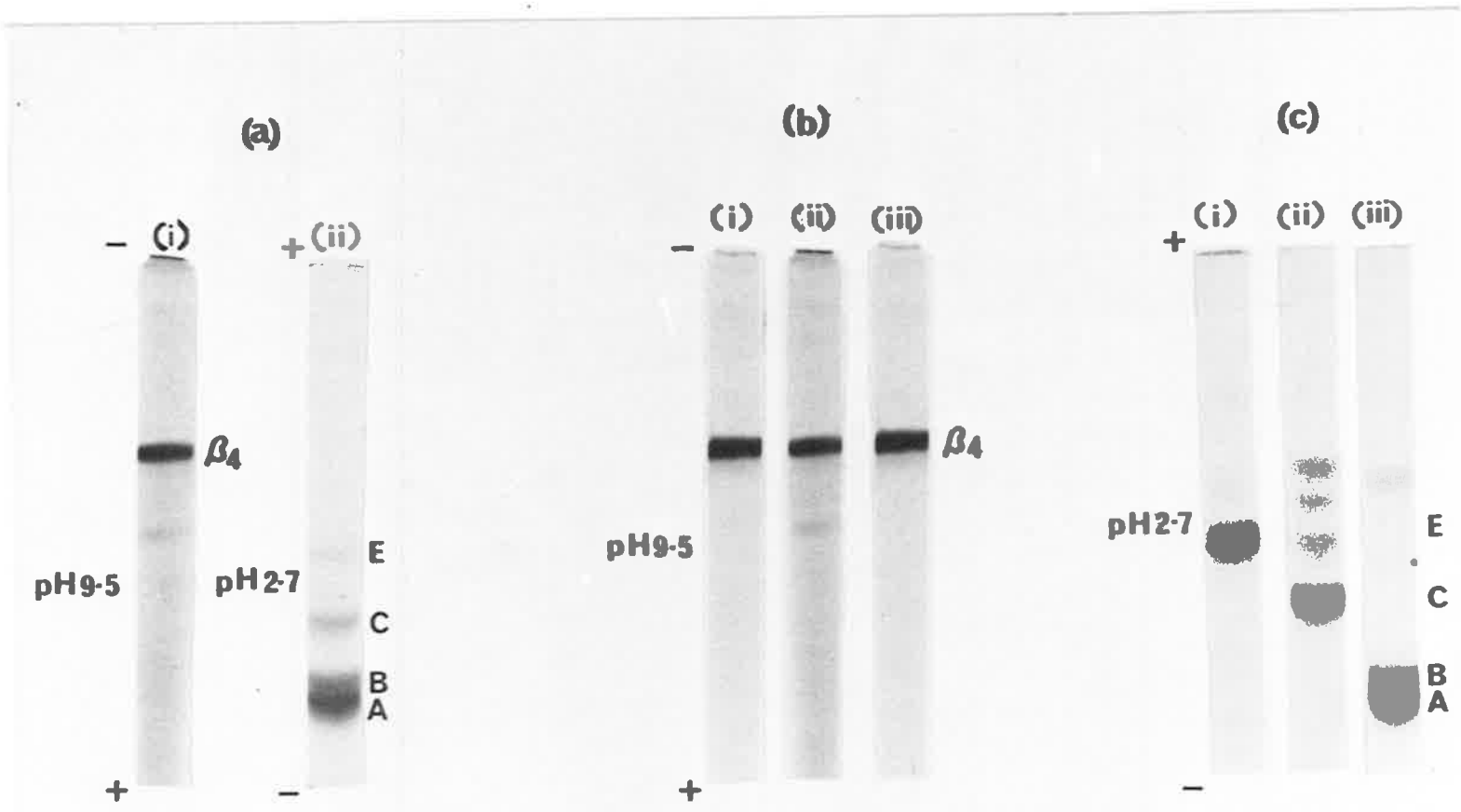
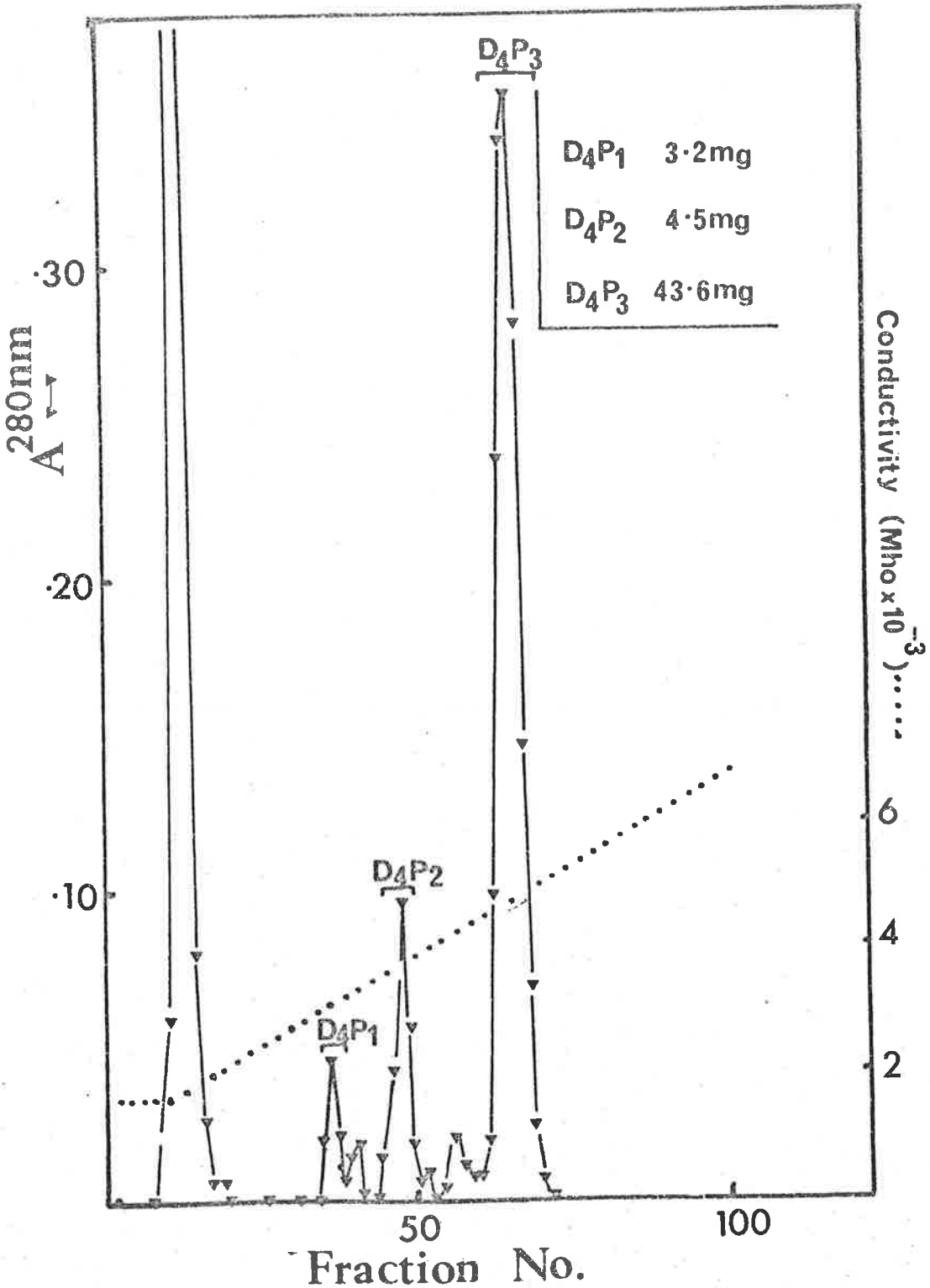

(a)
(i)
(ii)
β4
pH9-5
pH2-7
E
C
B
A
(b)
(i)
(ii)
(iii)
β4
pH9-5
(c)
(i)
(ii)
(iii)
pH2-7
E
C
B
A

FIGURE 7.2 Phosphocellulose chromatography of D4

D4 (60mg) was applied to phosphocellulose and elution carried out with a linear NaCl gradient (Methods). Absorbance at 280 nm was used to monitor the elution of protein. Fractions were pooled as shown.

Insert: amount of protein in each peak determined by measurement of radioactivity.


D4P3
D4P1 3·2mg
D4P2 4·5mg
D4P3 43·6mg
·30
·20
·10
A 280nm
D4P2
D4P1
50
100
Fraction No.
Conductivity (Mho x 10-3)
6
4
2

band E. These results accord with those of Chapter 4 in which pH 2.7 bands A, B, C and E were shown to constitute β_4.

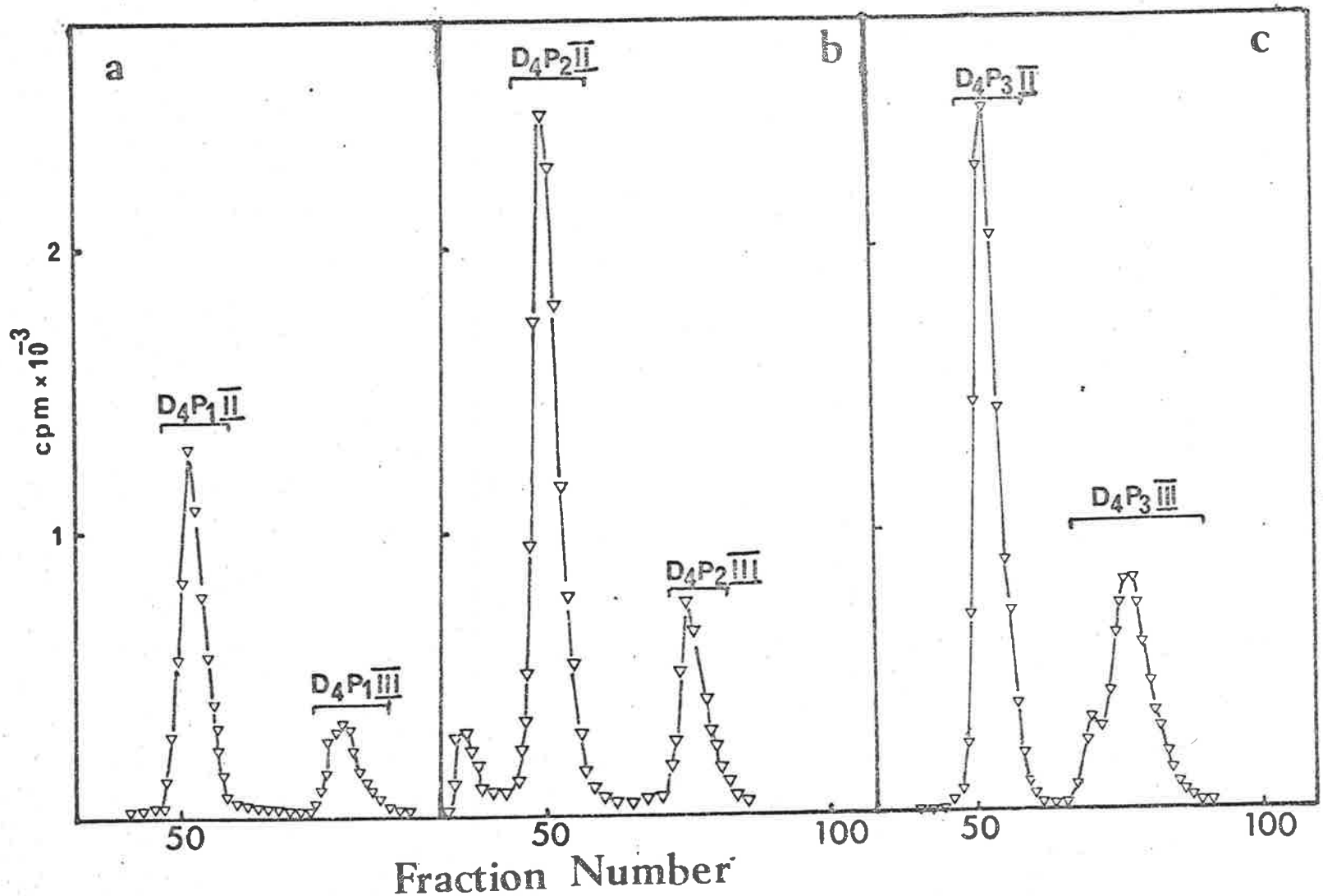
In an attempt to purify the four pH 2.7 subcomponents (bands A, B, C and E) of D4 for peptide mapping, phosphocellulose chromatography of D4 (60mg) was carried out as described in Methods (Fig. 7.2). The three peaks D4P1, D4P2 and D4P3 were pooled as shown, dialysed and lyophilised. The large absorbance peak which emerged first from the phosphocellulose column contained neither radioactivity nor protein as determined by the Lowry reaction and was subsequently discarded. Examination of the three phosphocellulose peaks D4P1, D4P2 and D4P3 by gel electrophoresis at pH 9.5 (Fig. 7.1b gels (i) (ii) and (iii)) revealed the expected single band (β_4) in each case. Gel electrophoresis at pH 2.7 of the three fractions (Fig. 7.1c, gels (i) (ii) and (iii)) showed that D4P1 contained only band E, D4P2 contained band C in a purity greater than 90% whereas D4P3 contained bands A and B in the same proportions as they occurred in D4. Each of the three fractions of D4 was entirely free of contamination by bands present in the other two phosphocellulose fractions as determined by pH 2.7 gel electrophoresis (Fig. 7.1c) and was used for peptide mapping without further purification.

(ii) Peptide mapping of D4P1 D4P2 and D4P3

2.1mg (1.0×10^6 cpm) each of D4P1, D4P2 and D4P3 were digested with trypsin (see Methods, Chapter 5) and the digests directly applied to Sephadex G-50 (Figure 7.3). Due

FIGURE 7.3 Sephadex G-50 chromatography of the tryptic digests of D4P1, D4P2 and D4P3

The tryptic digests of D4P1(a), D4P2(b) and D4P3(c) were applied to a Sephadex G-50 operated in PMA. 50μL aliquots were withdrawn from 3.0 ml fractions for radioactivity determination. Tubes were pooled as indicated.


a
b
c
$D_4P_1\overline{II}$
$D_4P_1\overline{III}$
$D_4P_2\overline{II}$
$D_4P_2\overline{III}$
$D_4P_3\overline{II}$
$D_4P_3\overline{III}$
cpm x 10^{-3}
2
1
50
50
100
50
100
Fraction Number

to the small amounts of peptide material used for this experiment, the elution of only radioactive peptides was monitored. All three radioactivity profiles showed a striking similarity to one another both in the positions of elution of the two major radioactive peaks (II and III) and in the relative distribution of radioactivity between them. All three profiles also resembled the radioactivity profile obtained from Sephadex G-50 chromatography of the tryptic digest of unfractionated SCM- feather keratin (Fig. 5.1). The profile of the D4P2 tryptic digest (Fig. 7.3b) contained one additional minor peak of radioactivity of molecular weight approximately 8900 (x). The presence of this peak may have been related to the presence in D4P2 of three minor impurities detected by gel electrophoresis at pH 2.7 (Fig. 7.1c gel (ii)). It is unlikely that peak (x) corresponded to a tryptic peptide of SCM- feather keratin because of its high molecular weight and it was not investigated further.

The similarity in elution position and relative amount of peaks II and III respectively of the three purified fractions (Fig. 7.3a, b, c) to peaks 2 and 3 of unfractionated SCM- feather keratin (Fig. 5.1) suggested that peak II in each case represented the amino-terminal tryptic peptide (T1) of D4P1 D4P2 and D4P3 and peak III contained peptides corresponding to T2 T4 and T5.

(a) <u>The amino-terminal peptides of D4P1 D4P2 and D4P3</u>

Peak II of each tryptic digest (Fig. 7.3), pooled as shown was subjected to HVPE at pH 6.5 and peptides detected by autoradiography (Fig. 7.4a). Each of the radioactive

FIGURE 7.4 Purification of T1 from peak II of D4P1, D4P2 and D4P3

The peptides of peak II of the Sephadex G-50 chromatograms of Fig. 7.3 were subjected to HVPE at pH 6.5 (3000 V, 1 h) and autoradiography carried out (a). Each peptide (T1) from (a) was cut out and HVPE at pH 2.7 carried out (3000 V, 3 h) followed by autoradiography (b).


pH 6·5
T1
D4P1
D4P2
D4P3
-
+
D4P1
D4P2
D4P3
pH 2·7

peptides(D4P1,T1,D4P2.T1,D4P3.T1)was cut out and HVPE at pH 2.7 was carried out followed by autoradiography (Fig. 7.4b). The relative mobilities of peptides D4P2 T1 and D4P3 T1 at pH 6.5 and pH 2.7 were very similar or identical to those properties of peptide T1b whose primary structure was established in the previous Chapter (Fig. 6.12). D4P1 T1 on the other hand, had a lower mobility at pH 2.7 than D4P2 T1 and D4P3 T1 and this suggested its possible identity to peptide T1c whose structure was also determined in the previous Chapter. To confirm the structure of these tryptic peptides, each was digested with chymotrypsin and the chymotryptic peptides purified by HVPE at pH 6.5 and then at pH 2.7. The autoradiograms of the pH 6.5 and pH 2.7 electrophoretograms are shown in Figures 7.5 a and b respectively. Digestion of the tryptic peptides with chymotrypsin produced three radioactive peptides in each case: each peptide was named to indicate the protein sample from which it was derived and its relative position within the amino-terminal sequence. The amino acid composition and relative mobilities of each of the nine purified peptides of Figure 7.5 are shown in Table 7.1. The relative mobilities and amino acid compositions of the three chymotryptic fragments of each of the amino-terminal tryptic peptides D4P2 T1 and D4P3 T1 are not significantly different to the same properties of the chymotryptic fragments of T1b (Table 6.4). Moreover, the identity in electrophoretic mobility at pH 6.5 and pH 2.7 of the intact peptides D4P2 T1 and D4P3 T1 to T1b further indicates the identity of these three peptides.

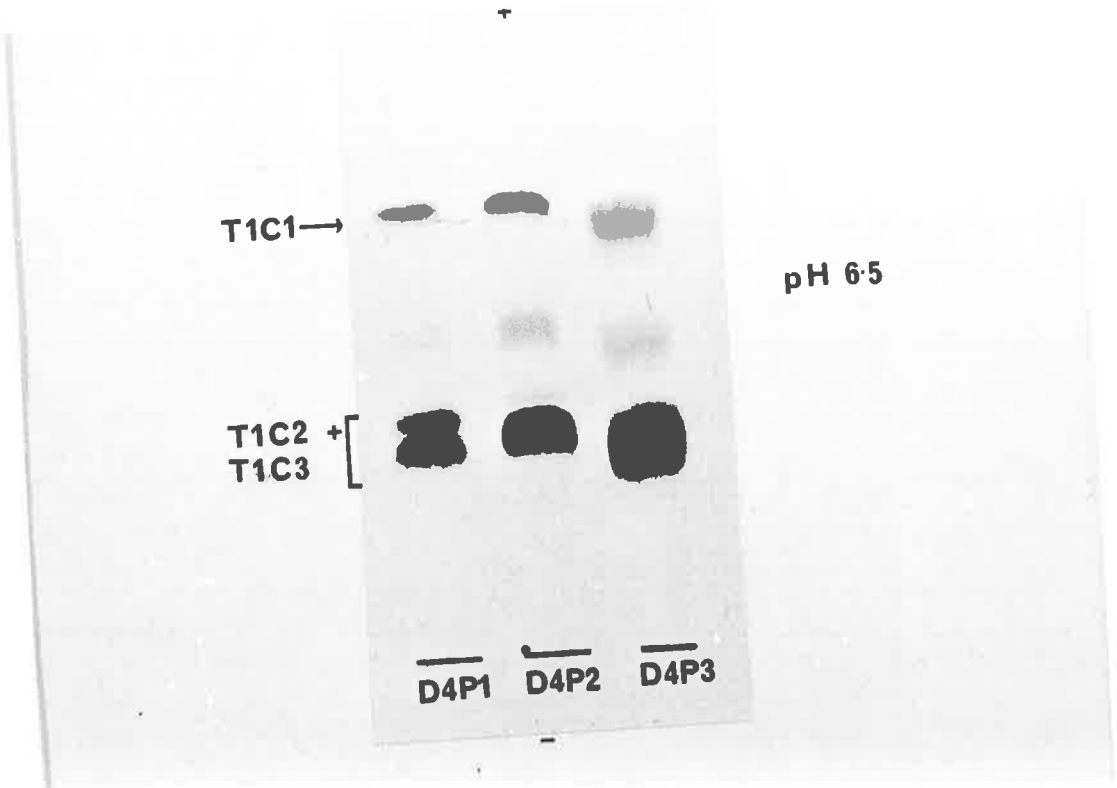
FIGURE 7.5 Purification of the chymotryptic peptides of T1 derived from D4P1, D4P2 and D4P3

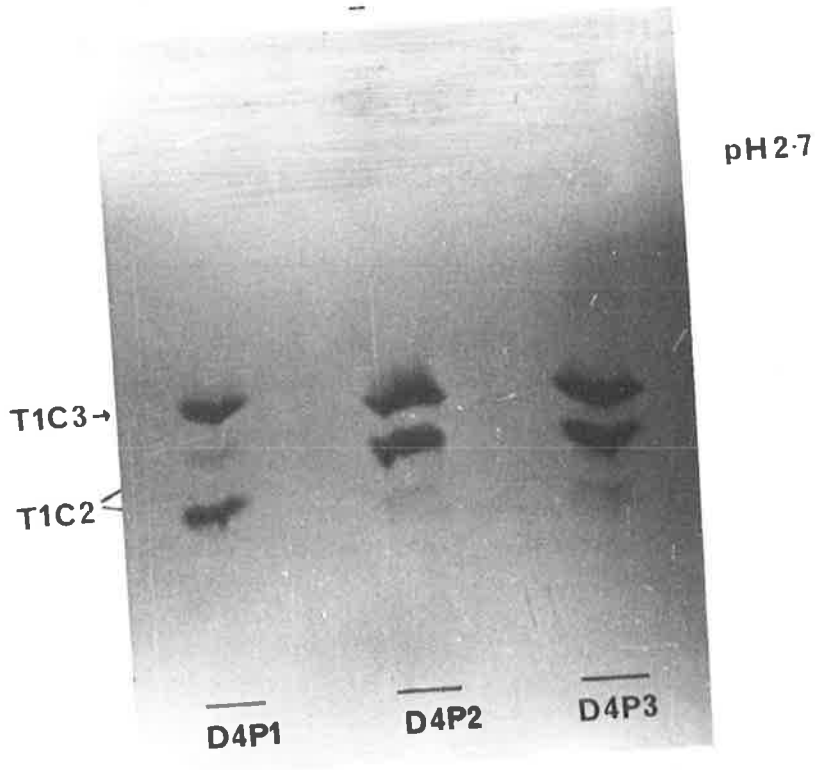
T1 from D4P1-3 were digested with chymotrypsin (20 μg) in 200 μL NEMA pH 8.3 and the digests subjected to HVPE at pH 6.5 (2000 V 1 h) (a).

Peptides T1C1 (D4P1-3) were cut out after detection by autoradiography, eluted from paper and hydrolysed for amino acid analysis.

T1C2 and T1C3 (D4P1-3) were cut out and completely separated by HVPE at pH 2.7 (3000 V, 1.5 h).

They were detected by autoradiography (b).


T1C1
T1C2 +
T1C3
pH 6·5
D4P1
D4P2
D4P3


pH 2·7
T1C3
T1C2
D4P1
D4P2
D4P3

The amino acid composition of the chymotryptic fragments of T4P1 T1 are very similar to the corresponding fragments of T1c. The only significant differences are the anomalously high content of glycine in peptide D4P1 T1C2 and a rather low content of SCM- cysteine as compared to T1c C2. The over-estimate of glycine is to some degree common to all the estimations of Table 7.1 and is probably due to the elution of small amounts of free glycine from the paper during peptide recovery. The value of glycine (1.71 residues) is consistent with one glycine residue but not with two since in no case was the glycine content of a peptide eluted from paper underestimated. The low content of SCM- cysteine is due to the partial destruction of this residue during acid hydrolysis since measurement of radioactivity in peptide D4P1 T1C2 before hydrolysis indicated 1.9 residues per mole. Furthermore the similarity in electrophoretic mobilities between D4P1 T1 and its chymotryptic fragments to T1c and its corresponding fragments indicate that the two tryptic peptides are identical.

(b) The internal and carboxyl-terminal peptides of D4P1, D4P2 and D4P3

Peak III peptides from G-50 chromatography of tryptic digests of D4P1, D4P2 and D4P3 (Fig. 7.3) were separated by chromatography on Dowex-50 in an exactly analagous manner to that used for the small (peak 3) peptides derived from unfractionated SCM- feather keratin (Fig. 6.9). Figure 7.6 shows the results of this experiment. The small tryptic

TABLE 7.1 Amino acid compositions and relative mobilities of the chymotryptic fragments of peptides D4P1-3 T1

The scale expander was used to facilitate the more accurate quantitation of residues in hydrolysates.

a : absolute μmoles

b : residue ratio

n.d. : loss of sample prevented arginine determination on the short column of the amino acid analyser.

TABLE 7.1

	D4P1T1C1		D4P2T1C1		D4P3T1C1		D4P1T1C2	
	a	b	a	b	a	b	a	b
SCM-cys	0.014	0.78(1)	0.026	0.72(1)	0.022	0.85(1)	0.018	1.28(2)
asp							0.014	1.00(1)
thr							0.016	1.14(1)
ser	0.018	1.00(1)	0.036	1.00(1)	0.032	1.23(1)	0.034	2.42(2)
glu							0.004	0.29(0)
pro							0.032	2.29(2)
gly							0.024	1.71(1)
ala							0.006	0.44(0)
val							-	-
met							-	
ile							-	
leu							0.028	2.00(2)
tyr								
phe	0.018	1.00(1)	0.036	1.00(1)	0.026	1.00(1)		
lys								
his								
arg							0.00	0 (0)

Mrel	0.89	0.89	0.86	0.44
Charge	-2	-2	-2	-2

TABLE 7.1 (Contd.)

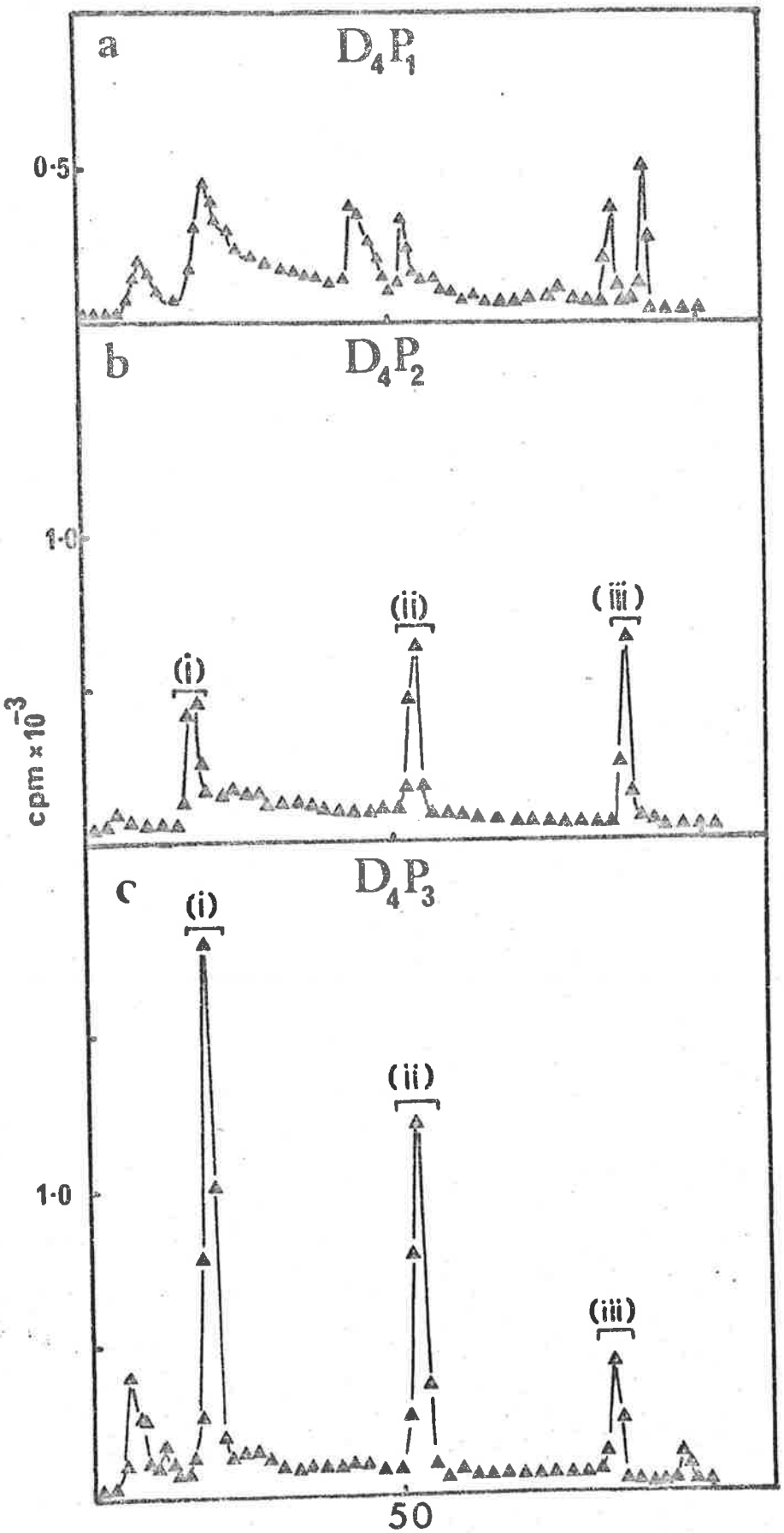
	D4P2T1C2		D4P3T1C2		D4P1T1C3	
	a	b	a	b	a	b
SCM-cys	0.064	1.60(2)	0.068	1.89(2)	0.022	1.57(2)
asp	0.040	1.00(1)	0.044	1.22(1)	0.026	1.86(2)
thr	0.034	0.85(1)	0.034	0.91(1)	0.006	0.43(0)
ser	0.010	0.25(0)	0.008	0.22(0)	0.018	1.29(1)
glu	0.008	0.20(0)	0.010	0.28(0)	0.018	1.29(1)
pro	0.102	2.55(3)	0.100	2.78(3)	0.010	0.71(1)
gly	0.048	1.20(1)	0.044	1.22(1)	0.006	0.43(0)
ala	0.008	0.20(0)	0.008	0.22(0)	0.018	1.29(1)
val	0.004	0.10(0)	0.004	0.12(0)	0.014	1.00(1)
met	-		-		-	
ile	-		-		-	
leu	0.080	2.00(2)	0.074	2.00(2)	0.004	0.30(0)
tyr	-					
phe	-					
lys	-				-	
his	-				-	
arg	0.034	0.85(1)	n d	(1)	0.014	1.00(1)
M rel	0.47		0.46		0.48	
Charge	-2		-2		-2	

TABLE 7.1 (Contd.)

	D4P2T1C3		D4P3T1C3	
	a	b	a	b
SCM-cys	0.064	1.88(2)	0.056	1.75(2)
asp	0.072	2.12(2)	0.062	1.94(2)
thr	-		0.008	0.25(0)
ser	0.030	0.88(1)	0.036	1.13(1)
glu	0.032	0.94(1)	0.032	1.00(1)
pro	0.024	0.71(1)	0.032	1.00(1)
gly	0.010	0.29(0)	0.014	0.44(0)
ala	0.034	1.00(1)	0.032	1.00(1)
val	0.032	0.94(1)	0.030	0.94(1)
met	-			
ile	-			
leu	-			
tyr				
phe				
lys				
his				
arg	0.040	1.18(1)	n d	(1)
M rel		0.47		0.46
Charge		-2		-2

FIGURE 7.6 Separation of internal and carboxyl-terminal tryptic peptides of D4P1-3

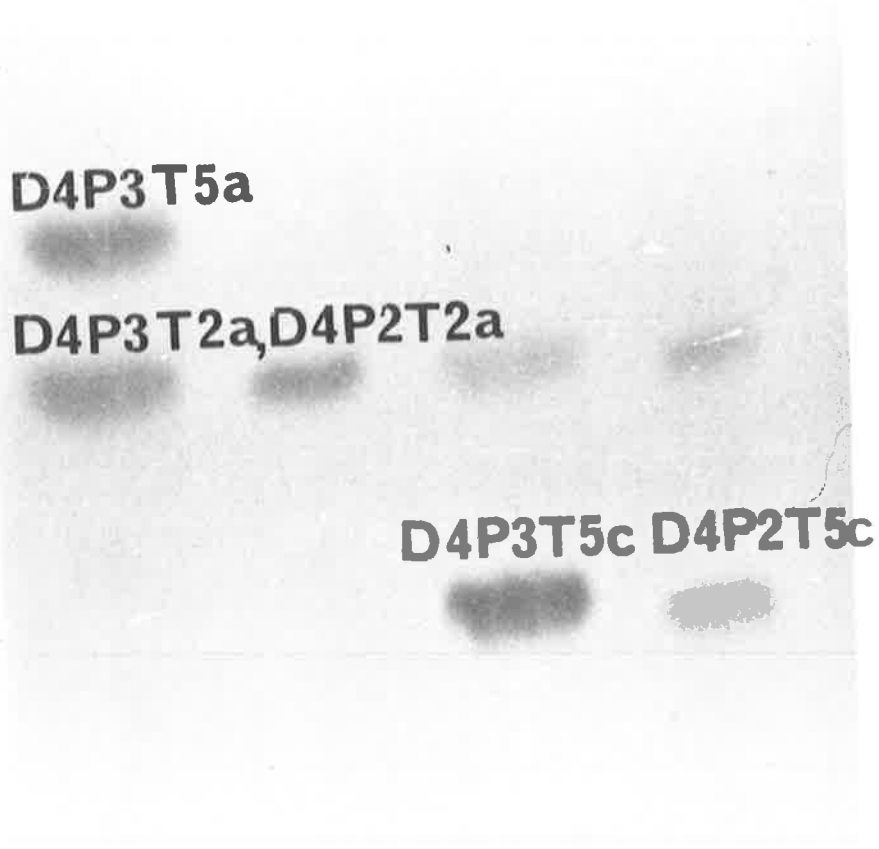
Peak III peptides of D4P1-3 were subjected to chromatography on Dowex-50 as described in Chapter 6. The conditions were identical to those of Fig. 6.9. 100 μL was withdrawn from the D4P2 (b) and D4P3 (c) fractions : μ200 L was withdrawn from D4P1 fractions (a) for determination of radioactivity.


a
D_4P_1
0·5
b
D_4P_2
1·0
(i)
(ii)
(iii)
cpm × 10^{-3}
c
D_4P_3
(i)
(ii)
(iii)
1·0
50
Fraction No

peptides of D4P1 (Fig. 7.6 (a)) were unexpectedly numerous, some six or seven radioactive peaks eluting from Dowex-50. The lack of material prevented further examination of these peptides. The number of radioactive tryptic peptides, seven in all, clearly established however that D4P1 was not a single polypetide chain (see Discussion). Only three prominent peaks were eluted for fractions D4P2 and D4P3 and two independent chromatography experiments with the small peptides of D4P2 and D4P3 established that the elution position of each peak (i) (ii) and (iii) in the two samples was extremely reproducible (± 1 tube). Tubes were pooled as shown in Figure 7.6 and peptides purified by HVPE at pH 6.5. Figure 7.7 shows the autoradiogram of this experiment. Due to loss of the peptides of peak (iii) of D4P2 and D4P3 before electrophoresis, only the data on peaks (i) and (ii) are available. The amino acid compositions and relative mobilities of the peptides eluted from the electrophoretogram (Fig. 7.7) are presented in Table 7.2. D4P3 T5a is identical in electrophoretic mobility to T5a (Table 6) and differs in amino acid composition only in the relative amount of SCM- cysteine which is half the expected value. However, determination of radioactivity indicated two residues of SCM- cysteine in D4P3 T5a. The amino acid compositions of D4P2 T2a and D4P3 T2a were not significantly different from that of T2 (Fig. 6.11, Table 6.7) but whereas T2 was neutral at pH 6.5, D4P3 T2a and D4P2 T2a carried charges of -2 at this pH (Table 7.2). The possible structure of these two peptides is considered in Discussion.

FIGURE 7.7 Purification of peaks (i) and (ii) of D4P2 and D4P3

HVPE at pH 6.5 (2000 V, 1h) of the peptides of peaks (i) and (ii) of D4P2 and D4P3 (Fig. 7.6). Peptides were detected by autoradiography.



(i) (i) (ii) (ii)

D4P3 D4P2 D4P3 D4P2

TABLE 7.2 Amino acid compositions of small tryptic peptides of D4P2 and D4P3

The scale expander was used to facilitate accurate quantitation of residues in hydrolysates .

a : absolute μmoles

b : residue ratio

TABLE 7.2 Amino acid compositions of small tryptic peptides of D4P2 and D4P3

	D4P3T5a		D4P2T2a		D4P3T2a	
	a	b	a	b	a	b
SCM-cys	0.103	1.00(2)	0.010	0.71(1)	0.010	0.71(1)
asp			0.016	1.14(1)	0.020	1.41(1)
thr						
ser	0.005	0.16(0)	0.018	1.29(1)	0.020	1.41(1)
glu			0.026	1.86(2)	0.024	1.71(2)
pro	0.0103	1.00(1)				
gly	0.0034	0.33(0)	0.006	0.43(0)	0.008	0.51(0)
ala						
val						
met						
ile						
leu	0.0103	1.00(1)				
tyr						
phe						
lys						
his						
arg		0	0.014	1.00(1)	0.014	1.00(1)
Mrel	0.76		0.62		0.62	
Charge	-2		-2		-2	

TABLE 7.2 (Contd.)

	D4P2T5b		D4P3T5b	
	a	b	a	b
SCM-cys	0.034	0.77(1)	0.030	0.79(1)
asp			0.002	0.05(0)
thr				
ser			0.004	0.15(0)
glu			0.004	0.15(0)
pro	0.046	1.05(1)	0.042	1.11(1)
gly			0.006	0.16(0)
ala			0.002	0.05(0)
val				
met				
ile				
leu	0.044	1.00(1)	0.042	1.11(1)
tyr	0.034	0.77(1)	0.038	1.00(1)
phe				
lys				
his				
arg		0		0
Mrel		0.40		0.40
Charge		-1		-1

The amino acid compositions and relative mobilities of D4P2 T5c and D4P3 T5c were identical to those properties of T5c (Table 6.7). The minor amounts of a faster moving tryptic peptide seen in peak (ii) of both the small peptides of D4P2 and D4P3 (Fig. 7.7) were insufficient for amino acid analyses and these peptides were not investigated further. Although no data was obtained on the peptide material of peak (iii) from D4P2 and D4P3 (Fig. 7.7), the elution position of this material from Dowex-50 coincides with that of T2 (Fig. 6.9).

DISCUSSION

(i) The amino-terminal sequences of D4P1 D4P2 and D4P3

Each of the three highly-fractionated samples of SCM- feather keratin D4P1, D4P2 and D4P3 only produced a single amino-terminal peptide on digestion with trypsin. D4P2 T1 and D4P3 T1 produced chymotryptic fragments indistinguishable from those of T1b with respect to amino acid composition and electrophoretic mobility and the sequences of peptides D4P2 T1, D4P3 T1 and T1b are probably identical. Similarly, the amino acid sequence of D4P1 T1 is probably identical to that of T1c. No trace of D4P1 T1 was found in tryptic digests of D4P2 or D4P3 and therefore the genes which encode the latter two protein samples are distinct from those which encode D4P1.

(ii) The internal and carboxyl-terminal tryptic peptides of D4P1, D4P2 and D4P3

The Dowex-50 chromatogram of the small, presumably

internal, peptides of D4P1 (Fig. 7.6a) indicates that there are at least six different peptides in this fraction. Since at most three such peptides (T2, T4 and T5) are expected for any sample of SCM- feather keratin containing a single polypeptide chain, D4P1 was clearly heterogeneous and despite the finding of a single amino-terminal tryptic peptide (D4P1 T1), D4P1 contained two or more distinct polypeptide chains.

For D4P2 on the other hand, there were only three prominent peaks corresponding to internal and carboxyl-terminal tryptic peptides (Fig. 7.6). Peak (i) contained D4P2.T2a whose amino acid composition suggested its identity to peptide T2 (Chapter 6). However its possession of a charge of -2 at pH 6.5 suggested that two of the 3 amides of T2 were carboxylates in T2a. It is conceivable that deamidation of two of the three amidated amino acids of T2 took place in the course of the isolation procedures used; storage of T2 was previously shown (Chapter 6) to result in cyclisation with concomitant deamidation of amino-terminal glutamine. Peak (iii) (Fig. 7.6b) probably corresponds to intact T2 as judged by its elution position from Dowex-50 (compare peak (iii) of Fig. 7.6b with peak 4 of Fig. 6.9) but its loss during isolation prevented further characterisation. The presence, however, of a peptide homologous in sequence to T2 and T2a but carrying a charge of -1 due to deamidation of only one of the amidated amino acids was not detected. The presence of such an intermediate would

reinforce the notion that T2a arose from T2 by two successive deamidation steps during isolation of peptides. Its absence suggests that T2a is not derived from T2 and that the protein sample D4P2 is composed of at least two keratin chains, one containing the sequence corresponding to T2 and the other containing the homologous sequence (T2a).

Peak (ii) of the D4P2 peptides contained one major peptide whose elution position from Dowex-50, amino acid composition and relative mobility were essentially identical to T5c, a carboxyl-terminal peptide identified in the previous chapter.

D4P2 on the basis of an examination of its tryptic peptides consists of at least two polypeptide chains. Despite the detection of a single amino-terminal peptide (D4P2 T1) and a single carboxyl-terminal peptide (D4P2 T5c) the properties of peak (iii) material (Fig. 7.6b) and D4P2 T2a suggest that they are homologous peptides derived from the same region of two different polypeptide chains.

The only significant difference between the small peptides of D4P2 and D4P3 was the occurrence in D4P3 of the additional peptide D4P3 T5a whose amino acid composition and relative mobility established its identity to T5a. The detection of two carboxyl-terminal peptides, T5a and T5c, in the tryptic digest of D4P3 and of D4P3 T2a and peak (iii) in Figure 7.6c (which is presumably T2), indicate that D4P3 consists of at least two polypeptide chains. Comparison of the

amino acid sequences of T5a and T5c (Chapter 6) indicates that the different polypeptide chains of D4P3 differ by at least one amino acid substitution, tyr^{100} in one for cys^{100} in the other, which cannot conceivably be due to partial deamidation of asparagine or glutamine residues during peptide isolation. The two bands, A and B, on pH 2.7 gels of D4P3 (Fig. 7.1c gel iii) indicated at least two polypeptide chains and examination of the tryptic peptides therefore established no additional heterogeneity.

In summary D4P1 although homogeneous when examined by gel electrophoresis at pH 9.5 and at pH 2.7 consisted of at least two different keratin chains as determined by the number of peptides produced by tryptic digestion. The isolation of both T2a and (presumably) T2 from tryptic digests of D4P2 suggested at least two polypeptide chains but the possible origin of T2a from T2 by deamidation was not eliminated - whether D4P2 is composed of one or more polypeptide chains is not absolutely clear. Examination of the tryptic peptides of D4P3 indicated at least two polypeptide chains and this accords with the two bands, A and B, observed on pH 2.7 gels. That two of the purified keratin samples, D4P1 and D4P2 are more heterogeneous than was suggested by gel electrophoresis indicates that the estimate of the number of keratin chains in down feather (Chapter 4) is an underestimate.

(iii) The Genetic basis for Heterogeneity

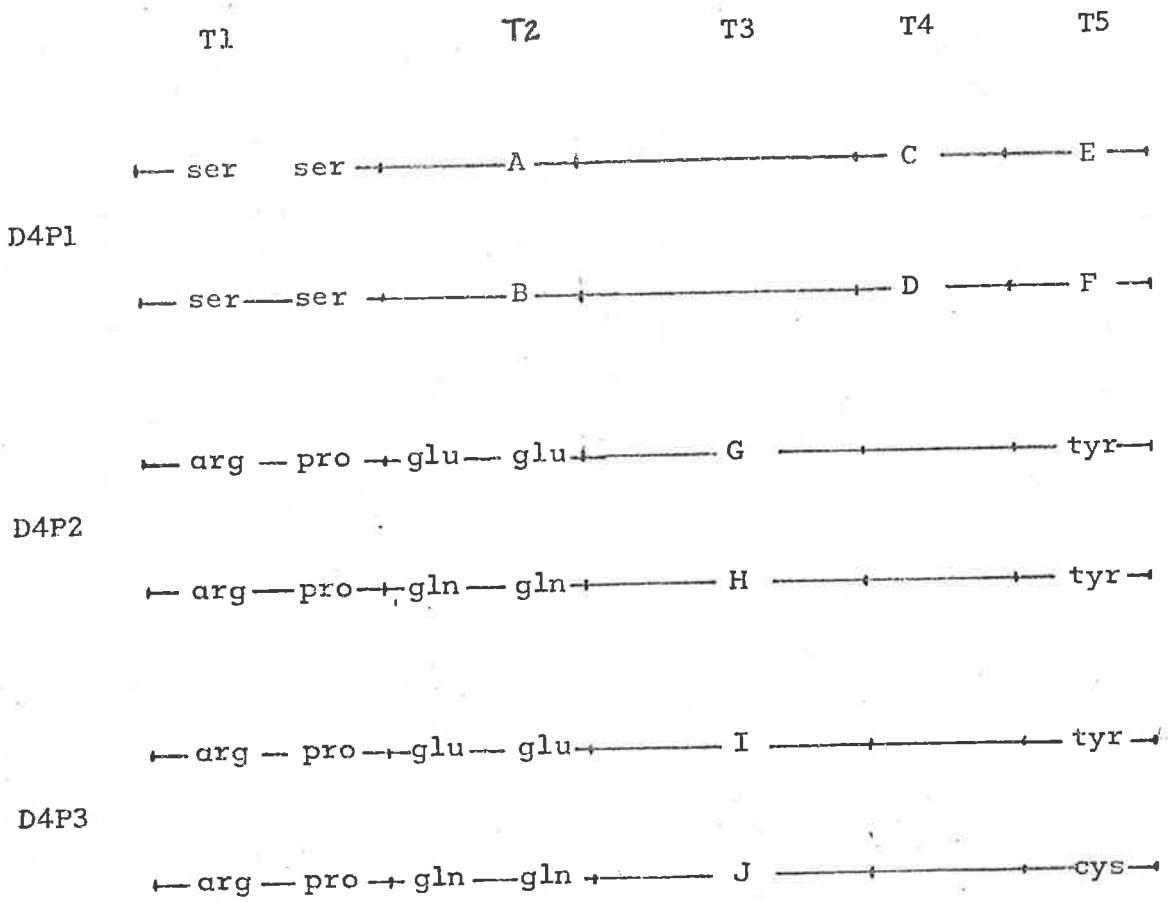
The genes which encode the proteins of D4P1 are entirely different from those which encode both D4P2 and D4P3

as can be seen by a comparison of the respective amino terminal tryptic peptides (Table 7.1). D4P3 consists of two or more keratin chains at least one of which is encoded by a gene which specifies cys^{100}. The genes which encode D4P2 <u>all</u> specify tyr^{100}. The gene(s) encoding the other keratin chain(s) of D4P3 also specifies tyr^{100} and the keratin chains of D4P2 must differ at least one amino acid substitution from those D4P3 keratin chains which have tyr^{100} in order to account for the difference in mobility on pH 2.7 gels of the two protein samples (Fig. 7.1). This amino acid substitution is possibly located in the portion of primary structure represented by T3 or T4 (Chapters 5 and 6) which were not examined in this work. The differences between the three SCM- keratin fractions and their sub-components are summarised in Figure 7.8.

FIGURE 7.8 The possible structures of the constituents of D4P1, D4P2 and D4P3

The nature and position of the amino acid residues shown within the six sequences were inferred from the electrophoretic properties and the amino acid compositions of isolated peptides. Pairs of capital letters (A,B) denote differences (one or more) in the amino acid sequences of two homologous tryptic peptides: the differences were inferred (see Results) and not identified directly.

FIGURE 7.8


T1
T2
T3
T4
T5
ser
ser
A
C
E
D4P1
ser
ser
B
D
F
arg
pro
glu
glu
G
tyr
D4P2
arg
pro
gln
gln
H
tyr
arg
pro
glu
glu
I
tyr
D4P3
arg
pro
gln
gln
J
cys

CHAPTER 8

THE AMINO-TERMINAL SEQUENCES OF DOWN FEATHER, ADULT BARB AND SCALE KERATIN CHAINS

INTRODUCTION

Previous chapters have established the existence in down feather of a large set of keratin chains most or all of which are closely homologous in primary structure and which are encoded by multiple genes. Kemp and Rogers (1972) found that SCM- keratin prepared from scales and different morphological of adult feathers were also resolved into multiple bands when examined by gel electrophoresis. Whereas adult barb and down feather keratin possessed multiple bands in common when examined by electrophoresis at pH 9.5 and pH 2.7 scale keratin possessed two prominent and unique bands at both pH values. The work of this Chapter was undertaken to establish that the keratin chains of the three types of keratinised tissue were encoded by different sets of genes.

METHODS

(i) Tissue

Down feathers were from newly hatched chicks. Barbs and scales were obtained from an adult bird approximately one year old.

(ii) Enzymic digestions

Digestions with chymotrypsin and thermolysin were carried out as described in previous Chapters. Phaseolain was a gift from Dr. W.H. Carey and digestions with phaseolain were carried out as described by Carey and Wells, (1971).

(iii) Isolation of blocked peptides by Dowex-50 Chromatography

Enzymic digests were adjusted to pH 5.0 by the dropwise addition of 5N acetic acid and applied to a column (1.5 x 10 cm) of Dowex AG 50W - X8, 100-200 mesh (H+ form) in H_2O (Narita, 1958; O'Donnell, 1971). Blocked peptides were eluted from the column with 40 ml of H_2O.

(iv) Chromatography of blocked peptides on Dowex-1

Dowex-1 chromatography of blocked peptides after their isolation by Dowex-50 was carried out as described by Kemp

(Ph.D thesis). The gradient used for elution was prepared in a Technicon Autograd. The composition of elution buffer in each of the nine chambers was:

Chamber	Molarity of Formic Acid
1	0.10 (30 ml)
2	0.25 "
3	0.25 "
4	1.00 "
5	1.00 "
6	1.00 "
7	2.50 "
8	2.50 "
9	5.00 "

RESULTS

(i) The preparation of the amino-terminal peptides scales, barbs and down feather keratin

Samples of $[^{14}C]$ - SCM- keratin were prepared from adult leg scales, adult feather barbs and down feather as described in Chapter 3 and subjected to chymotryptic digestion. Blocked amino-terminal chymotryptic peptides were selected by passage of each digest through a column of Dowex-50 (Narita, 1958; O'Donnell, 1971). Similar proportions, 8.4% and 8.9%, of the radioactivity present in the digests of $[^{14}C]$ - SCM- barb and down feather keratin respectively was recovered unbound from Dowex-50: 18% of the radioactivity of $[^{14}C]$ - SCM- scale keratin was recovered. Paper electrophoresis at pH 6.5 of the blocked peptides of either barbs or down feather produced only one

radiocative peptide whose relative mobility (asp=1) in each case was 0.89. Figure 8.1a shows the result of the experiment with $[^{14}C]$-SCM-barbs. No ninhydrin positive material was detected on the electrophoretograms of the blocked peptides of either barbs or down feather and chlorination revealed only the peptide in each case which was detected by autoradiography. To further purify the three sets of blocked peptides obtained by Dowex-50 chromatography, each set was subjected to Dowex-1 chromatography and the elution of the blocked peptides monitored by their radioactivity (Fig. 8.2).

(ii) Sequence analysis of the blocked peptides

(a) The blocked peptides of down feather keratin

The Dowex-1 chromatogram of the down feather peptides (Fig. 8.2a) displays three prominent peaks F1, F2 and F3 which were pooled as shown. The amino acid compositions of the three fractions is shown in Table 8.1. Peptides F2 and F3 represent the sequences N-acetyl-ser-SCM-cys-phe and N-acetyl-ser-SCM-cys-tyr. (Chapter 6 and D.J. Kemp, Ph.D. thesis). The amino acid composition of peptide F1 was considerably more complicated than that of F2 or F3 and was not investigated further.

(b) The blocked peptides of barb keratin

Peptides B2 and B3 (Fig. 8.2) eluted in precisely the same position from Dowex 1 as did their counterparts F2 and F3 of down feather. Their amino acid compositions (Table 8.1) and electrophoretic mobility (Fig. 8.1a) also indicated the

FIGURE 8.1 Paper electrophoresis at pH 6.5 of the blocked peptides of barbs and scale keratin chains

HVPE at pH 6.5 was carried out at 2000 V for 1h. and autoradiography used to detect radioactive peptides.

(a) Blocked peptides of (B) adult feather barbs after Dowex-50 isolation

(b) Blocked peptides (S) of adult scales after their isolation by Dowex-50 and subsequent fractionation on Dowex 1 (Fig. 8.2). Samples for electrophoresis were obtained as shown in Figure 8.2c.

(c) Thermolytic peptides of S1.

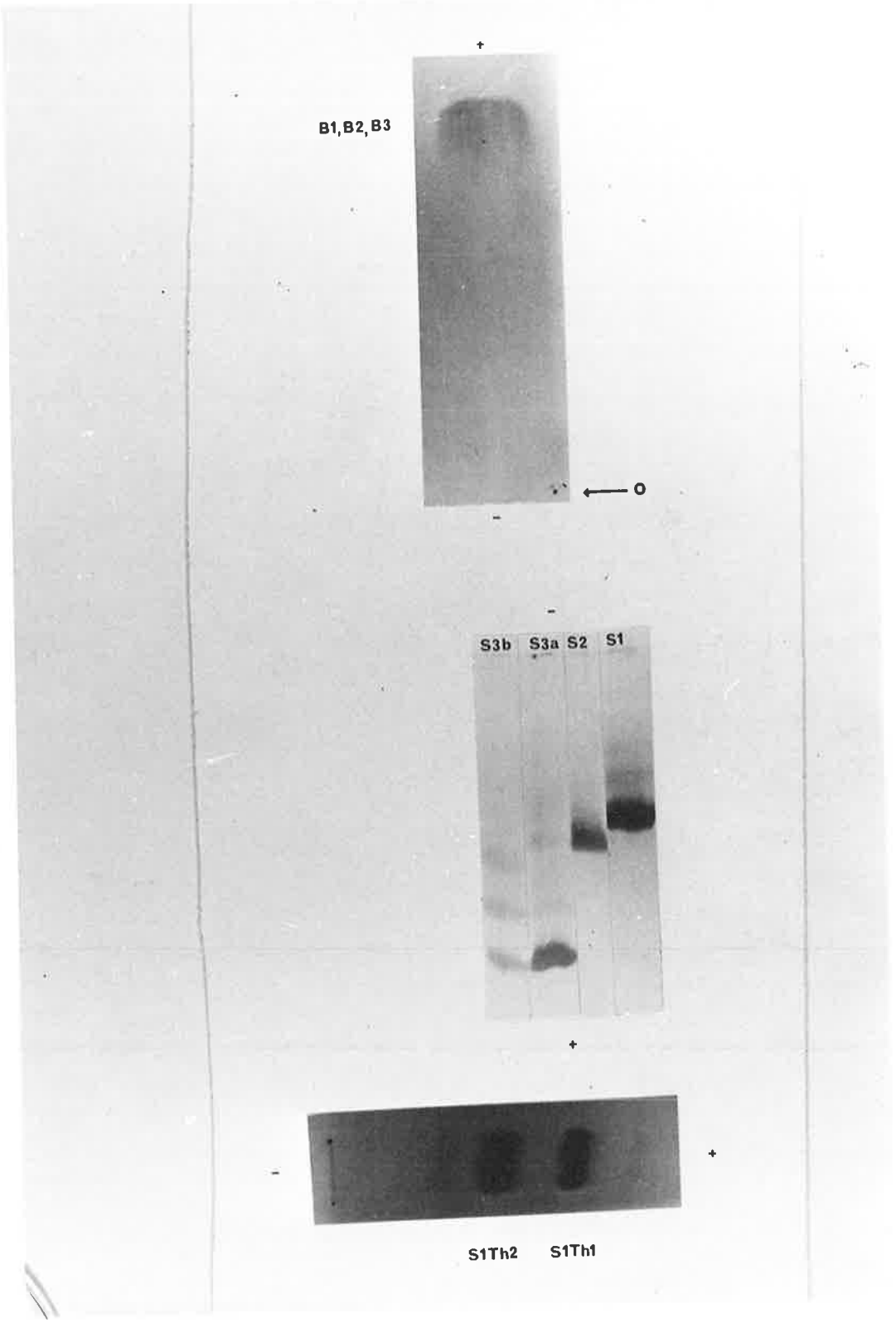

+
B1,B2,B3
O
-
-
S3b
S3a
S2
S1
+
-
+
S1Th2
S1Th1

FIGURE 8.2 Dowex-1 chromatography of blocked peptides

The blocked amino-terminal peptides from Dowex-50 were subjected to gradient elution from Dowex-1 (Methods).

(a) Down feather peptides (F)

(b) Adult barbs peptides (B)

(c) Adult scales peptides (S)

50 μL aliquots were withdrawn from fractions for radioactivity determination. Fractions were pooled as indicated.

The inserts to (a) and (b) represent the ratios of the radioactivity F2 : F3 and B2 : B 3 respectively.

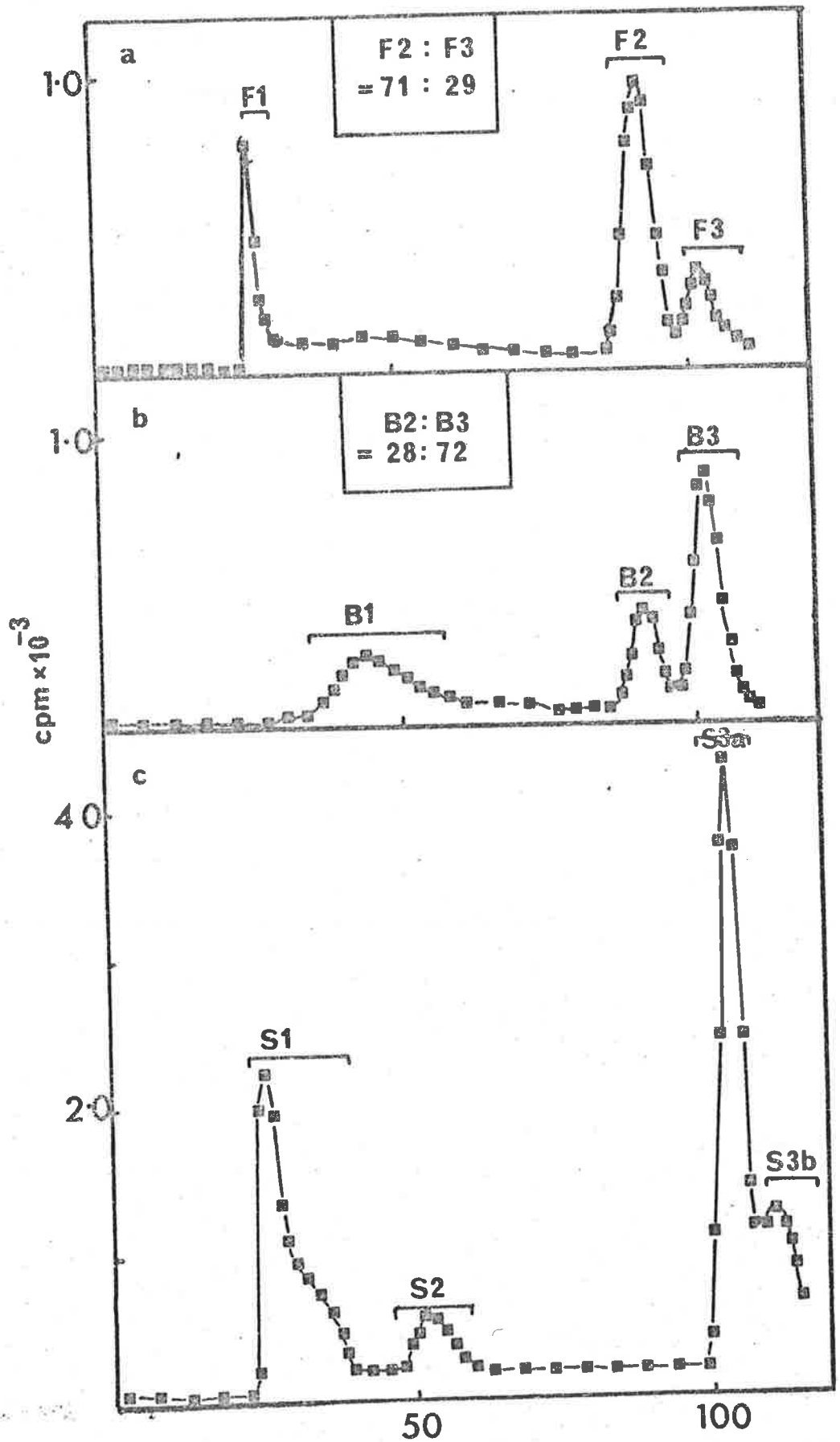

a
F2 : F3
= 71 : 29
F1
F2
F3
1·0
b
B2: B3
= 28: 72
B1
B2
B3
1·0
cpm ×10⁻³
c
S3a
S1
S2
S3b
4·0
2·0
50
100
Fraction No.

TABLE 8.1 Amino acid compositions of the purified amino-terminal peptides of barbs and down feathers

Samples of the peptides of down feather (F1 F2 and F3) and barbs (B1, B2 and B3) were subjected to amino acid analysis.

(a) Absolute μmoles

(b) Residue ratio

TABLE 8.1

	F1		F2		F3		B1	
	a	b	a	b	a	b	a	b
SCM-cys	0.020		0.080	0.80(1)	0.091	0.75(1)	0.093	2.66
asn+asp	0.008						0.038	1.09
thr							0.026	0.74
ser	0.025		0.092	0.92(1)	0.150	1.24(1)	0.039	1.11
gln+glu	0.011						0.110	3.14
pro	0.012						0.031	0.89
gly	0.016						0.011	0.31
ala	0.006						0.015	0.43
$\frac{1}{2}$cys	-						-	
val	0.005						-	
met	-							
ile								
leu	0.009						0.035	1.00
tyr	0.009				0.121	1.00(1)	-	
phe	0.012		0.100	1.00(1)				
lys								
his								
arg							-	

TABLE 8.1 (Contd.)

	B2		B3	
	a	b	a	b
SCM-cys	0.050	0.82(1)	0.247	0.90(1)
asn+asp				
thr				
ser	0.064	1.05(1)	0.325	1.18(1)
gln+glu				
pro				
gly				
ala				
½cys				
val				
met				
ile				
leu				
tyr			0.276	1.00(1)
phe	0.061	1.00(1)		
lys				
his				
arg				

identity of B2 and B3 to F2 and F3 respectively. The release by carboxypeptidase - A of phenylalanine and tyrosine respectively from B2 and B3 established the sequences N-acetyl-(ser-SCM-cys)-phe and N-acetyl-(ser-SCM-cys)-tyr and B2 and B3 were presumed to be identical in sequence to their down feather counterparts. The fraction B1 emerged from Dowex-1 as a broad peak (Fig. 8.2b) and its amino acid composition (Table 1) indicated that it was either a mixture of peptides or a peptide much larger than B2 or B3. Its amino acid composition differred significantly from that of F1 and the two fractions are probably not related.

(c) The blocked peptides of scale keratin

The radioactive blocked peptides of scales eluted as three major peaks S1, S2 and S3 and a minor peak, S3b (Fig. 8.2c). In order to purify the radioactive peptides of the three peaks, the pooled fractions indicated in Figure 8.2c were subjected to electrophoresis at pH 6.5 and peptides detected by autoradiography (Fig. 8.1b). Peak S2 contained only one radioactive peptide: S3a contained one radioactive species and peak S3b in addition to the prominent species observed in S3a contained two other minor constituents. The amino acid compositions and electrophoretic mobilities of the radioactive species designated S1, S2 and S3 are shown in Table 8.2. S3 was not a peptide and in two separate determinations, no amino acids were found in its hydrolysates. The amino acid composition of peptide S2 indicated the structure N-acetyl (ser_3, SCM-cys, leu) and consistent with this, its electrophoretic mobility

TABLE 8.2 Amino acid compositions of the amino-terminal peptides of scales

Samples of the blocked scale peptides S1, S2 and S3 and of the thermolytic fragments of S1 (S1Th1 and S1Th2) were subjected to amino acid analysis.

(a) Absolute μmoles

(b) Residue ratio

The values of (a) represent those of a single determination. The values of (b) represent averages of four determinations for S1, two for S2, two for S3, two for S1Th1 and one for S1Th2.

The molecular weight and charge of each peptide were estimated from the nearest integral number of residues and the relative mobility of each peptide (Offord, 1966).

TABLE 8.2

	S1		S2		S3		S1Th1		S1Th2	
	a	b	a	b	a	b	a	b	a	b
SCM-cys	0.030	1.84(2)	0.043	0.83(1)			0.094	0.83(1)	0.067	0.67(1)
asn+asp	0.00							0.14(0)		
thr	0.027	1.40(1)								
ser	0.050	2.80(3)	0.174	2.78(3)			0.305	3.08(3)	0.030	0.30(0)
gln+glu	0.005								0.029	0.29(0)
pro	0.048	2.38(2)							0.067	0.67(1)
gly	0.024	1.01(1)						0.25(0)	0.108	1.08(1)
ala	0.056	2.52(3)						0.21(0)	0.133	1.33(1)
½cys									0.025	0.25(0)
val	0.020	1.11(1)								
met										
ile										
leu	0.029	1.19(1)	0.064	1.00(1)					0.100	1.00(1)
tyr										
phe										
lys										
his										
arg										

MW	1,100	480		400	330
M(rel)	0.70	0.78	1.20	0.88	0.61
Charge	-3	-2		-2	-1

indicated a peptide with a net charge of -2 and a molecular weight of 480 (Offord, 1966). Carboxypeptidase A and phaseolain, a plant carboxypeptidase with a broader specificity than carboxypeptidase A (Carey and Wells, 1971) each released only leucine from peptide S2 with no trace of either serine or SCM- cysteine. The partial structure of S2 was thus N-acetyl-(ser_3, SCM-cys)-leu.

The amino acid composition of S1 (Table 8.2) indicated that the peptide contained 14 residues two being SCM- cysteine. Its electrophoretic mobility (0.70) indicated a peptide of molecular weight 1,100 possessing a net charge of -3 and this was consistent with its amino acid composition. The peptide was digested with thermolysin and two radioactive peptide fragments S1Th1 and S1Th2 (Fig. 8.1c) were obtained by electrophoresis at pH 6.5. The electrophoretic mobility of peptide S1Th1 indicated a species possessing a charge of -2 and of molecular weight 400 and its amino acid composition indicated the tetrapeptide N-acetyl-(ser_3, SCM-cys). The electrophoretic mobility of peptide S1Th2 indicated a singly charged species of molecular weight 330 and its amino acid composition indicated a pentapeptide (leu, SCM-cys, ala, pro, gly). Identification of the first four residues by the Dansyl-Edman procedure allowed the sequence leu-SCM-cys-ala-pro-gly to be deduced.

The partial sequence of S2, N-acetyl (ser_3 SCM-cys)-leu suggested that it was derived from S1 by partial chymotryptic cleavage and that peptides S1Th1 and S2Th2 were adjacent in the

primary structure of SCM- scale keratin chains whose amino-terminal sequence is presumably N-acetyl-(ser$_3$, SCM-cys)-leu-SCM- cys-ala-pro-gly (thr, pro ala$_2$, val).

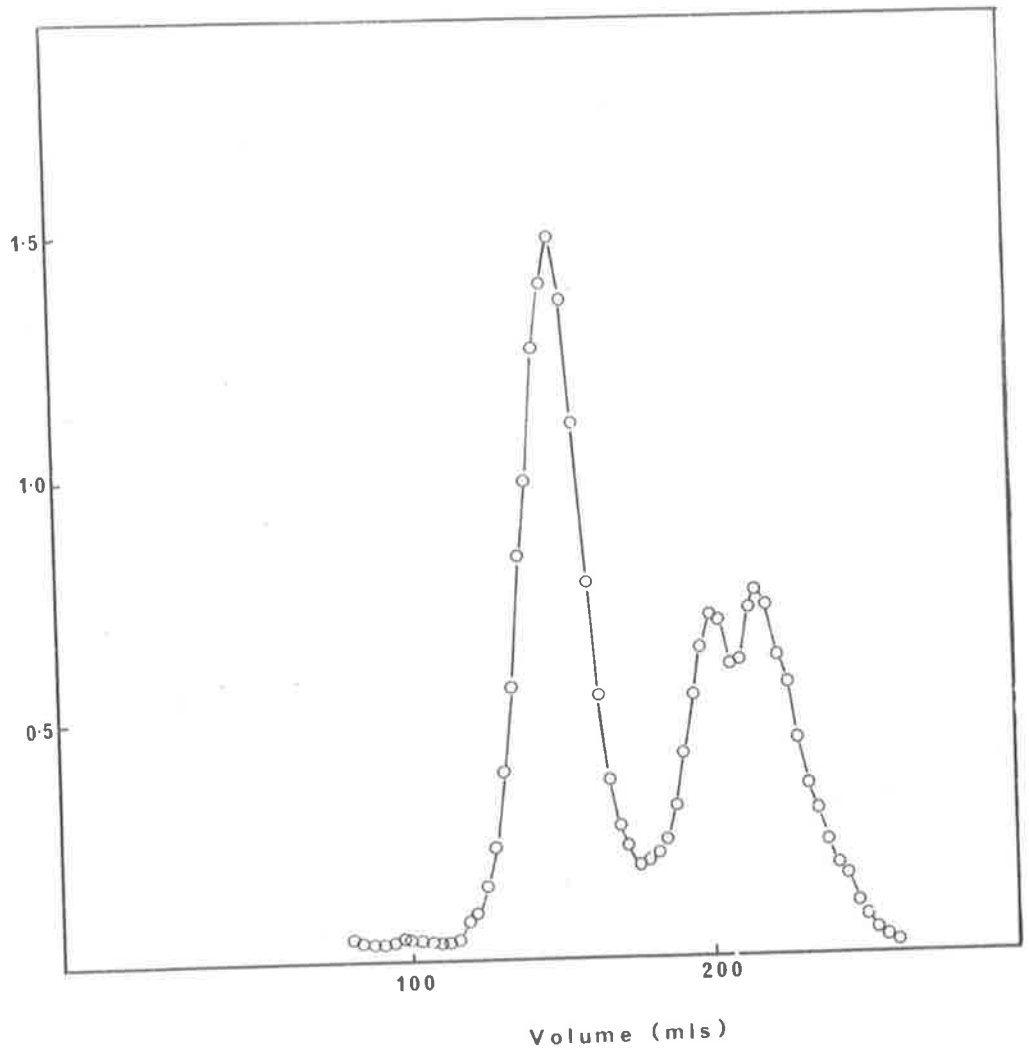
(iii) Peptide mapping of the $[^{14}C]$ - peptides of SCM- scale keratin

To test whether SCM- scale keratin possessed any cysteine containing sequences in common with down feather keratin, a tryptic digest of $[^{14}C]$- SCM- scale keratin chains was prepared and subjected to chromatography on Sephadex G-50 (Fig. 8.3a). Samples of peptides from the radioactive peaks were withdrawn and subjected to paper electrophoresis at pH 6.5 as described for the analogous experiment performed on down feather keratin (Chapter 5). The size-charge peptide map produced is shown in Fig. 8.3b. The distribution of radio-active material eluted from the Sephadex G-50 column broadly resembles the down feather profile (Fig. 5.1a) and two peaks of radioactive peptides representing two major size classes are evident in both profiles. Only four prominent radioactive peptides (ST1, ST2, ST3 and ST4) were detected in the size-charge peptide map (Fig. 8.3b) and their relative mobilities together with the relative mobilities of the eight prominent radioactive down feather peptides identified in Chapter 5 are given in Table 8.3. Electrophoresis at pH 6.5 separated three radioactive peptides (FT1, FT2 and FT3) in the high molecular weight peak in the down feather experiment whereas the high molecular weight peptide material of scale migrated as a single radioactive species (ST1) whose mobility (0.79) was significantly higher

FIGURE 8.3 Size-charge peptide map of the $[^{14}C]$ - SCM-cysteine containing tryptic peptides of scale keratin

(a) Separation of the tryptic peptides of $[^{14}C]$ - SCM-scale keratin by Sephadex G-50 chromatography. 50μL aliquots were withdrawn from each fraction for determination of radioactivity.

(b) 200μL samples of selected fractions were subjected to HVPE at pH 6.5 and peptides detected by autoradiography. The Sephadex G-50 elution profile (a) and the autoradiogram (b) have been aligned with respect to tube number.


1·5
1·0
0·5
100
200
Volume (mls)

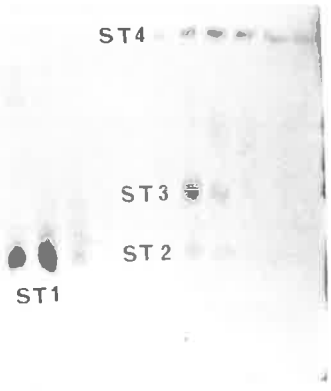

ST4
ST3
ST2
ST1

TABLE 8.3 The relative mobilities of the [^{14}C]-SCM-cysteine-containing peptides of down feather and scales

The mobilities of the prominent radioactive peptides in Figures 5.1 (FT1-FT8) and 8.3 (ST1-ST4) were calculated relative to aspartic acid (1.0). The scale peptides (S.T.) have been placed adjacent to their most closely similar feather counterpart.

TABLE 8.3

SCALE	DOWN FEATHER	
ST1(0.79)	FT1(0.73)	
	FT2(0.66)	LARGE PEPTIDES
	FT3(0.50)	
ST2(0.76)	FT4(0.78)	
	FT5(0.66)	
ST3(0.53)	FT6(0.41)	SMALL PEPTIDES
	FT7(0.34)	
ST4(0.00)	FT8(0.00)	

than any of its down feather counterparts. Of the small tryptic peptides of scale and down feather keratin, ST2 and FT4 (relative mobilities 0.76 and 0.78 respectively, Table 8.3) were of similar size and charge and could concievably have been identical in amino acid sequence. The neutral peptides FT8 and ST4 also possessed similar size properties although their neutrality at pH 6.5 is a poor criterion to suggest that they are identical.

The other major feature of the size-charge peptide maps (Figs. 8.3b and 5.1b) is that SCM- down feather keratin possessed many more $\left[^{14}C\right]$- tryptic peptides than did SCM- scale keratin and this presumably indicates that scale keratin is less heterogeneous than down feather keratin.

DISCUSSION

(i) The amino-terminal sequences of the keratin chains of scale, barb and feather

The results of this Chapter indicate that both down feather and adult feather barb keratin chains possess two amino terminal sequences N-acetyl-ser-cys-phe (F2, B2) and N-acetyl-ser-cys-tyr (F3, B3) in common although in remarkably different relative proportions (F2 : F3 = 71 : 29, B2 : B3 = 28 : 72, Insert to Fig. 8.1). In barb tissue, as in down feather tissue, multiple homologous genes are responsible for the synthesis of keratin chains. The ratio B2 : B3 is inconsistent with the existence of two equally efficient alleles encoding two keratin chains one of which has phe and the other tyr- a 50:50 ratio of

^{3} ^{3}

B2:B3 would be expected under this circumstance. Neither can the ratio be explained by multiple alleles within a population of birds since the barb tissue was derived from a single adult bird.

No scale keratin chains possessed either of the tripeptide sequences of barb or down feather chains and the partial amino terminal-sequence of scale keratin chains is N-acetyl-(ser$_3$ cys) leu-cys-ala-pro-gly (thr, pro ala$_2$, val).

The size-charge peptide map of the [^{14}C] - SCM-cysteine containing tryptic peptides of scale keratin (Fig. 8.3) indicate that at most two scale peptides ST2 and ST4 may have counterparts in down feather keratin (FT4 and FT8 respectively) - the scale peptide ST3 had no counterpart in down feather. The large size of peptide ST1 indicated that it was probably the amino-terminal tryptic peptide of scale keratin chains which was expected to have at least fifteen residues (fourteen residues in S1 plus at least arginine). That both peptides ST1 and ST3 were unique to scale keratin chains indicated that the differences between down feather and scale keratin chains were not only confined to the amino terminal portions of the two sets of protein molecules.

(ii) Gene expression in keratinising tissues

The amino-terminal sequences and size-charge peptide maps of down feather and scale keratin chains indicate that no

keratin gene is commonly expressed in the two different tissues. The differences which exist in the amino-terminal sequences of scale and down feather keratin cannot be accounted for by tissue specific post-synthetic modification of a common set of keratin chains and must result from the expression of two entirely different sets of genes.

The identity of the amino-terminal sequences of down feather and barb keratin represented by the peptides F2, B2 and F3, B3 respectively is consistent with the expression of at least some common genes in both tissues and this possibility can be tested by detailed primary structure analysis of the two sets of gene products. The different relative proportions of the two amino-terminal peptides (F2 : F3 = 71 : 29, B2 : B3 = 28 : 72) indicate that if the same genes are expressed in all cells of both adult barb and down feather tissue, then their synthesis is differentially regulated in a tissue specific manner. Alternatively, barb tissue and feather tissue may each be composed of different proportions of at least two cell types each of which selectively expresses some of a potentially much larger number of keratin genes.

CHAPTER 9

CONCLUDING DISCUSSION

The validity of the experimental data and its interpretation have been considered in the Discussion sections of Chapters 3-8. The objects of the present section are to relate the findings of these Chapters to other areas of Biology, to outline some of the major questions concerning the keratins of avian tissues together with potential experimental approaches for their solution.

(i) Keratin structure in relation to other proteins

(a) The nature of the heterogeneity of feather keratin molecules

The results of the isolation of nine fractions of SCM- down feather keratin have been presented in Chapter 4. The amino acid composition of eight of these fractions established their identities as feather keratin, the major synthetic product of feather cells. Although these eight fractions migrated essentially as single bands on pH 9.5 polyacrylamide gels, six fractions were further resolved into two or more bands by pH 2.7 polyacrylamide gels. The presence of at least nineteen prominent variants of down feather keratin was established by these techniques - most or all variants were found in individual chicks. The elution of only one symmetrical peak of radioactive protein by Sephadex G-100 chromatography of $[^{14}C]$- SCM- feather keratin of molecular weight 11,700 indicated that most or all keratin molecules were very similar in molecular weight. Peptide mapping of $[^{14}C]$ - SCM- feather keratin (Chapter 5) indicated the presence of at least twelve

prominent tryptic peptides, and the presence of common regions of primary structure in different keratin variants was inferred. Extensive sequence analysis of the tryptic peptides (Chapter 6) showed that individual keratin variants differred from one another by amino acid substitutions and preliminary examination of three highly purified electrophoretic variants of SCM-feather keratin (Chapter 7) suggested that at least one unique gene was required to encode each electrophoretic variants. Taken together, the evidence of these four Chapters compels the conclusion that multiple homologous genes exist in the chick genome each encoding a separate keratin variant. Multiple homologous genes have been proposed to account for the synthesis within one organism of the different C regions of both immunoglobulin heavy chains and light chains (Gally and Edelman, 1972). However, in feather keratin, no large regions of primary structure appear to be especially invariant or conversely, to be especially susceptible to sequence variation. On the other hand, immunoglobulin chains are composed of an amino-terminal region which, when compared with other chains of the same class, contain many amino acid substitutions and a carboxyl-terminal constant region containing far fewer differences (if any) between two chains of the same class.

Immunoglobulin chains are believed to result from the transcription and translation of a gene which is formed by the fusion of a specific C gene element with a specific V gene element both of which are selected from potentially larger

sets (Gally and Edelman, 1972). In comparison, keratin chains do not show a pattern of sequence substitutions compatible with such a scheme.

(b) Sequence heterogeneity in other proteins

The existence of multiple molecular forms of specific proteins within a single organism is not uncommon and has been demonstrated for a wide variety of proteins in bacterial and eukaryotic cells. Table 9.1 lists a number of examples of proteins in which heterogeneity has been demonstrated and a selected few of these will be discussed in this section.

Although multiple molecular forms of many enzymes have been identified, the structural differences between different forms often have not been studied and could conceivably result from the partial modification of a single gene product: The existence of variants of histones each differring in the extent to which specific lysine residues are acetylated (Candido and Dixon, 1972, Gorovsky, et al., 1973), of amidated and deamidated forms of rabbit muscle aldolase (Midelfort and Mehler, 1972) and of α1 collagen chains in which specific proline residues are hydroxylated to varying extents (Bornstein, 1971) are examples in which post-synthetic modification introduces heterogeneity.

In the course of sequence analysis of SCM- feather keratin, only two peptides, T1a (Chapter 6), and T2a (Chapter 7), were isolated, which could conceivably have arisen due to deamidation of asparagine or glutamine residues. However, it

TABLE 9.1 Heterogeneity of other proteins

[1] The structural difference between different molecular forms has not been identified directly but has been inferred.

1. Rosner and Paulus, (1971)
2. Engelbrecht and Sadoff, (1969)
3. Sadoff et al., (1970)
4. Summaria et al., (1972)
5. Sodetz et al., (1972)
6. Traugott and Massaro, (1973)
7. Barash et al., (1973)
8. Cheng et al., (1973)
9. Stenflo, (1973)
10. Johnson et al., (1973)
11. Alpert et al., (1972)
12. Sapolsky and Woessner, (1972)
13. Schmidt and Ebner, (1972)
14. Candido and Dixon, (1972)
15. Gorovsky et al., (1973)
16. Midelfort and Mehler, (1972)
17. Bornstein, (1967)
18. Dayhoff, (1969)
19. Croft, (1973)
20. see Introduction.

TABLE 9.1

PROTEIN	HETEROGENEITY DEMONSTRATED BY:	STRUCTURAL BASIS	REF.
Human α-fetoprotein	isoelectric focusing, ion exchange chromatography, agarose gel electrophoresis.	-	11
Bovine cathepsin-D	ion exchange chromatography, disc electrophoresis.	-	12
α-lactalbumin (Pig, Sheep, Goat)	gel electrophoresis, amino acid analysis, peptide mapping	-	13
Histones	sequence analysis, gel electrophoresis	partial acetylation of specific lysine residues	14 15
Rabbit muscle aldolase	sequence analysis	partial deamidation of specific asparagine residue	16
Rat α-1 collagen	sequence analysis	partial hydroxylation of specific prolyl residues	17
Human globin chains	sequence analysis	amino acid substitutions	18
Calf lens γ-crystalin	sequence analysis, ion exchange chromatography	amino acid substitutions	19
Mammalian α-keratin	ion exchange chromatography gel electrophoresis sequence analysis	amino acid substitutions	20
B. Subtilis aspartokinases	gel filtration	-	1
B. Cereus purine nucleotide phosphorylase	gel electrophoresis	post synthetic[1]	2
B. Cereus aldolase	gel electrophoresis gel filtration	post synthetic[1]	3
Human plasminogen and plasmin	isoelectric focussing gel electrophoresis	-	4

TABLE 9.1 (Contd.)

PROTEIN	HETEROGENEITY DEMONSTRATED BY:	STRUCTURAL BASIS	REF.
Rabbit plasminogen	isoelectric focussing gel electrophoresis	-	5
Rabbit creatine phosphokinase	starch gel electrophoresis	-	6
Oat leaf glutamine synthetase	gel electrophoresis	-	7
Rat nucleoside diphosphokinase	ion exchange chromatography	-	8
Rat prothrombin	gel electrophoresis peptide mapping	post synthetic[1]	9
Rat tyrosine amino transferase	ion exchange chromatography	post synthetic[1]	10

is equally likely that the sequences of these peptides were genetically determined by the substitution of an acidic amino acid for its amidated counterpart.

The existence of precursors of specific proteins which are subsequently modified by proteolysis is well known. Examples of this are precursors of bee mellitin (Kreil, 1973) and of immunoglobulin light chains (Milstein et al., 1972, Mach et al., 1973) which have additional amino-terminal sequences that are subsequently removed by tissue specific proteases. Feather keratin chains, however, are not synthesised as precursors; the protein product specified by mRNA from feather cells is indistinguishable from keratin prepared from the feathers of newly hatched chicks (Partington et al., 1973, Kemp and Rogers, unpublished results).

The existence of multiple molecular forms of many other proteins can be accounted for only by the existence of separate genes each encoding the synthesis of a different variant. The existence of at least six different forms of myosin light chains within the one organism has been established (Sarkar et al., 1971). Sequence analysis of selected peptides has indicated that at least three different genes are responsible for myosin light chain synthesis and the structures of two light chains (C1 and C3) are in part at least homologous (Taylor, 1972). Collagen from rat skin or rat tendon is composed of two types of polypeptide chain, $\alpha 1$ and $\alpha 2$, which are markedly different in primary structure although

both contain a high proportion of the repeating tripeptide sequence gly-pro-X (Gallop et al., 1972). The structure of both high sulphur and low sulphur proteins of wool and of immunoglobulins have been referred to previously; each group of proteins are families of protein molecules differring from each other by amino acid substitutions.

Calf lens γ-crystallin molecules are composed of at least five different variants and it has been established that the different forms are closely similar in physico-chemical properties and in particular in primary structure but are encoded by different genes (Bjork, 1970, Croft and Waley, 1971, Croft 1972, Slingsby and Croft, 1972). In two respects feather keratin chains most closely resemble the α-keratins of hair and the γ-crystallins of lens cells. Different genes are responsible for the synthesis of each variant in all three families of proteins, and the overall level of heterogeneity in each system is quite substantial. Although only four different chains are required to account for the chromatographic properties of γ-crystallin, microheterogeneity has been detected in one component (1Vb) previously believed to be homogeneous (Slingsby and Croft, 1972). The actual level of heterogeneity may therefore be much higher than that previously believed. All the variants of α-keratin are to be found in the one tissue type from a single animal and similarly all crystallin variants co-exist in a single tissue, the lens epithelium. In this respect also, feather keratin chains which are found in

a single tissue, the down feathers of a single chick, resemble the two former protein families.

Feather keratin is considerably more heterogeneous than myosin light chains, where only three are found in a single muscle tissue (Sarker, 1971), or than collagen chains, where at most three occur in the same tissue (Gallop et al., 1972). It is noteworthy that hair, feathers and lens epithelium are all epidermal tissues and the heterogeneity of their proteins may reflect some property common to these tissues.

(c) The salient features of the primary structure of feather keratin

All down feather keratin variants are indistinguishable in molecular weight. The fact that different variants all possess closely-homologous amino-terminal sequences (N-acetyl-ser-cys-) and closely homologous carboxyl-terminal sequences $\begin{matrix}\text{tyr}\\\text{cys}\end{matrix}$ - $\begin{matrix}\text{leu}\\\text{phe}\end{matrix}$ -pro-cys eliminates the possibility that small deletions of amino-terminal or carboxyl-terminal segments have taken place during the evolution of keratin genes. Furthermore, the incomplete sequence data of Chapter 6 is perfectly compatible with all down feather keratin variants having exactly the same number of amino acid residues.

All down feather keratin variants possessed amino; and carboxyl-termini rich in half cystine residues, whose distribution within the primary sequence was very similar in all of the variants and also in a purified fraction of emu rachis. In particular the work of Chapters 6 and 7 established that

half-cystine was the carboxyl-terminal residue of most, if not all, down feather keratin chains. The observation that this residue is carboxyl-terminal in all other keratin chains on which primary sequence studies have been carried out (including the high-sulphur proteins of mammalian hair) suggest an important role in the establishment or maintainence of the arrangement of native keratin chains in their respective tissues.

A large central portion of all down feather keratin chains and of emu rachis keratin is hydrophobic and contains most of the valine residues of keratin chains (Chapter 5). This segment has the amino acid composition expected for a β-pleated sheet region that is found in feather keratin. The complete amino acid sequence of this segment is important since it may represent an especially invariant part of all feather keratin molecules which has been conserved during evolution to fulfil a fundamental role in the organisation of keratin chains into fibrils.

(ii) The control of keratin synthesis

(a) The cellular distribution of keratin variants

All keratin variants need not be synthesised in every cell of the down feather. In each morphologically different cell type in down feather a different set of keratin genes could be active and the proportions of different keratin variants in a feather would reflect the relative proportions of cells of each type. Such a stringent limitation of gene expression has

been demonstrated in other terminally differentiated cell lines. For example, in amphibian red blood cells only one kind of haemoglobin, either adult or embryonic, but never both, is synthesised in the one cell (Rosenberg, 1970). Individual lymphocytes, the antibody producing cells of blood are also restricted to the production of a single species of immunoglobulin molecule whose chains (H and L) are encoded by only two genes, one for the light (L) chain and one for the heavy (H) chain. Furthermore, in an animal heterozygous for two allelic immunoglobulin genes only one of the alleles is expressed in any one lymphocyte. (Makela and Cross, 1970). It is important therefore to establish whether a single feather cell type produces <u>all</u> or only a small number of keratin variants selected from a much larger set.

Whether or not a stringent limitation of keratin gene expression exists at the level of individual cell-types, in different keratinising tissues, different keratin genes are expressed. Examination of the blocked amino-terminal peptides of scales, adult barbs and down feather keratin indicate that whereas the only prominent amino-terminal sequence of scale keratin was N-acetyl-(ser$_2$, cys)-leu-, (Chapter 8), barbs and down feather keratin each contain two amino-terminal sequences N-acetyl-ser-cys-phe and N-acetyl-ser-cys-tyr in markedly different amounts. These studies and the evidence of Kemp and Rogers, (1972) show that scale-specific keratin chains are not synthesised to a detectable degree in feather tissue. Common

keratin chains may be synthesised in down feather tissue and in adult barb tissue but if this is so their relative rates of synthesis differ markedly in the two tissues.

(b) The structure of keratin genes

Kemp et al., (personal communication) have shown that most or all down feather keratin mRNA molecules are homogeneous in size and the molecular weight of each (250 000) is sufficient to encode two keratin monomers of 10-11000 in molecular weight. Keratin chains however are synthesised on polysomes containing an average of 4-5 ribosomes and this argues against the presence of two keratin cistrons per mRNA molecule. If there is only one cistron as seems most likely, then the presence of a large untranslated segment in each feather keratin mRNA molecule can be inferred. Its length (460 nucleotides) is much greater than that of the poly A sequence (150 - 200 nucleotides) which is commonly present at the 3'-termini of eukaryotic mRNA molecules after their synthesis (Darnell et al., 1971. Shultz, 1973). Its existence suggests that in each keratin gene in addition to the polynucleotide sequence encoding the primary structure of keratin chains, an additional sequence is present which is important in their synthesis or its regulation. The homogeneity of keratin mRNA on formamide-acrylamide gels indicates that the size at least of the untranslated segment has been conserved as stringently during evolution as the size of the different keratin chains and therefore of the translated segment of mRNA.

(c) The arrangement of keratin genes in the chick genome

Multiple genes which have evolved recently in evolution from duplication events frequently exist in close proximity to one another on the chromosome. Thus, the existence of haemoglobin lepore whose amino-terminal region and carboxyl-terminal regions are identical to those of β-chains and δ-chains respectively (Baglioni, 1962) is believed to represent a rare crossover event between two non-allelic genes at adjacent loci, the β locus, and the δ locus which presumably arose by gene duplication and subsequent mutation of the β locus. Close linkage of the genes of two α globin variants in the deer (Taylor et al., 1972) and of ten haemoglobin variants in Chironomus tetans (Tichy, 1973) provide independent evidence indicating gene duplication and mutational divergence of adjacent gene copies as the origin of heterogeneity.

Three major possibilities exist with respect to the arrangement of keratin genes on the chick genome.

i) The genes which encode the synthesis of keratin variants whether of the same tissue or of different tissues are unlinked and distributed widely on different chromosomes. α and β chains of human haemoglobin are encoded by unlinked genes (Huehns and Shooter, 1965) believednonetheless to be derived from a common ancestral gene.

ii) Keratin genes encoding a set of tissue specific proteins, for example, scale keratin chains, may be closely linked, possibly in tandem array whereas gene sets encoding protein products for different tissues (barb and scale genes) are

unlinked. Immunoglobulin genes specifying C or V regions for a particular class of light chains or for heavy chains are closely linked although the genes specifying different classes of immunoglobulin chains are unlinked (Gally and Edelman, 1972). iii) All keratin genes are located in the same portion of the chick genome on a single chromosome.

The possible linkage of keratin genes could be studied by an analysis of the transmission between generations of at least two keratin loci at which identifiable alleles existed. The overall heterogeneity of down feather keratin however would make recognition of alleles difficult. The development of two dimensional acrylamide gel electrophoresis of SCM- feather keratin samples which allowed the rapid, reliable separation of the nineteen variants described in Chapter 4 would facilitate the recognition of alleles.

(iii) The role of multiple genes for feather keratin

The presence in the chick genome of at least nineteen genes each encoding different variants of down feather keratin must confer some selective advantage upon the organism. Kemp (Ph.D. thesis) suggested that multiple genes for feather keratin would allow an enhanced rate of keratin synthesis and that genes, evolved for this purpose, would accumulate mutations in parts of the structural genes encoding tolerant parts of the primary structure of keratin chains, the highly mutable portions (Fitch and Margoliash, 1967). Fraser et al.,

(1972) have proposed that keratin molecules generally may be more tolerant to primary sequence variation than other proteins since the cells in which they are synthesised are shortly destined to die and the keratin once synthesised is no longer involved in the biochemistry of the organism. Implicit in this theory of keratin diversity is that most or all feather keratin variants are functionally equivalent. Other terminally differentiated cell lines also produce large amounts of specialised protein products such as haemoglobin in red blood cells or silk fibroin in silk glands. However measurement of the re-iteration frequency of haemoglobin genes in red blood cells or in other somatic tissues has indicated that only 1-3 genes encoding globin chains are present in the genome of either duck (Bishop and Rosbach, 1973) or mouse (Harrison et al., 1972). Only 1-3 silk fibroin genes are present in the genome of Bombyx mori tissues[1]. These studies indicate that neither reiteration of genes nor their tissue specific amplification is necessary for the synthesis of large amounts of the protein products which they encode and there seems to be no reason per se to postulate reiteration of keratin genes as a means of increasing the synthesis of keratin chains. The only case so far established in which reiteration of the structural genes encoding a protein products has been established are the genes specifying histones in the sea urchin each of which is present as 1200 identical copies (Kedes and Birnstiel, 1971. Weintraub et al., 1972).

[1] Suzuki et al., (1972)

Furthermore, the assumption that the primary structure of keratin chains is less stringently conserved during evolution than that of other enzymes and proteins (Fraser et al., 1972) presumes that the structural requirements of keratin chains to form fibrils of determinate characteristics are less stringent than those structural requirements of other protein molecules which determine enzymic activity. The validity of this assumption is experimentally testable. A comparative study of the primary structures of homologous keratins from related species should establish the degree to which evolutionary variability can be tolerated. However insufficient data is at present available for a significant estimate of interspecies divergence to be made and any assumption regarding the degree to which sequence variation in keratin chains can be tolerated is premature.

The presence of multiple homologous variants of feather keratin all of which are alike in primary structure does not establish their functional equivalence and the unequal distribution of different variants in different morphological parts of the adult feather (Kemp and Rogers, 1972) suggests that different keratin genes have evolved to fulfil different structural requirements. The work of Chapter 8 demonstrated that the two tripeptides N-acetyl-ser-cys-phe and N-acetyl-ser-cys-tyr, although common to both adult barb and down feather keratin chains were present in different relative amounts in the proteins of the two tissues. Thus at least the quantitative

distribution between the two tissues of different keratin chains, differs markedly and this argues against the functional equivalence of all keratin chains. Examination of the number and nature of keratin chains within a single feather cell-type and their relation to the products of different cell types is important and may allow a more precise correlation of primary structure and function.

REFERENCES

Alpert, E., Drysdale, J.W., Isselbacher, K.J. and Schur, P.H. (1972). J. Biol. Chem. 247 3792.

Ambler, R.P. (1967). Methods. Enzymol. 11, 436.

Baglioni, C. (1962). Proc. Nat. Acad Sci. U.S. 48 1880.

Barash, I., Sadon, T. and Mor, H. (1973). Nature New Biol. 244, 150.

Bishop, J.O. and Rosbach, M. (1973). Nature New. Biol. 241, 204.

Bjork, I. (1970). Exp. Eye Res. 9, 152.

Blombock, B. (1967). Methods. Enzymol. 11, 398.

Bornstein, P. (1967). J. Biol. Chem., 242, 2572.

Bruton, C.J. and Hartley, B.S. (1970). J. Mol Biol., 52, 165.

Bryce, G.F.A., and Crichton, R.R., (1971). J. Biol. Chem. 246, 4198.

Candido, E.P.M. and Dixon, G.H. (1972). Proc. Nat. Acad.Sci., 69, 2015.

Carey, W.F. and Wells, J.R.E. (1972). J. Biol. Chem., 247, 5573.

Cheng, Y., Robison, B. and Parks, R.E. (1973). Biochemistry 12, 5.

Chrambach, A., Reisfeld, R.A., Wyckoff, M. and Zaccari, J. (1967). Anal. Biochem., 20, 150.

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

Crewther, W.G., Fraser, R.D.B., Lennox, F.G. and Lindley, H. (1965) in Anfinsen, C.B., Anson, M.L., Edsall, J.T., and Richards, F.M., (Eds.) Advances in Protein Chemistry New York, Academic, 20, 191.

Croft, L.R. (1972). Biochem. J., 121, 557.

Croft, L.R. (1973). Biochim. Biophys. Acta., 295, 174.

Croft, L.R. and Waley, S.G. (1971). Biochem. J., 121, 453.

Darnell, J.E. Philipson, L., Wall, R. and Adesnik, M. (1971). Science, 174, 507.

Darskus, R.L. (1972). J. Chromatogr., 69, 341.

Darskus, R.L. and Gillespie, J.M. (1971). Aust. J. Biol. Sci., 24, 515.

Dayhoff, M. (1969). Atlas of Protein Sequence and Structure.

Ein, D. (1968). Proc. Nat. Acad.Sci. U.S., 60, 982.

Elleman, T.C. and Dopheide, T.A. (1972). J. Biol. Chem., 247, 3900.

Engelbrecht, H.L. and Sadoff, H.L. (1969). J. Biol. Chem., 244, 6228.

Fairbanks, G., Levinthal, C. and Reeder, R.H. (1965). Biochem. Biophys. Res. Commun., 20, 393.

Fairbanks, G., Steck, T.L. and Wallach, D.F.H. (1971). Biochemistry, 10, 2606.

Fambrough, D.M., Fuyimura, F. and Bonner, J. (1968). Biochemistry, 7, 575.

Filshie, B.K. and Rogers, G.E. (1962). J. Cell. Biol., 13, 1.

Fitch, W.M. and Margoliash, E. (1967). Science, 155, 279.

Fraser, R.D.B., MacRae, T.P., Parry, D.A.D. and Suzuki, E. (1971). Polymer, 12, 35.

Fraser, R.D.B., MacRae, T.P., and Rogers, G.E., (1972) in Keratins, Their Composition, Structure and Biosynthesis, C.C. Thomas, Springfield, Ill.

Gallop, P.M., Blumenfeld, O.O. and Seifter, S. (1972). Ann. Rev. Biochem., 41, 617.

Gillespie, J.M. (1965) in Lyne A.G. and Short B.F. (Eds.) The Biology of Skin and Hair Growth, Angus and Robertson, Sydney.

Goddard, D.R. and Michaelis, L. (1935). J. Biol. Chem., 112, 361

Gorovsky, M.A., Pleger, G.L., Keevert, J.B. and Johmann, C.A. (1973). J. Cell Biol. 57, 773.

Gray W.R. (1972) in Methods Enzymol. 11, 121.

Gurd, F.R.N. (1967) Methods Enzymol, 11, 532.

Harding, H.W.J. and Rogers, G.E. (1971). Biochemistry, 10, 624.

Harrap, B.S. and Woods, E.F. (1964a). Biochem.J., 92, 8.

Harrap, B.S. and Woods, E.F. (1964b). Biochem. J., 92, 19.

Harris, J.I. (1967) Methods. Enzymol., 11, 390.

Harrison, P.R., Hell, A., Birnie, G.D. and Paul, J. (1972). Nature. (London) 239, 219.

Hartley, B.S. (1970). Biochem. J., 119, 805.

Haylett, T., Swart, L.S., Parris, D. and Joubert, F.J. (1971) Appl. Poly. Symp, No., 18, 37.

Heinrikson, R.L. (1966). J. Biol. Chem., 241, 1393.

Hess, M., Hilschmann, N., Rivat, L., Rivat, C. and Ropartz, C. (1971). Nature New. Biol., 234, 58.

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

Huehns, E.R. and Shooter, E.M. (1965). J. Med. Genet., 2, 48.

Inouye, M. (1971). J. Biol. Chem., 246, 4834.

Jeffrey, P.D. (1969). Biochemistry, 8, 5217.

Jeffrey, P.D. (1970). Aust. J. Biol. Sci., 23, 809.

Johnson, R.W., Robertson, L.E. and Kenney, F.T. (1973). J. Biol. Chem., 248, 4521.

Kedes, L.H. and Birnstiel, M.L. (1971). Nature. New. Biol., 230, 165.

Kemp, D.J., Dyer, P.Y. and Rogers, G.E. (1974). J. Cell. Biol. (in press).

Kemp, D.J., Partington, G.A. and Rogers G.E. (1974). Biochemistry (submitted for publication).

Kemp, D.J. and Rogers, G.E. (1972). Biochemistry, 11, 969.

Kotaki, A. (1962). J. Biochem., 51, 301.

Kreil, G. (1973). Eur. J. Biochem., 33, 558.

Laemli, U.K. (1970). Nature, (London), 227, 680.

Lai, C.Y., Chen, C. and Horecker, B.L. (1970). Biochem. Biophys. Res. Commun., 40, 461.

Lin, K.D. and Deutch, H.F. (1972). J. Biol. Chem., 247, 3761.

Lindley, H. and Elleman, T.C. (1972). Biochem.J., 128, 859.

Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall R.J. (1951). J. Biol. Chem., 193, 685.

Mach, B., Faust, C. and Vassalli, P. (1973) Proc. Nat. Aca. Sci., 70, 451.

Makela, O., and Cross, A.M. (1970). Prog. Allergy, 14, 145.

Michl, H. (1951). Monatsh. Chem., 82, 489.

Midelfort, C.F. and Mehler, A.H. (1972). Proc. Nat. Acad.Sci., 69, 1816.

Milne, P.R. and Wells, J.R.E. (1970). J. Biol. Chem., 245, 1566.

Milstein, C., Browlee, G.G., Harrison, T.M. and Mathews, M.B. (1972). Nature New Biol., 239, 117.

Moore, S. and Stein W.H. (1963). Methods Enzymol., 6, 818.

Narita, K. (1958). Biochim. Biophys. Acta., 28, 184.

O'Donnell, I.J. (1971). Aust. J. Biol. Sci., 24, 179.

O'Donnell, I.J. (1973a). Aust. J. Biol. Sci., 26, 401.

O'Donnell, I.J. (1973b). Aust. J. Biol. Sci., 26, 415.

O'Donnell, I.J. and Thompson E.O.P. (1962). Aust. J. Biol. Sci., 15, 740.

Offord, R.E. (1966). Nature (London), 211, 591.

Partington, G.A., Kemp, D.J. and Rogers, G.E. (1973). Nature. New. Biol., 246, 33.

Panyim, S. and Chalkely, R.C. (1969). Arch. Biochem. Biophys., 130, 337.

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

Porter, R.R. (1971). The Harvey Lectures. 65, 157., Academic., N.Y.

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

Rawles, M.E. (1965) in Lyne, A.G. and Short, B.F. (Eds.) "Biology of Skin and Hair Growth", Sydney, Angus and Robertson, p 105.

Rogers, G.E. and Filshie, B.K. (1963) in Borasky, R. (Ed.) "Ultrastructure of Protein Fibres". New York, Academic, p 123.

Rosenberg, M. (1970). Proc. Nat. Acad.Sci. U.S., 67, 32.

Rosner, A. and Paulus, H. (1971). J. Biol. Chem., 246, 2965.

Rydon, H.N. and Smith, P.N.G. (1952). Nature (London), 169, 922.

Sadoff, H.L., Celikkol, E. and Engelbrecht, H.L. (1970). Proc. Nat. Acad. Sci. U.S., 66, 844.

Sanger, F. and Tuppy, H. (1951). Biochem. J., 49, 481.

Sapolsky, A. and Woessner, J.F. (1972). J. Biol. Chem., 247, 2069.

Sarkar, S., Streter, F.A. and Gergely, J. (1971). Proc. Nat. Acad. Sci. U.S., 68, 946.

Schmidt, D.V. and Ebner, K.E. (1972). Biochim. Biophys. Acta., 263, 714.

Schroeder, W.A., Jones, R.T., Cormick, J. and McCalla, K. (1962). Anal. Chem., 34, 1570.

Shapiro, A.L., Vineula, E. and Mazel, J.V. (1967). Biochem. Biophys. Res. Commun., 28, 815.

Shechter, Y., Landau, J.W. and Newcomer, V.D. (1969). J. Invest. Derm., 52, 57.

Shultz, G.A. (1973). Devel. Biol., 31, 2.

Slingsby, C. and Croft, L.R. (1972). Biochem. J., 127, 609.

Smyth, D.G. (1967). Methods. Enzymol., 11, 216.

Sodetz, J.M., Brockway, W.J. and Castellino, F.J. (1972). Biochemistry 11, 4457.

Stenflo, J. (1973). J. Biol. Chem., 248, 6325.

Summaria, L., Arzadon, L., Bernabe, P. and Robbins, K.C. (1972). J. Biol. Chem., 247, 4691.

Suzuki, E. (1973). Aust. J. Biol. Sci., 26, 437.

Suzuki, Y., Gage, L.P. and Brown, D.D. (1972). J. Mol. Biol., 70, 637.

Swart, L.S. (1973). Nature New. Biol., 243, 27.

Taylor, E.W. (1972). Ann. Rev. Biochem., 41, 577.

Taylor, W.J., Easley, C.W. and Kitchen, H. (1972). J. Biol. Chem., 247, 7320.

Terry, W.D., Hood, L.E. and Steinberg, A.G. (1969). Proc. Natd. Aca. Sci. U.S., 63, 71.

Thompson, E.O.P. and O'Donnell, I.J. (1965). Aust. J. Biol. Sci., 18, 1207.

Tichy, H. Genetics, 74 Part 2. (June Sup.).

Traugott, C. and Massaro, E.J. (1973). Biochim. Biophys. Acta., 295, 549.

Tung, J. and Knight, C.A. (1972), Anal. Biochem., 48, 153.

Vithayathil, P.J. and Richards, F.M. (1960). J. Biol. Chem., 235, 2343.

Voitkevich, A.A. (1966). The Feathers and Plumage of Birds, London, Sidgwick and Jackson.

Vogel, A.I. (1962). Quantitative Inorganic Analysis. Longmans. 234.

Watterson, R.L. (1942). Physiol. Zool., 15, 234.

Weber, K. and Osborn, M. (1969). J. Biol. Chem., 244, 4406.

Weinberg, E.S., Birnstiel, M.L., Purdom, I.F. and Williamson, R. (1972). Nature (London), 240, 225.

Wessells, N.K. (1965). Devel. Biol., 12, 131.

Williams, J.G. and Gratzer, W.B. (1971). J. Chromatogr., 57, 121.

Woodin, A.M. (1954). Biochem. J., 57, 99.

Woods, E.F. (1971). Comp. Biochem. Physiol., 39A 325.

Woods, K.R. and Wang, K. -T. (1967). Biochim. Biophys. Acta. 133, 369.