THE AMINO ACID SEQUENCE OF CHICKEN HISTONE F3

by

W. F. Brandt

Submitted in fulfillment of the requirements

for the degree of

Doctor of Philosophy,

in the

Faculty of Science,

University of Cape Town.

December, 1972.

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my supervisor, Professor C. von Holt, for his enthusiastic encouragement given during all phases of this project. His valuable guidance and helpful assistance are greatly appreciated.

I also wish to thank Margaret Morgan for preparing all the diagrams in this thesis, and Valerie Green and Diana Loureiro for typing the manuscripts.

General assistance rendered by the staff of the Biochemistry Department, UCT, is gratefully acknowledged.

Finally, I wish to express my gratitude to the Council for Scientific and Industrial Research and the University of Cape Town, without whose financial assistance this work would not have been possible.

CERTIFICATION OF SUPERVISOR

In terms of paragraph 8 of "Regulations for the Degree of Ph.D." I, as supervisor of the candidate, W.F. Brandt, certify that I approve of the incorporation in this thesis of material that has already been published or submitted for publication.

Signed by candidate

Professor C. von Holt
Head of the Department of Biochemistry

ABBREVIATIONS AND SYMBOLS

BSA	=	N,O-bis(trimethylsilyl)acetamide
CNBr	=	cyanogen bromide
DNS Dansyl	=	1-dimethylaminonaphthalene-5-sulfonyl-
DMAA	=	N-dimethylallylamine
DMAP	=	3-dimethylamino-1-propyne
DNH	=	Deoxyribonucleohistone
DNP	=	Deoxyribonucleoprotein
EDTA	=	Ethylene diaminetetraacetic acid
HFBA	=	heptafluorobutyric acid
NBS	=	N-bromosuccinimide
PCMB	=	p-chloromercuribenzoate
PTC	=	phenylthiocarbamyl-
PTH	=	phenylthiohydantoin-
PITC	=	pehnylisothiocyanate
Quadrol	=	N,N,N',N'-tetrakis-(2-hydroxypropyl)-ethylenediamine
SDS	=	sodium dodecyl sulfate
TMS	=	trimethylsilyl-
Tris	=	tris-(hydroxymethyl)-amino methane
A_{230}	=	absorbance at 230 nm, 1 cm optical pathlength
$A_{230}^{1\%}$	=	absorbance of 1% (w/v) solution at 230 nm and optical pathlength of 1 cm
T	=	trace

SUMMARY

Histone F3 (III) from chicken erythrocytes was isolated by selective extraction from nucleoprotein with ethanolic-HCl and purified by a single gel filtration step. This protein was found to be homogeneous by the following criteria : gel filtration, electrophoretic mobility, N- and C-terminal amino acid residues and amino acid analysis.

The primary structure of this histone was established without resorting to the use of overlapping sequences. This has been achieved with specific chemical cleavages rather than enzymatic degradations chosen and applied, first to the original protein chain, and subsequently to the generated polypeptides, to yield sets of not more than 3 peptides in any single cleavage. Their relative position in the protein or polypeptides became evident after comparison of the N- and C-terminal amino acids in the cleavage products and the uncleaved starting material.

The simplicity of the peptide mixture after each cleavage, resulting in easy separation of the peptides, together with the highly efficient Edman degradation of automatic sequencing, allowed a rapid and relatively non-laborious primary structure determination.

Finally, the amino acid sequence is compared with those of protamines and other histones. The evolution and the structure of this protein in relation to DNA is briefly considered.

CONTENTS

INTRODUCTION

In recent years the remarkable advancements in molecular biology have been highlighted by the clarification of the role of DNA for the biosynthesis of messenger, ribosomal and transfer RNA. Now that the mechanism of gene action is fairly well understood the important question has been raised as to the mechanism of gene control in the eukaryotic cell. The cell nucleus contains the total genetic potential of the complete organism. In ways that are not yet clear, the nucleus responds to extracellular changes and stimuli by a differential gene activation which is the primary basis for the specialization of cells. Different regions of the chromosomal DNA are being transcribed into RNA in different cell types and in the same cell at different phases of its development. It follows that individual cells must repress the transcription of most of their DNA, while they selectively activate a relatively small number of genes for the synthesis of ribonucleic acids which direct the assembly of enzymes and structural proteins characteristic of the cell type.

A variety of substances have been implicated as candidates for the function of such gene repressors, noticeably histones (Stedman & Stedman, 1943; Huang & Bonner, 1962; Allfrey et al., 1963), and more recently nuclear acidic proteins (Mirsky & Ris, 1950; Stellwagen & Cole, 1969). Histones comprise a major fraction of the genetic apparatus - occurring in amounts comparable to that of DNA itself and were found to inhibit the activity of DNA as a primer for RNA synthesis (Allfrey et al., 1963; Huang, Bonner & Murray, 1964).

It is now, however, well established that histones which were initially thought to be highly complex consist of only five main groups of related proteins (Johns, 1967a), occurring in relatively constant amounts and uniform proportions in the chromatin of different cell types. In view of this striking uniformity in histone distribution and composition, it has been suggested (Panyim & Chalkley, 1969b; Littau et al., 1965) that histones are primarily structural proteins since

of histone (Hnilica & Bess, 1965; Phillips, 1967; Starbuck et al., 1968; Fambrough & Bonner, 1968). It has resisted purification and no structural work had been published. The presence of cysteine in histone F3 may well be of particular functional significance with respect to protein-protein interaction in the various functional stages of chromatin. Changes in protein sulfhydryl-disulfide content have been reported during gene activation (Stocken, 1966; Sadgopal & Bonner, 1970a,b). In view of this the primary structure of histone F3 seemed of special interest.

PART 1

THE ISOLATION, PURIFICATION AND CHARACTERIZATION OF F3 HISTONE FROM CHICKEN ERYTHROCYTES

1.1 INTRODUCTION

Kossel, in 1884, isolated a "basic substance" by extraction of goose erythrocyte nuclei with HCl and suggested the name "Histon". Since then histones have been shown to exist, associated with DNA, in all the somatic cell nuclei examined of all multicellular and more complex unicellular organisms (Johns, 1971b). Stedman and Stedman, in 1950, showed that histones consist of a number of proteins and postulated that histones might play a role in gene regulation. This started numerous groups to isolate, purify, characterize and study the function of histones. A large number of different techniques were used in the isolation and fractionation of these proteins resulting in the description of a larger number of different histones, pointing to an extreme complexity (Table 1.1).

This situation was considerably clarified by the application of gel electrophoresis and N-terminal group analysis for further characterization of histones (Neelin & Connell, 1959; Phillips, 1958). It then became obvious that histones consisted of five main groups of closely related proteins and that the complexity was caused by merely extracting different combinations. Further complication arose through the ease with which histones are degraded by proteolytic enzymes and the tendency of histones to aggregate (Panyim et al., 1968; Cruft, 1958).

In Table 1.1 the various histone nomenclatures are listed (Johns, 1971a). Throughout this investigation Johns' (1964) nomenclature is being used. According to this nomenclature the 5 main fractions have the following designations : F1, F2b, F2a1, F2a2 and F3.

TABLE 1.1

A COMPARISON OF THE VARIOUS HISTONE NOMENCLATURES

References	Fraction				
Johns (1964)	F_1	F_{2b}	F_{2a2}	F_{2a1}	F_3
Cruft et al. (1958)	α	γ		β	
Davidson (1957)	F_2	F_3			
Phillips & Johns (1959)	F_2				F_3
Luck et al. (1958)	Ia,Ib	IIa,IIb,IIc		III+IV	
Johns et al. (1961)	E_2	E_1			E_3
Mauritzen et al. (1967)			6,7,8	9	5
Fambrough et al. (1968)	1	IIa,IIb		IV	III
Starbuck et al. (1968)			AL	GAR	
Bellair et al. (1967)			β_6	β_7	β_5
Panyim & Chalkley (1969)	1,1',1''	3	4	5,5'	2,2',2''
Cruft et al. (1957)	$\alpha_1\alpha_2\alpha_3$	0.85γ	1.65γ and β		

(Johns, 1971a)

To prepare histones from any given tissue it is a necessary prerequisite to isolate deoxyribonucleoprotein (DNP) in a relatively pure form. For some tissues, e.g. calf thymus, because of a very small ratio of the volumes of cytoplasm to nucleus, this is a fairly simple procedure consisting merely of disrupting the cell and nuclear membrane, centrifuge down the

DNP and washing it a number of times in dilute salt solution.

An approach often adopted for the preparation of DNP from cells with a large cytoplasm-nucleus ratio commences with the isolation of nuclei. These are then mechanically disrupted and the sedimented DNP washed extensively with 0.14 M NaCl which removes most non-histone nuclear proteins. Wide variations in the details in these steps exist. Instead of saline some authors prefer the use of saline-citrate (Murray et al., 1968). These washings might be followed up with dilute tris-buffer (Hnilica, 1967) or a more concentrated NaCl solution to remove ribosomal proteins (Johns & Forrester, 1969).

From this DNP histones can be extracted with dilute acid which renders the DNA insoluble. The histones extracted are precipitated by alcohol or acetone. Histones isolated in this way were found to be undenatured and have been extensively used in chemical, conformational and recombination studies (Zubay & Wilkins, 1962). Alternatively, the DNP complex can be dissociated and solubilized using high salt concentration and the DNA and histone subsequently separated by ultra-centrifugation (Bauer & Johanson, 1966).

There are essentially two main approaches to the fractionation of histones :

(1) Histones are isolated as a group by one of the above procedures and then subjected to the various protein separation methods, for example, column chromatography on amberlite IRC-50 (Fambrough & Bonner, 1966), or carboxymethylcellulose (Johns, 1964), gel filtration (Mauritzen et al., 1967), countercurrent distribution (Butler et al., 1967) and precipitation techniques (Bijvoet, 1957). Due to the closely related properties of histones no single method is capable of a complete resolution of the histones into the five main groups.

(2) Alternatively, histones have been fractionated by selective extraction procedures from DNP. Increasing amounts of acid or NaCl have been used with limited success (Murray, 1966; Wilhelm & Champagne, 1969). A better selectivity in the extraction of histones is obtained by using mixtures of organic solvents and dilute acid as extracting solution,

taking advantage of the differences in the polarity of the various histones (Johns, 1967b). Further purifications of these fractions, however, is generally required. (For a more detailed review on histones and histone fractionation see Luck (1964) and Busch (1965)).

The application of these procedures resulted in the purification of histone F2a1 from calf thymus and the elucidation of its primary structure by the end of 1969 (DeLange et al., 1969a).

To judge from the number of publications, F3 histone has been studied rather less than other fractions. This may be due to the fact that this fraction forms dimers and aggregates and results were more difficult to interpret (Cruft et al., 1958; Phillips, 1967; Mauritzen et al.,1967; Neelin & Neelin, 1960; Bellair & Mauritzen, 1965).

Hnilica and Bess (1963) fractionated F3 histone into three distinct fractions. Results of the amino acid analysis, of the N-terminal analysis, and of peptide mapping indicated that these fractions were polymers or aggregates of a basic molecular unit. The molecular weight was found to be near 50,000 based on N-terminal analysis and gel filtration. More recent estimates also based on N-terminal analysis are in the range of 10,000 - 20,000 (Phillips & Simpson, 1969). Contradictory reports on the presence of cysteine in histone F3 from calf thymus appeared. Daly and Mirsky (1955) and Phillips (1965) observed the presence of cysteine in the arginine-rich histones (F2a, F3) and not in the lysine-rich histones (F2b, F1). In contrast, Hnilica and Bess (1965) and Fambrough and Bonner (1966) failed to find cysteine in histone F3 and considered it to be free of this amino acid. Panyim and Chalkley (1969b) showed that F3 histone after disaggregation in urea solutions could be resolved by very long polyacrylamide gels into three distinct electrophoretic bands.

Recently (Yokotsuka & Shimura, 1969) subfractionated F3 histone from calf thymus by ion exchange chromatography and countercurrent distribution into three fractions. Their amino acid composition varied widely and the main fraction

was estimated to contain 101 residues (MW 11,000) including 1 cysteine and a N-terminal alanine.

In spite of the conflicting results on the homogeneity of histone F3 and its molecular size, it seemed, however, well established that histone F3 from calf thymus is rich in arginine, possesses alanine as both N- and C-terminal amino acid (Phillips & Simpson, 1969), and is the only cysteine containing histone (Phillips, 1965; Sadgopal & Bonner, 1970b).

Histone F3 from pea shoots was shown by gel electrophoresis, amino acid analysis and peptide mapping to be closely related to that of calf thymus, except that its cysteine content was lower (Fambrough & Bonner, 1968).

Although chicken erythrocyte histones have been studied, little work has been done on subfractionation of the various histones. Attempts to subfractionate the histones by selection and extraction (Murray et al., 1968) and gel exclusion chromatography (Bellair & Mauritzen, 1967) did not seem to yield homogeneous fractions.

Since the results on the properties of histone fraction F3 described by various authors varies widely, the first step in the determination of the primary structure was to develop a method which would yield highly purified protein on a large scale

1.2 PREPARATION OF CALF THYMUS HISTONES

Calf thymus histones have been intensively studied (Table 1.1) and have, for this reason, been isolated to serve as an arbitrary standard and aid in the identification of histone fractions from other species.

DNP was isolated from calf thymus (4.1.1), histones extracted from the DNP with dilute HCl and recovered by precipitation with acetone (4.1.3). No attempt was made to exclude O_2 from the various solutions used in the isolation.

Gel electrophoresis (4.3.1) of calf thymus histone results in 4 main fractions (Fig. 1.1, gel 1). After reducing this preparation with mercaptoethanol prior to the electrophoresis a fifth band appeared (Fig. 1.1, gel 2), and simultaneously the concentration of the fraction with the mobility of F2b diminishes. This electrophoretic pattern corresponds to

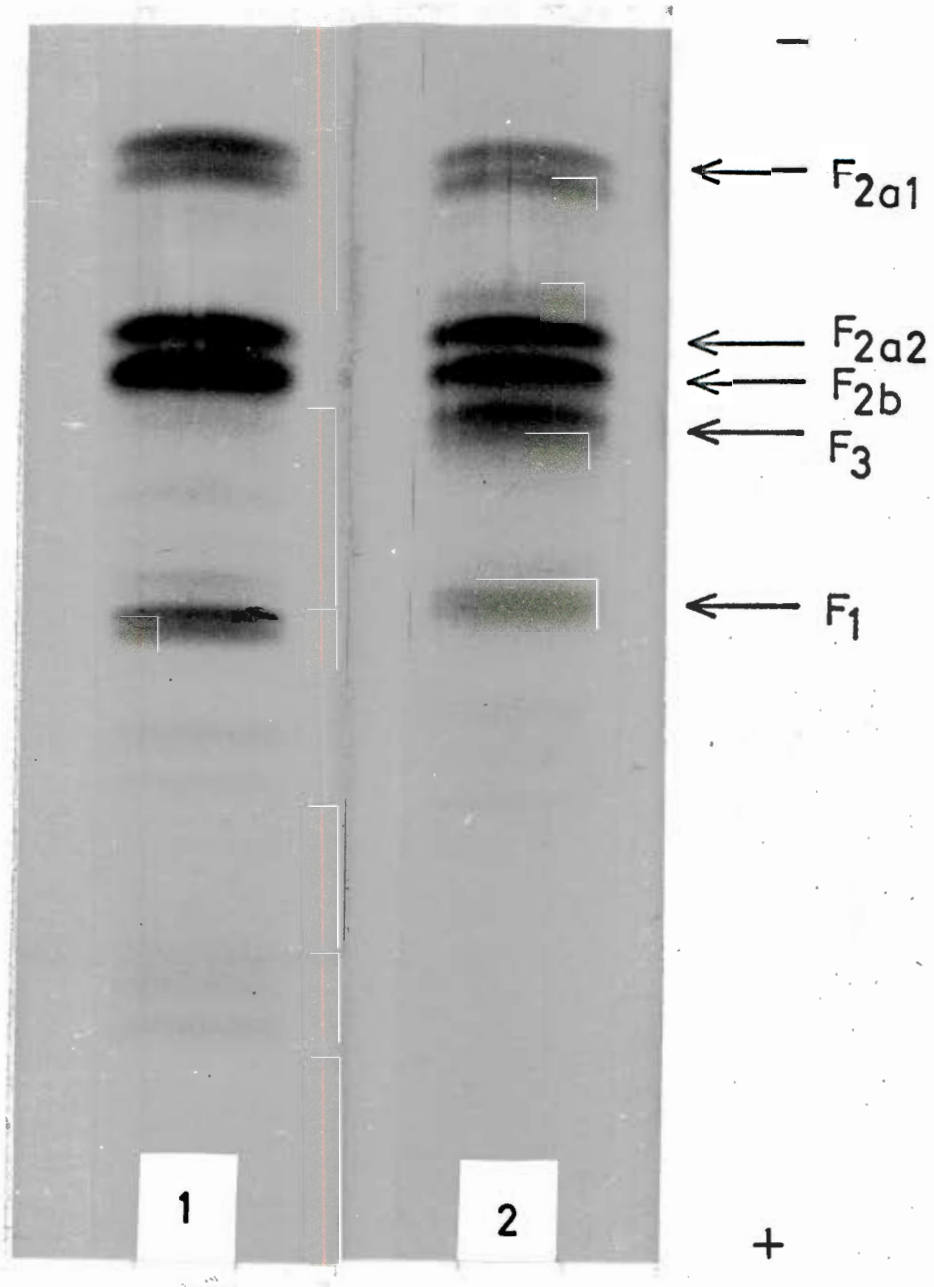


Fig. 1.1 : Polyacrylamide gel electrophoresis of calf thymus histones : (1) Isolated without excluding atmospheric O_2 and (2) after reduction with mercaptoethanol. Gels were run for 3.5 hours in 15% polyacrylamide gels (4.3.1).

that described by Panyim and Chalkley (1969a,b).

From Fig. 1.1 it is evident that histone F3 in its oxidized state forms a product which has an increased electrophoretic mobility identical to that of histone F2b. Only sometimes the presence of small amounts of a fraction with lower mobility, possibly a dimer, were observed. This may indicate that calf thymus histone F3 contains at least two cysteines sterically arranged in such a way that on oxidation an internal disulphide bond is formed resulting in a conformational change of the molecule increasing its electrophoretic mobility in polyacrylamide gels. Similar observations were made by Chalkley (1970).

In contrast, Fambrough and Bonner (1968) found that calf thymus histone F3 forms a series of polymers. This polymerisation behaviour has, however, never been repeated in other laboratories (Panyim et al., 1970) and it seems likely that the histone fraction used by these authors was contaminated with sulfhydryl containing non-histones (Sadgopal & Bonner, 1970b).

1.3 PREPARATION OF CHICKEN ERYTHROCYTE HISTONES

The large erythrocyte cytoplasm, filled predominantly with haemoglobin, makes it necessary to isolate nuclei prior to the deoxyribonucleoprotein (DNP). Erythrocytes were therefore lysed with digitonin, the nuclei centrifuged down and washed extensively with saline-citrate (4.1.2.1).

The crude DNP was obtained by mechanically disrupting the nuclei, sedimenting the insoluble fraction by centrifugation and repeated washings of the sediment with saline-citrate to which bisulphite had been added to prevent proteolytic degradation of histones (4.1.2.2).

The histones were isolated by extracting the DNP with dilute acid (4.1.3). The electrophoretic pattern of chicken erythrocyte histones in comparison to that of calf thymus histones is shown in Fig. 1.2.

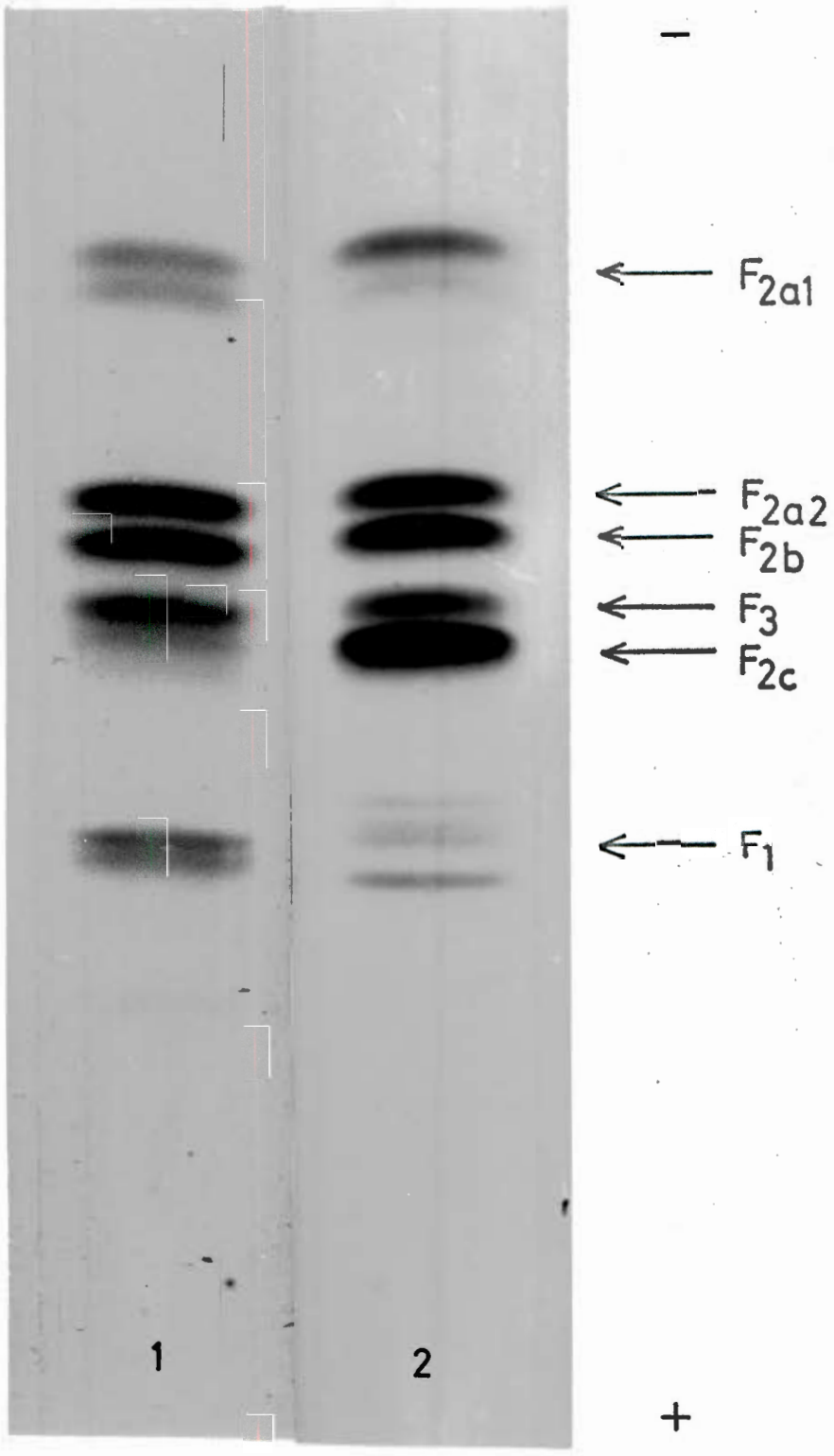


Fig. 1.2 : Polyacrylamide gel electrophoretic pattern of the histones from (1) calf thymus and (2) chicken erythrocytes. Both samples were reduced with mercaptoethanol and run for 3.5 hours in 15% polyacrylamide gels (4.3.1).

From Fig. 1.2 it is evident that the histone pattern in the two animals is very similar. The corresponding histone fractions have identical electrophoretic mobilities, although F1 in chicken is more complex and F3 is partly obscured by an additional band which is believed to be the avian erythrocyte specific histone F2c, reported by Hnilica (1964) and Vidali and Neelin (1968).

1.4 PREPARATION AND PURIFICATION OF HISTONE F3 FROM CHICKEN ERYTHROCYTES

1.4.1 Preparation of crude histone F3

Histones, with the exception of F1, have closely related properties. This is due to the fact that they all have very similar molecular weights and similar amounts of polar amino acids (Phillips, 1963; Butler et al., 1968). The binding strength of the various histones to DNA, however, varies considerably. For the purification this difference in binding was exploited by using a selective extraction described by Johns (Johns, 1967b; Johns et al., 1960). The arginine-rich histones (F2a1, F2a2, F3) were extracted from the crude deoxyribonucleoprotein with ethanolic-HCl (4.1.4). A further fractionation is achieved by selectively precipitating the histones from this extract with increasing amounts of ethanol and acetone.

The histone F3 fraction was isolated by dialysing the ethanolic-HCl extract against three changes of 2 volumes of absolute ethanol. This fraction was found to be contaminated by substantial amounts of fraction F2a2, and to a lesser extent, by fraction F2a1. A considerably purer fraction was obtained by harvesting the precipitate after a shorter dialysis period (4.1.4), although the yields were lower.

Fig. 1.3 shows the electrophoretic patterns of the various fractions obtained from the ethanolic-HCl extract by selective precipitation with ethanol and acetone (4.1.4).

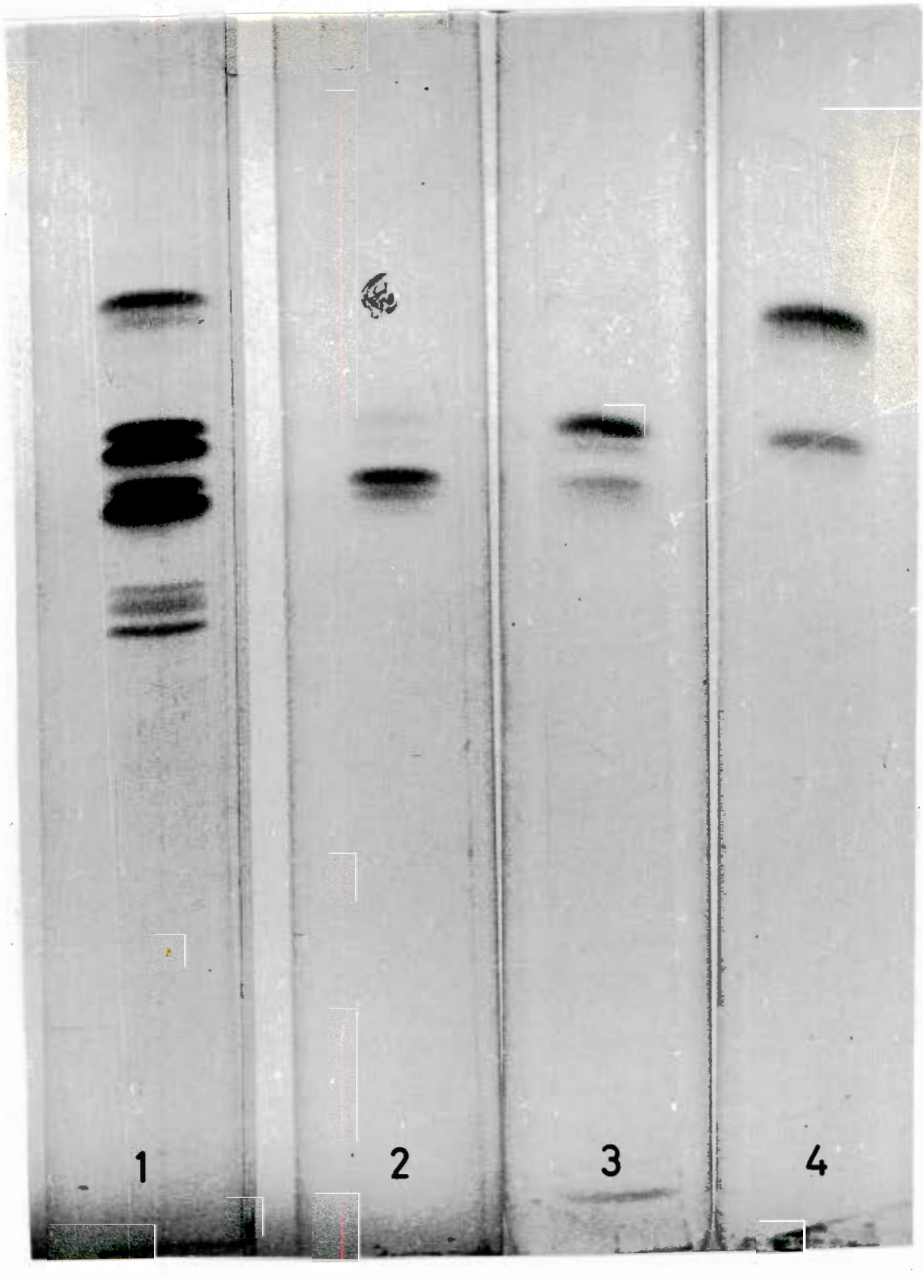


Fig. 1.3 : Polyacrylamide gel electrophoretic pattern of histone fractions obtained by selective extraction with ethanolic-HCl (80% (v/v) ethanol in 0.25 N HCl) followed by selective precipitation. (1) Chicken erythrocyte histones, (2) fraction obtained after 2 x 4 hour dialysis against 2 volumes of ethanol (F3), (3) after a further 18 hours against fresh ethanol, (4) and finally after 3 volumes of acetone. All samples were reduced with mercapto-ethanol and run for 3.5 hours in 15% polyacrylamide (4.3.1).

1.4.2 Purification of F3 histone

It is evident from the gel electrophoresis that the crude histone F3 fraction is mainly contaminated by histone F2a2 (Fig. 1.3, gel 2). An attempt to repeat the selective precipitation with ethanol was not very successful since most of the crude fraction did not dissolve in ethanol-HCl.

The crude fraction was also subjected to gel filtration on a Sephadex G-100 column. Histones are known to aggregate to large complexes in solution (Cruft et al., 1958; Hnilica & Bess, 1965; Phillips, 1965; Johns, 1968). The protein preparation was therefore dissolved in 6 M urea before it was applied to the column. All the protein was eluted in a single fraction (Fig. 1.4a) and gel electrophoresis confirmed that no separation had been achieved.

During these studies it was noticed that when mercaptoethanol, which was routinely used, was occasionally omitted in the electrophoresis of the crude F3 histone an additional band with lower mobility appeared. This seemed to indicate that F3 histone might contain a single cysteine residue which readily establishes a disulphide bridge under oxidative conditions resulting in the formation of a dimer.

Dimerization will result in an increase in the molecular dimension. After oxidation it should thus be possible to separate the contaminating proteins with lower molecular weights from histone F3 by gel filtration.

In a pilot experiment several mg of fraction F3 were oxidized in 6 M urea at pH 8 by passing O_2 through the solution. This solution was applied to the Sephadex G-100 column (Fig. 1.4b).

Two fractions were eluted. Peak (A) was found on electrophoresis to be pure F3 dimer. When this oxidation was repeated on a large scale yields of the dimer were found to be low. However, o-iodosobenzoate, an oxidation reagent thought to be specific in converting sulfhydryl groups to disluphide bonds (Hellerman et al., 1941, 1943) nearly quantitatively converts F3 histone to the dimer (4.2.3). The dimer was then purified on a Sephadex G-100 14.5 x 100 cm

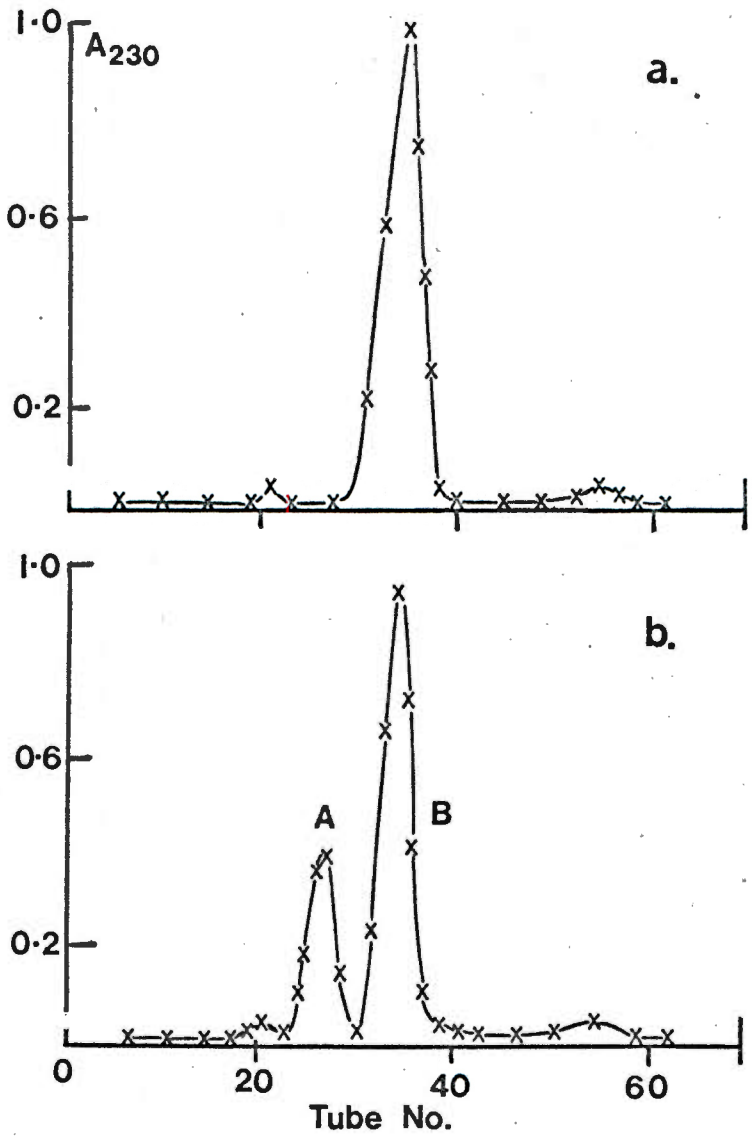


Fig. 1.4 : Elution pattern of a) unoxidized and b) oxidized crude histone F3 from a Sephadex G-100 1.5 x 90 cm column. Samples (6 mg) were applied in urea to the column. The oxidation was performed by adjusting the urea solution to pH 8 and introducing O_2. The eluent used was 0.01 N HCl, the fraction volume 1.3 ml and the flow rate 6 ml/h (4.2.1).

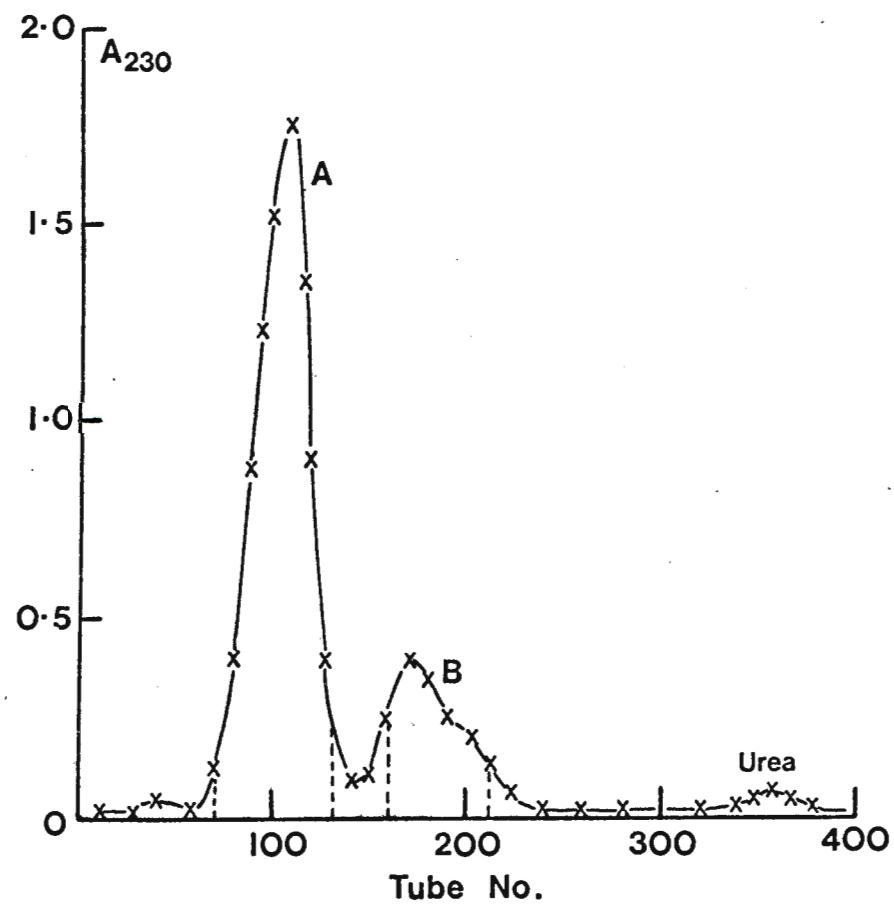


Fig. 1.5a : Elution pattern of crude histone F3 (700 mg) from a Sephadex G-100 14.5 x 100 cm column after oxidation with o-iodosobenzoate (4.2.3). The eluant used was 0.01 N HCl , the fraction volume 25 ml and the flow rate 200 ml/h.

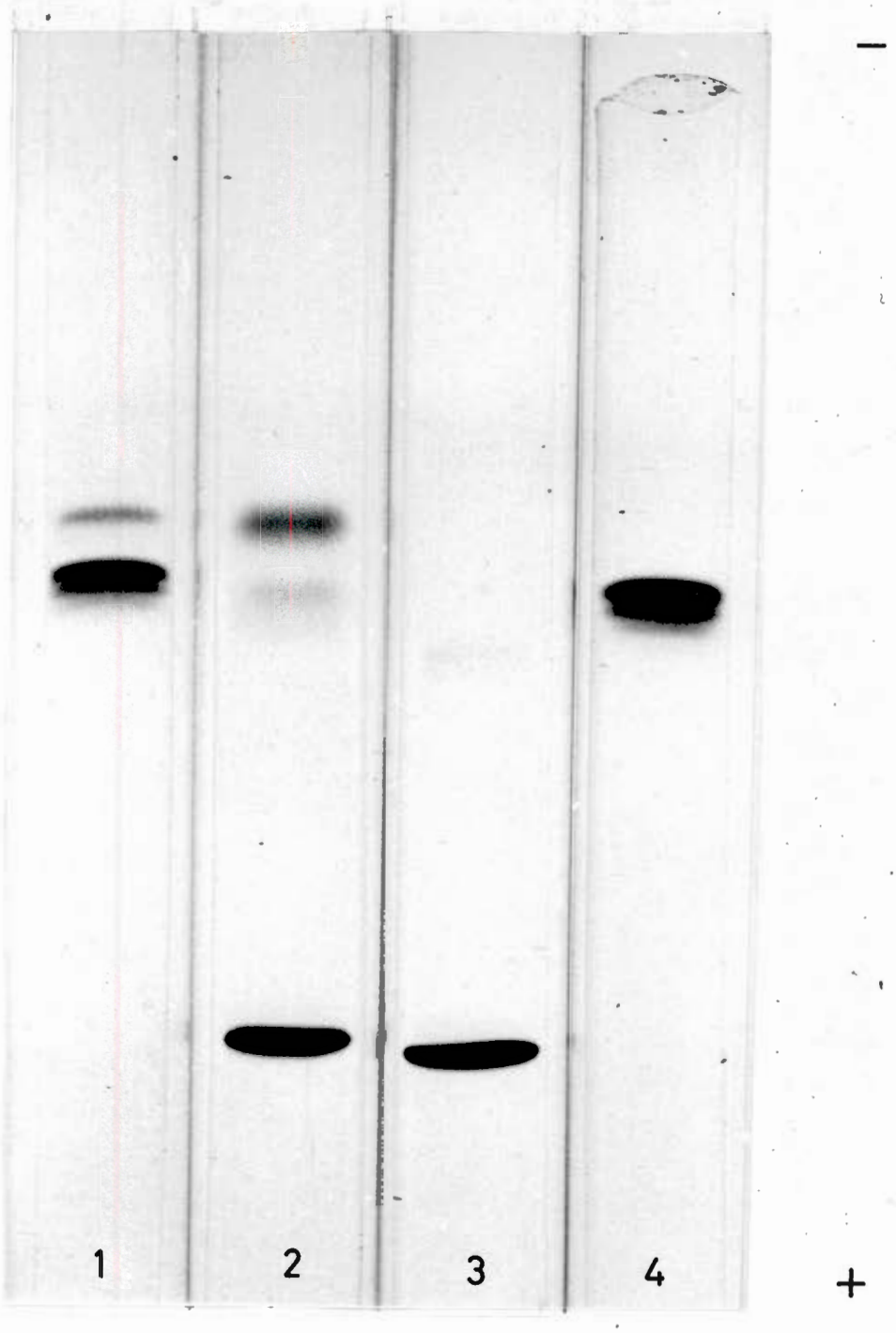


Fig. 1.5b : Electrophoretic patterns of the various fractions obtained in Fig. 1.5a. Gel 1 corresponds to crude histone F3, 2 to crude histone F3 after treatment with o-iodosobenzoate, 3 to fraction A (Fig. 1.5a) and 4 to fraction A (Fig. 1.5a) reduced with mercaptoethanol. Fraction B (Fig. 1.5a) consisted of histone fraction F2a2 and traces of F3 histone (gel 2).

on a gram scale (Fig. 1.5).

From Fig. 1.5b, gel 3, it is evident that this protein is in a highly purified form and no further purification seemed necessary. Fig. 1.6, gel 5, shows the gel electrophoretic pattern of F3 histone after the dimer had been reduced with mercaptoethanol. The electrophoretic microheterogeneity reappears and is very similar to that of the corresponding fraction from calf thymus histone. This microheterogeneity can either be due to the presence of a family of closely related proteins with small variations of their amino acid sequences or to a fractional modification of the amino acid side chains of one sequence-homogeneous protein. It will be shown later that fractional acetylation is the cause of the microheterogeneity (1.5.3 and 1.5.4).

1.5 CHARACTERIZATION OF HISTONE F3 FROM CHICKEN ERYTHROCYTES

1.5.1 Endgroup analysis

N-terminal group analysis was performed by dansylating the protein followed by releasing the labelled amino acids by acid hydrolysis. C-terminal amino acids were determined by hydrazinolysing the protein and dansylating the liberated free amino acids. In both cases the only amino acid that could be detected by thin layer chromatography was DNS-Ala (4.3.6, 4.3.7).

1.5.2 Molecular weight determination

The electrophoretic mobility of sodium dodecyl sulphate (SDS)-complexes of F3 histone was compared to the SDS-complexes of F2a1 and a number of standard proteins (Fig. 1.7a) (4.3.3.1). The estimated molecular weight of histone F3 was between 18 - 20,000. The molecular weight of F2a1 was found to be between 15 - 17,000 which exceeds the actual value of 11,200 by 30% (DeLange et al., 1969a). It appears therefore that the electrophoretic mobility of SDS-histones behave anomalously. Such behaviour of SDS-proteins in general is not fully understood (Dunker & Rueckert, 1966). Possible reasons might be

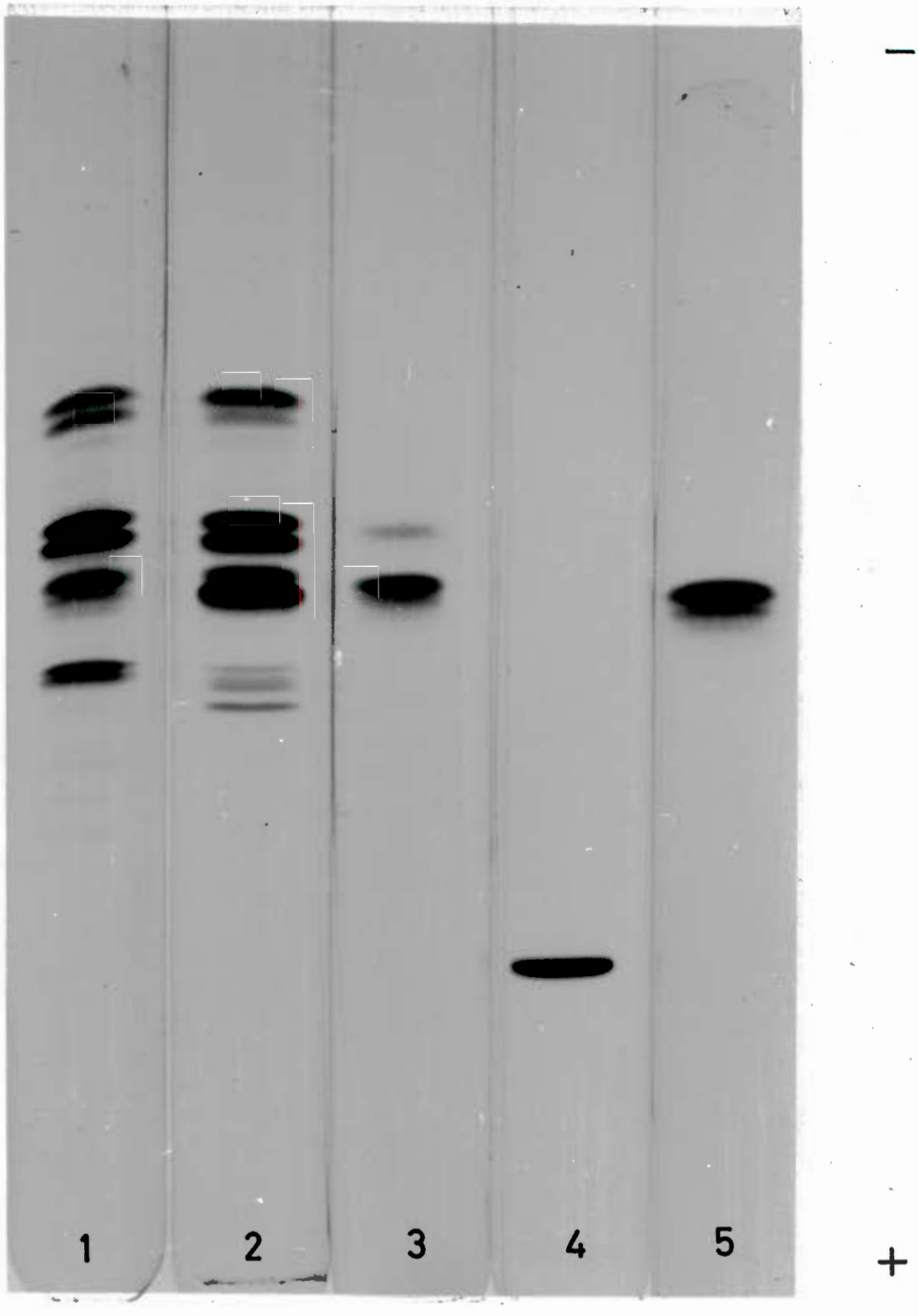


Fig. 1.6 : Electrophoretic pattern of chicken histone F3 at the various purification stages. (1) Calf thymus histones, (2) Chicken erythrocyte histones, (3) Crude histone F3 obtained by selective extraction of DNP and selective precipitation, (4) F3 dimer obtained after gel filtration, (5) Histone F3 obtained after reduction of the dimer with mercaptoethanol. All gels were run for 3.5 hours in 15% polyacrylamide (4.3.1).

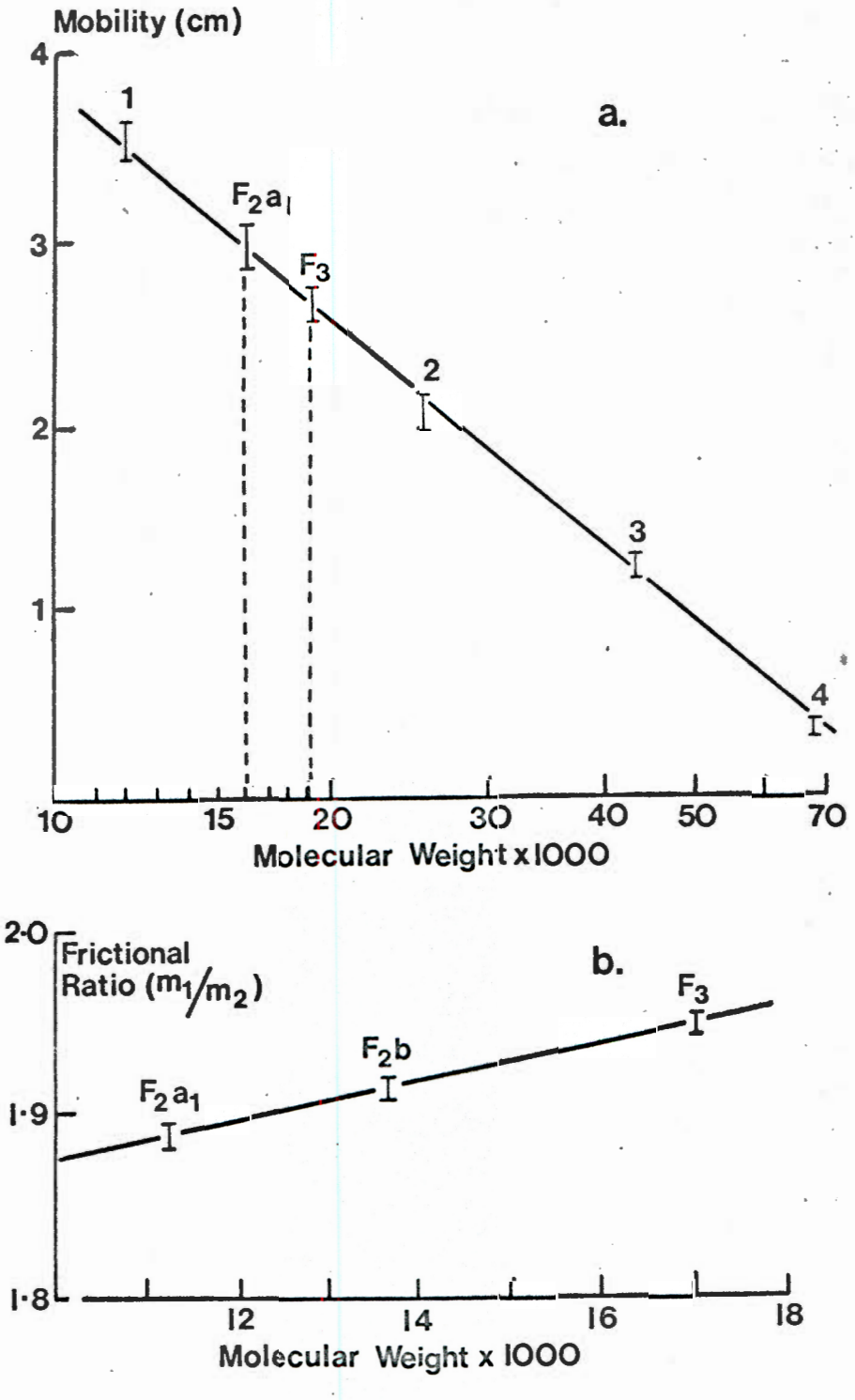


Fig. 1.7 : Determination of the molecular weight of histone F3.

a) Plot of the logarithm of the molecular weights of (1) cytochrome C (2) chymotrypsin (3) ovalbumin (4) ovalbumin dimer versus their relative electrophoretic mobility of the SDS-complexes (4.3.3.1).

b) Plot of the molecular weight of various histones versus their frictional coefficient in 10% and 15% polyacrylamide gels (4.3.3.2).

that histones contain about 15% more basic amino acids than the average protein, resulting in the reduction of the negative charge of the SDS-complex. Alternatively, it could be due to the high axial ratios of histones. Molecular weight estimates by gel filtration (Fig. 1.4) are in the region of 50,000 (Hnilica & Bess, 1965). After applying a correction based on the results obtained for F2a1 one would arrive at an estimate for the molecular weight of histone F3 of 13,000 - 15,000.

The frictional coefficients of histone F2a1 and F2b of known molecular weight on different percentage gels were compared to the frictional coefficient of F3 (4.3.3.2). The molecular weight deduced for histone F3 was near 17,000 (Fig. 1.7b).

1.5.3 Acetyl content of histone F3

It is very well established that calf thymus histones contain acetyl groups (Phillips, 1963). In order to determine the presence and the amount of acetyl groups in F3 histone the protein was hydrolysed in 5.7 M HCl followed by extracting the liberated acetic acid with ether. This extract was subjected to gas chromatographic analysis (4.3.8). Acetic acid was shown to be present (Fig. 1.8). Correcting for the incompleteness of the extraction, approximately 0.4 moles acetic acid per 16,000 g of protein (approximately 1 mole) were found to be present in the hydrolysate. This agrees very well with values found in calf thymus F3 histone by Phillips (1963). The presence of acetic acid in the hydrolysate suggests the possibility that microheterogeneity of histone F3 (Fig. 1.6) may be due to fractional acetylation.

1.5.4 Enzymatic deacetylation of histone F3

DeLange et al. (1969a) found, during sequence studies on histone F2a1 (calf), that the lysine in position 16 is 50% acetylated. The two bands of F2a1 evident after gel electrophoresis are probably due to histone F2a1 and acetylated F2a1 molecules (Panyim & Chalkley, 1969b) (Fig. 1.2). A similar situation might pertain to histone F3. Acetylation of a single lysine residue would reduce the positive charge

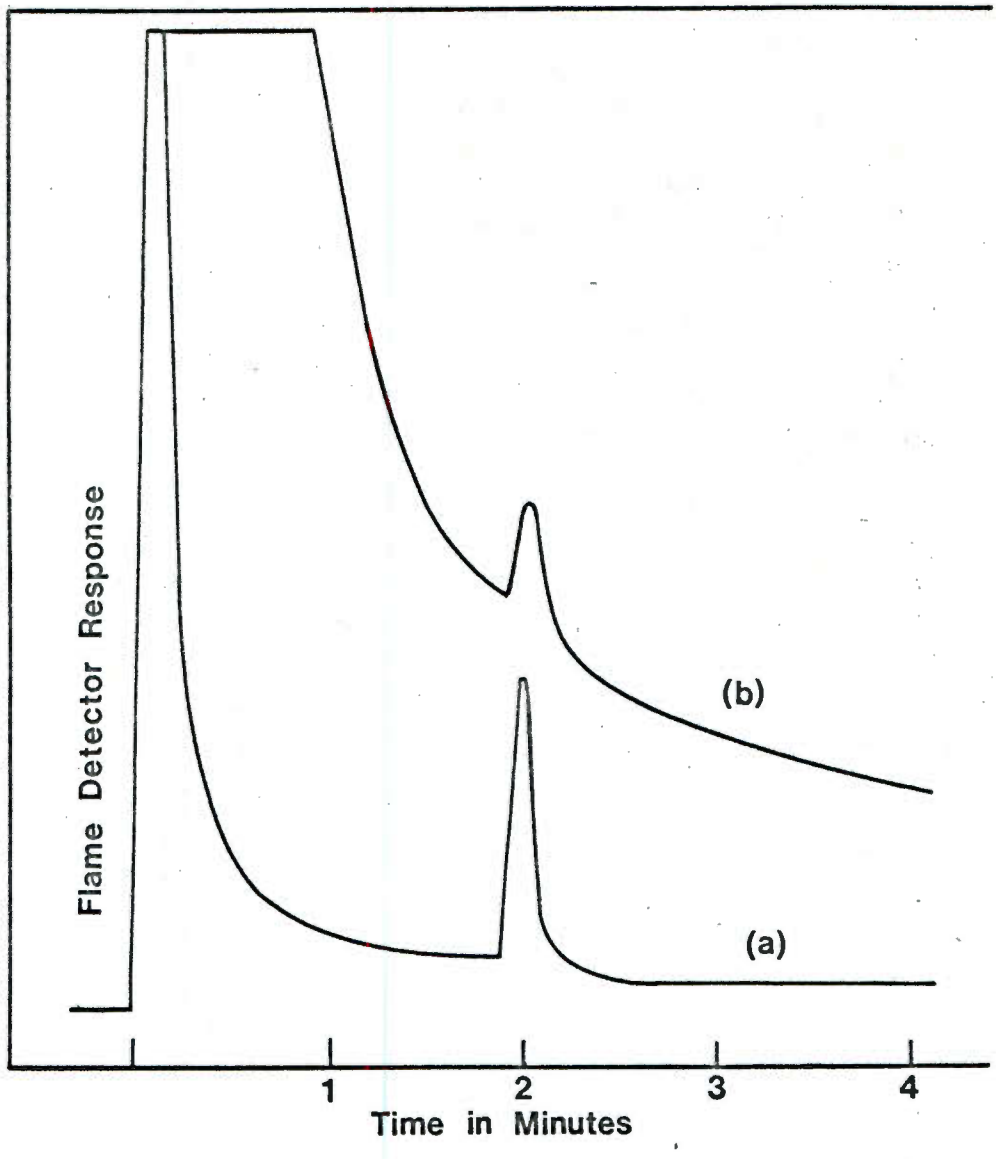


Fig. 1.8 : Gas chromatographic identification of acetic acid. a) 43 nmoles acetic acid at a detector range of 1 x 10^{-9} amp. b) 1.6% of the ether extract of an acid hydrolysate of histone F3 (10 mg) at 3 x 10^{-11} amp. detector range (4.3.8).

of the molecule by one and assuming the presence of 30 basic residues, reduce the mobility for every acetylated residue present by 3%. The differences in electrophoretic mobilities of the histone F3 subfractions (Fig. 1.6, gel 5) are in the order of 2 - 3%. To prove that the microheterogeneity is caused by acetyl groups the histone was incubated with calf thymus histone deacetylase (4.3.9).

To the incubation mixture a small amount of histone F2a2 was added to monitor for proteolytic activity of the enzyme preparations. In Fig. 1.9a the electrophoretic pattern of histone F3 is compared to that of the enzyme treated histone. After an incubation period of one hour only a few percent of the slower moving electrophoretic bands remain (Table 1.2). Together with the gas chromatographic evidence (1.5.3), this proves that the electrophoretic microheterogeneity of histone F3 is caused by acetylation of lysine residues. The two minor bands represent probably mono and diacetylated histone F3 molecules.

TABLE 1.2

ENZYMATIC DEACETYLATION OF HISTONE F3

Sub-fraction	Histone F3 untreated		Histone F3 treated with histone deacetylase	
	Peak area[1]	% of total	Peak area[1]	% of total
F_3^0	1350	67	1720	88
F_3^1	450	23	174	9
F_3^2	198	10	54	3
Total	1998		1948	

1. The electrophoretic bands from Fig. 1.9a were quantitated from the densiometric tracing of Fig. 1.9b and 1.9c.

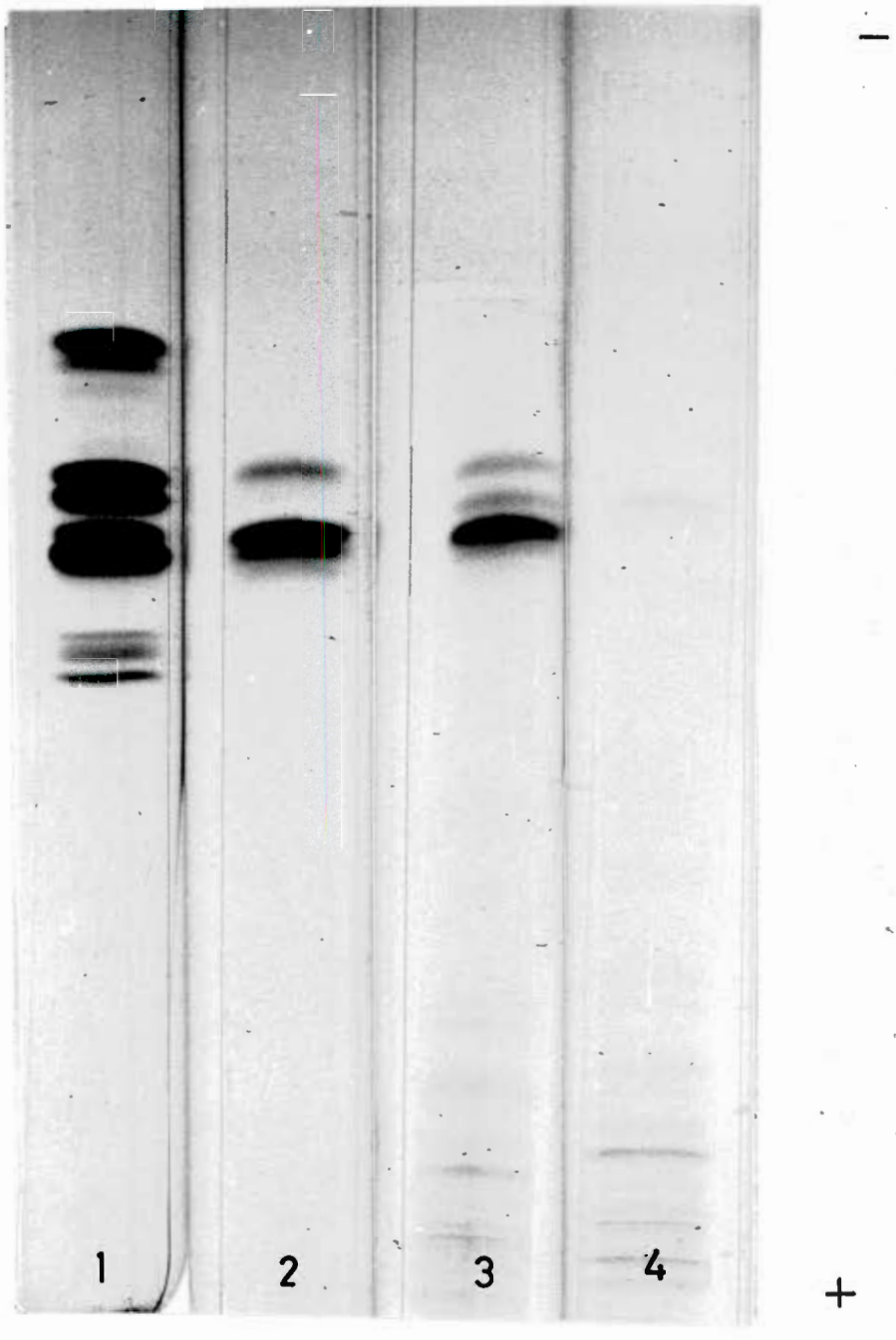


Fig. 1.9a : Enzymatic deacetylation of histone F3. (1) Total chicken erythrocyte histones, (2) histone F3 and F2a2 before (3) and after treatment with histone deacetylase. Histone F2a2 was added as internal standard. (4) Histone deacetylase (4.3.9). All samples were reduced with mercaptoethanol and electrophoresed for 3.5 hours in 15% polyacrylamide gels (4.3.1).

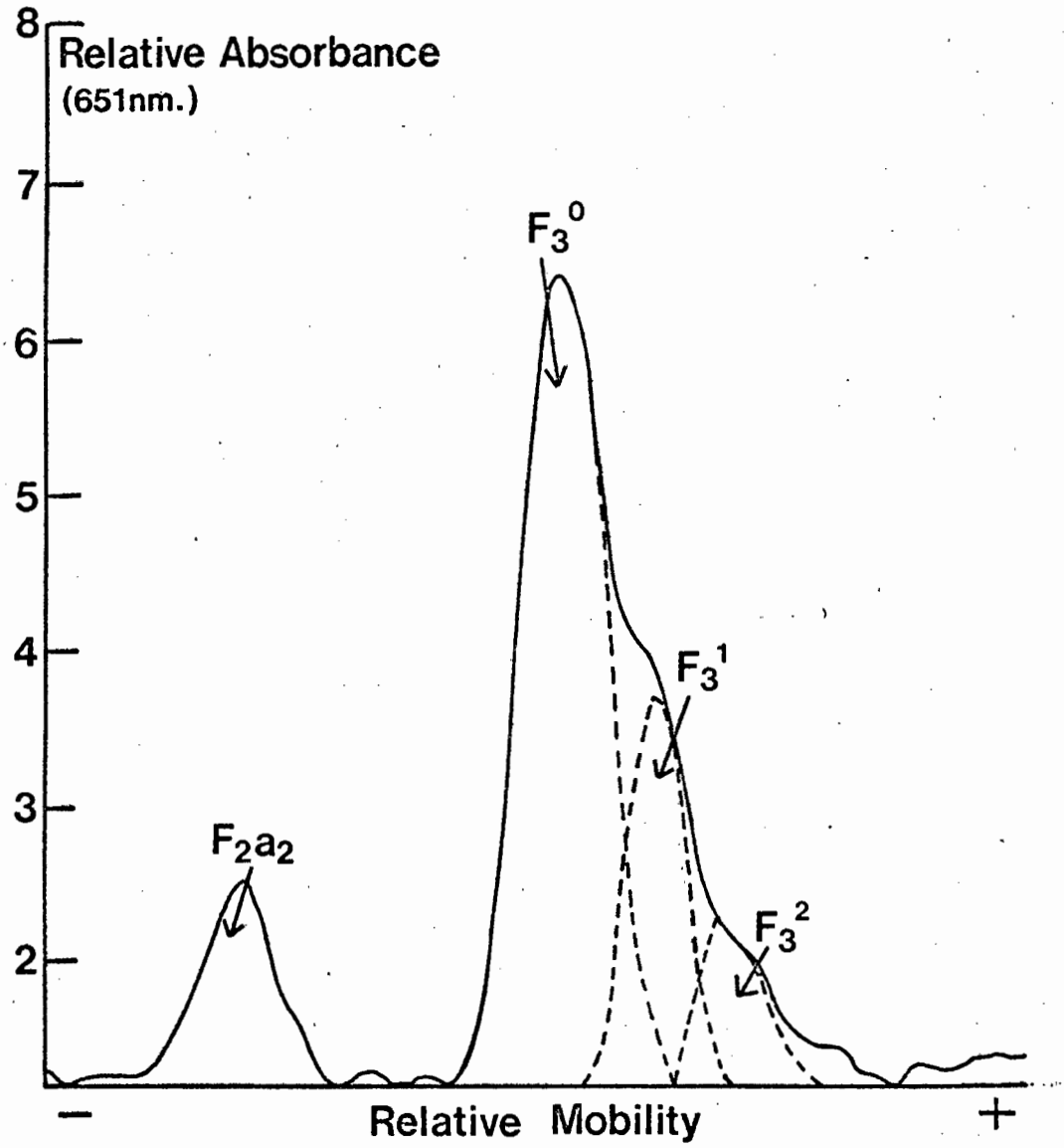


Fig. 1.9b : Densitometer tracing of histone F3 electrophoresed on polyacrylamide gel (Fig. 1.9a).

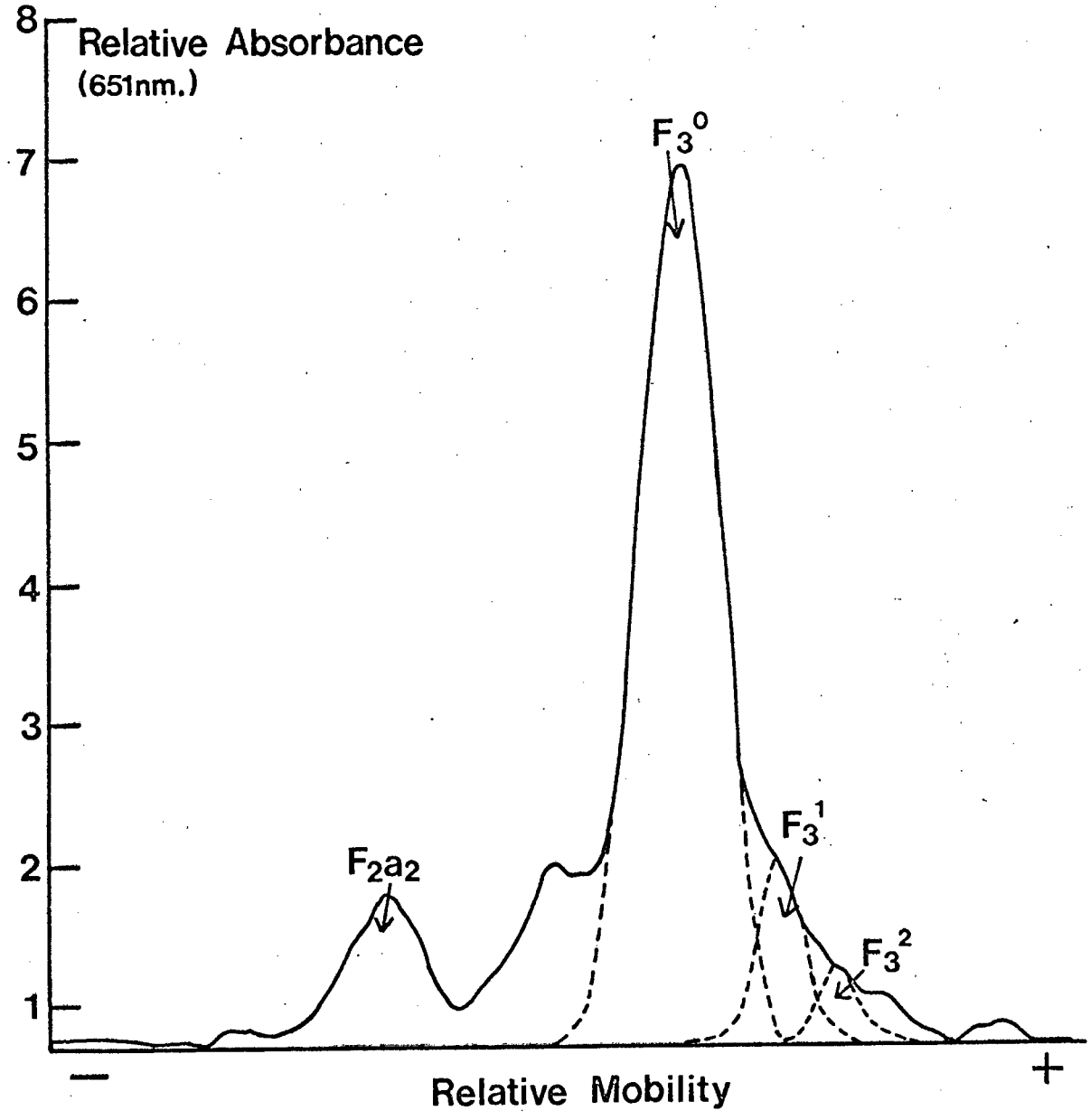


Fig. 1.9c : Densitometer tracing of histone F3 that had been incubated with histone deacetylase and electrophoresed on polyacrylamide gel (Fig. 1.9a). Histone F2a2 served as internal standard.

1.5.5 Amino acid composition of histone F3

The F3 dimer was hydrolysed in 5.7 M HCl for 24, 36 and 72 hours. Values for serine, threonine, tyrosine and cysteine were obtained by extrapolating their recoveries to zero time (Table 1.3) (4.3.2). Values obtained are expressed as moles per 100 moles of all amino acids recovered. When these values were expressed as moles per 135 moles all amounts of amino acids had near integral values indicating that this is the minimum molecular composition of the molecule. The calculated minimum molecular weight is 15,500 which is close to the experimental value, indicating that the F3 histone consists of 135 amino acids. (An error of ± 1 amino acid in the amino acids present in larger amounts could occur).

The presence of a single cysteine in F3 histone from chicken erythrocytes was confirmed by spectrophotometric titration against p-chloromercuribenzoate (4.3.4) (Fig. 1.10). From the ultraviolet spectrum it can be deduced that tryptophane is absent since the three tyrosine present account for the absorbancy at 280 nm (Fig. 1.11).

1.6 SUMMARY

Histone F3 from chicken erythrocytes has been isolated in a highly purified form. This protein was characterized and seemed homogeneous since no evidence to the contrary had been obtained. The electrophoretic microheterogeneity observed on polyacrylamide gels is due to acetylated lysine residues.

The protein consists of a single polypeptide chain with alanine as both N- and C-terminal residues. The amino acid composition was estimated to be : Lys 13, His 2, Arg 18, Asp 5, Ser 6, Thr 10, Glu 15, Pro 5, Gly 7, Ala 18, Cys 1, Val 6, Met 2, Ile 7, Leu 12, Tyr 3, and Phe 4 giving a total of 135 residues. Approximately 1.5 lysine out of the 13 lysine residues are ε-N-methylated. A single cysteine is present in this histone. The calculated molecular weight for the protein is near 15,500.

TABLE 1.3

AMINO ACID COMPOSITION OF F3 HISTONE

Amino acids	μmoles	Moles per 100 moles	Moles per 135 mole	Moles per mole F3
Lys	.1069	8.41	12.90	13
ε-N-Me-Lys[1]	.0136	1.07		
His	.0178	1.40	1.91	2
HN_4	.0848	6.67	(9.07)	(9)
Arg	.1667	13.11	17.83	18
Asp	.0486	3.82	5.19	5
Thr	.0915	7.20	9.79	10
Ser	.0534	4.20	5.71	6
Glu	.1433	11.27	15.33	15
Pro	.0590	4.64	6.31	6
Gly	.0669	5.26	7.15	7
Ala	.1726	13.57	18.45	18
Cys ½	.0102	.80	1.08	1
Val	.0571	4.49	6.11	6
Met	.0170	1.34	1.82	2
Ile	.0643	5.03	6.84	7
Leu	.1131	8.89	12.09	12
Tyr	.0302	2.37	3.23	3
Phe	.0370	2.91	3.96	4
				135

Values are corrected for destruction during hydrolysis (4.3.2).

1. ε-N-methylated lysines residues were estimated by the method of Gershey et al. (1969) but using a shorter (10 cm) column. ε-N-dimethyllysine predominated over the mono and trimethyl derivative.

Tryptophan was found to be absent as estimated from the UV spectrum (Fig. 1.11).

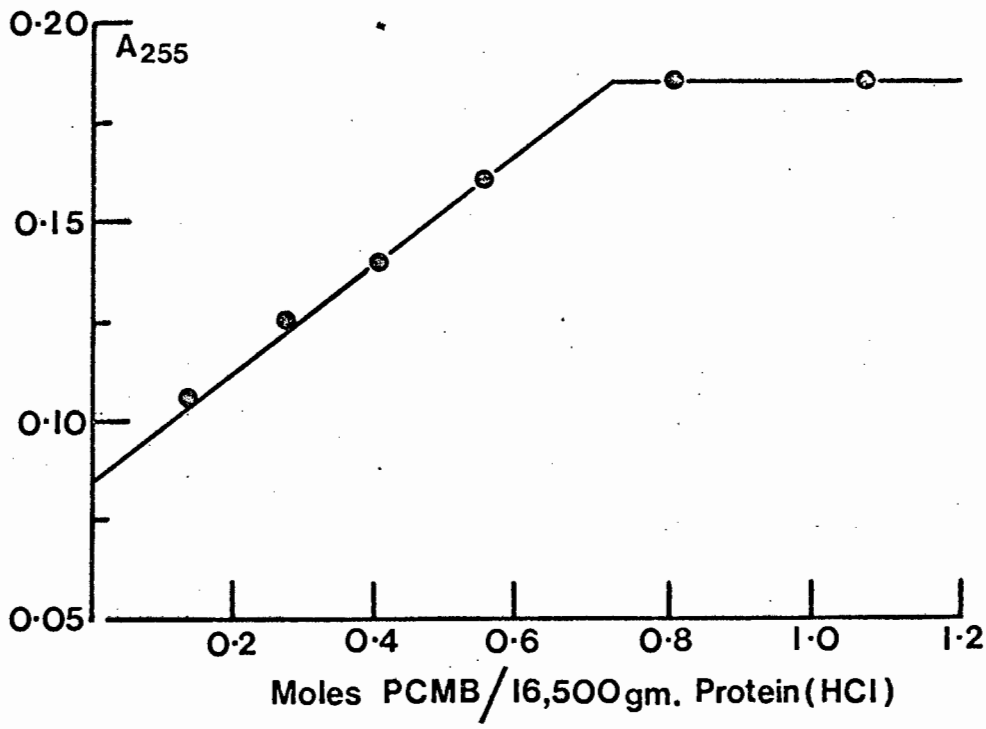


Fig. 1.10 : Spectrophotometric titration of histone F3 with p-chloromercuribenzoate (4.3.4).

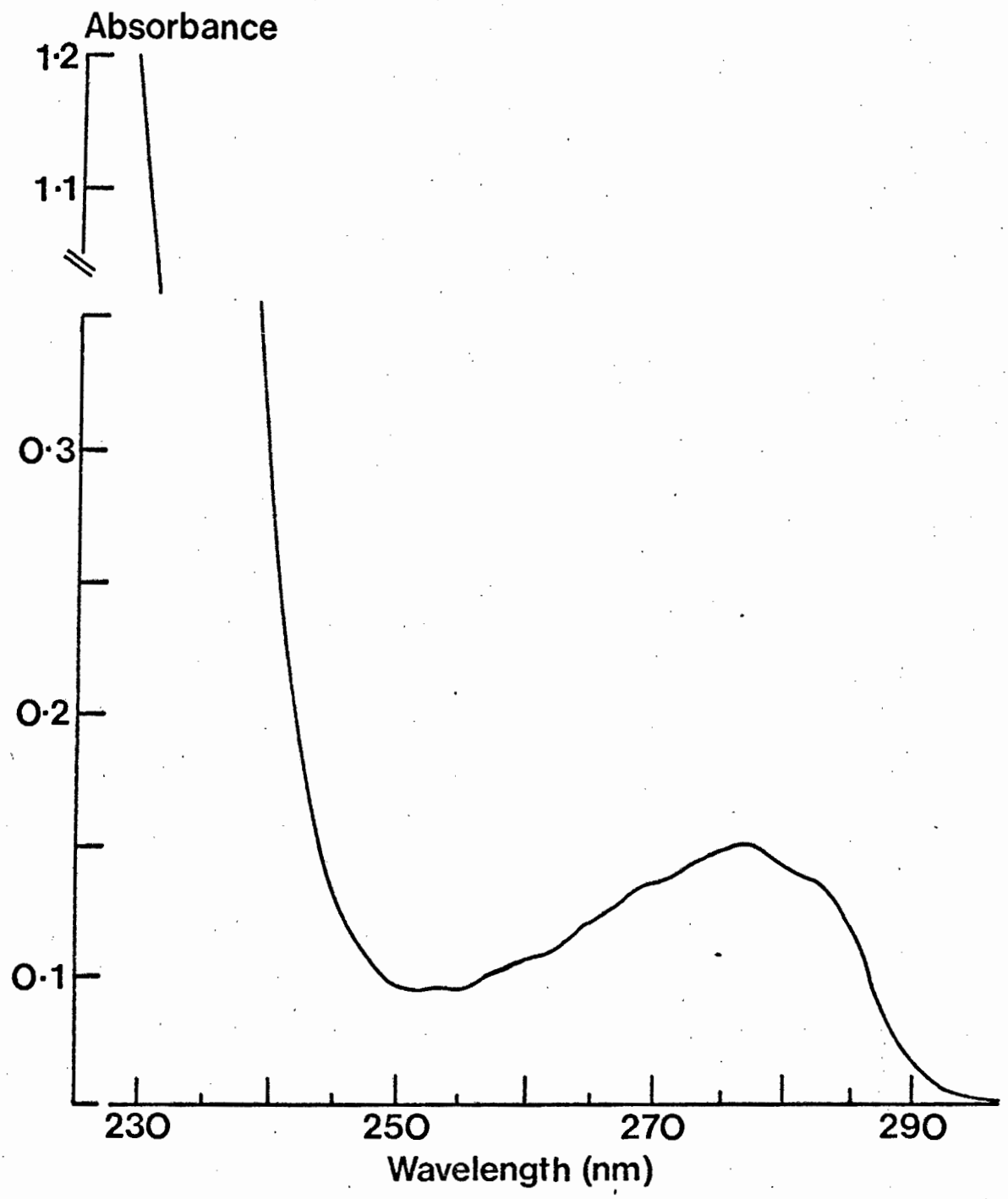


Fig. 1.11 : The ultraviolet spectrum of histone F3 in 0.01 N HCl (0.47 mg/ml). The optical pathlength : 10 mm. ($A^{1\%}_{230}$ = 2.43, $A^{1\%}_{275}$ = 0.213).

PART 2

ELUCIDATION OF THE PRIMARY STRUCTURE OF HISTONE F3

2.1 INTRODUCTION

Since Sanger (1954) elucidated the primary structure of insulin a large number of new and improved techniques have been introduced. Quoting Edman (1970) : "Anyone entering the field of protein structure determination will be faced with a bewildering choice of techniques for sequence determination". Fortunately, however, well established approaches exist.

The technique that had the greatest impact on sequencing methodology was the Edman degradation (Edman, 1950) which, until now, remained the most useful method yet devised that is potentially able to degrade proteins and peptides sequentially. Although modern methods for the identification and quantitation of amino acids or their derivatives are very well developed, the approach to the determination of the primary structure is dictated by the limited number of amino acids that can be step-wise removed from the N-terminal end of the protein. Non-quantitative yields in the individual reactions of the degradation are the main cause for this limitation. This has resulted in the situation that reliable information as to the sequence can be obtained only up to 10 degradation steps. Recently, however, the number of degradation steps has been considerably extended by Edman and Begg (1967) who have described a "protein sequenator"; a machine in which the various operations of the Edman degradation have been optimized with respect to yields and automated (4.5.2).

Despite this improved methodology increasing the ease and the speed of the investigation, the basic principle used to determine the amino acid sequence of a protein has not altered much since Sanger (1954) determined the sequence of insulin. Due to the limited number of Edman degradations that can be performed it is still necessary to fragment the

protein by some method and examine the peptides. A second fragmentation by another method, with subsequent examination of the generated peptides, is necessary to align the two series of fragments on the basis of their partial identities. In practice this is complicated in that the complete sequence of all peptides, because of their size, cannot be determined. Larger fragments might therefore have to be cleaved again while others are not obtainable in a pure form. Finally, the structure of the protein is reconstructed from all the peptides, the structure of which may have been completely or partially determined, by logical and combinational examination of overlapping fragments. Computer programs have been devised (Goldstone & Needleman, 1970) to spot the unique amino acid sequence with the least possible amount of information. Although these techniques and their complications are very well documented, the uniqueness of the individual protein removes the elucidation of its primary structure from the routine, and each protein poses a specific and new problem.

A very important part of the amino acid sequence determination is thus the fragmentation of the protein into peptides. The rather unspecific partial acid hydrolysis of peptide bonds has nearly been totally replaced by more specific cleavages. From recent literature on primary structure determination it becomes apparent that the most frequent approach in degrading proteins or peptides consists of the use of highly specific proteolytic enzymes (Smyth, 1967a; Kasper, 1970b). The protein is degraded by at least two enzymes with variable specificity, e.g. trypsin and chymotrypsin resulting in two sets of overlapping peptides. A difficulty in this approach arises in some cases where large parts of the molecule are not attacked by trypsin and chymotrypsin (cores). Pepsin, papain, pronase and other proteolytic enzymes can aid in overcoming this problem, but their specificity is very much lower compared to that of trypsin (Smyth, 1967b; Witkop, 1961).

Considerable time and effort is spent on the separation and purification of the smaller fragments produced, especially as the size of the protein increases. It was realized that

the primary structure elucidation would be considerably simplified if the protein is first degraded into a few larger fragments (Witkop, 1961). These could then be fairly easily tackled by the above mentioned methods. Up to this stage no enzymes have been described which could perform such a task. However, a number of nonenzymic methods have been discovered which cleave the peptide bond next to amino acids with certain functional groups in their side chains (Witkop, 1961). Chemical methods have been described which cleave the polypeptide chain next to tryptophan, tyrosine, histidine, cysteine, serine and methionine (Spande et al., 1970), to mention some of them. Although a large variety of chemical cleavages have been proposed, most have not been tried except on model compounds and on peptides of known structure. A noticeable exception is the cyanogen bromide cleavage (Gross, 1967) at methionine residues which is fairly frequently used with excellent results. Hydrolysis with dilute acid at aspartic acid residues, N-bromosuccinimide cleavage at tryptophan and tyrosine and cleavage at the hydroxy amino acids by an acyl shift (Schultz, 1967; Iwai, 1967; Witkop, 1961; Spande et al., 1970) have only occasionally been used in sequence studies judging from recent literature.

It appears that the main obstacles in many of these chemical degradations at their present stage of development are low yields and undesirable side reaction.

The introduction of these techniques has, however, not changed the basic approach of structure determination of proteins which relies on the establishment of partial or complete amino acid sequences of relatively small and pure overlapping fragments.

A new strategy has been proposed by Gray (1968) in which the sequential Edman degradation is applied to mixtures of peptides produced by selective cleavages. Analysis of the amino acids removed at each step provides information on the "nearest neighbours" which as such is not sufficient to deduce the structure of a particular peptide. However, cross-correlation between results from two or more digests would enable one to deduce the sequence of the individual peptides

and then finally the complete structure.

In a theoretical example ribonuclease (124 residues) has to be cleaved consecutively at Arg, Lys, Cys, Met, Tyr and His residues. Only ten Edman degradations on each of the six mixtures would enable one to deduce the complete amino acid sequence of this protein except for a single residue (Gray, 1968).

Judging from recent literature, this approach, however, has not been applied yet and this is probably due to quite a number of problems which this method would pose in practice. It seems, however, feasible that this technique could be used to supplement information obtained in the conventional approach, e.g. on simple peptide mixtures that cannot readily be purified.

2.1.1 Strategy of sequence determination

The well established approach to primary structure determination requires that the protein is cleaved into at least two sets of overlapping fragments.

Digesting histone F3 with trypsin would yield at least 32 peptides (Fambrough & Bonner, 1968). Even if the tryptic action were to be restricted to arginine residues by blocking lysine residues (Kasper, 1970b) at least 19 fragments would be generated. Considerable time and effort would have to be spent on the isolation and purification of these fragments. In addition, difficulties could be expected during their sequential analysis by the automated Edman degradations. It is well established that the automated procedure performs well on proteins, however, difficulties are experienced with peptides due to their higher solubility in organic solvents used during washing and extraction cycles (4.5.2) (Niall et al., 1969; Braunitzer et al., 1970).

Because automated sequence analysis is capable of performing many more useful Edman degradations as compared to the manual procedure, a small number of larger fragments would have several advantages over a large number of small peptides.

At this stage it was realized that if a protein or polypeptide were to be split into two or three fragments only,

their relative position in the peptide chain would become obvious by comparing the terminal amino acids of the uncleaved chain and its fragments. Thus, by choosing the correct sequence of degradation methods it should be possible to cleave a protein into a series of fragments, the relative position of which, in the protein molecule, will become apparent after endgroup analysis of the parent polymer and its daughter fragments. A second set of overlapping fragments for alignment purposes would thus not be required.

Since only a few fragments are produced at any one cleavage the separation and purification should be a relatively simple task.

If, however, certain fragments should pose solubility problems as they become progressively smaller towards the C-terminal end during the Edman degradation, one could resort to the production of a fragment by another cleavage method in which this C-terminal region becomes a penultimate region in another peptide (an overlapping fragment).

From the amino acid composition of histone F3 (Table 1.3) it is evident that Cys, Met, Tyr and His are potential sites for such a fragmentation. Since only two methionine residues are present and cyanogen bromide cleaves the peptide bond at this amino acid in high yields (4.4.1), it is the obvious method to start the fragmentation of histone F3 into three fragments.

The type of subsequent fragmentation of these three polypeptides will depend on their amino acid composition. The necessity of other primary cleavages of the protein and their nature in later stages of the structure elucidation will depend on results obtained with the cyanogen bromide cleavage products.

2.1.2 Nomenclature of fragments

The following scheme for the naming of fragments has been adopted. The first two letters indicate the type of cleavage by which the fragments had been generated, e.g. CN denotes a cyanogen bromide cleavage. This is followed by a number which indicates its relative elution order on

exclusion chromatography. Additional purification procedures and elution order will be denoted by a bracketed number. Subsequent cleavages and purification steps are coded in the same way and are added to the name of the parent fragment.

Example : CN-1 NB-1(2)

This fragment has been generated by cleaving fragment CN-1 with N-bromosuccinimide and subjecting the products to exclusion chromatography. The first peak eluted was subjected to ion exchange chromatography (CMC-cellulose) and the fragment corresponded to the second peak eluted. Fragment CN-1 had previously been generated by cleaving histone F3 with cyanogen bromide and corresponded to the first peak eluted after subjecting the mixture to exclusion chromatography.

2.2 CYANOGEN BROMIDE CLEAVAGE OF HISTONE F3

2.2.1 Cleavage and purification of fragments

Histone F3 dimer was cleaved at methionine residues by dissolving 200 mg protein in 30 ml 70% (v/v) formic acid and adding a 130 fold excess CNBr (4.4.1). The freeze-dried cleaved protein was subjected to gel filtration on a Sephadex G-100 column (Fig. 2.1).

The elution volume of fraction A falls into the outer volume of the column. Fraction B corresponds to the position of the dimer and C to the molecular weight of the monomer.

In a first approximation one can assume that the integral over the ultraviolet absorption at 230 nm of a fraction is directly proportional to the total number of peptide bonds and that the elution volume is proportional to the size of a fragment. Therefore fraction A, B and C represent only a small contamination of large molecular weight uncleaved or partially cleaved F3 histone, whereas fraction D, E and F would probably correspond to the expected 3 fragments. This assumption was born out by the subsequent amino acid analysis (Table 2.2).

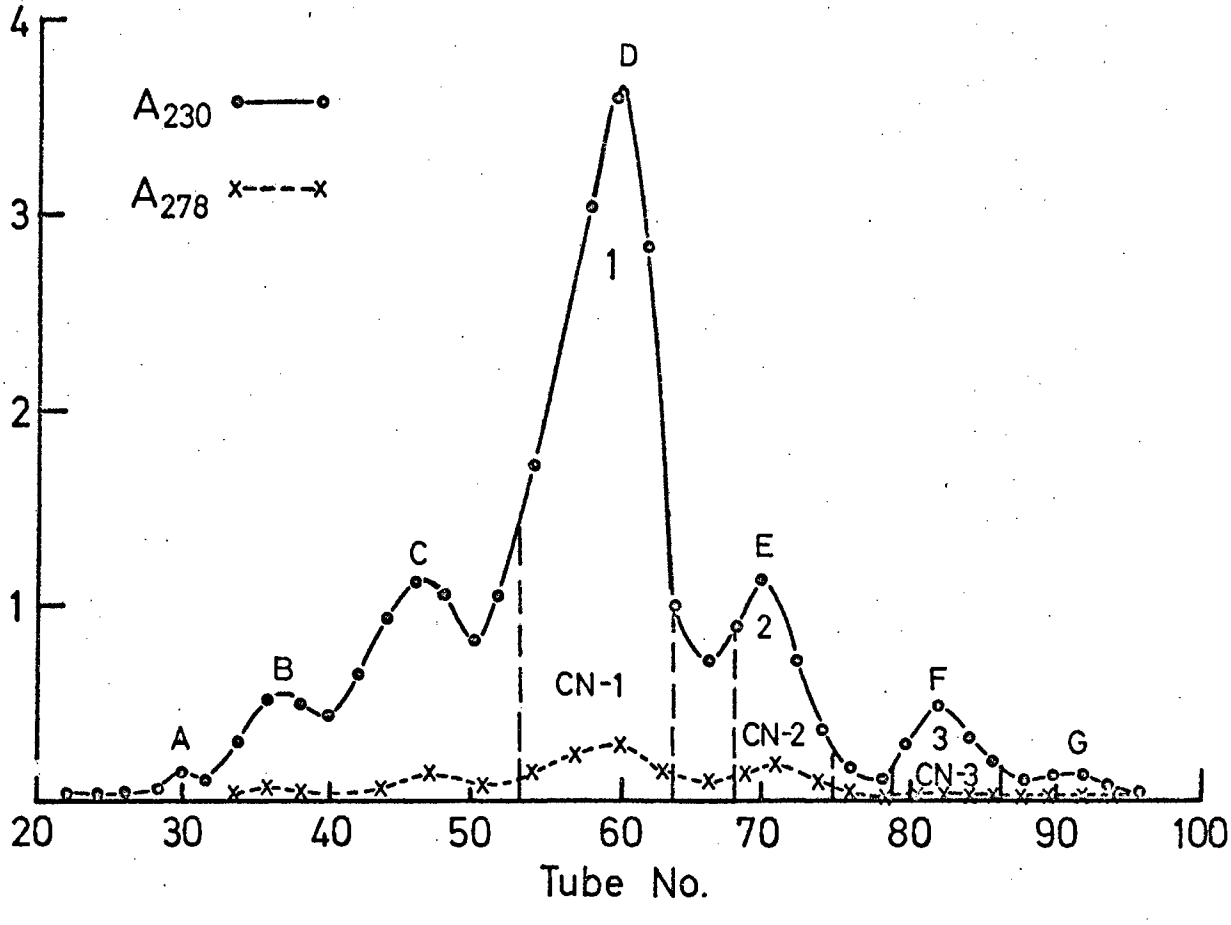


Fig. 2.1 : Elution pattern obtained on Sephadex G-100 2.5 x 100 cm column after cleaving F3 histone dimer with CNBr at methionine residues (4.4.1). The eluent used was 0.01 N HCl, the flow rate 12 ml/h and the fraction volume 5.5 ml. Peak A, B and C correspond to uncleaved and partially cleaved histone; D, E and F to the three expected fragments, and G to urea.

TABLE 2.1

APPROXIMATE SIZE OF CNBr FRAGMENTS OF HISTONE F3

Fraction (Fig. 2.1)	Fragment	Relative peak area[1]	Relative size of the fragments
1	CN-1	243	103
2	CN-2	56	23
3	CN-3	24	10

1. The areas were estimated from the elution pattern (Fig. 2.1).

All three fragments were obtained in a pure form as judged from the electrophoretic pattern in Fig. 2.2. Fragment CN-1 shows the same electrophoretic microheterogeneity as histone F3 (Fig. 1.6) indicating that all acetylated lysine residues occur in this piece. Fragment CN-2 contains the cysteine residue since on reduction with mercaptoethanol its molecular weight becomes smaller as indicated by an increase of its retention volume on gel chromatography.

From the ultraviolet spectra it was evident that fragment CN-1 and CN-2 contained equal proportions of aromatic amino acids while CN-3 did not contain any at all (Fig. 2.1).

All three fragments were also characterized by their N-terminal amino acid and amino acid composition (Table 2.2). These results indicated that the three fragments were homogeneous and accounted for all the amino acid residues present in the uncleaved protein.

2.2.2 Alignment of fragments

The absence of homoserine lactone in fragment CN-3 (Table 2.2) places it in the C-terminal position. In order to align fragment CN-1 and CN-2, both having Ala as N-terminal amino acid (Table 2.2), a number of Edman degradations were performed on these peptides and the uncleaved protein (4.5.1) (Table 2.3).

TABLE 2.2

AMINO ACID COMPOSITION AND TERMINAL RESIDUES OF HISTONE F3 AND CNBr FRAGMENTS

Fragment	F3 Histone uncleaved			Fraction 1 (Fig. 1.2) CN-1		
Amino acid	mole %	residues		mole %	residues	
Lys	8.41			11.90		
$Lys(Me)_{1,2}^{0}$	1.07	12.90	(13)	11.14	10.72	(11)
His	1.40	1.91	(2)	1.42	1.15	(1)
NH_3	6.67	(9.07)		10.61	(8.63)	
Arg	13.11	17.83	(18)	16.40	13.35	(13)
Asp	3.82	5.19	(5)	2.51	2.03	(2)
Thr	7.20	9.79	(10)	9.76	7.94	(8)
Ser	4.20	5.71	(6)	5.94	4.84	(5)
Glu	11.27	15.33	(15)	11.25	9.15	(9)
Pro	4.64	6.31	(6)	5.90	4.83	(5)
Gly	5.26	7.15	(7)	6.20	5.04	(5)
Ala	13.57	18.45	(18)	13.57	11.05	(11)
Cys ½	0.80	1.08	(1)	-		
Val	4.49	6.11	(6)	4.89	3.98	(4)
Met*	1.34	1.82	(2)	0.53	0.45	(1)
Ile	5.03	6.84	(7)	3.72	3.03	(3)
Leu	8.89	12.09	(12)	8.87	7.22	(7)
Tyr	2.37	3.23	(3)	2.40	1.96	(2)
Phe	2.91	3.96	(4)	3.64	2.97	(3)
Residues/ mole			(135)			(90)
N-Terminal (4.3.6)	Ala			Ala		
C-Terminal (4.4.1)	Ala			homoSer		

TABLE 2.2 /cont'd.....

Fragment	Fraction 2 (Fig. 1.2) CN-2			Fraction 3 (Fig. 1.2) CN-3			SUM 1,2 + 3
Amino acid	mole %	residues		mole %	residues		residues
Lys	3.93	1.26	(1)	6.57	1.01	(1)	13
Lys(Me)$_{1,2}$ [0]	-			-			
His	4.24	1.25	(1)	-			2
NH_3	12.80	3.77		7.27	1.12		
Arg	3.73	1.10	(1)	25.81	3.97	(4)	18
Asp	6.89	2.03	(2)	6.51	1.00	(1)	5
Thr	6.54	1.92	(2)	0.02	0		10
Ser	3.77	1.11	(1)	0.03	0		6
Glu	14.46	4.25	(4)	14.31	2.2	(2)	15
Pro	-			6.73	1.04	(1)	6
Gly	4.24	1.25	(1)	6.64	1.02	(1)	7
Ala	16.86	4.96	(5)	13.78	2.12	(2)	18
Cys ½	2.25	0.66	(1)	-			1
Val	6.21	1.83	(2)	0.02	0		6
Met*	2.77	0.82	(1)	-			2
Ile	6.02	1.77	(2)	12.26	1.89	(2)	7
Leu	12.86	3.78	(4)	6.73	1.04	(1)	12
Tyr	3.16	0.93	(1)	-			3
Phe	3.20	0.94	(1)	-			4
Residues/ mole			(30)			(15)	135
N-Terminal (4.3.6)	Ala			Pro			
C-Terminal (4.4.1)	homoSer			Ala			

* After CNBr-cleavage Met was analysed as homoserine lactone.

[0] Mono and dimethyl-ε-N-lysine were determined by the method of Gershey et al. (1969) but using only a 10 cm column. The dimethyl derivative predominated.

C-terminal residues of fragments are evident from the amino acid composition and the nature of the cleavage (4.4.1).

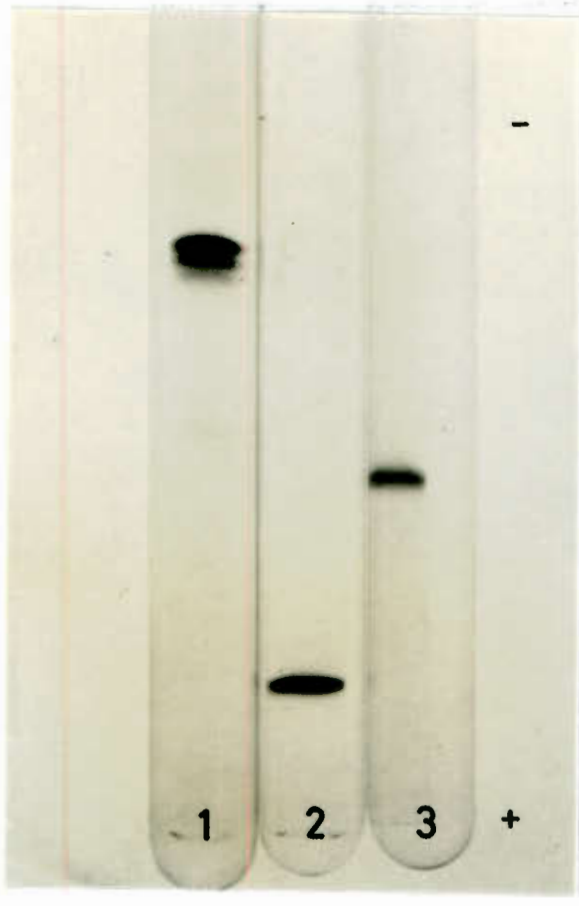


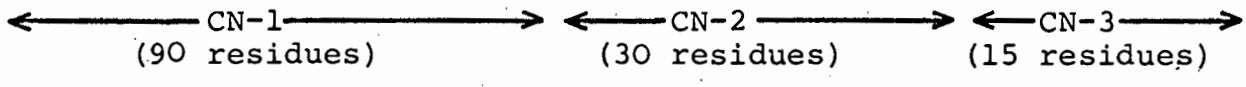
Fig. 2.2 : Polyacrylamide gel electrophorteic pattern of the three fragments obtained after CNBr cleavage of the F3 dimer. Gel 1, 2 and 3 correspond to fraction 1, 2 and 3 (CN-1, CN-2, CN-3) in Fig. 2.1. Gel 1 was run for 3.5 h and the remaining two for 1 h (4.3.1).

TABLE 2.3

MANUAL EDMAN DEGRADATION OF HISTONE F3 AND CNBr FRAGMENTS

Material	Step 1	Step 2	Step 3
Histone F3	Ala	Arg	Thr
CN-1	Ala	Arg	Thr
CN-2	Ala	Leu	Gln

The N-terminal sequence of CN-1 is identical to that of histone F3 and is therefore situated in the N-terminal position of the protein. This, together with the amino acid composition of the fragments, allows one to draw the following arrangement of the single polypeptide chain of histone F3 :



Ala-Arg-Thr----------Met-Ala-Leu-Gln------Met-Pro--------Ala

All the acetylated and methylated lysine residues occur in fragment CN-1 (Table 2.2, Fig. 2.2).

2.3 N-BROMOSUCCINIMIDE CLEAVAGE OF CYANOGEN BROMIDE CLEAVAGE FRAGMENTS

Histone F3 has been cleaved into three fragments which could readily be purified and aligned in the protein molecule (2.2.2). Since it will probably be impossible to establish the complete sequence of fragment CN-1 (90 residues) by sequential degradation of the uncleaved protein, fragment CN-1 had thus to be further fragmented. From the amino acid composition it is evident that only two tyrosine residues are present in fragment CN-1 (Table 2.2). N-Bromosuccinimide under certain conditions selectively cleaves peptide bonds at tyrosine residues (4.4.2). Although this cleavage has been mainly tried on model peptides, it may yield the three desired fragments.

Since it is impossible to predict how many Edman degradations could be performed on a particular peptide it was decided to cleave fragment CN-2 as well which contained one Tyr residue (Table 2.2). Apart from yielding new peptides for sequential analysis it would also yield information about the distribution of amino acids in the fragment which would aid in the sequence analysis of the whole fragment. NBS-cleavage would yield only two fragments. It is evident from the amino acid composition of fragment CN-2 that it possesses quite a few additional potential cleavage sites (Cys, Asp, Lys and Arg) which would yield two or three fragments only.

2.3.1 CN-2 : Cleavage and purification of fragments

Fragment CN-2 (6 mg) was dissolved in 4 ml 50% (v/v) acetic acid and cleaved by titrating the solution with NBS (4.4.2). From the titration curve it is evident that approximately 7 moles of NBS per mole of peptide are required to obtain a constant absorption at 260 nm (Fig. 2.3). Since the oxidation of tyrosine requires only 3 moles NBS the additional consumption is probably due to the oxidation of cystine to cysteic acid (3 moles) and the bromination of the histidine residue (compare Table 2.2 and Table 2.4). The yield of peptide cleaved by NBS based on the absorption at 260 nm (spirolactone $\varepsilon 260$ = 11,000) is near 100% (Fig. 2.3). The cleaved peptide was freeze-dried and subjected to gel filtration on a Sephadex G-50 column (Fig. 2.4). From the elution pattern it is evident that a few percent uncleaved peptide are still present (Fig. 2.4, Peak A).

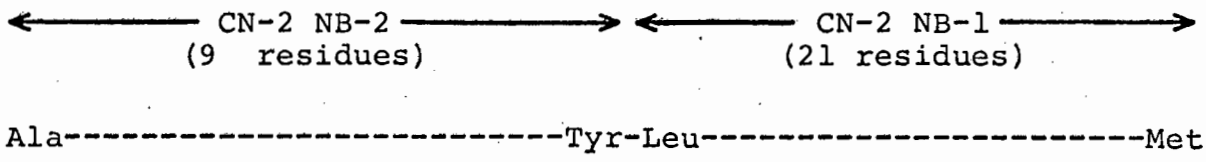
The two fractions (Peak B and C) were characterized by their ultraviolet spectra (Fig. 2.5), amino acid composition and N-terminal amino acids (Table 2.4). The C-terminal amino acids became apparent from the amino acid composition of the fragments and the nature of the cleavages (Table 2.4). Both fractions were slightly cross-contaminated, but from the amino acid composition it is obvious that the contamination is only in the order of a few percent (Table 2.4).

From the ultraviolet spectra of the fragments (Fig. 2.5) or the comparison of A230 nm and A260 nm elution pattern (Fig. 2.4) it can be deduced that fraction 2 contains the spirolactone residue (i.e. tyrosine) (4.4.2).

2.3.2 CN-2 : Alignments of fragments

Fragment CN-2 NB-2 and the uncleaved peptide both had Ala as their N-terminal amino acid (Table 2.4). This, together with the fact that this fragment contains the spirolactone (Fig. 2.5), places it in the N-terminal position. The fact that only two fragments are produced by the cleavage and that fragment CN-2 NB-1 contains the homoserine lactone (Table 2.4) clearly places it in the C-terminal position. This, together with the amino acid composition, allows one to assign

to fragment CN-2 the following tentative structure :



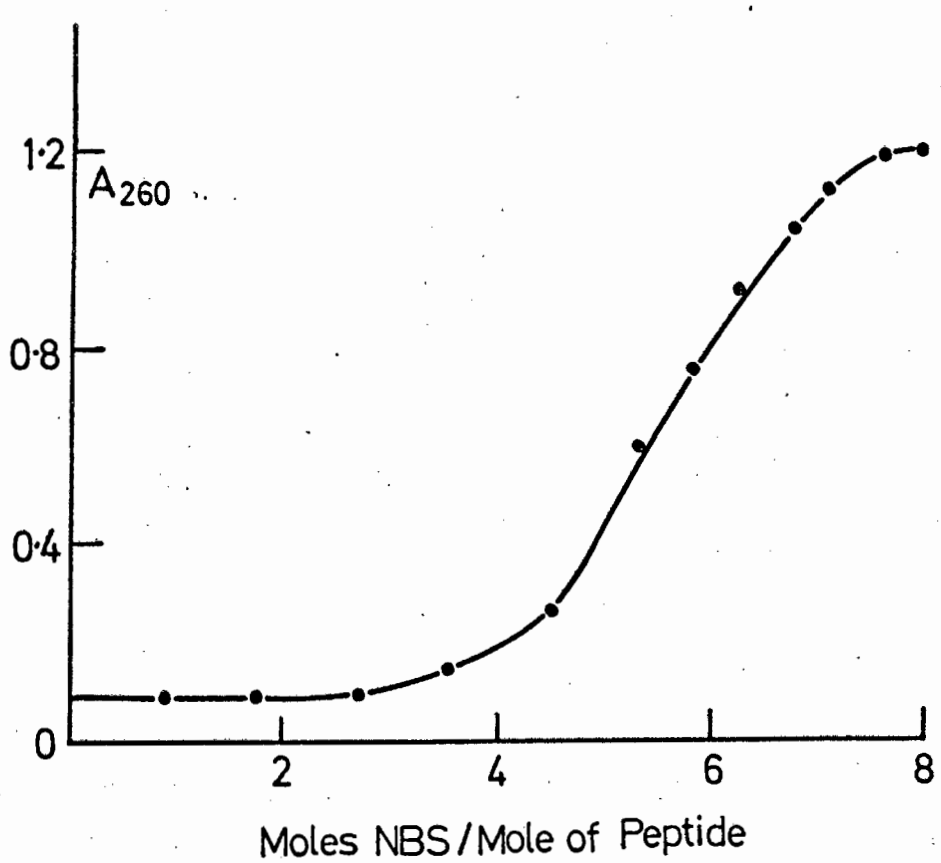


Fig. 2.3 : Titration of fragment CN-2 (5.6 mg) in 50% acetic acid (4 ml) with 0.093 M N-bromosuccinimide (4.4.2.1) (optical path length : 0.2 cm).

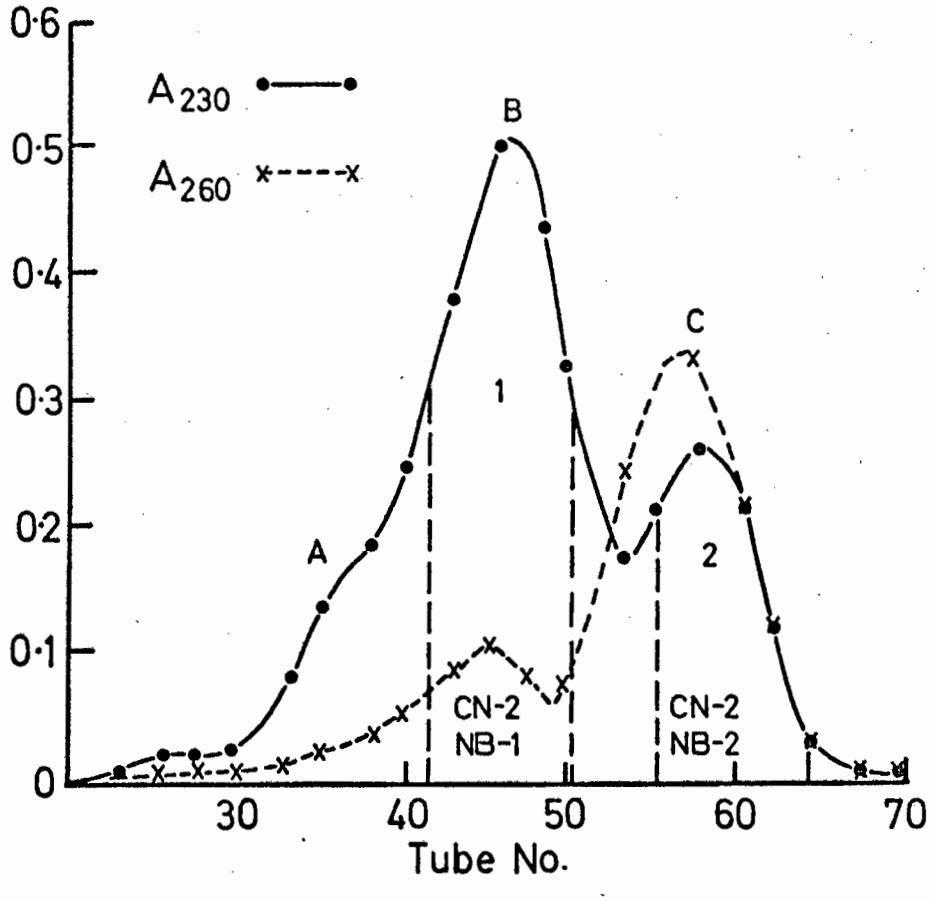


Fig. 2.4 : Elution pattern obtained on a Sephadex G-50 1.5 x 90 cm column after cleaving fragment CN-2 with NBS (4.4.2.1) at tyrosine residues. The eluent used was 0.01 N HCl, the flow rate 8 ml/h and the fraction volume 1.8 ml. Peak A probably corresponds to uncleaved fragment because it has the same elution volume as CN-2.

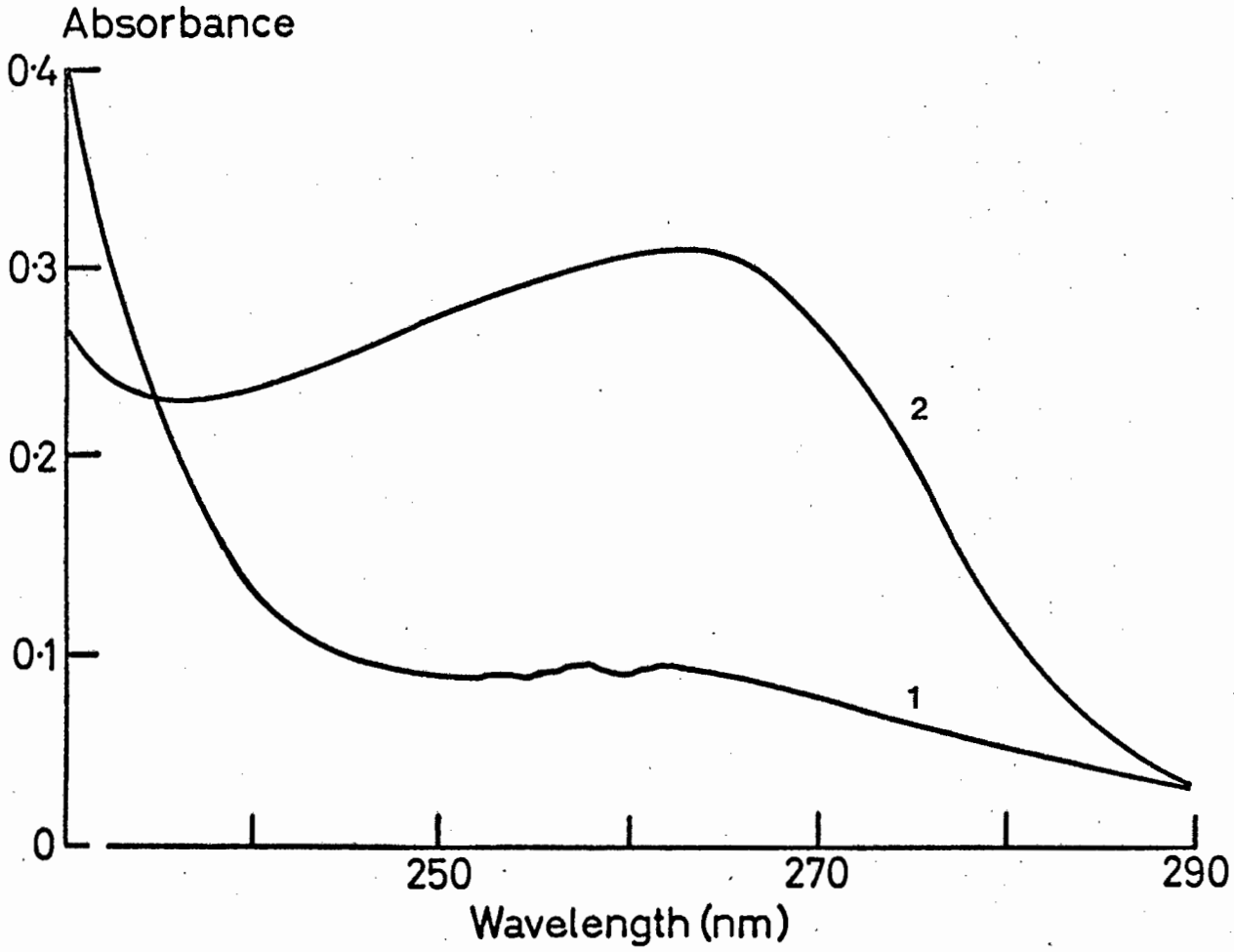


Fig. 2.5 : Ultraviolet spectra of fraction 1 and 2 (CN-2 NB-1 (0.18 mg/ml) and CN-2 NB-2 (0.12 mg/ml)*) in 0.01 N HCl eluted from Sephadex G-50 (Fig. 2.4).

2.3.3 CN-2 NB-1 : Cleavage and alignment of fragments

Fragment CN-2 NB-1 contains a single histidine residue (Table 2.4, 2.2) and should therefore split into two fragments when cleaved with NBS under certain conditions (4.4.2). One mg of this fragment was oxidized with NBS in pyridine-buffer at 100°C (4.4.2.2).

The cleaved peptide was applied to a Sephadex G-50 1.5 x 90 cm column. A single major peak near the total volume of the column was eluted. All of this fraction was submitted to endgroup analysis. Besides DNS-ε-Lys, DNS-Leu and DNS-Ala was identified (4.3.6).

TABLE 2.4

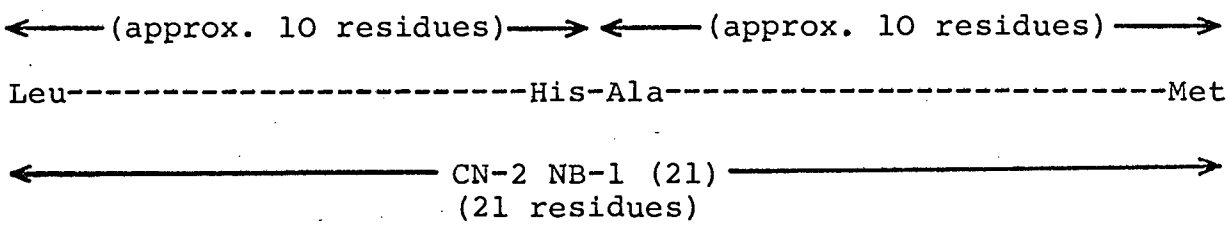
AMINO ACID COMPOSITION AND TERMINAL RESIDUES OF NBS CLEAVAGE FRAGMENTS OF CN-2

Fragment	Fraction 1 (Fig. 2.4)			Fraction 2 (Fig. 2.4)		
	CN-2 NB-1			CN-2 NB-2		
Amino acid	mole %	residues		mole %	residues	
Lys	5.42	1.08	(1)	0.56	0.06	
His[0]	+		(1)	-		
NH_3	15.00			30.00		
Arg	5.56	1.11	(1)	0.51	0.05	
Asp	10.68	2.14	(2)	0.53	0.05	
Thr	10.09	2.02	(2)	0.34	0.03	
Ser	0.07	0.01	-	10.32	1.03	(1)
Glu	7.29	1.46	(1)	31.70	3.17	(3)
Pro	-			-		
Gly	4.96	0.99	(1)	0.70	0.07	
Ala	12.10	2.40	(2)	32.30	3.23	(3)
$CySO_3H$	1.1	0.20	(1)	-		
Val	9.18	1.84	(2)	0.31	0.03	
homoSer	4.50	0.90	(1)	-		
Ile	9.18	1.84	(2)	0.30	0.03	
Leu	15.30	3.06	(3)	11.40	1.14	(1)
Tyr*	-			+		(1)
Phe	4.80	0.96	(1)	1.06	0.10	
Residues/ mole			(21)			(9)
N-terminal (4.3.6)	Leu Small amounts Ala			Ala Small amounts Leu		
C-terminal (4.4.1, 4.4.2)	homoSer			Tyr*		

[0] Brominated histidine was detected by the Pauly reaction (4.4.2.1).

* After the N-bromosuccinimide cleavage tyrosine is converted to the spirolactone. Its presence is revealed by its characteristic ultraviolet absorption (Fig. 2.5, 4.4.2.1).

The similar elution volume of the two peptides, together with the N-terminal analysis, indicates the following arrangement of fragments :



2.3.4 CN-1 : Cleavage and purification of fragments

Fragment CN-1 (67 mg) was cleaved at tyrosine residues by dissolving it in 50% (v/v) acetic acid and titrating it with NBS (4.4.2.1, Fig. 4.6). Approximately 8 moles of NBS per mole of peptide were consumed in the course of the titration which can be accounted for by the oxidation of two tyrosine residues (6 moles) and the bromination of a histidine residue (Table 2.2).

The cleaved fragment was freeze-dried and applied to a Sephadex G-100 column (Fig. 2.7). From the elution pattern and the calculated amount of spirolactone formed (Fig.2.6) it is apparent that the fragmentation yield was fairly high.

Endgroup analysis revealed that fraction 2 consisted of a single fragment while fraction 1 contained at least two pieces.

Fraction 2 was applied to a column of Sephadex G-25 (2.5 x 30 cm). A single peak was obtained. Polyacrylamide gel electrophoresis revealed the presence of two major bands (Fig. 2.8). The fragment was further characterized by its ultraviolet spectrum (Fig. 2.9), N-terminal residue and amino acid composition (Table 2.5, column 4) which indicated that it was homogeneous. Further investigation indicated that the electrophoretic heterogeneity is caused by the spirolactone. This conclusion is based on the observation that when this fragment (CN-1 NB-2) was treated with a NaOH solution the ultraviolet spectrum, due to the spirolactone, and the electrophoretic pattern of the fragment changed (Fig. 2.8, 2.9), the major double band becomes homogeneous.

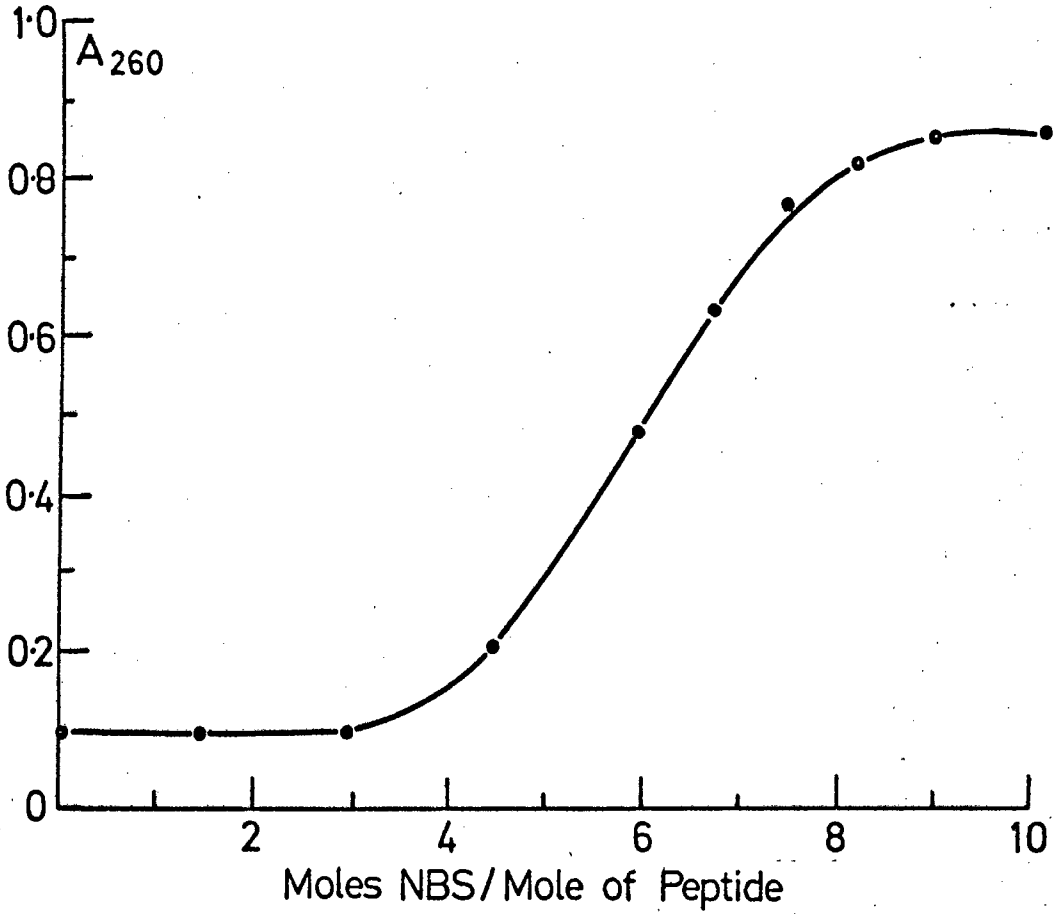


Fig. 2.6 : Titration of fragment CN-1 (67 mg) in 50% (v/v) acetic acid (30 ml) with 0.1 M N-bromosuccinimide (4.4.2.1) (optical path length : 0.2 cm).

Fraction 1 (Fig. 2.7) was applied to a CMC-column buffered to pH 4.4 with acetate. The column was first eluted with a NaCl gradient and finally with 0.2 N HCl (Fig. 2.10).

Fractions were pooled as indicated, freeze-dried and desalted on a Sephadex G-25 2.5 x 30 cm column.

All three fractions were essentially pure as judged from the electrophoretic pattern (Fig. 2.11). Fragment CN-1 NB-1(2) has a similar heterogeneity as histone F3 indicating that it contains all the acetylated lysine residues (Fig. 2.11). The relative quantities of the three bands (Fig. 2.11, Gel 2) are probably altered by the unknown effect of spirolactone on the electrophoretic mobility. All three fractions were

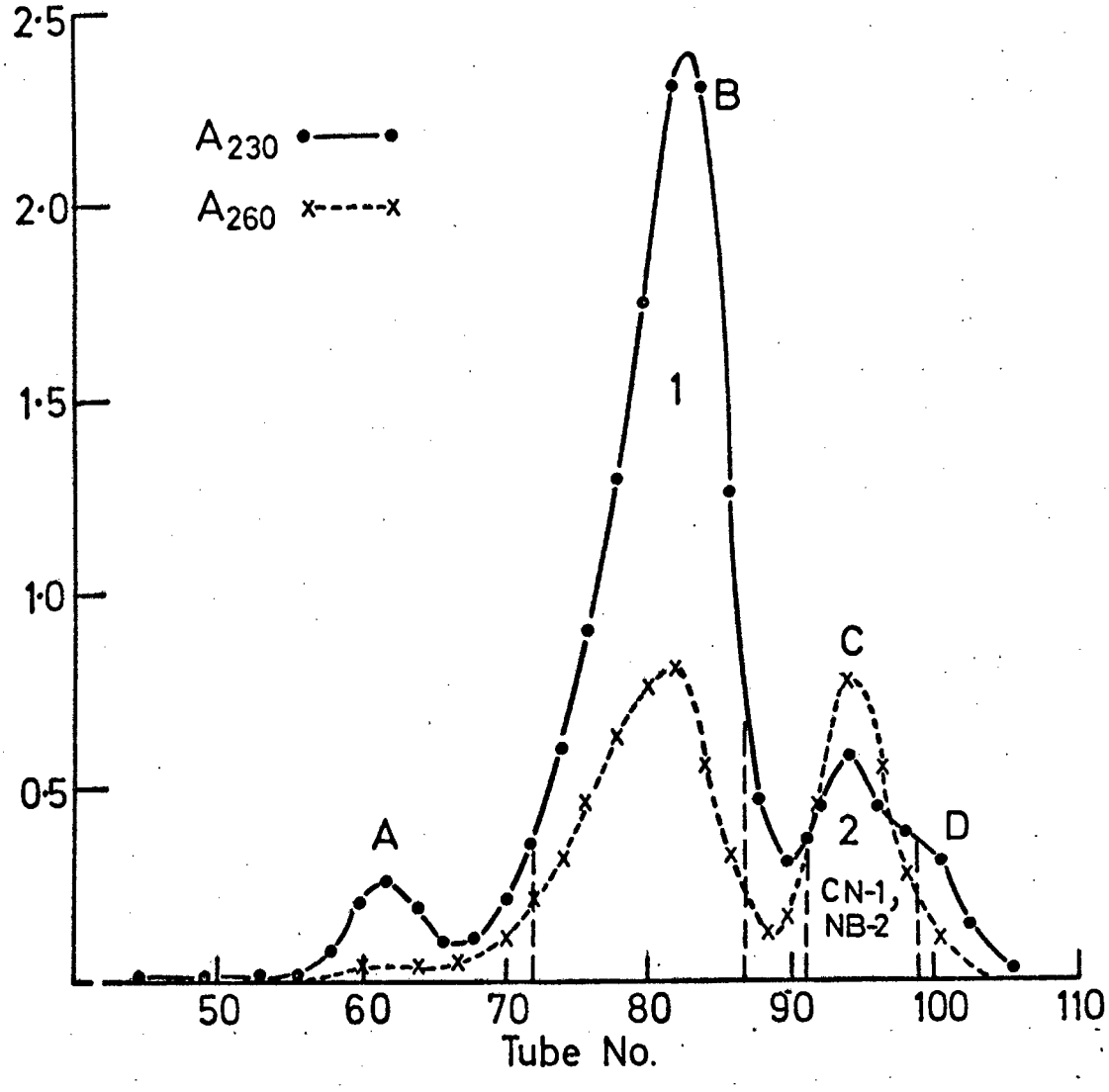


Fig. 2.7 : Elution pattern obtained on a Sephadex G-100 2.5 x 100 cm column after cleaving fragment CN-1 (67 mg) with N-bromosuccinimide at tyrosine residues. The eluent used was 0.01 N HCl, the flow rate 17 ml/h and the fraction volume 5.5 ml. The elution volume of peak A corresponds to uncleaved fragment (Fig. 2.1) and peak D to urea used to dissolve the fragments.

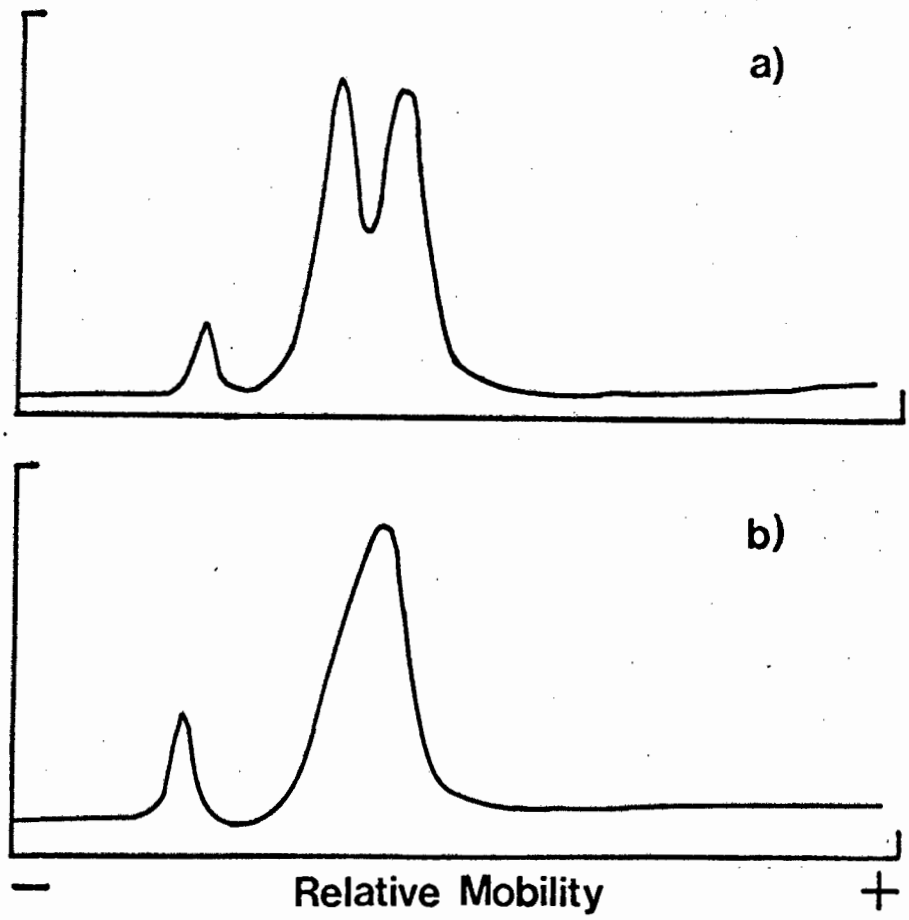


Fig. 2.8 : Electrophoretic pattern of fraction 2 (CN-1 NB-2, Fig. 2.7) (a). Fraction 2 freeze-dried and applied in urea onto the gel (4.3.1) (b) after fraction 2 had been treated with a NaOH solution (Fig. 2.9).

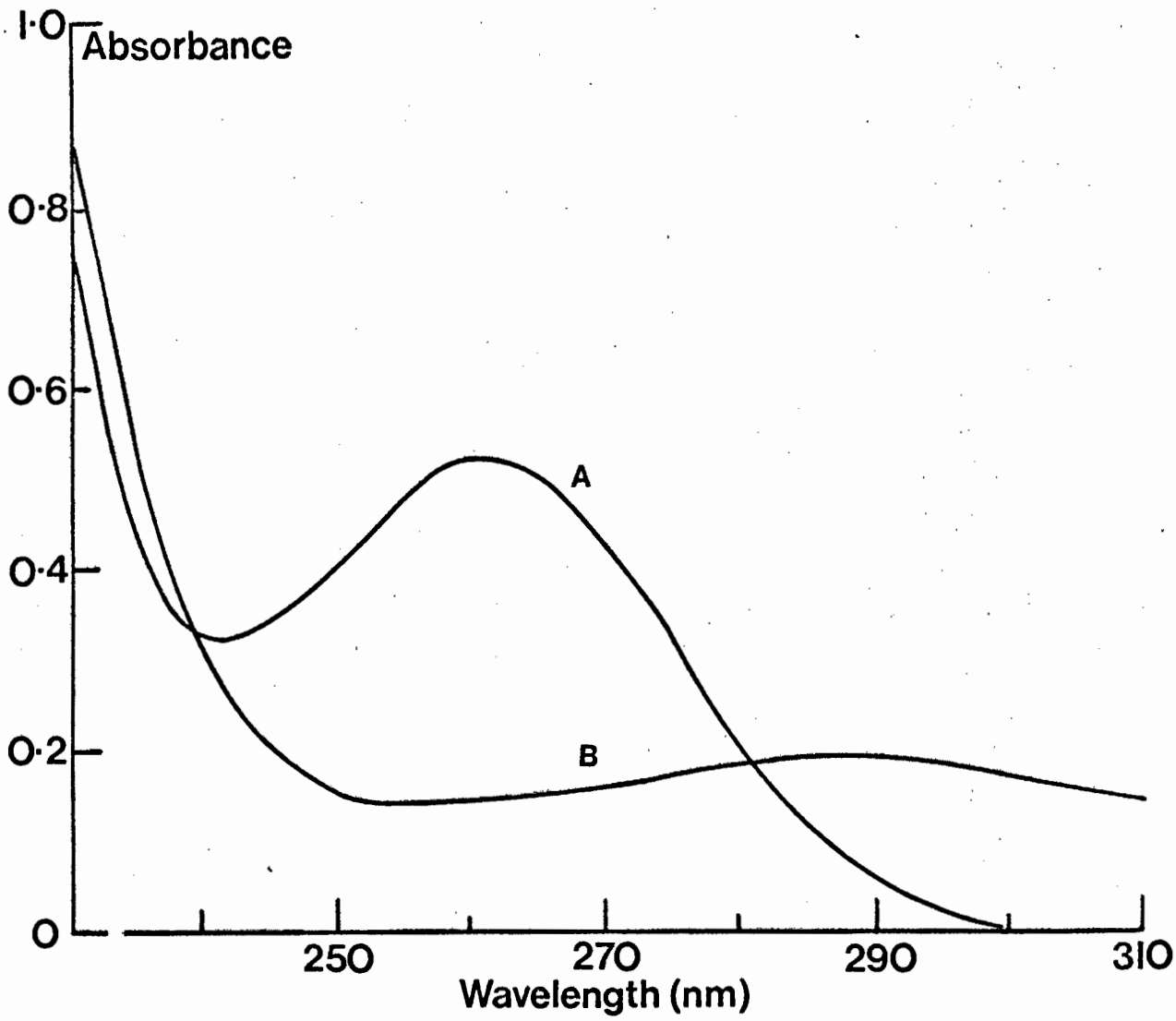


Fig. 2.9 : Ultraviolet spectra of fraction 2 (CN-1 NB-2, Fig. 2.7). (A) Fraction 2 in 0.01 N HCl after elution from a Sephadex G-25 column and (B) after the solution had been raised to pH 9 with NaOH.

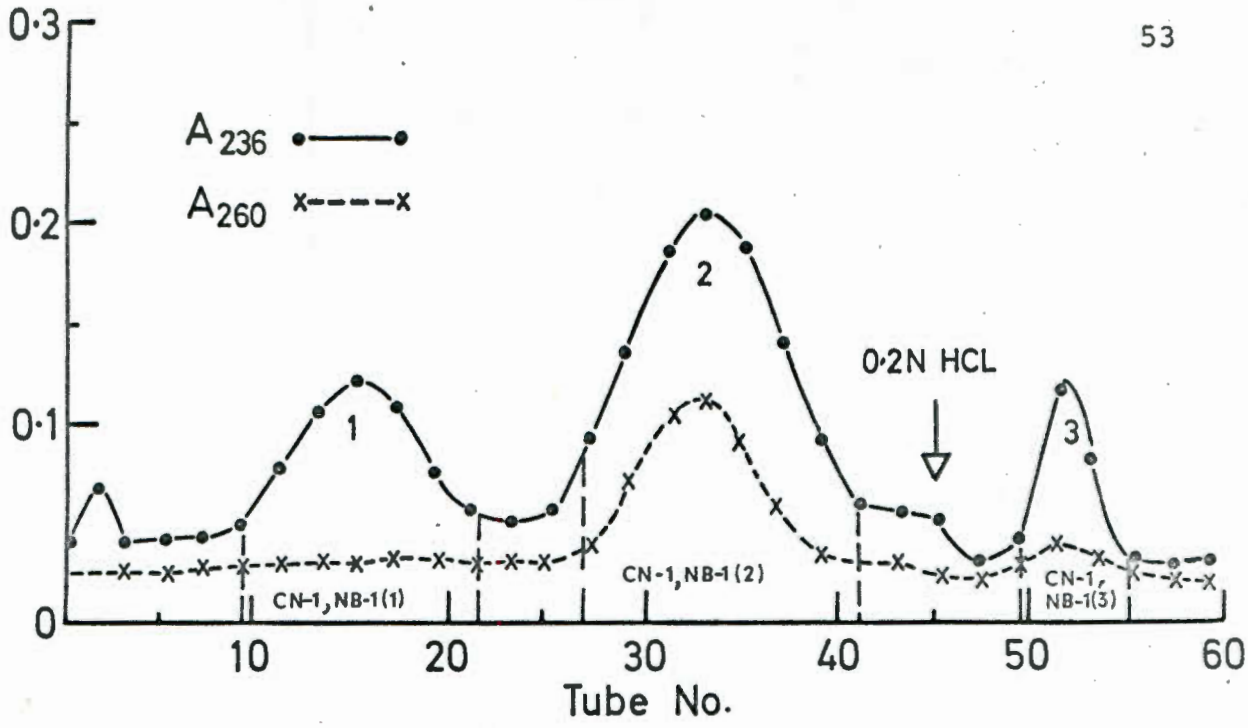


Fig. 2.10 : Elution pattern obtained by subjecting fraction 1 (Fig. 2.7) to CMC-column chromatography. The fractions were eluted at pH 4.4 (acetate buffer) with a linear NaCl gradient and finally with 0.2 N HCl (4.4.4). The flow rate was 30 ml/h and the fraction volume 5.5 ml.

Fig. 2.11 : Electrophoretic pattern of fractions obtained by cleaving fragment CN-1 with N-bromosuccinimide at tyrosine residues. Gel 1, 2 and 3 correspond to fractions 1, 2 and 3 eluted from a CMC-column (Fig. 2.10). All gels were run for 2 h (4.3.1).

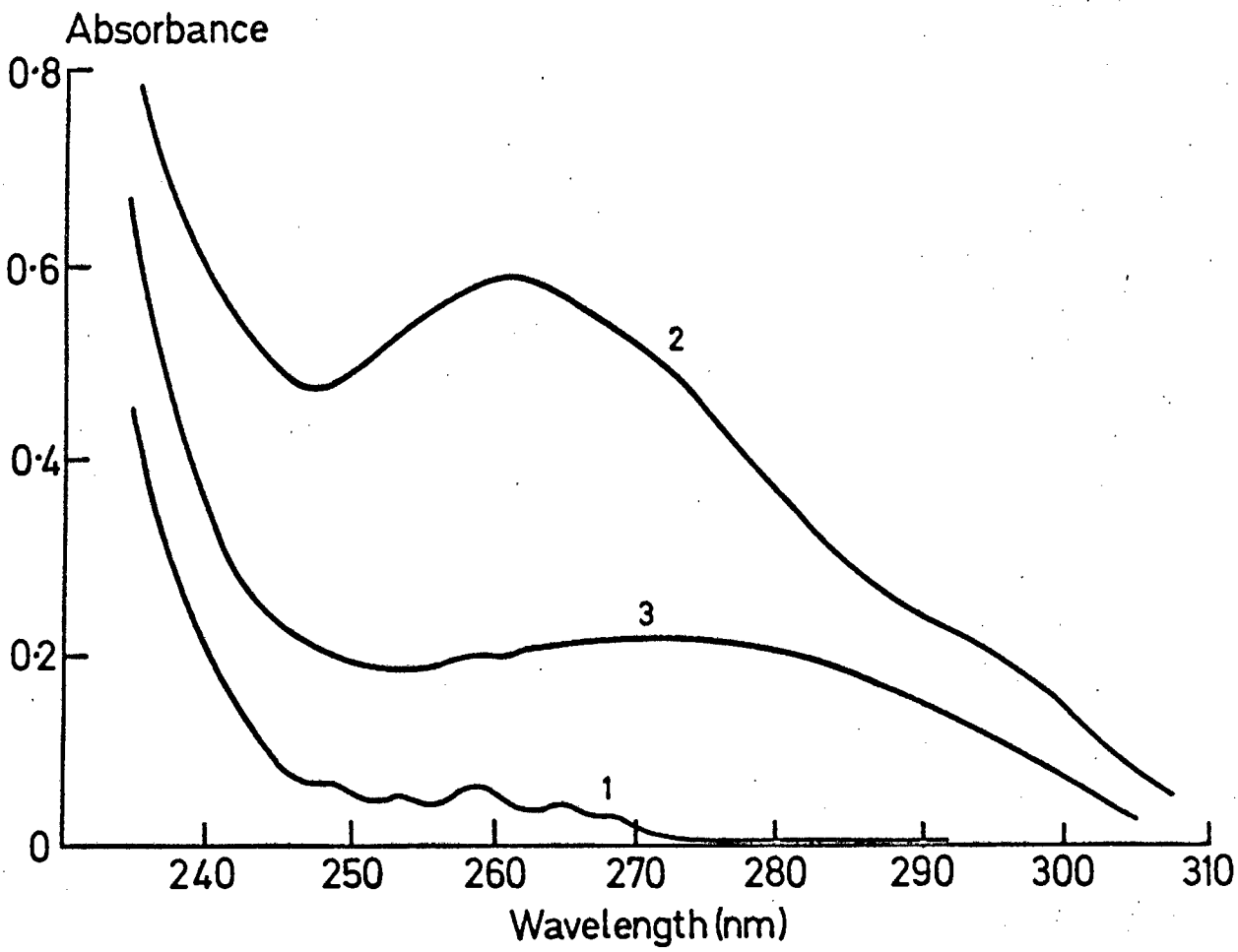


Fig. 2.12 : Ultraviolet spectra of desalted fractions 1, 2 and 3 (CN-1 NB-1(1), CN-1 NB-1(2) and CN-1 NB-1(3) eluted from the CMC-column (Fig. 2.10). The respective amounts are 0.43, 0.52 and 0.61 mg/ml.

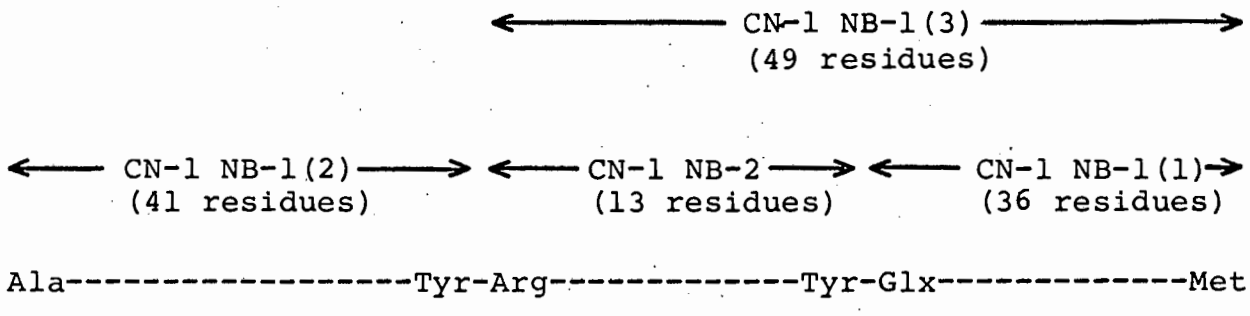
characterized by their ultraviolet spectra (Fig. 2.12), N-terminal residues and amino acid composition (Table 2.5) which revealed that they were homogeneous.

The yields for the fragments were CN-1 NB-2:3 mg, CN-1 NB-1(1):12 mg, CN-1 NB-1(2):15 mg and CN-1 NB-1(3):2 mg.

2.3.5 CN-1 : Alignment of fragments

The amino acid composition of fragment CN-1 NB-1(1), CN-1 NB-1(2) and CN-1 NB-2 (Table 2.5) accounted for all the amino acids present in the uncleaved fragment CN-1 (Table 2.2). Fragment CN-1 NB-1(3) arises from the incomplete cleavage at tyrosine residue between fragment CN-1 NB-2 and CN-1 NB-1(1) (Table 2.5). This indicates that the Tyr-Glx bond is less susceptible to the cleavage compared to the Tyr-Arg bond (Gross, 1967). The absence of spirolactone in fragment CN-1 NB-1(1) (Fig. 2.12) and the presence of homoserine lactone (Table 2.5) places it in the C-terminal position of fragment CN-1. Fragment CN-1 NB-1(2) is positioned in the N-terminal region since it has the same N-terminal residue as the uncleaved fragment viz. Ala (Table 2.5). Fragment CN-1 NB-2 is therefore placed in the middle portion of fragment CN-1 which is confirmed by the isolation of fragment CN-1 NB-1(3) in which CN-1 NB-2 is still attached to the N-terminus of fragment CN-1 NB-1(1) (Table 2.5).

The following tentative structure of fragment CN-1 can therefore be drawn :



Fragment CN-1 NB-1(3) could become useful in the elucidation of the primary structure of peptide CN-1 NB-2 since it is less likely to pose solubility problems in the automated Edman degradation.

TABLE 2.5

AMINO ACID COMPOSITION AND TERMINAL RESIDUES OF NBS CLEAVAGE FRAGMENTS OF CN-1

Fragment	Fraction 2 (Fig. 2.7) CN-1 NB-2		Fraction 1 (Fig. 2.10) CN-1 NB-1(1)	
Amino acid	mole %	residues	mole %	residues
Lys	0.10	0.01	8.31	3.02 (3)
His[0]	-		-	
NH_3	9.10	1.18	9.70	3.52
Arg	30.30	3.98 (4)	11.91	4.33 (4)
Asp	0.10	0.01	5.38	1.96 (2)
Thr	7.54	0.98 (1)	6.01	2.19 (2)
Ser	0.10	0.01	7.92	2.88 (3)
Glu	9.57	1.24 (1)	16.97	6.17 (6)
Pro	7.71	1.00 (1)	2.70	0.98 (1)
Gly	7.71	1.00 (1)	0.67	0.23
Ala	9.40	1.22 (1)	6.33	2.30 (2)
Cys	-	-	-	
Val	8.05	1.04 (1)	5.43	1.97 (2)
homoSer	-		1.14	0.38 (1)
Ile	7.20	0.93 (1)	5.33	1.94 (2)
Leu	8.22	1.07 (1)	13.88	5.05 (5)
Tyr*	+	0.75 (1)	-	
Phe	-		8.02	2.97 (3)
Residue/ mole		(13)		(36)
N-terminal (4.3.6)	Arg		Glx	
C-terminal (4.4.2)	Tyr*		homoSer	

[0] Histidine was qualitatively determined by the Pauly reaction (4.4.2.1).

* After NBS cleavage tyrosine residues are converted to the spirolactone which possesses a characteristic ultraviolet spectrum (Figs. 2.9, 2.12) (4.4.2). C-terminal residues

TABLE 2.5 /cont'd.....

Fragment / Amino acid	Fraction 2 (Fig. 2.10) CN-1 NB-1(2) mole %	residues	Fraction 3 (Fig. 2.10) CN-1 NB-1(3) mole %	residues
Lys	19.48	7.91 (8)	7.74	3.67 (3)
His^0	+	(1)	-	
NH_3	14.40	5.70	10.67	5.05
Arg	13.71	5.40 (5)	16.33	7.76 (8)
Asp	0.40	0.15	3.94	1.87 (2)
Thr	13.01	5.20 (5)	6.80	3.23 (3)
Ser	4.41	1.77 (2)	6.01	3.13 (3)
Glu	4.84	1.94 (2)	14.56	6.91 (7)
Pro	7.74	3.10 (3)	4.21	1.99 (2)
Gly	9.99	4.00 (4)	2.75	1.30 (1)
Ala	20.96	8.38 (8)	7.97	3.75 (3)
Cys	-		-	
Val	2.67	1.07 (1)	5.87	2.78 (3)
homoSer	-		1.10	0.52 (1)
Ile	0.70	0.25	5.35	2.54 (3)
Leu	2.46	0.98 (1)	11.93	5.67 (6)
Tyr*	+	(1)	+	(1)
Phe	0.70	0.25	5.46	2.59 (3)
Residue/ mole		(41)		(49)
N-terminal (4.3.6)	Ala		Arg Small amount Ala	
C-terminal (4.4.2)	Tyr*		homoSer	

become apparent from the amino acid composition and the nature of the cleavage (4.4.2).

The molecular composition of fraction 3 is deduced by assuming that it is contaminated by fraction 2 (CN-1 NB-1(2)) (compare N-terminal groups).

Since the contaminant is very rich in Lys and Ala only 3 Lys and 3 Ala residues have been assigned to this fraction instead of 4.

2.4 CLEAVAGE OF FRAGMENT CN-1 NB-1(1) WITH DILUTE ACID

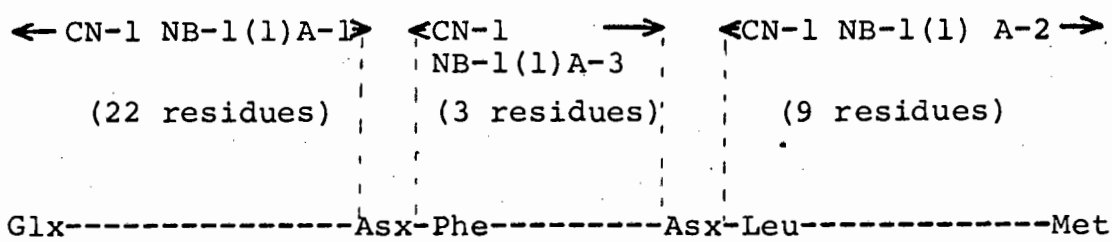
2.4.1 Cleavage and purification of fragments

Fragment CN-1 NB-1(1) (4 mg) was cleaved at aspartic acid residues in 0.03 M HCl at 105°C for 17 h (4.4.3). The optimum time of cleavage was established from the appearance of free amino acid during the hydrolysis (Fig. 2.13).

The resulting fragments were separated on a Sephadex G-25 column (Fig. 2.14). The fractions were characterized by their endgroups and amino acid composition (Table 2.6). Although fraction 1 was fairly pure as judged from N-terminal analysis fraction 2 and 3 were cross-contaminated. They might also contain smaller fragments produced by unspecific cleavage of peptide bonds. Further purification was not attempted at this stage.

2.4.2 Alignment of fragments

The presence of homoserine lactone in CN-1 NB-1(1) A-2 (fraction 2, Table 2.6) places it in the C-terminal position of the uncleaved fragment. Since the N-terminal groups of the uncleaved fragment (CN-1 NB-1(1)) and CN-1 NB-1(1) A-1 (fraction 1) are identicalviz. Glx (Table 2.6) the following tentative structure of fragment CN-1 NB-1(1) can be drawn :



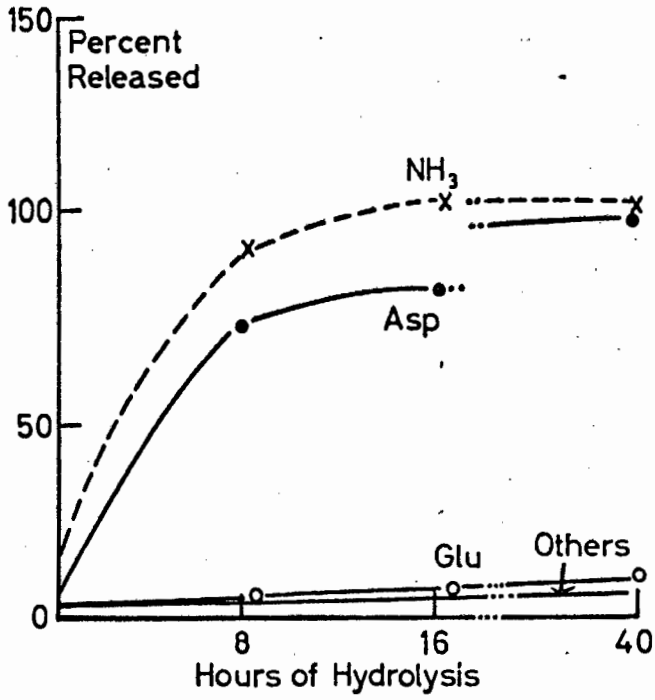


Fig. 2.13 : Appearance of free amino acids during hydrolysis of fragment CN-1 NB-1(1) in 0.03 N HCl at 105°C (4.4.3). The amounts of free amino acids were expressed as percentage of total amount of each present in the fragment.

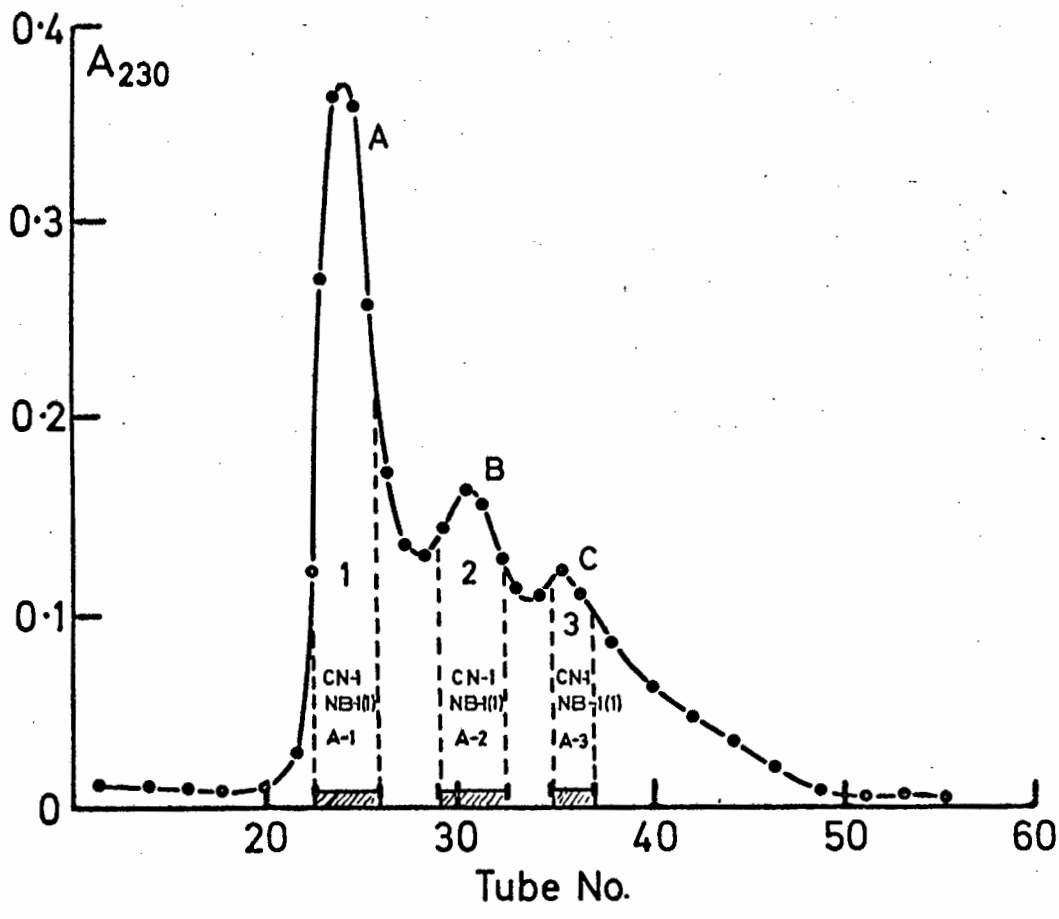


Fig. 2.14 : Elution pattern obtained on Sephadex G-25 1.5 x 90 cm column after cleavage of fragment CN-1 NB-1(1) with dilute HCl at aspartic acid residues (4.4.3). The eluent used was 0.01 N HCl, the fraction size was 1 ml and the flow rate 8 ml/h.

TABLE 2.6

AMINO ACID COMPOSITION OF FRACTIONS OBTAINED BY CLEAVING FRAGMENT CN-1 NB-1(1) WITH DILUTE ACID

	Fraction 1 (Fig. 2.14)			Fraction 2 (Fig. 2.14)			Fraction 3 (Fig. 2.14)	
Fragment	CN-1 NB-1(1) A-1			CN-1 NB-1(1) A-2			CN-1 NB-1(1) A-3	
Amino acid	mole %	residues		mole %	residues		mole %	residues
Phe	4.98	1.13	(1)	7.90	0.8	(1)	15.1	
Lys	10.22	2.30	(2)	3.20	0.3	(0)	12.1	
NH_3	17.00			10.10	1.1		30.5	
Arg	14.80	3.36	(3)	11.70	1.2	(1)	9.1	
Asp	-			0.40	0.0	(0)	7.4	
Thr	4.28	0.97	(1)	3.30	0.3		10.6	
Ser	3.60	0.81	(1)	16.50	1.8	(2)	7.2	
Glu	20.46	4.65	(5)	16.00	1.7	(1)	14.5	
Pro	4.50	1.02	(1)	-			-	
Gly	-			-			-	
Ala	3.91	0.89	(1)	9.40	1.0	(1)	7.9	
Val	4.37	0.99	(1)	10.30	1.1	(1)	6.0	
homoSer	-			3.90	0.4	(1)	-	
Ile	10.50	2.30	(2)	4.30	0.4	(0)	3.0	
Leu	18.48	4.20	(4)	12.50	1.3	(1)	7.5	
Residue/ mole			(22)			(9)		
N-terminal (4.3.6)	Glx			Leu Smaller amount of others especially Phe			Phe Smaller amount of others especially Leu	

Fraction 2 seems to be contaminated by fraction 1, which is rich in Glu residues. Therefore, instead of 2 Glu residues only 1 was assigned to this fraction on the basis of Lys ratios. Fraction 3 contains free amino acids especially Asp and probably small fragments produced by unspecific cleavage, and it is therefore not possible to deduce the molecular composition. The composition can, however, be deduced indirectly to be Phe_1, Lys_1, Thr_1 (i.e. the difference between and amino acid composition of CN-1 NB-1(1) and fraction 1, fraction 2 and two aspartic acid residues.

2.5 N-BROMOSUCCINIMIDE CLEAVAGE OF HISTONE F3

2.5.1 Cleavage and purification of fragments

Cyanogen bromide cleavage of histone F3 followed by N-bromosuccinimide and dilute acid cleavage had produced a total of 11 fragments and two free amino acids. Endgroup determination, amino acid composition of the fragments and the specificity of the cleavage reaction had made it possible to align all fragments.

Subsequent sequence determinations could, however, pose difficulties towards the C-terminal end of the various fragments, due to their increasing solubility as a result of shortening of the peptide chains. Therefore, a further series of fragmentation of the protein was undertaken. The aim was to produce new fragments in which amino acid sequences in the C-terminal position of the peptides of the first series, became positioned more towards the centre in the fragments to be produced in the second series.

Performic acid oxidized histone F3 (4.3.5) was dissolved in 50% (v/v) acetic acid and titrated with NBS until no further increase in the A260 reading occurred (4.4.2.1). After the cleaved protein had been freeze-dried it was dissolved in 6 M urea and applied to a Sephadex G-100 column (Fig. 2.15). Fractions were pooled as indicated and characterized by gel electrophoresis (Fig. 2.16). From previous results it was expected that all fragments must be contained in fraction 1 and 2, whereas peak A and B therefore would correspond to uncleaved and partially cleaved histone.

Fraction 2 was desalted on a Sephadex G-25 column. Its electrophoretic heterogeneity, due to the spirolactone (Fig. 2.16), N-terminal amino acid and amino acid composition was identical to that of fragment CN-1 NB-2 (compare Tables 2.5 and 2.7).

Fraction 1 was freeze-dried and applied to a CMC-column and eluted with a NaCl gradient at pH 4.4 (Fig. 2.17) (4.4.4). Fractions were pooled as indicated and desalted on a Sephadex G-25 x 30 cm column.

All four fractions were characterized by their electrophoretic mobility, endgroups and amino acid composition (Fig. 2.18, Table 2.7). Fraction 2 (NB-1(2)) (Table 2.7)

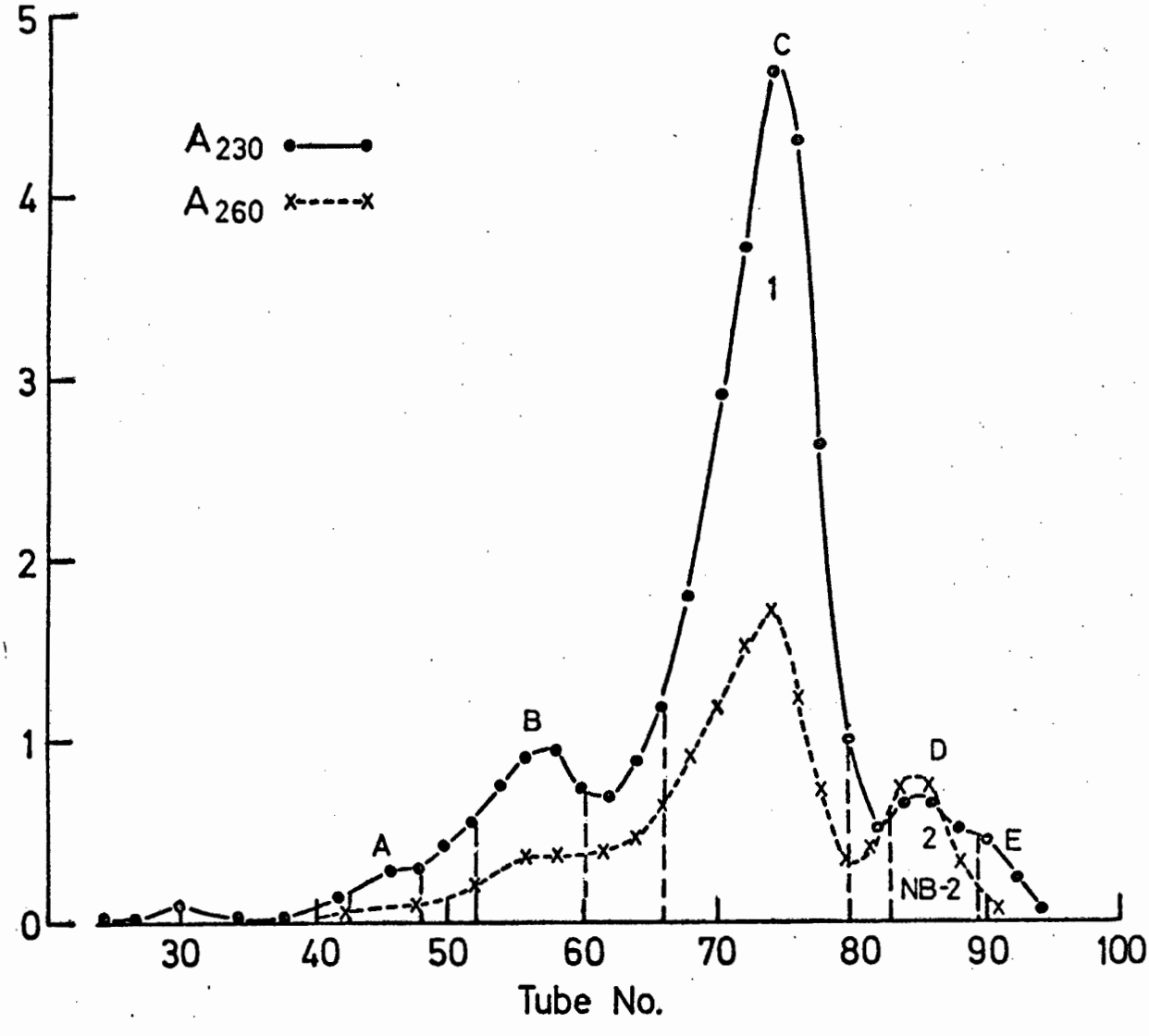


Fig. 2.15 : Elution pattern obtained on a Sephadex G-100 2.5 x 100 cm column after performic acid oxidized histone F3 (100 mg.) had been cleaved with NBS at tyrosine residues (4.4.2.1). The eluent used was 0.01 N HCl, the flow rate 17 ml/h and the fraction size was 5.5 ml. Peak E corresponds to urea used to disaggregate the fragments (4.2.1).

Fig. 2.16 : Electrophoretic pattern of fractions obtained after cleaving performic acid oxidized histone F3 with N-bromosuccinimide. Gel 1 corresponds to F3 dimer that had been reduced with mercaptoethanol and gel 2, 3, 4 and 5 to peaks A, B, C and D in Fig. 2.15. All gels were run for approximately 1.5 h (4.3.1).

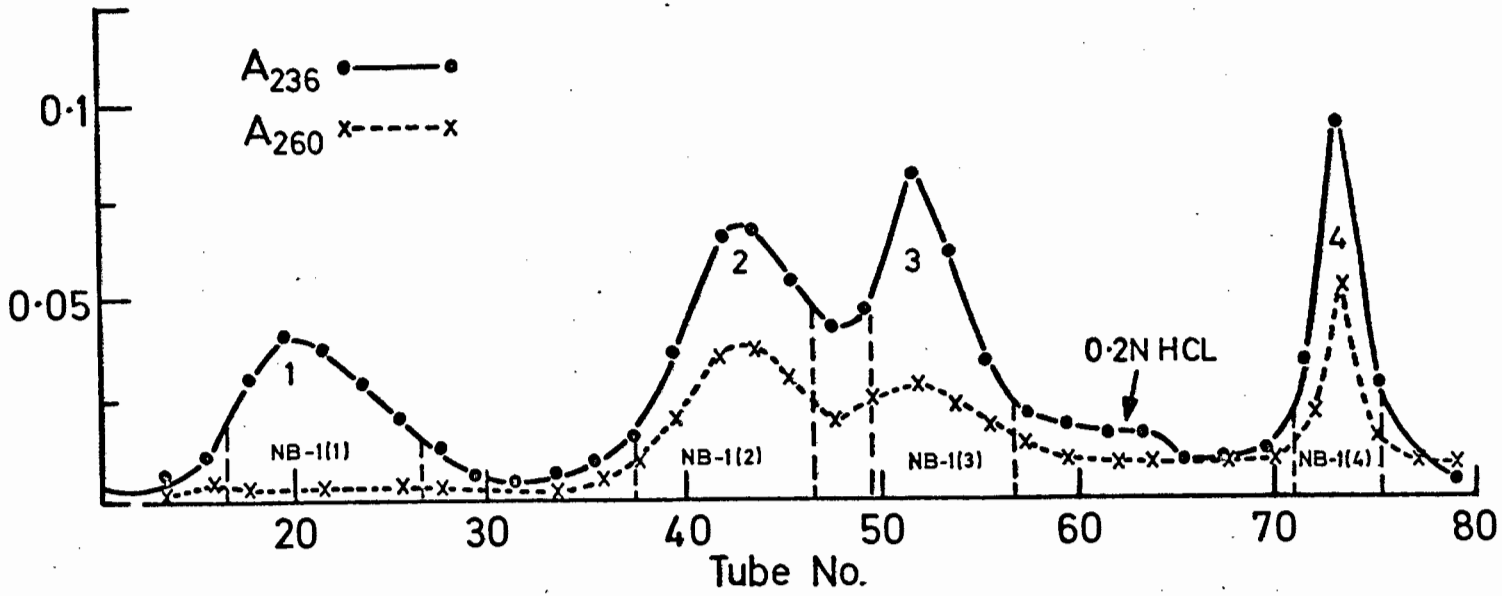


Fig. 2.17 : Elution pattern obtained by subjecting fraction 1 (Fig. 2.15) to CMC-column chromatography. The fractions were eluted at pH 4.4 (acetate buffer) with a linear NaCl gradient, and finally with 0.2 N HCl (4.4.4). The fraction size was 3.5 ml and the flow rate 20 ml/h.

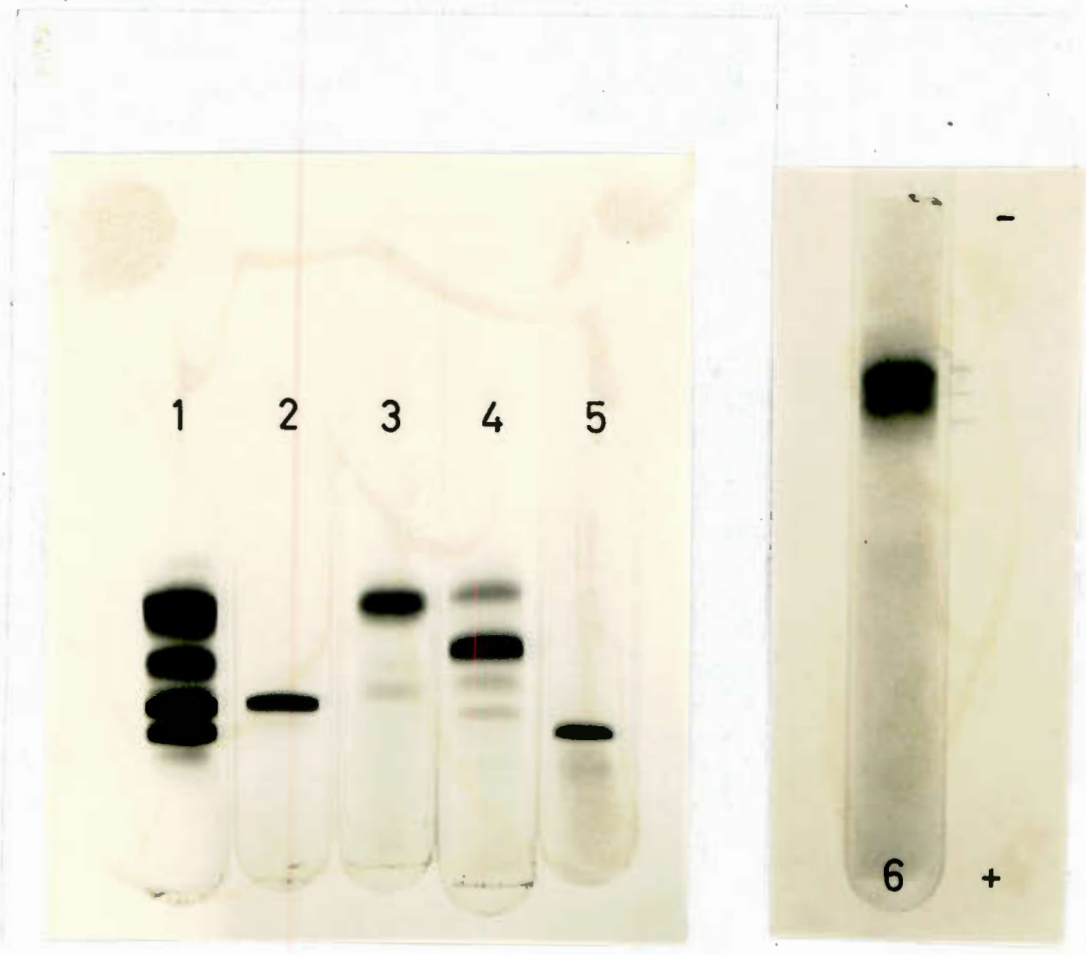


Fig. 2.18 : Electrophoretic pattern of fractions obtained after subjecting fraction 1 (Fig. 2.15) to CMC-column chromatography. Gel 1 corresponds to fraction 1 (Fig. 2.15) and gels 2, 3, 4 and 5 to fractions 1, 2, 3 and 4 (NB-1(1), NB-1(2), NB-1(3) and NB-1(4)) eluted from the CMC-column (Fig. 2.17). Gel 6 is the same as gel 3 but was electrophoresed for 2.5 h instead of 1.5 h (4.3.1).

TABLE 2.7

AMINO ACID COMPOSITION AND TERMINAL RESIDUES OF N-BROMOSUCCINIMIDE CLEAVAGE FRAGMENTS OF HISTONE F3

Fragment	Fraction 2 (Fig. 2.15)		Fraction 1 (Fig. 2.17		Fraction 2 (Fig. 2.17)	
	NB-2		NB-1(1)		NB-1(2)	
Amino acid	mole %	residues	mole %	residues	mole %	residues
Lys	0.04	0.05	5.40	1.90 (2)	19.46	7.78 (8)
His[0]	-		+	(1)	+	(1)
NH_3	9.12		9.80		14.67	
Arg	31.98	4.05 (4)	14.34	5.03 (5)	13.38	5.35 (5)
Asp	0.10	0.07	9.10	3.20 (3)	0.68	0.31
Thr	7.71	0.98 (1)	5.42	1.90 (2)	11.36	4.74 (5)
Ser	-		-		4.79	1.92 (2)
Glu	9.12	1.15 (1)	9.15	3.23 (3)	6.35	2.40 (2)
Pro	7.01	0.89 (1)	3.00	1.06 (1)	7.96	3.18 (3)
Gly	8.59	1.08 (1)	5.74	2.03 (2)	9.43	3.77 (4)
Ala	8.07	1.02 (1)	11.90	4.20 (4)	19.36	7.74 (8)
$CySO_3H$	-		2.60	0.92 (1)	-	
Val	8.42	1.06 (1)	5.47	1.92 (2)	2.89	1.16 (1)
$MetSO_2$	-		2.00	0.70 (1)	-	
Ile	7.36	0.93 (1)	11.26	3.97 (4)	0.68	0.27
Leu	10.17	1.28 (1)	11.90	4.20 (4)	3.42	1.36 (1)
Tyr*	+	(1)	-		+	(1)
Phe	0.05	0.05	2.70	0.95 (1)	0.49	0.20
Residues/ mole		(13)		(36)		(41)
N-terminal (4.3.6)	Arg		Leu		Ala (mainly)	
C-terminal (4.4.2)	Tyr*		Ala		Tyr*	

[0] Histidine was qualitatively determined by the Pauly reaction (4.4.2.1).

* After NBS cleavage tyrosine residues are converted to the spirolactone which possesses a characteristic ultraviolet spectrum (Fig. 2.9, 2.12) (4.4.2). C-terminal residues

TABLE 2.7 /cont'd.....

Fragment / Amino acid	Fraction 4 (Fig. 2.17) NB-1(4) mole %	residues	Fraction 3 (Fig. 2.17) NB-1(3) mole %	residues
Lys	7.00	3.05 (3)	5.97	3.43 (3)
His[0]	-		-	
NH_3	10.00		13.10	
Arg	11.84	5.15 (4-5)	15.36	7.77 (8)
Asp	4.67	2.03 (2)	5.23	2.98 (3)
Thr	4.07	1.78 (2)	5.80	3.30 (3)
Ser	4.62	2.01 (2-3)	6.50	3.73 (4)
Glu	20.50	9.00 (9)	16.03	9.16 (9-10)
Pro	2.08	0.92 (1)	3.86	2.20 (2)
Gly	1.68	0.73 (0)	3.16	1.81 (2)
Ala	12.60	5.48 (5)	11.24	6.40 (6-7)
$CySO_3H$	-		-	
Val	4.75	2.06 (2)	5.61	3.21 (3)
$MetSO_2$	1.46	0.64 (1)	0.70	0.40 (1)
Ile	5.29	2.30 (2)	5.09	2.91 (3)
Leu	13.66	5.90 (6)	10.88	6.20 (6)
Tyr*	+	(1)	+	(1-2)
Phe	5.88	2.53 (3)	4.56	2.61 (3)
Residues/ mole		(43-45)		(55-58)
N-terminal (4.3.6)	Ala		Arg (mainly)	
C-terminal (4.4.2)	Tyr*		Tyr*	

become apparent from the amino acid composition and the nature of the cleavage (4.4.2).

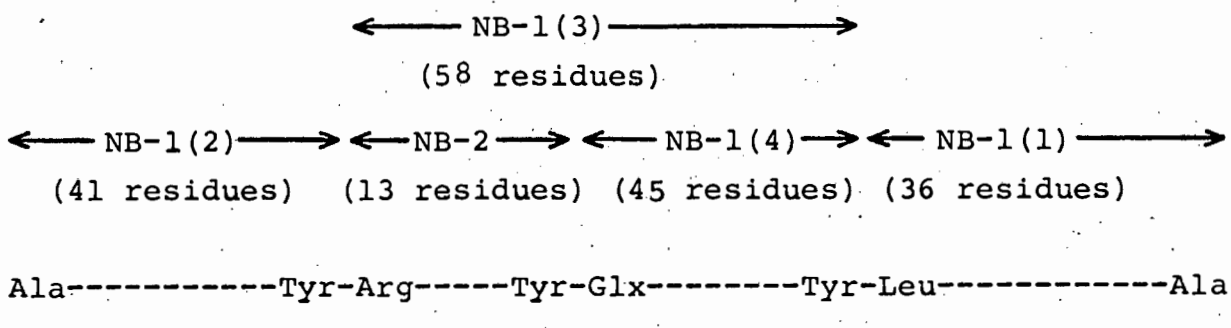
A certain amount of ambiguity exists in the number of residues in fraction 3 and 4 due to cross contamination between the various fractions.

was not entirely pure, but, since it had been obtained in a pure form by cleaving fragment CN-1 with NBS (2.3.4), no attempts were made to purify it further.

Fragments NB-1(1), NB-1(2), NB-1(4) and NB-2 accounted for all the amino acids present in F3 histone. Fragment NB-1(3) corresponds to fragment NB-2 still being attached to fragment NB-1(4) (Table 2.7). No attempts had been made to purify it further.

2.5.2 Alignment of fragments

The absence of spirolactone (260 nm absorption) in fragment NB-1(1) (Fig. 2.17) places it in the C-terminal position of the protein. Fragment NB-1(2) must be in the N-terminal part of the protein since both possess Ala as their N-terminal residue. Fragment NB-1(2) is followed by fragment NB-2 and then by NB-1(4). This is deduced from the fact that in fragment NB-1(3), NB-2 is attached to the N-terminus of NB-1(4) (Table 2.7).



This sequence was expected from the results obtained from the NBS cleavage of the CNBr-fragments (2.3.5). Note that fragment NB-1(2) possesses the electrophoretic heterogeneity also apparent in F3 histone and thus contains all the acetylated lysine residues.

2.6 FRAGMENTATION OF HISTONE F3 : SUMMARY AND CONCLUSIONS

For the purpose of sequence analysis histone F3 was fragmented into a series of smaller fragments (Fig. 2.19). Specific chemical cleavages were chosen and applied first to the original protein chain and subsequently to the generated polypeptides to yield sets of not more than 3 peptides in any single cleavage. Their relative position in the protein or polypeptides became evident after comparison of the N- and C-terminal amino acids in the cleaved products and the uncleaved starting material. The simplicity of the peptide mixture after each cleavage resulted in an easy separation of the fragments.

The fragmentation of the protein and fragments is summarized in Fig. 2.19.

To provide for the eventuality that the C-terminal sections of some peptides can not be satisfactorily sequenced due to solubility problems, an additional set of larger fragments were prepared by subjecting the protein directly to N-bromosuccinimide cleavage at the tyrosine residues (Fig. 2.19).

From the results it is also evident that all four fragments obtained by cleaving histone F3 with N-bromosuccinimide can be readily aligned due to the fact that the tyrosine peptide bonds possess different susceptibilities to the cleavage (2.3.2, 2.5.2). It seems therefore possible, if this should be required, to align four or more fragments obtained in a single cleavage by characterizing the partially cleaved fragments. In this case, the relative position of the fragments also became apparent from results obtained previously.

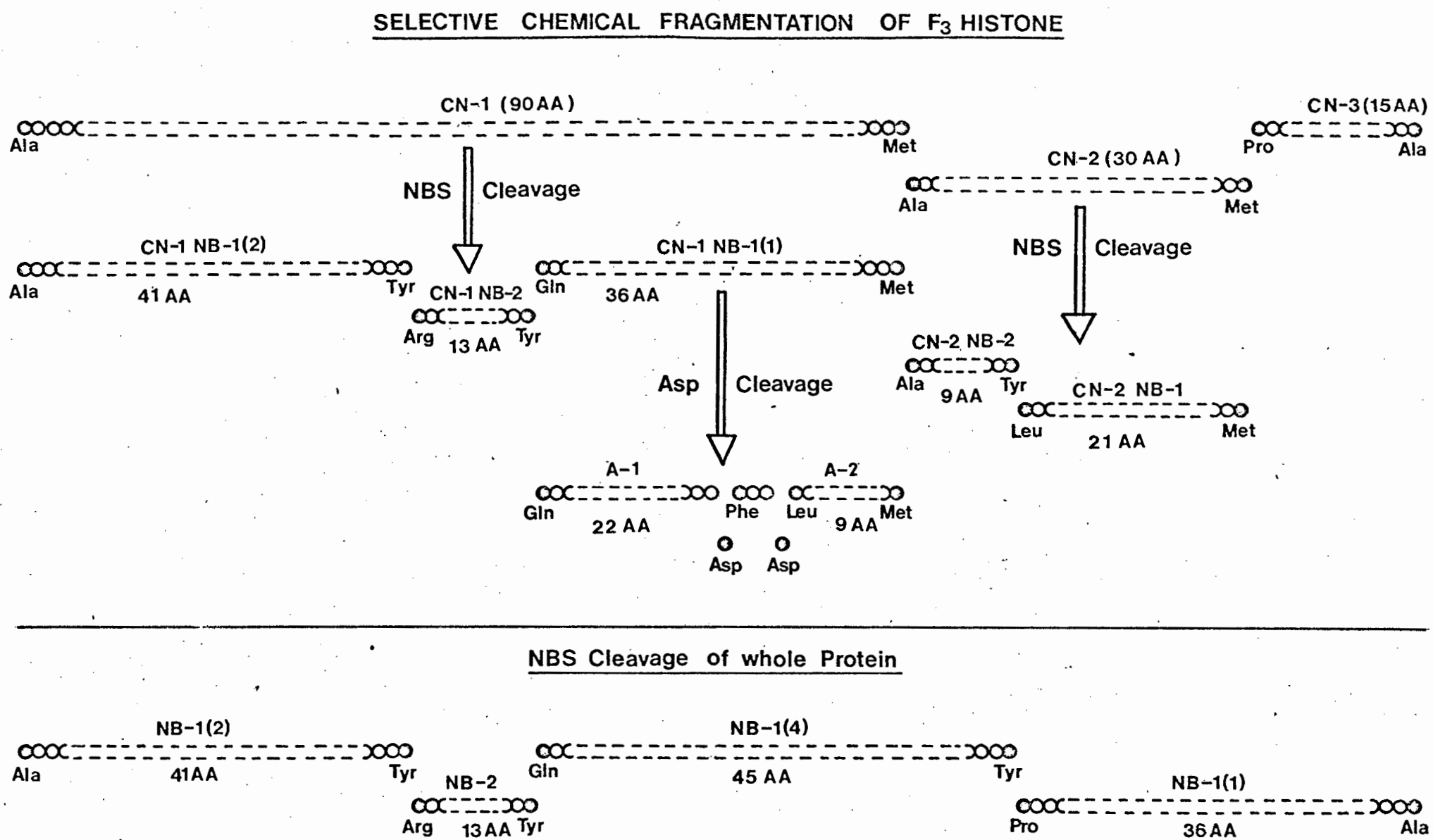


Fig. 2.19 : Summary of the selective chemical fragmentation of histone F3 and the alignment of fragments

2.7 SEQUENCE ANALYSIS OF HISTONE F3

2.7.1 Outline of approach

Histone F3 has been degraded into a series of fragments the position of which in the protein molecule is known (Fig. 2.19).

The elucidation of the amino acid sequence of this protein was now attempted by subjecting, at first, the uncleaved protein to the automatic Edman degradations. Depending on how far the stepwise degradation could be carried, the next fragment in the protein chain would be selected to continue with the sequence determination from the N-terminal end towards the C-terminus. The available fragments could be used in the following order : CN-1 NB-2, CN-1 NB-1(1), CN-1 NB-1(1) A-2, CN-2, CN-2 NB-1 and CN-3 (Fig. 1.19). If the amino acid sequence of a particular fragment became apparent from the degradation of the preceding fragment, the degradation of the former becomes unnecessary, e.g. should it be possible to align the first 54 amino acids in the structure of the uncleaved protein, fragment CN-1 NB-2 would not be required for sequence studies. On the other hand, should it not be possible to elucidate the complete amino acid sequence of a fragment, additional cleavages would have to be performed.

If losses of peptides with advancing degradation due to increasing solubility should occur, another fragment in which the C-terminal part had been shifted into the N-terminal region could be sequenced. For example, fragment CN-2 could be substituted by fragment NB-1(1) (Fig. 2.19). One could also resort to chemical procedures to decrease the hydrophobicity of the peptides (Braunitzer et al., 1970).

2.7.2 Preliminary amino acid sequence studies

Uncleaved histone F3 was subjected to sequential Edman degradation. The procedure used for the sequential degradation was similar to that described by Edman and Begg (1967) (4.5.2).

The dried fractions containing the amino acid thiazolinones were converted to the corresponding PTH-

derivatives with 1 N HCl at 80°C for 10 minutes (4.5.2.4). All fractions were subjected to gas chromatographic analysis and selected fractions to thin layer chromatography and spot tests (4.5.3.1.3). Results are shown in Table 2.8. Only the two major amino acids identified together with their quantities are given in the first two columns (Table 2.8). The amino acid present in lower amounts (second column) is invariably due to "carry over" from the previous step. The appearance of this "carry over", which gets more pronounced with increasing number of steps, is probably due to incomplete cleavage (Edman & Begg, 1967; Edman, 1970) and to a lesser extent incomplete coupling and extraction.

As the stepwise degradation proceeds, gradually more amino acids appear contributing to a general increase of the background. This is thought to be due to unspecific cleavage of the protein chain (Edman, 1970) since their relative amounts are generally related to the amino acid composition of the protein (Table 2.13, Page 84).

To assign an amino acid identified in the chromatogram to a specific position in the sequence requires, therefore, a quantitative comparison at each cycle. Only if, after one degradation cycle, the concentration of a particular amino acid increases over its concentration in the previous step can an assignment be made. For this reason column 3 in Table 2.8 was included which gives the concentration of that amino acid which, in the subsequent step, increases significantly, e.g. position 24 (Table 2.8) can be assigned to Ala since there is a three fold increase in the amount of alanine over the background concentration of this amino acid in cycle 23. The assignment gets progressively more difficult due to the cumulative effect of a few percent decrease in yield in each cycle and the gradual appearance in the chromatograms of an increasing general background of other PTH-amino acids which eventually make the identification impossible. Therefore, assignments of amino acid positions in degradation cycle 29 and 30 can only be done with reservation (Table 2.8).

TABLE 2.8

SEQUENTIAL DEGRADATION OF HISTONE F3

Yields of some of the PTH-amino acids recovered after sequential degradation of 7 mg histone F3. Quadrol was used as the coupling buffer in degradation procedure (4.5.2.2) (Edman & Begg, 1967)

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no		Yield in nmoles		Yield in nmoles		Yield in nmoles
1	Ala	320	-		*Arg	-
2	*Arg	+	-		Thr	-
3	^{0}Thr	60,200	-		*Lys	-
4	*Lys	+	-		Glu	-
5	Gln	100	Glu	30	Thr	2
6	^{0}Thr	54,67	Glu	?	Ala	10
7	Ala	200	xLeu	6	*Arg	-
8	*Arg	+	Ala	6	*Lys	-
9	*Lys	+	xLeu	10	Ser	-
10	Ser	20	xLeu	12	Thr	4
11	^{0}Thr	30,63	Ala	12	Gly	-
12	Gly	65	Thr	14	Gly	65
13	Gly	90	xLeu	10	*Lys	-
14	*Lys	+	xLeu	11	Ala	10
15	Ala	106	xLeu	10	Pro	5
16	Pro	35	Ala	8	*Arg	-
17	*Arg	+	Pro	22		
18	X		Pro	15	Glu	4

TABLE 2.8 /cont'd.....

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no		Yield in nmoles		Yield in nmoles		Yield in nmoles
19	Glu	24	Gln	19	xLeu	4
20	Leu	54	Ala	9	Ala	9
21	Ala	44	Leu	32	Thr	5
22	Thr	25	Ala	21		
23	X		xLeu	10	Ala	12
24	Ala	32	xLeu	8	Ala	32
25	Ala	45	xLeu	10	*Arg	±
26	*Arg	+	Ala	7		
27	X		Ala	4		
28	X		xLeu	2	Ala	2
29	Ala	7	xLeu	3	Pro	3
30	Pro	7	Ala	4		

* PTH-Arg has been identified by a spot test and PTH-Lys by thin layer chromatography (4.5.3.1.3).
All other PTH-amino acids have been identified and quantitated by gas chromatography.

xLeu denotes the sum of Leu and Ile (4.5.3.1.3).

⁰ Double peak due to Thr and Δ-Thr (4.5.3.1.3).

Comparing column 1 and 2 in Table 2.8 it is obvious that in a number of cases the assignment of positions becomes difficult. In addition, with increasing number of cycles peaks appeared in the gas chromatogram which could not be assigned to amino acids. It was found that considerable amounts of Quadrol, which had not been extracted in earlier steps in the Edman degradation cycle, contaminated the thiazolinone fractions and added to the gas chromatographic background. The sequence analysis was repeated on the protein for a number of times with prolonged ethyl acetate extraction, vacuum steps, and also addition of acetic acid (0.01 M) to the ethyl acetate (Edman & Begg, 1967). Also, performic acid oxidized histone was used instead of histone F3 dimer. None of the modifications brought improvements over the results in Table 2.8.

During these studies it was noticed that yields consistently dropped at step 16 followed by an increase in carry over of amino acids in all subsequent fractions (e.g. Leu 20 and Leu 21 - Table 2.9). This may indicate either anomalous coupling or cleavage at this proline residue. Therefore, at degradation cycle 16 the coupling reaction with phenylisothiocyanate was performed twice manually (4.5.2.2). The program was then allowed to proceed normally to the end of that cycle. No improvement was evident after the quantitation of proline. In a further experiment the protein was again subjected to normal degradation and at cycle 16 to two double cleavages (4.5.2.2). The gas chromatography of the PTH-amino acid of each of the double cleavage extracts revealed that substantial amounts of Pro has not been cleaved off the protein in one single double cleavage (Table 2.10). This result is not entirely unexpected because of the particular features of the proline peptide bond (Scheraga, 1972). Not only does phenylthiocarbamyl derivative involving proline in certain cases cyclize less readily in peptides, but this property can also be shown on free proline, which couples easily, in fact, several times faster than alanine to phenyl-isothiocyate at pH 8.6, but takes very much longer to cyclize to the PTH-amino acids than alanine in ethanol containing small amounts of HCl (Fig. 2.20).

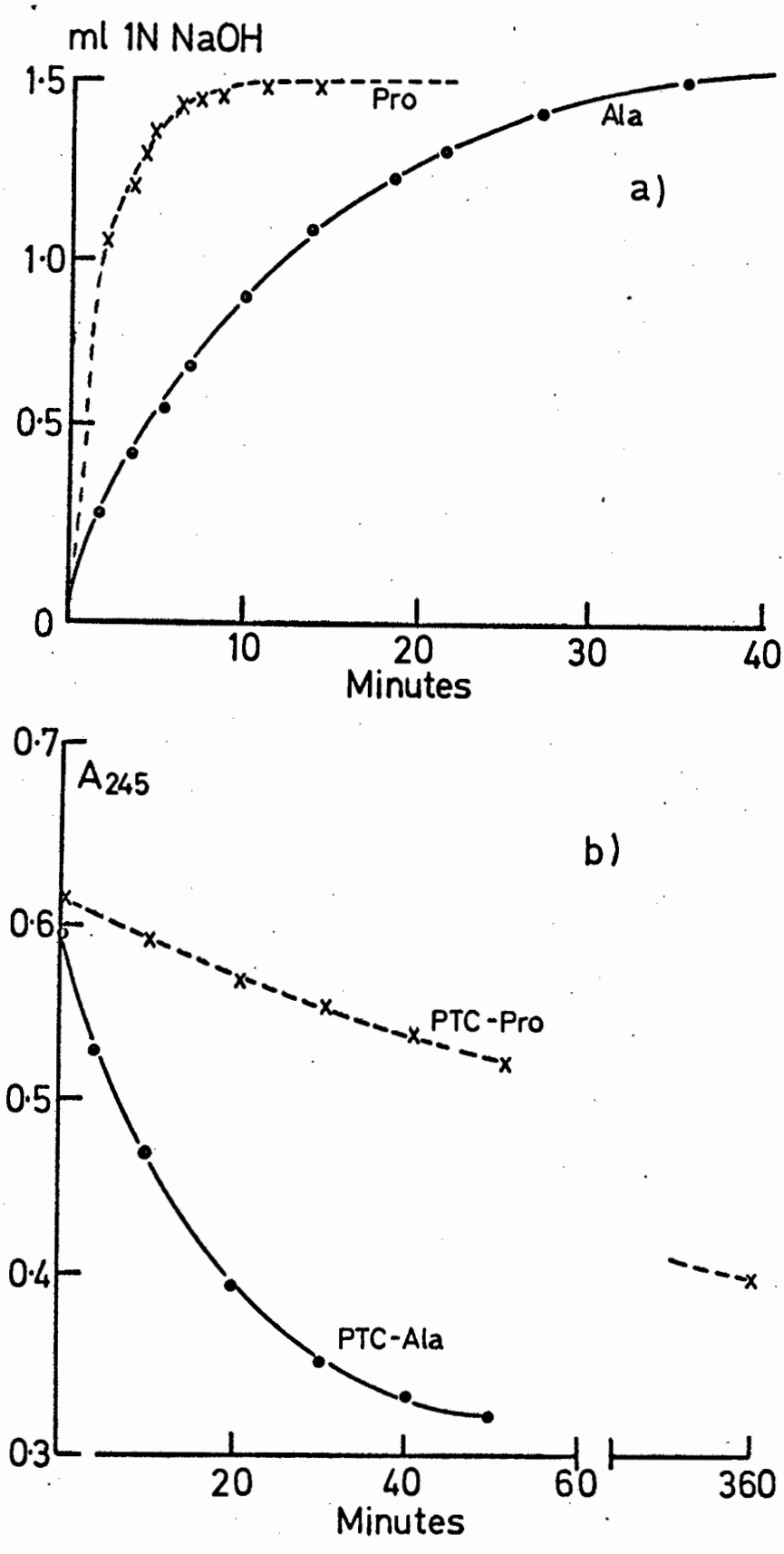


Fig. 2.20 : a) Coupling of phenylisothiocyanate to proline (1.49 nmoles) and alanine (1.69 nmoles) at 40°C in pyridine-H_2O (1:1, v/v) buffer pH 8.6. The rate was measured by the NaOH consumption (Edman, 1970).

b) Cyclization of PTC-alanine and PTC-proline to the PTH-derivative at room temperature in 0.3 N HCl in 96% ethanol. The rate was measured by the decrease in the 245 nm absorption (Ilse & Edman, 1963).

It appears therefore probable that the proline peptide bond to certain amino acids containing large side chains is less readily cleaved in the anhydrous acid cleavage reaction in the automated Edman degradation. From molecular models it is apparent that for steric reasons the phenylthiocarbamyl derivative might less readily cyclize to form the amino acid anilino-thiazolinone. The improvement in yields in subsequent fractions by the multiple application of the anhydrous acid at step 16 is apparent from Table 2.9 and Table 2.10.

TABLE 2.9

THE SEQUENTIAL DEGRADATION OF HISTONE F3

Results are the same as in Table 2.8 and in addition all quantities of Ala, Leu, and Pro from degradation cycle 15-22 have been included.

Step no.	Residue assigned	Yields in nmoles		
		PTH-Ala	PTH-Leu	PTH-Pro
15	Ala	106	10	4
16	Pro	8	7	35
17	Arg	8	8	22
18	Lys	6	10	15
19	Gln	6	10	12
20	Leu	8	54	5
21	Ala	44	32	5
22	Thr	21	20	5

Basic amino acids are not very well quantitated by gas chromatography, spot tests or thin layer chromatography. Fractions thought to contain these PTH-amino acids were therefore hydrolyzed in HCl and then identified and quantitated by amino acid analysis (4.5.3.2, Table 2.11).

TABLE 2.10

THE SEQUENTIAL DEGRADATION OF HISTONE F3

Conditions are the same as in Tables 2.8 and 2.9 except that a second double cleavage with heptafluorobutyric acid had been applied at cycle 16 (b)

Step no.	Residue assigned	Yields in nmoles		
		PTH-Ala	PTH-Leu	PTH-Pro
15	Ala	104	13	9
16a		21	10	42
16b	Pro	16	8	20
17	Arg	14	5	10
18	Lys	9	8	6
19	Gln	10	6	6
20	Leu	13	75	5
21	Ala	78	23	6
22	Thr	15	10	6

From Table 2.11 it is apparent that lysine 9 is partially and lysine 27 nearly completely ε-N-methylated. The occurrence of these derivatives in the N-terminal region was expected from earlier results (Table 2.2). In the degradation leading to the results in Table 2.11 two double cleavages were applied in cycle 16 resulting in larger yields after step 16 compared to yields obtained earlier (Table 2.8).

From these preliminary results the importance of establishing the relative amounts of every amino acid in all fractions became apparent. Since considerable difficulties are experienced in quantitating the more polar PTH-amino acids by gas chromatography (4.5.3.1.3), it was necessary to supplement the information obtained from gas chromatography with results of the re-hydrolysis of the PTH-derivatives followed by amino acid analysis (4.5.3.2). In subsequent experiments the following procedure was adopted for the

TABLE 2.11

SEQUENTIAL DEGRADATION OF HISTONE F3

Yields of basic amino acids determined after re-hydrolysis of the PTH-derivatives and amino acid analysis (4.5.3.2). The sequential degradation was done with two double cleavages at cycle 16

Step no.	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
		Yield in nmoles		Yield in nmoles		Yield in nmoles
2	Arg	340	-		-	
4	Lys	247	-		Glu	2
	ε-N-Me-Lys	Trace				
5	Glu	217	(NH_2	300)	Thr	Trace
8	Arg	148	Glu	30	Lys	8
					ε-N-Me-Lys	-
9	Lys	120	Ala	12	Ser	-
	ε-N-Me-Lys	19				
14	Lys	105	Leu	10	Ala	11
	ε-N-Me-Lys	Trace				
17	Arg	122	Ala	34	Lys	24
					ε-N-Me-Lys	Trace
18	Lys	100	Arg	14	Glu	15
	ε-N-Me-Lys	Trace				
23	Lys	92	Arg	29	Ala	30
	ε-N-Me-Lys	Trace				
26	Arg	54	Ala	20	Lys	14
					ε-N-Me-Lys	Trace
27	Lys	19				
	ε-N-Me-Lys	37	Ala	20		
28	-		Lys	9	Ala	20
			ε-N-Me-Lys	4		

identification of the amino acid derivatives : all fractions were subjected to gas chromatographic analysis. If an unambiguous assignment could not be made because the yield in a particular cycle was lower than expected, the PTH-extract and the corresponding aqueous phase of this and the previous cycle were subjected to amino acid analysis (4.5.3.2). Because, during sequencing, the yields progressively decrease and gas chromatographic identification becomes increasingly difficult against the increasing background, due to Quadrol and other non-amino acid contaminants, all fractions obtained during the later stages of a particular sequential degradation were analysed by both gas chromatography and amino acid analysis. Only a large increase in the amount of an amino acid with no concomitant increase in any of the other amino acids was taken as a positive identification. Serine gives very low PTH yields and the hydrolysis of the latter leads to complete destruction. Therefore, in most cases, it was placed by virtue of the complete absence of any increase of amino acids in that cycle together with amino acid composition data of the respective fragment.

2.7.3 Modified sequential degradation procedure

The careful quantitative evaluation of the results of each cycle could, however, not overcome the progressively decreasing yields which did not allow degradation beyond cycle 28. Even up to that step assignments could only be made with reservation. It was observed that different batches of Quadrol not only yielded varying gas chromatic background but also different yields of amino acid PTH derivatives. Even after lengthy purification steps Quadrol can still contain a substantial amount of aldehydes (Edman & Begg, 1967) which can block terminal-NH_2. Furthermore, it was found (as previously mentioned, page 80) that the Quadrol buffer was not completely removed in the washing steps in the Edman degradation cycle. As a result Quadrol is present in the anhydrous acid cleavage reaction and finally contaminates the amino acid thiazolinone fractions leading to increasing difficulties in

the identification.

Recently Braunitzer and Schrank (1970) described the purification of 3-dimethylamino-1-propyne (DMAP) for the preparation of a volatile buffer to be used for degradation of peptides. This DMAP was accordingly purified as well as the other compounds necessary for the preparation of the buffer. Particular attention was paid to the complete removal of aldehydes (Edman & Begg, 1967) (4.5.2.3).

The DMAP-buffer was substituted for the Quadrol-buffer in the protein degradation procedure. The programme needed slight modifications to compensate for the higher volatility of the DMAP-buffer (4.5.2.2., Table 4.2).

Histone F3 dimer was again subjected to sequential degradation (Tables 2.12, 2.13). The modifications resulted in a considerable improvement in the repetitive yields of the amino acids. Furthermore, the gas chromatograms showed very little unspecific background due to impurities (Fig. 2.21). As a result the first 48 amino acids could unequivocally be identified (Tables 2.12, 2.13) as compared with only 27 residues in the Quadrol procedure (Tables 2.8, 2.11). If necessary the degradation could have been carried further as the high yield in cycle 48 indicates (Fig. 2.21).

Residue 28 was inferred to be Ser because no other amino acid could be identified. Fragment CN-1 NB-1(2) (Table 2.5) contains two serine residues and only one had been identified positively (Step 10).

The following N-terminal amino acid sequence of histone F3 was deduced from the results obtained (Tables 2.8 - 13). The sequence accommodates all the amino acid present in the fragment CN-1 NB-1(2) (Table 2.5)

```
                        5                  Me  10
H-Ala-Arg-Thr-Lys-Gln-Thr-Ala-Arg-Lys-Ser-Thr-Gly-Gly-Lys
15                      20                      25      Me
Ala-Pro-Arg-Lys-Gln-Leu-Ala-Thr-Lys-Ala-Ala-Arg-Lys-Ser-Ala-
30                      35                      40
Pro-Ala-Thr-Gly-Gly-Val-Lys-Lys-Pro-His-Arg-Tyr-Arg-Pro-Gly-
45
Thr-Val-Ala-Leu-OH.
```

TABLE 2.12

SEQUENTIAL DEGRADATION OF HISTONE F3

Yields of PTH-amino acids recovered after sequential degradation of 8 mg of histone F3 dimer with 3-Dimethylamino-1-propyne of the coupling buffer (4.5.2.2), and two double cleavages at step 16.

	Yields nmoles									
Step no.	Ala	Ser	Gly	Thr	Val	Pro	Leu	Tyr	Most likely residue	Step no.
1	349	-	-	-	-	-	-	-	Ala	1
6	12	-	-	141,78	-	-	10	-	Thr	6
7	286	-	-	6	4	-	10	-	Ala	7
10	20	30	-	6	6	-	14	-	Ser	10
11	16	-	-	127,58	5	-	14	-	Thr	11
12	16	-	174	8	7	-	16	-	Gly	12
13	16	-	198	-	7	6	15	-	Gly	13
14	16	-	13	-	5	6	16	-	X	14
15	248	-	9	-	5	7	18	-	Ala	15
16	10	-	8	-	6	70	16	-	Pro	16
19	20	-	9	-	8	10	26	-	X	19
20	28	-	8	-	10	7	210	-	Leu	20
21	149	-	9	-	8	8	43	-	Ala	21
22	33	-	9	25,52	8	7	28	-	Thr	22
23	20	-	6	6	10	10	20	-	X	23
24	120	-	3	-	12	10	27	-	Ala	24
25	161	-	4	-	10	9	37	-	Ala	25
26	40	-	4	-	9	8	30	-	X	26
27	30	-	8	-	14	9	32	-	X	27
28	28	-	8	-	10	9	33	-	X	28
29	81	-	10	-	10	10	32	-	Ala	29
30	32	-	8	-	8	79	32	-	Pro	30

TABLE 2.12 /cont'd.....

Step no.	Yield nmoles								Most likely residue	Step no.
	Ala	Ser	Gly	Thr	Val	Pro	Leu	Tyr		
31	106	-	6	-	14	22	33	-	Ala	31
32	33	-	10	27,35	14	18	33	-	Thr	32
33	34	-	77	-	18	14	47	-	Gly	33
34	33	-	72	-	10	12	40	-	Gly	34
35	20	-	20	-	98	10	35	-	Val	35
36	20	-	10	-	28	10	37	-	X	36
37	18	-	8	-	14	12	34	-	X	37
38	18	-	6	-	14	40	35	-	Pro	38
39	17	-	8	-	11	20	35	-	X	39
40	14	-	8	-	8	14	36	8	X	40
41	19	-	6	-	12	12	38	30	Tyr	41
42	15	-	6	-	9	9	32	10	X	42
43	17	-	6	-	10	35	34	8	Pro	43
44	15	-	19	-	10	24	31	7	Gly	44
45	14	-	12	12,26	10	20	29	5	Thr	45
46	13	-	7	-	35	19	28	8	Val	46
47	36	-	6	-	20	16	24	6	Ala	47
48	29	-	8	-	21	12	53	4	Leu	48

All PTH-and TMS-PTH-amino acids were identified and quantitated by gas chromatography (4.5.3.1.3).

Thr partly decomposes to ΔThr(dehydro) giving rise to two characteristic peaks in the gas chromatograms (4.5.3.1.3). Leu and Ile elute together but can be distinguished after silylation (4.5.3.1.3).

TABLE 2.13

SEQUENTIAL DEGRADATION OF HISTONE F3

Yields of amino acids recovered in steps 23 - 42 determined after re-hydrolysis of the PTH-derivatives by amino acid analysis (4.5.3.2). Results were obtained in parallel with results in Table 2.12.

Step no.	Yield in nmoles															Likely residue
	Lys	Lys(Me)	His	NH_3	Arg	Asp	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr & Phe	
23	120	T	11	60	32	9	22	22	14	44	8	3	8	27	10	Lys
24	21	T	8	70	30	11	21	14	10	130	8	2	6	30	12	Ala
25	17	T	7	60	32	11	22	13	10	128	9	2	6	30	12	Ala
26	18	8	12	90	98	8	21	11	19	44	8	3	6	16	12	Arg
27	25	107	11	100	58	9	18	11	20	30	10	3	7	20	11	Lys (Me)
28	20	50	10	100	49	15	28	19	36	44	15	3	9	32	13	Ser
29	21	11	10	50	45	12	24	15	24	102	13	2	12	30	14	Ala
30	19	T	12	95	43	13	24	106	25	47	11	3	6	30	15	Pro
31	26	-	12	90	47	17	33	50	25	137	18	4	7	40	22	Ala
32	22	-	8	90	39	15	28	29	35	52	11	3	6	33	20	Thr
33	27	-	14	100	50	16	34	22	128	42	14	4	11	39	20	Gly
34	25	-	8	145	41	16	33	15	124	51	19	4	12	39	19	Gly
35	15	-	5	75	27	12	27	11	49	38	72	3	9	30	14	Val
36	67	-	17	80	40	16	29	7	36	36	27	3	10	35	19	Lys
37	61	-	10	73	40	12	24	7	26	29	13	3	9	27	15	Lys
38	20	-	9	54	22	10	17	40	22	20	10	2	6	18	11	Pro
39	12	-	66	50	29	8	16	19	16	17	7	2	6	17	9	His
40	18	-	50	75	78	11	21	20	24	26	12	3	9	29	16	Arg
41	22	-	40	73	60	13	28	19	24	31	13	2	11	28	38	Tyr
42	21	-	30	75	80	10	24	17	24	25	10	4	12	25	30	Arg

All fractions contained 0 - 5 nmoles of Ser and Thr.

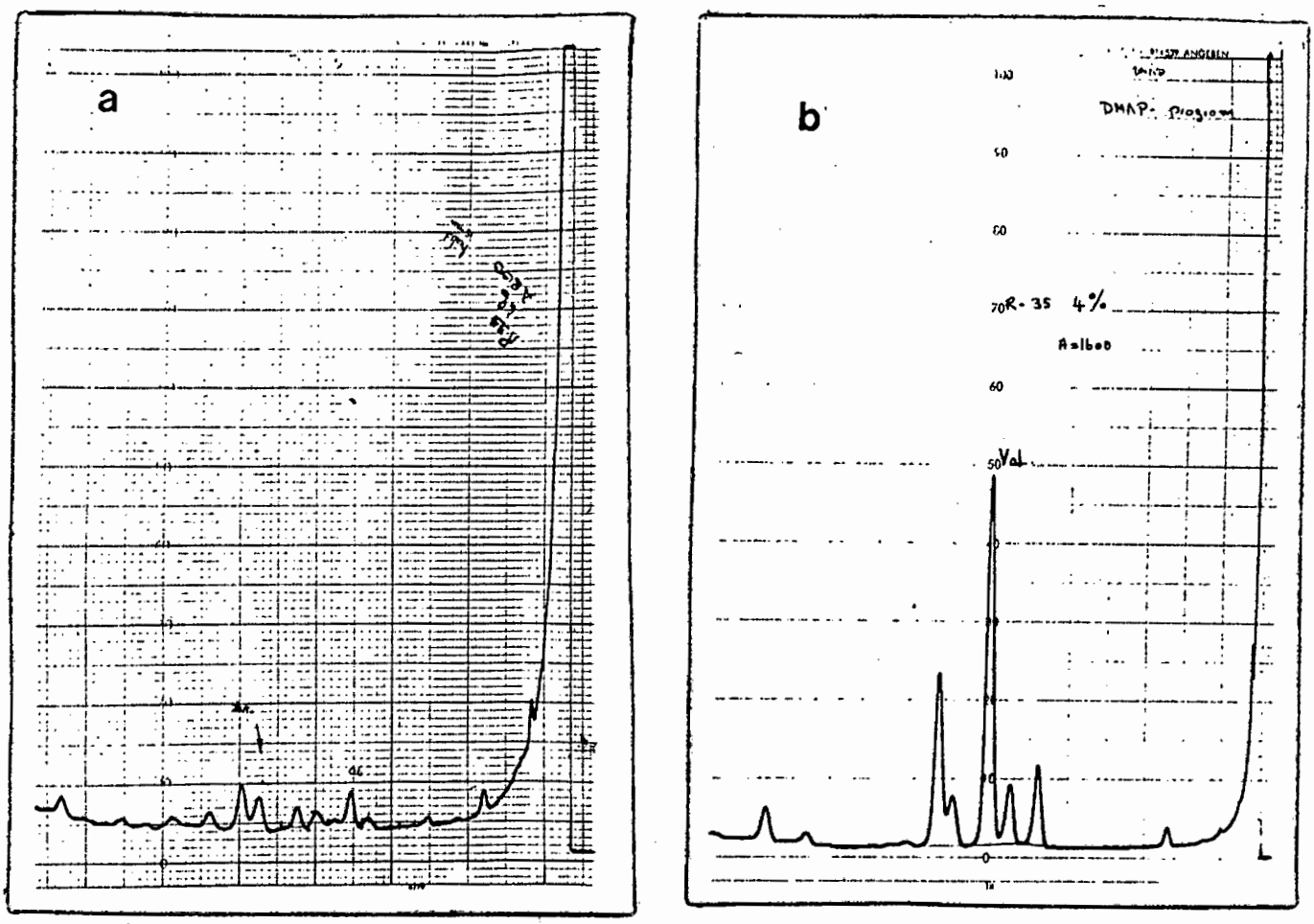


Fig. 2.21 : Gas chromatograms obtained during the degradation of histone F3.

a) 35th degradation cycle of histone F3 (7 mg) using Quadrol buffer. 6% of the sample was injected at an attenuation of 800.

b) 35th degradation cycle of histone (8 mg) using the dimethylamino propyne buffer. 4% of the sample was injected at an attenuation of 1600.

c) Same as (b) at degradation cycle 48 using 6% of the sample at an attenuation of 800. (All chromatograms were obtained on a Beckman GC-45 (4.5.3.1)).

The assignment of 48 residues in the uncleaved protein extends well into the next fragment CN-1 NB-2. It was therefore decided to continue the degradation with that peptide.

2.8 SEQUENCE ANALYSIS OF FRAGMENTS

2.8.1 Methodology

Though the automatic degradation procedure devised by Edman and Begg (1967) works very well for proteins, problems appear when attempts are made to degrade peptides. Because peptides tend to be more soluble in organic solvents, prolonged solvent extractions, necessary to remove the non-volatile Quadrol buffer, cause large losses of peptide material from the sequencer cup (4.5.2.1). Shorter and more hydrophobic peptides are almost quantitatively lost during the ethyl acetate extraction (Braunitzer et al., 1970; Niall et al., 1969). Improvements have been achieved by combining the use of volatile reagents with changes in instrument design, thereby minimizing peptide losses (Niall et al., 1969). The dimethylallylamine-buffer which replaces the Quadrol-buffer in the Edman degradation of peptides can be removed by evaporation and makes prolonged solvent extractions unnecessary (4.5.2.2).

Further advances in the sequence analysis of peptides have been made by the introduction of reagents which increase the hydrophilicity of peptides and therefore reduce losses during the extraction steps. Braunitzer et al. (1970, 1971) reported that lysine-containing peptides reacted with 4-sulfophenylisothiocyanate could be degraded with good yields because the presence of the sulfonate group decreases the solubility in organic solvents used during the extraction. In view of the extensive use of tryptic digestion which produces C-terminal lysine peptides (Inman et al., 1972) this modification may become very useful.

Although this procedure is not restricted to peptides containing a C-terminal lysine it partly loses its usefulness in the complete structure elucidation of a peptide which does not posses a C-terminal lysine, due to the fact that the residual non-lysine containing peptide will again be soluble in the organic solvent.

For the Edman degradation of peptides DMAA-buffer (dimethylallylamine-pyridine-trifluoroacetic-H_2O) is generally used (Niall et al., 1969). I always found, however, that this buffer contained yellowish contaminants and contributed substantially to the gas chromatographic background, complicating the identification of certain PTH-amino acids (Fig. 2.22).

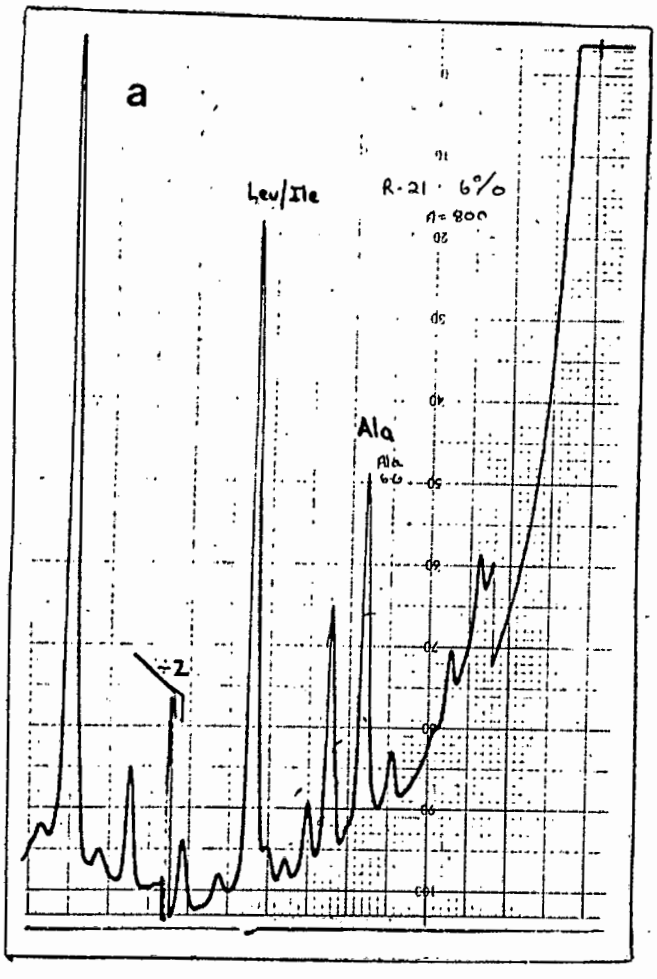
Therefore, instead of the DMAA-buffer, highly purified DMAP-buffer (3-dimethylamino-1-propyne in propanol-trifluoroacetic acid-H_2O (4.5.2.3)) was used in the Edman degradation of peptides (Braunitzer & Schrank, 1971). This DMAP-buffer, like the DMAA-buffer, can largely be removed by evaporation making prolonged solvent extractions unnecessary. In order to reduce the number of extractions a single anhydrous acid cleavage, instead of a double cleavage, was used (Niall et al., 1969). Furthermore, all the organic solvent extractions in the degradation cycle were preceded by a precipitation step (4.5.2.2). The complete degradation cycle for peptides used throughout this study is listed in Table 4.3.

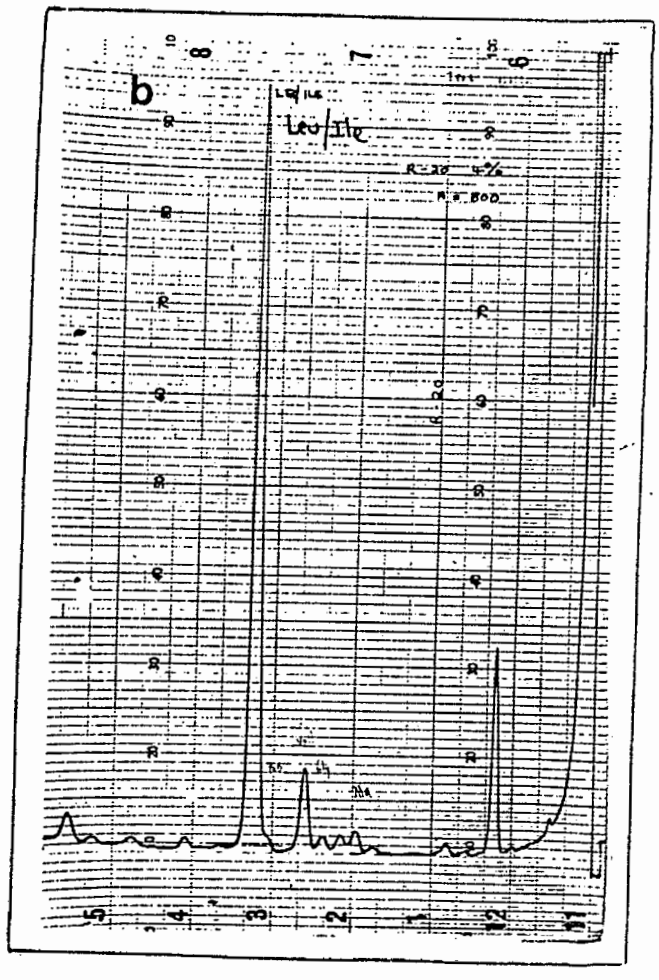
The gas chromatographic background in this methodology was drastically reduced and the yields of the PTH-derivatives substantially increased (Fig. 2.22). Generally about 1000 nmoles of the freeze-dried fragment were dissolved in H_2O and applied to the sequencer cup in the same way as the protein (4.5.2.2.2). The weight of fragment subjected to the sequential degradation was not corrected for salt and moisture content.

The quantitation of thiazolinone fractions and the assignment of residues to the sequence was done as outlined earlier (2.7.2).

2.8.2 Fragment CN-1 NB-2

Fragment CN-1 NB-2 was subjected to the Edman degradation and from the results it became apparent that an amino acid once it appeared as a cleavage product at the end of a cycle reappeared in large amounts in subsequent cycles and normal background levels of that particular amino acid





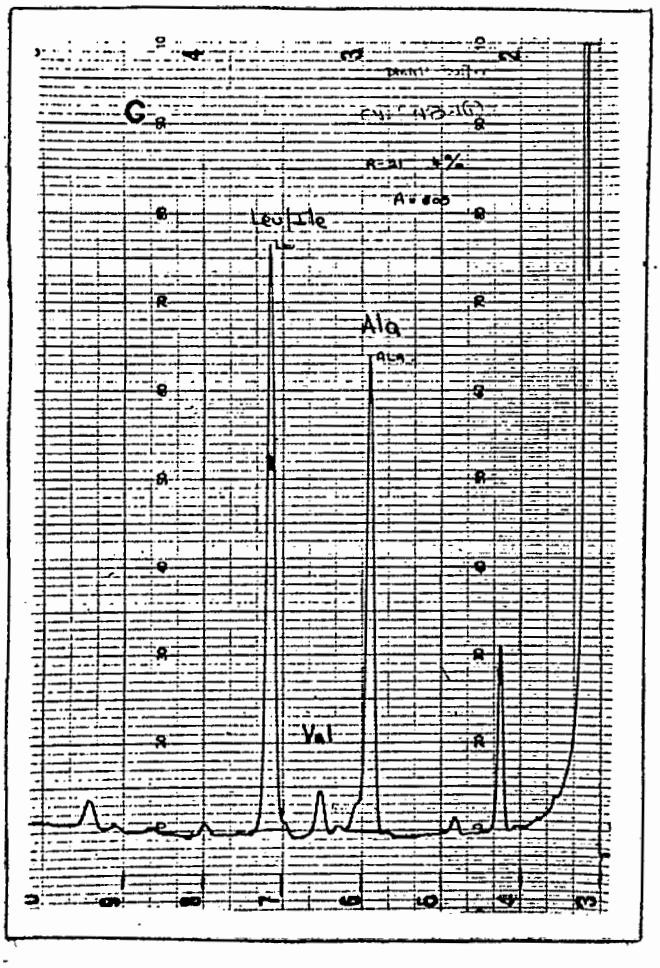


Fig. 2.22 : Gas chromatograms obtained during the degradation of fragment CN-1 NB-1(1) (3.7 mg).

a) 21st degradation cycle (the last that could be assigned to a particular amino acid) using DMAA-buffer. 6% of the sample was injected at an attenuation of 800.

b) 20th degradation cycle on a fresh sample using DMAP-buffer. 4% of the sample was injected at an attenuation of 800.

c) Same as (b) at degradation cycle 21 (all chromatograms were obtained on a Beckman GC-45 (4.5.3.1)).

TABLE 2.14

SEQUENTIAL DEGRADATION OF FRAGMENT CN-1 NB-2

Yields of PTH-amino acids recovered during the sequential degradation of 2.3 mg fragment CN-1 NB-2 (4.5.2.2). PTH-amino acids were quantitated by gas chromatography (4.5.3.1.3) and amino acid analysis after acid hydrolysis (4.5.3.2)

Step no.	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
		Yield in nmoles		Yield in nmoles		Yield in nmoles
1	*Arg	450	Leu	20	Pro	-
2	Pro	440	Val	30	Gly	-
3	Gly	200	Pro	51	Thr	-
4	[0]Thr	110,60	Pro	36	Val	-
5	Val	290	-		Ala	-
6	Ala	230	Val	60	Leu	4
7	Leu	260	Ala	40	Arg	36
8	Arg	130	Leu	50	Glu	6
9	Glu	120	Ala	14	Leu	20
10	Ile	160	Glu	60	Arg	50
11	Arg	120	Leu	50	Arg	120
12	Arg	150	Leu	30		
13			Arg	90		

* The Arg recovered in step 1 was obtained by two anhydrous acid cleavages yielding 225 nmoles Arg each.

Fractions 1 - 3 and 7 - 13 were also subjected to amino acid analysis (4.5.3.2).

[0] Double peak due to Thr and ΔThr (4.5.3.1.3).

were only reached after about 3 degradation cycles.

A similar observation was made during the degradation of histone F3 after Pro-16 (Page 75) and was found to be due to an incomplete cleavage of this residue from the residual protein. Because this phenomena occurred for the first time between step 1 and step 2 the first degradation cycle on a fresh sample was followed by another cleavage (without prior coupling) with the anhydrous acid and another butyl chloride extraction. The yields of Arg-1 present in the two fractions doubled. From Table 2.14 it is apparent that the sequential degradation is no longer out of phase after the initial double cleavage.

Fragment CN-1 NB-2 was subjected to a total of 13 Edman degradation cycles (Table 2.14) from which the first 12 residues could readily be assigned to the sequence. In cycle 13 no positive identification, neither gas chromatographically nor by amino acid analysis, could be made. However, from the amino acid composition of this fragment (Table 2.5) it is evident that residue 13 must be the C-terminal spirolactone, i.e. tyrosine. The complete sequence of fragment CN-1 NB-2 could thus be deduced :

$$\overset{5}{\text{H-Arg-Pro-Gly-Thr-Val-Ala-Leu-Arg-Glu-}}\overset{10}{\text{Ile-Arg-Arg-Tyr-OH}}$$

2.8.3 Fragment CN-1 NB-1(1)

Fragment CN-1 NB-1(1) was subjected to 37 cycles of the Edman degradation. The results obtained by gas chromatographic identification and quantitation are given in Table 2.15. It is apparent that several residues in the sequence remain unassigned. Therefore the unused remainder of the organic solvent layer (containing PTH-derivative of the amino acid) of every fraction together with its aqueous layer, containing unextracted PTH-derivatives (4.5.2.4) were subjected to acid hydrolysis and amino acid analysis (Table 2.16, 4.5.3.2). All amino acids except Ser and Thr can be readily assigned to the sequence by considering the quantitative variation of all

TABLE 2.15

SEQUENTIAL DEGRADATION OF FRAGMENT CN-1 NB-1(1)

Yields of some PTH-amino acids recovered during the sequential degradation of 3.7 mg of fragment CN-1 NB-1(1) (4.5.2.2). PTH-amino acids were quantitated by gas chromatography. (4.5.3.1.3).

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no.		Yield in nmoles		Yield in nmoles		Yield in nmoles
1	Glu Gln	300 150	All	< 4	Lys	-
2	Lys	150	Glu	64	Ser	-
3	Ser	20	All	< 2	Thr	-
4	*Thr	40,150	All	< 2	Glu	?
5	Glu	354	Thr	8,16	xLeu	13
6	Leu	310	Glu	90	xLeu	310
7	Leu	280	Glu	17	xLeu	280
8	Ile	263	All	2	-	
9	X		xLeu	50	-	
10	X		All	<10	xLeu	10
11	Leu	214	All	< 5	Pro	-
12	Pro	150	xLeu	60	Phe	5
13	Phe	130	Pro	28	Glu	?
14	Glu Gln	108 20	Phe	35	-	
15	X		Phe	20	xLeu	6
16	Ile	105	Phe	20	Val	-
17	Val	102	xLeu	44	-	
18	X		Val	46	Glu	13
19	Glu	41	Val	20	xLeu	12
20	Ile	71	Glu	30	Ala	2

TABLE 2.15 /cont'd.....

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no.		Yield in nmoles		Yield in nmoles		Yield in nmoles
21	Ala	55	xLeu	55	Glu	34
22	Glu	28	Ala	35	Asp	-
23	Asp	15	Glu	34	Phe	4
24	Phe	40	Asp	20	-	
25	X		Phe	30	Thr	-
26	*Thr	7,10	Phe	21	Asp	-
27	Asp	16	Thr	10,12	xLeu	5
28	Leu	30	Asp	31	-	
29	X		xLeu	30	Phe	5
30	Phe	29	xLeu	23	-	
31	X		Phe	25	All	< 4
32	X		Phe	15	All	< 4
33	X		Phe	11	Ala	4
34	Ala	9	Phe	6	Val	2
35	Val	5	Ala	3	All	< 4
36	X		All	< 6		
37	X		All	< 6		

xLeu denotes the sum of Ile and Leu (4.5.3.1.3).

* = Double peak due to Thr and ΔThr (4.5.3.1.3).

TABLE 2.16

SEQUENTIAL DEGRADATION OF FRAGMENT CN-1 NB-1(1)

Yields of amino acids recovered during the sequential degradation of 3.7 mg of fragment CN-1 NB-1(1) (4.5.2.2). Amino acids were quantitated by amino acid analysis after re-hydrolysis of the PTH-derivatives (4.5.3.2).

Step no.	Yields in nmoles												Most likely residue	Step no.
	Phe	Lys	NH_3	Arg	Asp	Glu	Pro	Ala	Val	Met	Ile	Leu		
1	2	10	290	60	3	589	10	20	5	-	6	13	Gln	1
2	8	525	280	58	6	89	20	20	5	-	6	13	Lys	2
3	7	78	134	17	8	32	<10	15	5	-	2	11	X	3
4	4	28	220	14	16	20	<10	12	14	-	2	27	X	4
5	2	7	117	14	5	400	<10	11	14	-	2	27	Glu	5
6	2	7	115	17	3	82	<10	3	3	-	2	349	Leu	6
7	2	8	139	19	2	34	<10	12	4	-	3	390	Leu	7
8	6	15	145	37	5	60	<10	5	6	-	289	82	Ile	8
9	3	20	240	216	9	56	<10	13	9	-	69	50	Arg	9
10	8	227	100	72	6	24	<10	5	2	-	24	17	Lys	10
11	7	115	140	82	7	24	36	6	2	-	11	360	Leu	11
12	8	49	144	67	6	32	270	7	2	-	10	127	Pro	12
13	226	7	116	21	4	11	54	6	2	-	3	26	Phe	13
14	94	8	156	23	5	92	34	6	2	-	2	18	Gln	14
15	45	12	138	98	9	72	12	4	11	-	4	24	Arg	15

TABLE 2.16 /cont'd.....

Step no.	Yields in nmoles												Most likely residue	Step no.
	Phe	Lys	NH_3	Arg	Asp	Glu	Pro	Ala	Val	Met	Ile	Leu		
16	24	14	120	69	9	45	<10	6	7	-	2	209	Leu	16
17	9	4	110	49	4	32	<10	3	140	-	2	84	Val	17
18	8	4	100	93	7	28	<10	5	73	-	2	50	Arg	18
19	8	6	85	53	2	56	<10	5	37	-	3	26	Glu	19
20	6	8	115	44	6	75	<10	10	19	-	99	16	Ile	20
21	6	9	130	54	8	66	<10	90	10	-	60	10	Ala	21
22	10	10	176	32	10	92	<10	65	9	-	41	13	Gln	22
23	7	5	130	24	25	69	<10	52	3	-	26	9	Asx	23
24	76	5	130	15	30	51	<10	20	5	-	9	8	Phe	24
25	46	35	216	18	26	40	<10	7	3	-	<2	6	Lys	25
26	46	37	200	14	30	43	<10	5	2	-	<2	7	X	26
27	28	24	220	15	51	36	<10	9	4	-	<2	14	Asx	27
28	30	13	170	23	44	29	<10	5	4	-	<2	75	Leu	28
29	27	17	170	40	45	27	<10	7	4	-	<2	72	Arg	29
30	65	17	170	35	43	25	<10	7	3	-	<2	43	Phe	30
31	59	10	210	27	32	36	<10	7	3	-	<2	35	Glx	31
32	40	8	200	20	26	38	<10	7	2	-	<2	18	X	32
33	30	6	178	20	24	36	<10	8	5	-	<2	15	X	33
34	16	6	168	14	14	30	<10	30	8	-	<2	11	Ala	34
35	6	3	147	12	14	26	<10	29	27	-	<2	10	Val	35
36	6	4	190	11	14	27	<10	31	28	-	3	13	X	36
37	8	3	85	10	13	28	<10	26	25	-	3	10	X	37

All fractions contained between 0-5 nmoles Ser and Thr

amino acids obtained from consecutive Edman degradation cycles.

The results obtained by the two identification techniques compare very well except that yields are consistently lower in the gas chromatographic quantitation compared to those obtained by amino acid analysis (Tables 2.15, 2.16). This difference can probably be ascribed to incomplete extraction of PTH-amino acids from the 1 N HCl used in the conversion of the amino acid thiazolinones to the PTH-derivatives (4.5.2.4), gas chromatographic quantitation by peak height instead of peak area and the assumption that the detector response is linear for all concentrations of PTH-amino acids (4.5.3.1.3).

Furthermore, Table 2.16 reveals a number of phenomena that were generally observed in the degradation of peptides and are thought to be caused by the cumulative effect of various side reactions (Edman & Begg, 1967). A general background of all amino acids appear in the chromatograms which generally increases with the number of degradations performed on the peptide. This background also appears in the degradation of proteins (Table 2.13). The relative amounts of amino acids are generally related to the amino acid composition of the peptide under investigation. This seems to indicate that this background is caused by a small amount of non-specific cleavage of the peptide chain. Because the increase can also be observed in the gas chromatography results and not only in those of the amino acid analysis a partial solubility of the residual peptide in butyl chloride used for the extraction of the thiazolinone is not the sole reason for the appearance of this background although it might account for the generally higher yields obtained by amino acid analysis as the degradation approaches the C-terminal end of the peptide. (Residual peptide in the thiazolinone fraction will be hydrolysed to free amino acids and therefore contribute to the recoveries of amino acids). In absolute terms yields of all amino acids during the sequence determination gradually decreases and a gradual increase in overlap of amino acid residues in consecutive cycles becomes evident (Table 2.16). Edman and Begg, (1967) suggested that the cleavage reaction is not quantitative under the conditions due to an unfavourable equilibrium.

This view is supported by the observation that the degradation of peptides gets more readily out of phase compared to proteins. In the former a single cleavage is used, whereas in the latter a double cleavage is applied (Tables 4.2, 4.3). Furthermore, it became evident that not all peptide bonds cleave with equal ease, e.g. peptide bonds involving proline are sometimes less readily cleaved than others (2.7.2, 2.8.3).

From the results in Table 2.15 and Table 2.16 it was inferred that position 3, 32 and 33 are occupied by serine since no other amino acid could be identified by gas chromatography, nor did an increase in any amino acid above the previous cycle occur after hydrolysis. This assignment is supported by the amino acid composition of fragment CN-1 NB-1(1) (Table 2.5) and its fragments (Table 2.6) obtained by a cleavage at aspartic acid residues (2.4, page 58). Out of the three serines present one is situated in the first 22 residues and the other two in the remainder of the fragment.

Because fragment CN-1 NB-1(1) contains 36 residues, position 36 must be occupied by the C-terminal homoserine lactone arising out of methionine (Table 2.5). The amino acid composition of fragment CN-1 NB-1(1) A-1 and CN-1 NB-1(1) A-2 support results obtained by the sequential analysis.

This, therefore, is the complete amino acid sequence of fragment CN-1 NB-1(1) :

```
                    5                   10                  15
H-Gln-Lys-Ser-Thr-Glu-Leu-Leu-Ile-Arg-Lys-Leu-Pro-Phe-Gln-Arg

                   20                  25                  30
  Leu-Val-Arg-Glu-Ile-Ala-Gln-Asx-Phe-Lys-Thr-Asx-Leu-Arg-Phe-

                   35
  Glx-Ser-Ser-Ala-Val-Met-OH
```

2.8.4 Fragments CN-1 NB-1(1) A-3 and CN-1 NB-1(1) A-2

These two fragments were sequenced to confirm the C-terminal amino acid sequence of fragment CN-1 NB-1(1) determined earlier (2.8.3). All of the freeze-dried fractions

obtained from the Sephadex G-25 column (Fig. 2.14) were subjected to the sequential Edman degradation. Fractions obtained from the degradation were directly hydrolyzed for amino acid analysis (4.5.3.2).

Tables 2.17 and 2.18 show that both fractions are cross-contaminated with each other and also contain small amounts of peptides produced by unspecific cleavage of fragment CN-1 NB-1(1) during the cleavage at aspartic residues in the dilute acid (2.4). Nevertheless, from the sequential degradations (Table 2.15, 2.16) and the amino acid composition of the peptides (Table 2.6), the sequence of amino acids in the fragments can be determined :

CN-1 NB-1(1) A-3

Phe-Lys-Thr

CN-1 NB-1(1) A-2

Leu-Arg-Phe-Glx-Ser-Ser-Ala-Val-Met

The structure of fragment CN-1 NB-1(1) A-3 and CN-1 NB-1(1) A-2 deduced agrees with the C-terminal sequence of fragment CN-1 NB-1(1) deduced earlier.

2.8.5 Fragment CN-2

Performic acid oxidized fragment CN-2 was subjected to 30 cycles of the Edman degradation (Table 2.19). The recoveries of PTH-amino acids were not affected by performing the sequential degradation on the performic acid oxidized fragment as compared to a pilot run on the CN-2 dimer. However, in step 20, where previously there had been a gap in the sequence, cysteic acid could be identified (Table 2.17).

From the results in Table 2.19 it is evident that in steps 16 - 18 (Asn-Thr-Asp) a considerable drop in yields occurs. The reason for this is not clear.

TABLE 2.17

SEQUENTIAL DEGRADATION OF FRAGMENT CN-1 NB-1(1) A-3

Yields of amino acids recovered during sequential degradation (4.5.2.2) of fraction 3 (Fig. 2.14). Amino acids were quantitated by amino acid analysis after re-hydrolysis of the complete unconverted fractions (4.5.3.2).

Step no.	Yields in nmoles										Most likely residue
	Phe	Lys	Arg	Asp	Glu	Gly	Ala	Val	Ile	Leu	
1	420	22	17	-	21	26	76	46	46	145	Phe
2	51	350	78	-	26	12	99	42	11	30	Lys
3	129	93	42	-	25	15	66	40	11	11	X

TABLE 2.18

SEQUENTIAL DEGRADATION OF FRAGMENT CN-1 NB-1(1) A-2

Yields of amino acids recovered during sequential degradation (4.5.2.2) of fraction 2 (Fig. 2.14). PTH-amino acids were quantitated by amino acid analysis after re-hydrolysis of the complete unconverted fractions (4.5.3.2).

Step no.	Yields in nmoles										Most likely residue
	Phe	Lys	NH_3	Arg	Glu	Gly	Ala	Val	Ile	Leu	
1	52	28	138	24	19	21	19	19	7	302	Leu
2	38	35	164	228	50	14	40	31	5	93	Arg
3	259	14	136	37	62	12	52	60	5	28	Phe
4	45	16	91	18	178	-	14	24	5	40	Glx
5	14	14	125	ND	127	5	16	16	5	52	X
6	9	13	160	ND	72	5	14	21	28	45	X
7	9	9	44	ND	28	5	62	19	14	19	Ala
8	ND	ND	ND	ND	45	7	57	53	21	16	Val
9	ND	ND	ND	ND	60	14	43	45	21	16	X

TABLE 2.19

SEQUENTIAL DEGRADATION OF FRAGMENT CN-2

Yields of some of the PTH-amino acids recovered during the sequential degradation of 4.4 mg fragment CN-2 (4.5.2.2). PTH-amino acids were quantitated by gas chromatography (4.5.3.1.3). Fractions 17 to 29 were also subjected to amino acid analysis after acid hydrolysis of the PTH-derivatives

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no.		Yield in nmoles		Yield in nmoles		Yield in nmoles
1	Ala	610	-		xLeu	-
2	Leu	640	-		Glu	-
3	Gln Glu	350 200	Leu	30	Glu	200
4	Glu	500	-		Ala	-
5	Ala	330	-		Ser	-
6	Ser	40	-		Glu	?
7	Glu	350	Ala	9	Ala	9
8	Ala	290	-		Tyr	-
9	Tyr	210	Ala	30	xLeu	8
10	Leu	320	Ala	14	Val	2
11	Val	216	xLeu	35	Gly	4
12	Gly	170	Val	60	Leu	35
13	Leu	195	Gly	35	Phe	4
14	Phe	195	xLeu	57	Glu	?
15	Glu	180	Phe	66	Asp	4
16	Asn Asp	40 90	Glu	12	Thr	4

TABLE 2.19 /cont'd.....

	Prominent amino acid		Amino acid present in 2nd largest amount		Background concentration of amino acid becoming prominent in next cycle	
Step no.		Yield in nmoles		Yield in nmoles		Yield in nmoles
17	Thr	75	Asp	72	Asp	72
18	Asp	57	Ala	21	Leu	13
19	Leu	100	Asp	47	Cy-SO_3H	-
20	Cy-SO_3H	30	Leu	57	Ala	8
21	Ala	90	Leu	28	Ile	6
22	Ile	90	Ala	40	His	7
23	His	50	Ile	38	Ala	17
24	Ala	57	His	43	Lys	12
25	Lys	38	Ala	44	Arg	13
26	Arg	41	Lys	22	Val	18
27	Val	43	Arg	29	Thr	7
28	Thr	24	Val	39	Ile	17
29	Ile	30	Val	25	-	
30	X		Ile	20		

From the amino acid composition of fragment CN-2 (Table 2.2), fragment CN-2 NB-1 and fragment CN-2 NB-2 (Table 2.4), it is evident that position 30 is occupied by the C-terminal homoserine lactone. This, together with the results obtained from the sequential degradation (Table 2.17), allowed the complete amino acid sequence of fragment CN-2 to be deduced.

```
                    5                   10
H-Ala-Leu-Gln-Glu-Ala-Ser-Glu-Ala-Tyr-Leu-Val-Gly-Leu-Phe-

 15                  20                  25
 Glu-Asn-Thr-Asx-Leu-Cys-Ala-Ile-His-Ala-Lys-Arg-Val-Thr-

    30
 Ile-Met-OH
```

2.8.6 Fragment CN-3

From the amino acid analysis, which shows the presence of proline, difficulties in the sequential degradation were expected. A pilot sequential degradation on this fragment revealed that all residues once they had appeared in the sequence reappeared in large amounts in subsequent cycles. The amount of a particular residue diminished again to background levels after 2 to 3 degradations. This indicated that the sequential degradations came out of phase as from step one on, which had released some proline.

A new sample of fragment CN-3 was degraded. Neither a second phenylisothiocyanate coupling reaction in the first degradation cycle nor a second extraction increased the yield of proline in cycle one substantially (Table 2.20). However, when two acid cleavage reactions were performed on the residual peptide after the first Edman degradation cycle yields drastically increased. After each cleavage reaction in cycle one the amino acid thiazolinones were extracted and quantitated (Table 2.20). From the results it is apparent that at this stage nearly all of the proline had been recovered.

TABLE 2.20

SEQUENTIAL DEGRADATION OF FRAGMENT CN-3

Yields of PTH-amino acids recovered during the sequential degradation (4.5.2.2) of 2.5 mg of fragment CN-3. All PTH-derivatives were quantitated by gas chromatography (4.5.3.1.3) except Arg and Lys which were determined by amino acid analysis (4.5.3.2)

	Yields in nmoles										
Step no.	Ala	Gly	Pro	Leu	Ile	Asp	Glu	Gln	Arg	Lys	Most likely residue
1a	-	-	260	-	-	-	-	-	-	-	-
1b	-	-	20	-	-	-	-	-	-	-	-
1c	-	-	260	-	-	-	-	-	-	-	-
1d	-	-	100	-	-	-	-	-	-	-	-
Σ1	-	-	640	-	-	-	-	-	-	-	Pro
2	-	-	60	-	-	-	-	-	-	500	Lys
3	-	-	20	-	-	430	-	-	-	60	Asp
4	-	-	T	-	337	180	-	-	?	?	Ile
5	-	-	T	-	280	180	325	190	?	?	Gln
6	-	-	T	300	-	100	230	20	?	?	Leu
7	280	-	T	270	30	40	111	4	25	20	Ala
8	230	-	T	230	18	20	50	-	160	20	Arg
9	60	-	T	30	10	15	20	-	290	20	Arg
10	35	-	T	25	266	7	17	-	218	15	Ile
11	10	-	T	20	205	-	20	-	242	15	Arg
12	9	108	T	15	74	-	36	-	170	10	Gly
13	10	100	4	15	34	-	132	-	121	10	Glu
14	8	82	5	-	24	-	95	-	68	5	X
15	25	84	4	-	18	-	50	-	-	-	Ala

1a Normal peptide Edman degradation cycle (4.5.2.2).

1b The ethyl acetate extraction was repeated.

1c A second cleavage followed by extraction.

1d A third cleavage followed by extraction.

The fragment was then subjected to another 14 cycles of the Edman degradation carried out in the normal way (Table 2.20). The results show that a rapid increase in overlap of amino acid between consecutive degradation cycles occurs. The probable reasons for this have been discussed earlier (2.8.3). The incomplete cleavage is probably more pronounced due to the fact that a larger quantity of peptide than usual had been subjected to the degradation.

Results obtained in the sequential degradation allow all amino acids except the penultimate residue to be assigned to the structure. By comparing the sequence determination with the amino acid composition of this fragment it becomes apparent that one Arg remains unassigned (Table 2.2). Considering, however, the Arg yields during the sequential degradation it is not possible to assign any of the positions between 1 and 13 to the additional Arg. The most likely position would therefore be position 14.

The complete amino acid sequence of fragment CN-3 therefore is :

H-Pro-Lys-Asp-Ile-Gln-Leu-Ala-Arg-Arg-Ile-Arg-Gly-Glu-
Arg-Ala-OH

(5 and 10 mark positions 5 and 10; 15 marks position 15)

2.9 COMPLETE SEQUENCE AND CONCLUSIONS

The complete amino acid sequence of histone F3 from chicken erythrocytes has been established without resorting to the use of overlapping sequences. To achieve this, specific chemical cleavages rather than enzymatic degradations were chosen and applied, first to the original protein chain, and subsequently to the generated polypeptides to yield sets of not more than 3 peptides in any single cleavage. Their relative position in the protein or polypeptides became evident after comparison of the N- and C-terminal amino acids in the cleavage products and the uncleaved starting material (Fig. 2.19).

Existing automatic Edman degradation procedures for proteins and peptides were slightly modified utilizing a volatile N,N-dimethylamine-1-propyne-buffer (Braunitzer & Schrank, 1971). This allowed the determination of the first 48 residues in the uncleaved protein compared with 27 using the non-volatile buffer Quadrol (Edman & Begg, 1967).

The limitations of the Quadrol-buffer in the degradation were found to be supported by results obtained by Olson et al. (1972) who also only succeeded in assigning the first 26 residues of uncleaved calf thymus histone F3 using this non-volatile buffer.

Subsequently the peptides CN-1 NB-2, CN-1 NB-1(1), CN-2 and CN-3 (Fig. 2.23) were subjected to the automatic degradation. The procedure differed from that for proteins in that the length of the washing and extraction steps were reduced to minimize peptide losses due to their higher solubility in the organic solvents used in the automatic Edman degradation (Table 4.3).

In all cases it was possible to degrade and assign all residues to all the above mentioned peptides to at least the penultimate amino acid. Assignment of the penultimate and ultimate amino acid residues were aided by the amino acid composition of the fragment and the identity of the C-terminal residue.

Since the position of all fragments in the protein is known the complete structure is evident (Fig. 2.24). The structure has in some regions been confirmed by degrading fragment CN-1 NB-1(1) A-2 and CN-1 NB-1(1) A-3. Fragment NB-1(1) was also attempted (results not shown) but yielded little information above that obtained by degrading fragment CN-2. A possible reason for this has been discussed (2.8.5).

Through the large number of degradation cycles performed on the protein and peptides the amide derivative of Glu and Asp was partly or completely hydrolysed. It was thus not possible to determine with certainty whether these amino acids occur as the free acid or the amide derivative in the later stages of the degradation. In addition, PTH-glutamine and PTH-asparagine cannot readily be identified by gas chromatography (4.5.3.1.3).

STEPS IDENTIFIED BY AUTOMATIC SEQUENTIAL DEGRADATION OF F_3 HISTONE FRAGMENTS

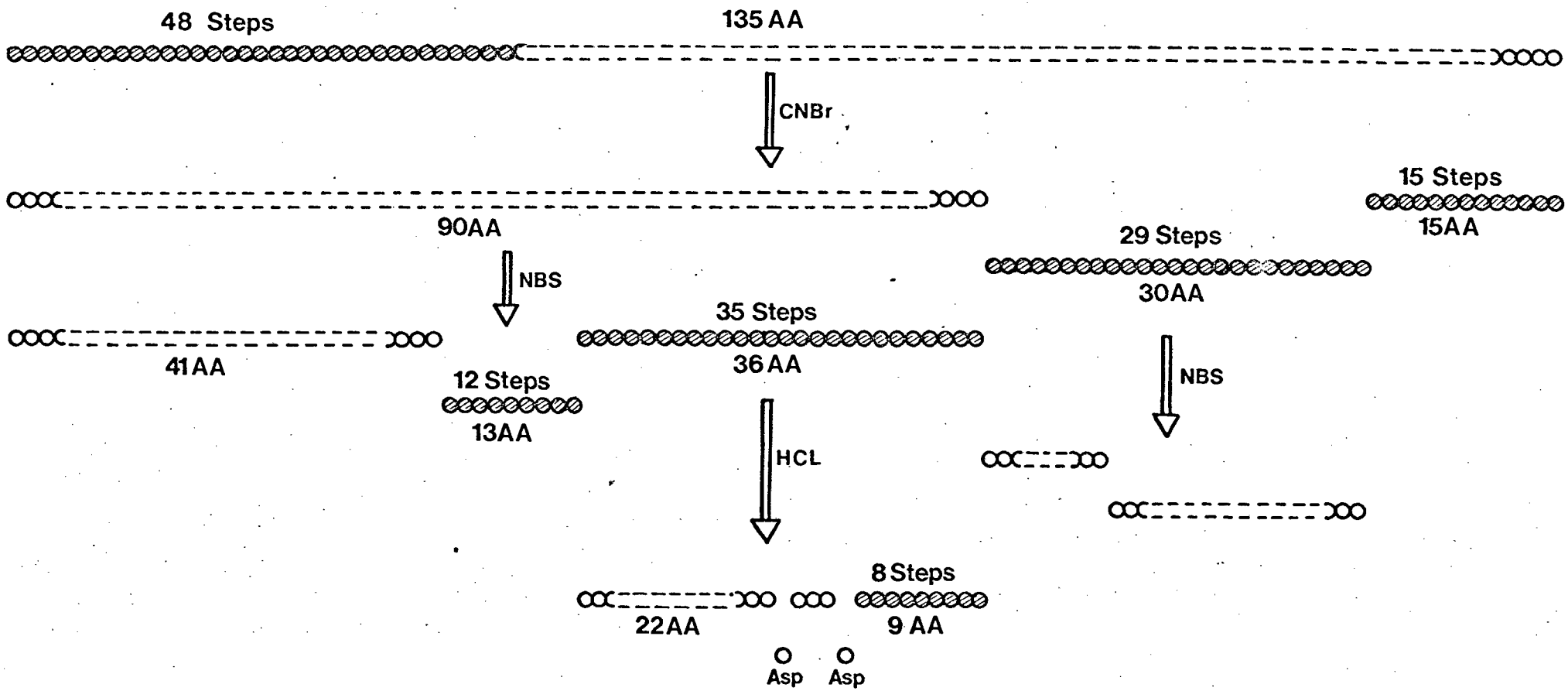


Fig. 2.23 : Diagrammatic representation of the sequential analysis of histone F3. Hatched spheres correspond to steps identified by automatic sequential degradation of histone F3 and fragments (for nomenclature of fragments see Fig. 2.19).

```
                                   (Me1-3)x
                    5                  |                              15
H-Ala-Arg-Thr-Lys-Gln-Thr-Ala-Arg-Lys-Ser-Thr-Gly-Gly-Lys-Ala-

                                                  (Me1-3)y
                   20                  25             |              30
Pro-Arg-Lys-Gln-Leu-Ala-Thr-Lys-Ala-Ala-Arg-Lys-Ser-Ala-Pro-

                   35                  40                            45
Ala-Thr-Gly-Gly-Val-Lys-Lys-Pro-His-Arg-Tyr-Arg-Pro-Gly-Thr-

                   50                  55                            60
Val-Ala-Leu-Arg-Glu-Ile-Arg-Arg-Tyr-Gln-Lys-Ser-Thr-Glu-Leu-

                   65                  70                            75
Leu-Ile-Arg-Lys-Leu-Pro-Phe-Gln-Arg-Leu-Val-Arg-Glu-Ile-Ala-

                   80                  85                            90
Glx-Asx-Phe-Lys-Thr-Asx-Leu-Arg-Phe-Glx-Ser-Ser-Ala-Val-Met-

                   95                 100                           105
Ala-Leu-Gln-Glu-Ala-Ser-Glu-Ala-Tyr-Leu-Val-Gly-Leu-Phe-Glu-

                  110                 115                           120
Asn-Thr-Asn-Leu-Cys-Ala-Ile-His-Ala-Lys-Arg-Val-Thr-Ile-Met-

                  125                 130                           135
Pro-Lys-Asp-Ile-Gln-Leu-Ala-Arg-Arg-Ile-Arg-Gly-Glu-Arg-Ala-


OH
```

Fig. 2.24 : Complete amino acid sequence of histone F3 (x is near 0.2 and y near 1 (Table 2.11)).

Serine residues could not, in most cases, be directly identified due to the low recoveries of PTH-serine and had to be identified by indirect evidence, i.e. whenever a gap appeared in the sequence, and the amino acid composition of the fragment permitted this, a serine assignment was made.

Another observation during the sequence analysis of histone F3 was that proline residues followed or preceded by a basic residue were less susceptible to the cleavage reaction in the Edman degradation. The recoveries could be increased at these positions by repeating this cleavage reaction several times. Whether this is a general phenomena remains to be shown.

Simultaneously the complete amino acid sequences of chicken histone F3 (Brandt & von Holt, 1972) and calf thymus histone F3 (DeLange et al., 1972) have been elucidated. The latter has been achieved by the conventional approach utilizing trypsin, chymotrypsin, pepsin, thermolysin and CNBr for the fragmentation of the protein.

Comparing the two amino acid sequences it is apparent that the expected additional Cys residues (1.2) in calf histone F3 occurs in position 96 replacing a serine residue in the chicken F3 histone.

Otherwise both sequences are identical. The same lysine residues are ε-N-methylated (Lys 9 & Lys 27). In chicken histone F3 the extent of methylation of residue 9 is fairly small while Lysine 27 is nearly completely methylated. Of all the methylated lysines present the ε-N-dimethyl lysine predominates over the mono- and trimethylated derivative.

It was not possible to assign the acetyl groups

to particular lysine residues. It was, however, established that the acetyl-ε-N-lysine occurs in the first 40 residues of the molecule. In view of the similarity between chicken protein and that from calf it is likely that this derivative also occurs at lysine residues 14 and 23 like in calf thymus (DeLange et al., 1972). The extent of acetylation is again very similar (Fig. 1.6).

Initially one additional Glu residue had been assigned to the sequence of fragment CN-1 NB-1(1) (Table 2.16) and thus to the complete sequence (Brandt & von Holt, 1972). This has been traced to a low concentration of Glu in the commercially available amino acid calibration mixture. This illustrates a shortcoming in the method of sequential analysis of non-overlapping fragments in which penultimate and ultimate amino acid residues are assigned relying on the amino acid composition of the particular fragment. A faulty amino acid analysis may result in a wrong assignment of residues towards the C-terminal end of a peptide.

In conclusion, it can however be said that the new approach to the sequence elucidation of proteins offers several advantages over the conventional approach. Due to the limited number of fragments obtained these are readily separated and, together with the automated Edman degradation, allow a relatively non-laborious and rapid amino acid sequence determination.

Furthermore, the limited fragmentation can be applied according to ones needs and the results obtained in earlier stages of the structure elucidation studies. This eliminates a large amount of unnecessary work and

thus offers a more direct route to the complete amino acid sequence determination.

PART 3

EVOLUTION, STRUCTURE AND FUNCTION OF HISTONE F3

3.1 INTRODUCTION

Though most of the mechanisms for gene control in prokaryotes have been elucidated during the last decade it is unlikely that a simple extrapolation of this knowledge will suffice to explain gene regulation in eukaryotes, in particular, in view of the highly complex structure of chromatin. In recent years a considerable amount of research has been performed on histones and chromatin which resulted in a better understanding of the molecular structure of the deoxyribonucleoprotein complex. From these studies it has emerged that eukaryotic chromosomes are composed of DNA, histones and non-histone proteins (acidic & residual proteins) (Tuan & Bonner, 1969). DNA and histones occur in relatively equivalent and constant amounts per cell nucleus (Feughelman et al., 1955). The close association of histones and DNA has led to two suggestions concerning their function.

The first of these is based on the original hypothesis of Stedman and Stedman (1950) that histones are involved in control mechanisms of genetic regulation. The second hypothesis assigns to the histones a structural role in maintaining and controlling the conformations of the chromosomes through the cell cycle (Bradbury & Crane-Robinson, 1971b).

The relatively small number of different histones seem to indicate a more structural role and involvement in the gross repression of inactive genes than for a role in the precise control of active genes.

In eukaryotic cells, molecules of DNA several cm in length are condensed into chromosomes of dimensions of the order of microns. This process is reversible and must involve a series of precise interactions between the histones and the DNA molecules and between the histones themselves in order to control the complex conformational changes which DNA undergoes during chromosomal condensation. Both views do not necessarily exclude each other and histones may play a more direct role in the control of the genetic potential of

the cell. However, models on histone function are, at this stage, mainly speculative and the exact role of histones in the cell nucleus is still not clear.

In view of the increasing understanding of the relationship between function, structure and information transfer at the molecular level, it is instructive to consider the structure of histone F3 and its possible evolution. In this way insight may be gained into the interaction of histones and DNA and more light may be thrown on the role of histones in the cell nucleus.

3.2 EVOLUTION OF F3 HISTONE

When Stedman and Stedman (1950) implicated histones in gene control it was believed that these proteins consisted of a complex mixture. In recent years, however, it has been established that the heterogeneity of histones is limited. In general, five major fractions with very similar properties have been identified both electrophoretically and chemically (Panyim & Chalkley, 1969; Johns, 1967b).

Evidence has accumulated that the same or very similar groups of histones are present not only in vertebrates but also for example in insects and plants. Panyim and Chalkley have concluded from electrophoretic studies that a large number of different organisms have very similar histone fractions and that especially histone F2a1 and F3 have an identical electrophoretic mobility no matter from which source they have been isolated (Panyim et al., 1970, 1971). The striking similarity of histone F2a1 from calf thymus and pea embryo has recently been demonstrated by the elucidation of their primary structure (DeLange et al, 1969a; Ogawa et al., 1969). The two proteins have been shown to be identical except for two conservative replacements. Thus only two changes have taken place in the 1.5 billion years since the divergence of the pea and bovine lines, and this degree of conservation of sequence is unique for two homologous proteins separated so widely on the evolutionary scale (Dickerson, 1972).

A similar situation pertains to histone F3. From electrophoretic studies and peptide mapping Fambrough and Bonner (1968) suggested that histone F3 from pea embryo and calf thymus might also be very similar. Furthermore,

DeLange et al. (1972) have recently elucidated the primary structure of F3 histone from calf thymus. Therefore, histone F3 from calf and from chicken can now be compared. The two proteins are identical, except for a replacement of serine by a cysteine in calf histone.

The ε-N-methylated lysine residues occur in the same positions in similar amounts which possibly also applies to the acetyl groups of calf and chicken histone F3 (Fig. 1.6).

In the course of the evolution of such divergent organisms, in other proteins many amino acid replacements are being observed though those amino acids constituting the active site of an enzyme remain unchanged. Therefore, a strong evolutionary pressure must exist to preserve the archetypal amino acid sequence in histones. The rigid conservation of the sequence of amino acids in the active centre of an enzyme can be understood in terms of reaction specificity. If, in histones, the totality of the amino acid sequence is being preserved through evolution, then this indicates a highly specific function of the proteins involving precise conformation of the entire protein. The identical position of modified lysine residues suggests that the conformational changes through amino acid modification have a particular significance on structure and function of histone F3.

3.2.1 Internal sequence homology in histone F3

It has been suggested that proteins have evolved from shorter and less complex peptides. In the earliest stages of prebiotic evolution the primitive proteins arose out of repetitive amino acid sequences composed of a number of short and similar segments. Thus, a large protein could be built up by repeating a small number of primordial sequences as structural building blocks. In primitive replicating systems the repetitive amino acid sequence may have been coded for by a repetitive DNA sequence, which in turn had been formed by repeated duplication of short oligo nucleotides. The

bacterial ferredoxins (Benson et al., 1967) have been cited as evolved along these lines (Jukes, 1966). In them the second half of the amino acid sequence is almost an exact duplicate of the first and each half has a regular periodic arrangement of cysteine residues. More recently Dus et al. (1968) noticed repetitive pieces in the sequence of cytochrome C_2 and suggested that these features were survivals of a primitive ancestral repeating sequence. Such sequences have also been noted in the protamines (Black & Dixon, 1967).

The repetition of amino acid stretches may be partly or completely obscured by other important mechanisms which produce alterations in the amino acid sequence, namely : replacements, deletions and insertions, the latter two occur by partial duplication and recombination in the cistron coding for that particular protein (Dixon, 1966).

From such an analysis Black and Dixon (1967) established that a fundamental sub-unit exists in all protamines represented by X - Arg - Arg - Arg - Arg where X may be Ala, Gly, Ile, Pro, Ser, Thr or Val. This sub-unit directly accounts for two-thirds of the amino acid sequences in protamines (Table 3.1). Apparently, in the evolution of the protamines occasional deletions during partial duplication of the cistrons which coded for the ancestral sub-unit lead to gaps in the fundamental pentapeptide. In the amino acid sequence of the histone sequence a sub-unit of the structure is far less apparent. Nevertheless, an alignment of groups of amino acid residues has been attempted for known histone sequence by Phillips (1971b) (Table 3.1, column 2) who established that certain sequences are repetitive and occur also in other histones (F1, F2a2). Phillips suggests, for example, in the case of F2a1, that the peptide Gly-Gly-Gly-Arg-Pro might have been a precursor to this histone. A different repetitive pentapeptide is obtained for each histone F3 and F2a1 if it is assumed that protamines and histones had the same precursor. As a basis for an alignment a pentapeptide with a hydrophobic amino acid followed by as many as possible basic residue was selected.

From Table 3.1, column 3 and 4, it is apparent that such a peptide, although not as basic as the "protamine peptide", nevertheless repetitive occurs in both proteins. A peptide in which two basic residues are sandwiched between two hydrophobic amino acid residues occur fairly frequently in both

TABLE 3.1

Homologies within the amino acid sequences of protamines, histone F2a1 and F3. Stretches of the amino acid sequences of clupeine Z, histone F2a1 from calf thymus and F3 from chicken have been aligned to obtain maximum homology.

Res. No.	Clupeine Z+	Res. No.	F2a1* histone
1-	Ala-Arg-Arg-Arg-Arg	1-	Ser-Gly-Arg-Gly-Lys
6-	Ser-Arg-Arg . .	6-	. Gly-Gly-Lys .
9-	Ala	9-	Gly-Leu-Gly-Lys .
10-	Ser-Arg-Pro . .	13-	Gly-Gly-Ala-Lys-Arg
13-	Val-Arg-Arg-Arg-Arg	18-	. . His-Arg-Lys
18-	Pro-Arg-Arg . .	28-	Gly-Ile-Thr-Lys .
21-	Val	32-	Pro-Ala-Ile-Arg-Arg
22-	Ser-Arg-Arg-Arg-Arg	37-	Leu-Ala . Arg-Arg
27-	Ala-Arg-Arg-Arg-Arg	41-	Gly-Gly-Val-Lys-Arg
		76-	. Ala-Lys-Arg-Lys
		89-	. Ala-Leu-Lys-Arg

Res. No.	F2a1 histone	Res. No.	F3 histone
15-	Ala-Lys-Arg . .	1-	Ala-Arg-Thr-Leu-Gln
18-	His-Arg-Lys-Val-Leu	7-	Ala-Arg-Lys-Ser-Thr
34-	Ile-Arg-Arg . Leu	16-	Pro-Arg-Lys . Gln
38-	Ala-Arg-Arg-Gly-Gly	21-	Ala-Thr-Lys-Ala-Ala
42-	Val-Lys-Arg-Ile-Ser	25-	Ala-Arg-Lys-Ser-Ala
54-	Thr-Arg-Gly-Val-Leu	35-	.Val-Lys-Lys-Pro-His
66-	Ile-Arg-Asp-Ala-Val	51-	Ile-Arg-Arg-Tyr-Gln
76-	Ala-Lys-Arg-Lys-Thr	63-	Ile-Arg-Lys-Leu-Pro
90-	Leu-Lys-Arg-Gln-Gly	114-	Ala-Lys-Arg-Val-Thr
		126-	Ala-Arg-Arg-Ile-Arg

+ Black and Dixon, (1967)

* Phillips, (1971b)

histone F2a1 and F3. The pentapeptides listed in Table 3.1, column 4, are highly homologous with the fundamental protamine pentapeptide. Approximately 33 base changes out of the 147 bases, making up the codons for the peptides listed (Table 3.1, column 4), are required to produce all the 10 peptides from the fundamental protamine pentapeptide.

A slightly different but still highly repetitive sequence is evident by aligning the basic amino acids in the C-terminal half of a pentapeptide (Table 3.2). It is, of course, apparent that this peptide partly overlaps with the pentapeptides listed in Table 3.1, column 4.

The results in Table 3.1 indicate that at the N- and C-terminal end the repetitive sequences in the histones occur at approximately every 10 - 12 residues, whereas any homology is absent in the central non-basic part of the histone F3. The spacing of 10 - 12 may indicate that the parent polypeptide has been larger than a pentapeptide.

This would fall in line with the hypothesis that histone F2a1 and other proteins have been derived from a dodecapeptide. Bauer (1971) on translating the calf thymus histone F2a1 sequence into mRNA code, revealed the existence of two groups of internal homologies A and B. Amino acid sequences in a number of proteins were found to be homologous to these two peptides.

From these homologies an ancestral histone IV peptide A (AHAP) could be deduced which is not only repetitive in this histone but apparently also occurs in proteins like ferredoxin, cytochrome C and immunoglobin L chains, to name some of them.

Bauer proposes that histone F2a1 and the other mentioned proteins were derived from an ancestral dodecapeptide (Table 3.3). To be able to distinguish true homologies from random similarities Bauer uses random sequences as a reference. The degree of identity of base positions is expressed by the quotient (f) (identical base positions/all base positions). This homology quotient for random sequences was found to be $f_{RS} = 0.43 \pm 0.1$.

Applying this analysis to histone F3 it is evident that homologies exist between various stretches of F3 histone and AHAP (Table 3.3).

TABLE 3.2

INTERNAL SEQUENCE SIMILARITIES OF HISTONE F3

Gly - Gly . Lys
Gly - Gly - Val - Lys - Lys
Ile - His - Ala - Lys - Arg
Arg - Gly - Ile - Arg - Arg
Gln - Leu - Ala - Arg - Arg
Lys - Ala - Ala - Arg - Lys - Ser
Lys - Ala - Pro - Arg - Lys
Gln - Thr - Ala - Arg - Lys - Ser
Leu - Ile - Ile - Arg - Lys
Leu - Arg - Gly - Ile - Arg
Gln - Arg - Leu - Val - Arg
Thr - Val - Ala - Leu - Arg
Pro - His - Arg - Tyr - Arg

Generally only a few base changes are required to produce the various peptides.

However, the observations by Bauer (1971) have to be interpreted with a certain amount of caution especially for shorter peptides, e.g. peptide 128 - 134 (Table 3.3), can be shifted one position to the left or completely inverted and it still shows homology. Furthermore, polyarginine is also highly homologous according to this criterion (Table 3.3). Alternatively, it may indicate a close relationship between the ancestral histone IV A peptide and a very Arg-rich peptide, e.g. protamine. McLachlan (1971) maintains that many of the repeats noted in protein sequences may easily have occurred by chance. Parallel duplication of a gene is considered as an event in evolution which leads to two distinct new genes (Dixon, 1966). Serial duplication in which the new gene codes for a single polypeptide chain of twice the original length, has also occurred sometimes. The constant parts of the antibody heavy chains are good examples (Milstone & Pink, 1970). However, in spite of a number of examples, serial

TABLE 3.3

COMPARISON OF ANCESTRAL HISTONE IV A PEPTIDE (BAUER, 1971) WITH HISTONE F3

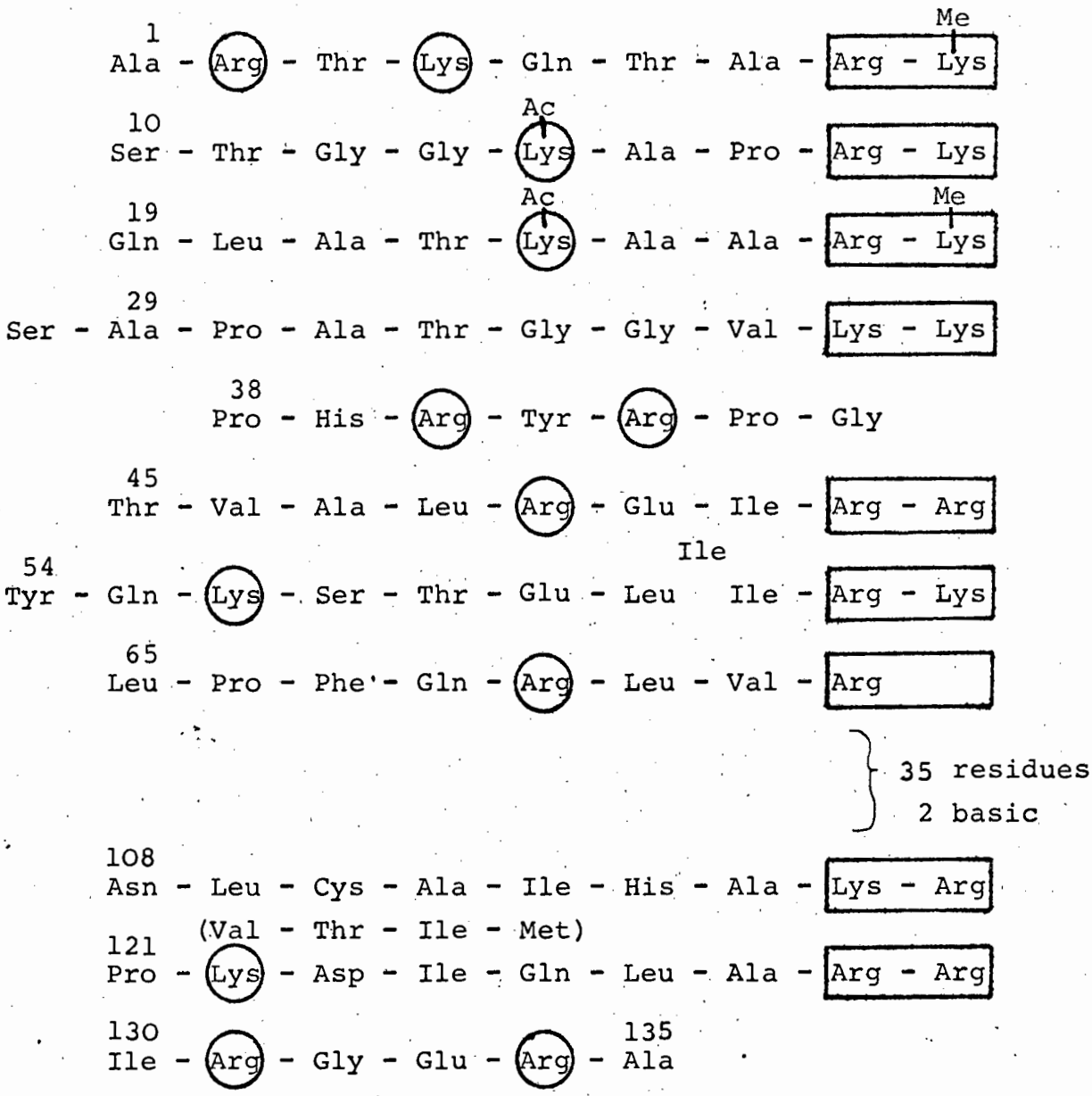
Amino acid residue No.	1	2	3	4	5	6	7	8	9	10	11	12	
AHAP — Amino acid residue	Ser	Gly	Leu Arg Ile	Gly Thr Ala	Lys Arg Phe	Gly Pro	Gly Val	Ala Val Leu	Lys Ile	Arg	His	Arg	
mRNA codon	UCn	GGn	CU$_n$ AG$_n$	GG$_n$ AC$_n$	CG AAAG UUCU	GG$_n$ CC	G^G_Un	GC$_n$ CU$_n$	A^A_UA	CGn	C CAU	CGn	Common RNA bases f
residue No. 128-134		Ala GGn	Arg CGn	Arg CGn	Ile A AUC U	Arg GGn	Gly A CAG	Gln					$^{18}/_{21}$ = .85
41-49	Tyr C UAU		Arg CGn	Pro CCn	Gly GGn	Thr ACn	Val GUn	Ala GCn	Leu CUn	Arg CGn			$^{22}/_{27}$ = .81
28-40	Ser UCn	Ala GCn	Pro CCn	Ala GCn	Thr ACn	Gly GGn	Gly GGn	Val GUn	Lys A AAG	Lys Pro CCn	His C CAU	Arg CGn	$^{32}/_{36}$ = .88
7-17	Ala GCn	Arg CGn	Lys A AAG	Ser UCn	Thr ACn	Gly GGn	Gly GGn	-	Lys A AAG	Ala GCn	Pro CCn	Arg GGn	$^{25}/_{33}$ = .75
1-6		Ala GCn	Arg CGn	Thr ACn	Lys A AAG	Gln A CAG	Thr ACn	Ala GCn	-	Arg CGn	Lys A AAG		$^{22}/_{27}$ = ,81

TABLE 3.3. /cont'd...

Amino acid residue No.	1	2	3	4	5	6	7	8	9	10	11	12	
AHAP — Amino acid residue	Ser	Gly	Leu Arg Ile	Gly Thr Ala	Lys Arg Phe	Gly Pro	Gly Val	Ala Val Leu	Lys Ile	Arg	His	Arg	
mRNA codon	UCn	GGn	$\begin{matrix}CU_n\\AG\end{matrix}$	$\begin{matrix}GG_n\\AC\end{matrix}$	CG AAAG UUCU	$\begin{matrix}GG_n\\CC\end{matrix}$	$\begin{matrix}G_Gn\\U\end{matrix}$	$\begin{matrix}GC_n\\CU\end{matrix}$	$\begin{matrix}A_A\\U\end{matrix}A$	CGn	$\begin{matrix}C\\CAU\end{matrix}$	CGn	Common RNA bases f
residue No.	Ala GCn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	Arg CGn	$28/36 = .78$
134-128 (inverted)		Gln A CAG U	Gly GGn	Arg CGn	Ile A AUC U	Arg CGn	Arg CGn	Ala GCn					$15/21 = .72$

TABLE 3.4

INTERNAL SEQUENCE HOMOLOGY IN HISTONE F3



duplication is considered to be a rare genetic event (McLachlan, 1972).

On the basis of internal homologies one can divide histone F3 into three regions depending on whether or not the region contains a repetitive sequence. Both basic N- and C-terminal regions contain a repetitive nonapeptide (Table 3.4). The middle non-basic portion (residues 73 to 114) of the molecule which also contains the cysteine is devoid of any repetitive sequences. From Table 3.4 it is evident that most of the basic residues are regularly spaced

in the molecule especially the dibasic peptides.

The symmetry of the distribution of the modified lysines is remarkable. These homologies are supporting evidence in favour of an ancestral peptide gene which, by duplication and other mechanisms modifying its original genetic information, gave rise to at least part of the structure of histone F3.

3.2.2 Direct comparison of F3 and F2a1

Phillips (1971b) compared the amino acid sequences of histone F2a1, F2b and F1. All three seem to be structurally related, although the relationship may be more distant between F1 and F2a1. It is maintained that although "they evolved in detail very different, the results reveal common parentage for these proteins". From Table 3.5 it is apparent that overlapping of residues also exist in F3 and F2a1. The complete absence of homology in the non-basic region of the two proteins is evident.

These results indicate that histone F2a1 and F3 are related. They may have been derived from a common and less complex ancestral peptide. Before the amino acid sequence of the two histones became "frozen" over a billion years ago the sequence of certain regions and sites must have been less rigidly conserved. In this way the specific nature of the interaction between the ancestral histone and DNA, characteristic for chromatin, could evolve in the common precursors to all differentiated organisms.

One of the selective factors in the evolution of histones may have been the requirement to bind to DNA in order to stabilize a higher structure. As the protamines illustrate, this function can be achieved satisfactorily with a much simpler structure.

The primitive histone may therefore have originated from a highly basic, protamine like polypeptide stabilizing the pre-nucleoprotein. As the primitive organism evolved non-basic residues may have been added to the basic peptide to exert some regulatory control on the increasing amount of genetic material. In the basic part only few changes were

TABLE 3.5

SEQUENCE SIMILARITIES BETWEEN HISTONE F3 AND F2al

```
                1                       12                      19
F3        Ala-Arg-Thr-Lys-  -  -Gly-Gly-Lys-Ala-  -  -Gln-Leu-Ala-Thr-Lys
          1                       6                       27
F2al  Ser-Gly-Arg-Gly-Lys       Gly-Gly-Lys-Gly       Gln-Gly-Ile-Thr-Lys
      ---------------------------------------------------------------------
          23                                                      33
F3    Lys-Ala-Ala-Arg-Lys-Ser-Ala                             Gly-Gly-Val-Lys-Lys-Pro
          16                                                      41
F2al  Lys-Arg-His-Arg-Lys                                     Gly-Gly-Val-Lys-Arg-Ile
          31                                                      13
      Lys-Ala-Ile-Arg-Arg-Leu-Ala-                            Gly-Gly-Ala-Lys-Arg-His
      ---------------------------------------------------------------------
          39                  47
F3    His-Arg-Tyr-Arg-  -Ala-Leu-Arg- . -Glu-Ile-Arg-Arg-Tyr
          18                  33
F2al  His-Arg           Ala-Ile-Arg-Arg-Leu-Ala-Arg-Arg-Gly
          25                                  34
      His-Ala-Lys-Arg                     Ile-Arg-Lys-Leu
      ---------------------------------------------------------------------
          58
F3    Thr-Glu-Leu-Leu-Ile-Arg-Lys-  -  -  -  -  -  -  -  -  -  -  -  -
          73
F2al  Thr-Glu-His-Ala-Lys-Arg-Lys
      ---------------------------------------------------------------------
F3      -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
F2al
      ---------------------------------------------------------------------
          112                             120
F3    Ile-His-Ala-Lys-Arg-Val-Thr-Ile-Met-  -  -  -  -  -  -  -  -  -  -
          74
F2al  Glu-His-Ala-Lys-Arg-Lys-Thr-Val
                              81
                        Val-Thr-Ala-Met
      ---------------------------------------------------------------------
          125
F3    Gln-Leu-Ala-Ala-Arg-Ile-Arg-Gly-Glu-Arg-Ala
               37
F2al      Leu-Ala-Arg-Arg
               90
          Leu-Lys-Arg-Gln-Gly-Arg-Thr
```

permissible without interfering with the precise stabilizing fit between DNA and histone, while other stretches, through extensive mutations, acquired different functions. The intriguing question, however, as to the nature of the function of these molecules, for which the perfect structure had been found over a billion years in the common ancestor, remains unanswered.

An alternative possibility of the evolution of the histones, although less probable, is that they have evolved from completely different proteins and the similarity and repetition has arisen due to the requirements of binding to the double stranded DNA which possesses a highly repetitive topography of regularly spaced, negatively charged phosphate groups and a major and minor groove with a hydrophobic interior.

3.3 ANALYSIS OF THE PRIMARY STRUCTURE OF F3 HISTONE FOR POSSIBLE CONFORMATION

The complete amino acid sequence determinations, the glycine-rich arginine-rich histone fraction F2a1 (DeLange et al., 1969a; Ogawa et al., 1970) have revealed the striking feature that the distribution of amino acid along the polypeptide chain is extremely irregular.

The distribution of amino acids in different regions of the F3 molecule and F2a1 are shown in Table 3.6. Also included in the Table are the amino acid compositions of a globular protein and an enzyme. It is clear from the Table 3.6 that histone F2a1 (and F2b, which is not shown) is polar with a heavy density of basic residues in one portion of the molecule. In the case of histone F2a1 the amino terminal half of the molecule contains a very high proportion of basic residues and glycine. The distribution had led to the suggestion that the amino terminal half of the molecule is the primary site of interaction with DNA (DeLange et al., 1969a). The C-terminal half of F2a1 has a composition which is similar to that of enzymes and globular proteins and is thought to possess a secondary structure which is not imposed by its interaction with DNA.

A similar situation pertains to histone F3. Inspection of the amino acid sequence reveals that the molecule can be divided into three distinct regions (Table 3.6). A basic N-terminal region which is also rich in proline is followed by a region (approximately 41 residues) which is rich in acidic and apolar amino acids, which in turn is followed by a basic C-terminal section which is very similar to the N-terminal region.

Both the basic C- and N-terminal regions are rich in helix-destabilizing residues like Pro, Gly, hydroxyamino acids and adjacent similarly charged residues, e.g. Arg - Lys, and one would expect it to be devoid of any secondary structure, i.e. existing in solution as a random coil. In contrast, the middle non-basic portion contains more non-polar and acidic residues and the composition is not unlike that of globular proteins and clearly contains the potential for the formation of secondary structures.

Furthermore, the sequence of a protein can be displayed to obtain information on those regions which have the potential for helix formation. One way is to construct "helical wheels" (Schiffer & Edmundson, 1967) which are projections of the side chains on to a plane perpendicular to the helix axis, assuming the residues to lie on an α-helix. Alternatively, the features of the side-chain distribution become apparent if the helix is opened out, as illustrated in Fig. 3.1, displaying a two dimensional helical surface (Bradbury & Crane-Robinson, 1971).

From Fig. 3.1 it is apparent that the basic residue (hatched spheres) occur in clusters which would probably destabilize an α-helical conformation in these regions. Potential regions for the formation of a helical structure exist in the non-basic region which also contains the cysteine residue.

Although this is a somewhat naive way of examining protein sequences it does give qualitative information on which of the polypeptide chain regions might be expected to form helical conformations. It is, however, possible that short helical segments exist in the basic regions.

TABLE 3.6

COMPARISON OF THE AMINO ACID DISTRIBUTION OF VARIOUS REGIONS IN F3 AND F2al HISTONES

	F3			F2al		Lactalbumin	Ribonuclease
	1-72	72-113	114-135	0-51	51-102		
	mole %			mole %		mole %	mole %
Pro	6.9	0.0	4.6	2.0	0.0	1.7	3.0
Gly	6.9	2.4	4.6	24.0	9.7	5.0	2.5
Acidics	8.4*	29.3	13.8	2.0	12.0	8.0	8.0
Basics	32.0	4.9	31.9	31.3	17.6	11.0	11.0
Apolars	32.0	46.5	36.5	29.4	35.0	31.0	26.5
Aromatics	5.6	9.8	4.6	4.0	12.0	12.0	10.0
Thr, Ser	14.9	12.2	4.6	6.0	12.0	11.0	20.0
B/A	3.8	0.17	2.3	15.6	1.5	1.4	1.4

* Out of the 6 Glu, 4 are in the amide form.

It seems, therefore, likely that the basic N- and C-terminal parts of the F3 molecule are rather extended and are the primary site for the interaction with the DNA-phosphate-ester chain. The central non-basic portion of the molecule probably possesses a secondary structure which is not imposed by its interaction with DNA.

This conformation is supported by ORD studies on calf histone F3 which indicate a helix content of near 15% in neutral aqueous solution. Helix content is increased to values of between 25% and 30% for the salt-induced conformational changes and is 17% for pH induced changes (Bradbury et al., 1967). Furthermore, histone F3 should be a rather extended molecule and possess a very high axial ratio. This is borne out by molecular weight estimates based on elution volume on exclusion chromatography using globular proteins as standards. The apparent molecular weight is near 50,000 (Hnilica & Bess, 1963) which exceeds the actual value of 15,500 by a factor 3.

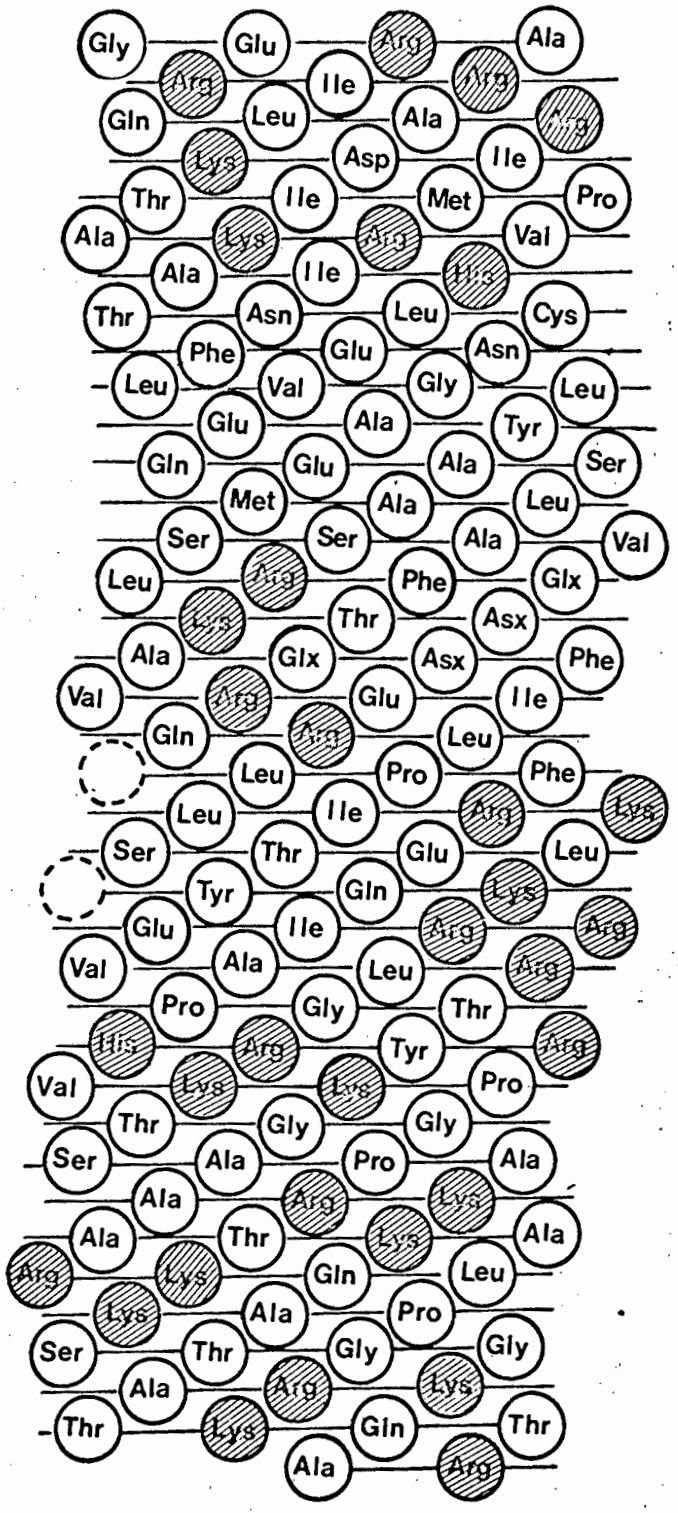


Fig. 3.1 : Helical surface of F3 histone.

3.4 STRUCTURE AND FUNCTION OF DEOXYRIBONUCLEOHISTONE

3.4.1 Introduction

The presence of histones in association with DNA is now well established. The evidence for chromosomal localization of the basic proteins range from microspectrophotometric determination of histone-dye complexes in situ (Swift, 1964) to the direct analysis of isolated chromosomes and chromatin fractions (Sadgopal & Bonner, 1970a; Marushige & Bonner, 1966).

Furthermore, it is certain that due to their proximity and quantity the proteins associated with DNA in the chromosomes of higher organisms influence both the structure and function of the genetic material. It is very well established, e.g. that the addition of histones to DNA suppresses DNA dependent RNA synthesis (Huang & Bonner, 1962).

The physical state of chromatin correlates closely with its biosynthetic activity. A classic example is the morphological changes in giant chromosomes of Dipteran insects which show that certain segments undergo a characteristic "puffing" which represent swelling of those chromosomal regions which are sites of most intensive RNA synthesis (Pelling, 1964). Such change in state and function of pre-existing DNA is temporal and has been correlated with alterations in chromosomal proteins.

It has been concluded that the state of association and cross-linking of DNA-protein fibrils in chromosomes can be modified and that such modifications are likely to involve changes in the amount, nature, and configuration of DNA-association proteins. It has been suggested that histone structural modification, such as phosphorylation, acetylation, methylation and changes in thiol/disulfide content may be mechanisms affecting the structure and reactivity of the DNA-histone complex (Allfrey, 1971d).

In mammalian cells there is both chemical and morphological evidence that only a small fraction (5 - 10%) of DNA in the chromatin is functional as a template for the synthesis of RNA (Paul & Gilmour, 1966).

Electron microscopy revealed that chromatin is seen to be distributed in two easily distinguishable forms - dense clumps of compact fibrils containing most of the DNA and

diffuse regions containing loosely extended filaments of about 100 - 150 Å diameter (Bram & Ris, 1971). High resolution autoradiography has revealed that the diffuse regions are engaged in ribonucleic acid synthesis (Littau et al., 1965; Huxley & Zubay, 1961).

Experimental evidence has accumulated indicating that conformation of DNA in the chromatin of eukaryotic cells differs from the conformation of isolated protein free DNA (Simpson, 1972). The conclusion from hydrodynamic, spectroscopic, microscopic and other measurement has been that DNA in chromatin is most probably supercoiled into a tertiary helix, likely with a pitch of 120 Å and a diameter of about 100 Å (Richards & Pardon, 1970). The details of stabilization of such a supercoil are not known though. It became evident that the supercoiled structure will depend on the nature and conformation of the histones, their internal distribution, and their interaction with other components of the chromatin. For example, at very low concentrations of deoxyribonucleoprotein in solutions lacking bivalent ions supercoiling appears to be absent. As the deoxyribonucleoprotein concentration increases, or on the addition of Mg^{2+} ions, the supercoil appears, indicating that some of the forces responsible for the generation of this configuration may be intermolecular. That the conformation of histones is very important in maintaining the supercoil is indicated by the observation that partial removal of histones by dissociation of the nucleohistone complex in sodium chloride at above 1.2 M concentration results in the loss of supercoiling and successful reconstitution of the complex from histones is dependent on the method by which the histones had been prepared (Richards & Pardon, 1970).

Differences in the DNA-binding by different histone fractions have been observed (Johns & Forrester, 1970). The lysine-rich histones (F1) are only loosely bound to DNA, they are the easiest to remove in dilute acid and are the first to dissociate when the salt concentration is raised (Murray et al., 1968; Murray, 1969; Wilhelm & Champagne, 1969). This dissociation demonstrates that their removal is not accompanied by any striking change in the DNA template activity (Bonner et al., 1968).

Suggestions have been made that the very lysine-rich histones tend to form cross-links between DNA molecules, whereas the arginine-rich histones align themselves along separate DNA molecules (Littau et al., 1965). This is in accord with electron microscopic observations indicating that the state of chromatin is altered from a compact to a more diffuse configuration when the lysine-rich histones are selectively removed (Littau et al., 1965).

From whole histone prepared by salt dissociation or from histones less fraction F1, nucleohistone can be reconstructed with an X-ray diffraction pattern indistinguishable from native nucleoprotein. Fraction F1 alone never gives the supercoil configuration with DNA under conditions which for a mixture of F3 and F2a2 produce the supercoil (Richards & Pardon, 1970).

While these findings indicate that histones are involved in supercoiling the DNA in heterochromatic regions and in the functional repression of the chromosomes, the complexity of the situation is emphasized by the failure to detect any major difference in the nature and content of the histones extracted from genetically active and inactive chromatin (Palotta et al., 1970).

3.4.2 The primary structure of histone F3 in relation to the structure of deoxyribonucleohistone

In the light of the considerable amount of information available on histones and deoxyribonucleohistone complex the possible interaction between histone F3 and DNA will be examined. The nature of the DNA-histone complex, the forces controlling the fit of histones onto the double stranded DNA are clearly of great importance in understanding the structure and function of the gene and their organization in the chromosome.

Based on the properties on the individual amino acids and their sequences in the histones a model of the nucleohistone complex structure can be built. Suggestions for the structural arrangements between histone F2a1 and F2b and DNA in the DNH complex have been made (Phillips, 1971c; Richards & Pardon, 1970). Because these histones are structurally

related to histone F3 one would expect that the histone F3 - DNA complex would be very similar. On building such a model certain facts have to be accepted and a number of assumptions to be made :

A general attribute of histones is their ability to complex with DNA. It is still uncertain whether the various histones show any specificity in their combination with DNA. In the light of present knowledge it seems, however, likely that the specificity of combination is very low (Johns & Butler, 1964; Garret, 1968). The attachment of histones to DNA takes place mainly through salt linkages of positively charged lysine- and arginine residues to negative phosphate groups. From a detailed titration study on nucleohistone and its protein component (Walker, 1965) it was concluded that 80% of the basic groups in nucleohistone are involved in ionic bonds to phosphate groups. The carbonyl, tyrosine and histidine groups were freely titrated.

The strength of binding of histones seems to be related to the number of Arg residues, the arginine-rich histones (F2a1 and F3) are the most difficult to remove with salt or acid solutions (Wilhelm & Champagne, 1969; Murray et al., 1968).

It appears, however, that hydrophobic and possibly hydrogen bonds might also be involved though to a lesser extent (Lewin & Jacobson, 1967), e.g. the arginine-rich histones are readily extracted with ethanolic-HCl (Fig. 4.1 page 138).

The groups involved in this type of bonding are not yet known (Bartley & Chalkley, 1972). The histone to DNA weight ratio in calf thymus deoxyribonucleohistone is about 1.1 - 1.7/1 (Tuan & Bonner, 1969; Dounce et al., 1972). Therefore the number of positive charges contributed by the five histone groups, present in approximately equal amounts, are nearly equal to the number of phosphate groups in the associated DNA. F3 histone, like the other histones, contains about 30 basic residues per 15,000 daltons. A stretch of DNA with 10^5 nucleotide pairs possesses 30 negative phosphate groups, and an average weight of 9300 daltons.

It has earlier been mentioned that histone F1, which shows no α-helical structure (Bradbury, 1969) in nucleohistone, is not bound in the same detailed manner as the other histones.

It has been suggested that it might cross-link several nucleohistone strands by attaching itself to a few phosphate groups of the DNA and perhaps also to carboxyl groups of other histones (Littau et al., 1965). Thus, the F1 histone will add to the weight of histone per unit weight of DNA but will add little to the length of DNA required to accommodate the histones.

The average amount of α-helix in nucleohistone is 30% (Bradbury, 1969). The inter-phosphate distance on each strand of the DNA-dyad helix is 7 Å and the inter-residue distance in the non-helical peptide chain is 3.5 Å, whereas α-helical sections have an effective length of 1.5 Å per amino acid residue.

Protamines are thought to straddle the narrow groove of DNA in nucleoprotamine (Wilkins, 1959). Similarly, those parts of histones involved in salt linkages to phosphates in nucleohistones might also cover the narrow groove (Phillips, 1971c).

In contrast, Richards and Pardon (1970) suggest that the basic region of the histones lies in and follows the large groove in the DNA double helix. (The absence of the first layer line reflection of DNA in the X-ray diffraction patterns of DNH supports this possibility).

It appears from titration studies on deoxyribonucleohistone with cationic dyes (Kurashina et al., 1970) and poly lysine (Li et al., 1972) that approximately 25% of the phosphate groups are not involved in bonds with histones. Values of 50% free phosphate groups (e.g. Clark & Felsenfeld, 1971) have been claimed to be erroneous due to the incorrect assumption that poly lysine will only bind to DNA free of histones (Li et al., 1972).

These requirements demand that the maximum number of basic amino acids reach negative groups, the non-basic amino acids forming loops away from the general structure so that the next basic residue in the protein chain returns to the next phosphate group in either DNA strand. From molecular models it becomes apparent that one or two non-basic amino acids cannot form a complete loop. Therefore a phosphate group in such a region would be bypassed without becoming neutralized by a basic amino acid. However, three or more

non-basic residues can form a loop without bypassing a phosphate residue.

A model for the histone F3 - DNA complex, incorporating these features is given in Fig. 3.2. The DNA helix has been unwound in such a way that the phosphate groups of the minor groove are arranged in one plane. The protein fits neatly into the groove held there mainly by salt-linkages between basic amino acids (red groups) and phosphate groups (hatched spheres) (Fig. 3.2), while stretches of non-basic amino acids bulge out from the structure possibly in a helical form. The DNA histone ratio in this particular model is 1.1/1 and 27% of the phosphate groups are free. This is near the value expected for the nucleohistone complex without fraction F1.

The model in Fig. 3.2 represents the most compact complex possible between histone F3 and DNA. The histone could also be fitted onto the DNA by accommodating some or all the non-basic regions into one of the grooves. This, however, would lower the histone/DNA ratio and increase the number of unbound phosphates.

Histone F3 contains homologous stretches of amino acid residues (Table 3.4). The repetitive unit is in most cases a nonapeptide which possesses two consecutive basic residues in its C-terminus. This unit is, however, absent in the middle non-basic portion of the molecule (Table 3.4).

The spacing of basic groups may well result in a very tight bonding of these regions to DNA while the amino acid stretches in between will interact less strongly with DNA. This regularity in the protein structure may in turn impose some regularity onto the structure of DNA and contribute to the supercoil.

In the light of the present resolution of the chromatin structure at the electron microscopic level, together with evidence from X-ray diffraction, studies indicate that DNA can be considered to exist in at least three forms of increasing degrees of compactness, viz. the double helix of packed DNA (diameter 30 Å) the first supercoil of the protein covered helix (elementary fibre, diameter 100 Å) and a more densely packed second supercoil of the protein-DNA supercoil fibre (super-supercoil, diameter 200 Å) (Bram & Ris, 1971; Richards & Pardon, 1970; Bahr, 1970).

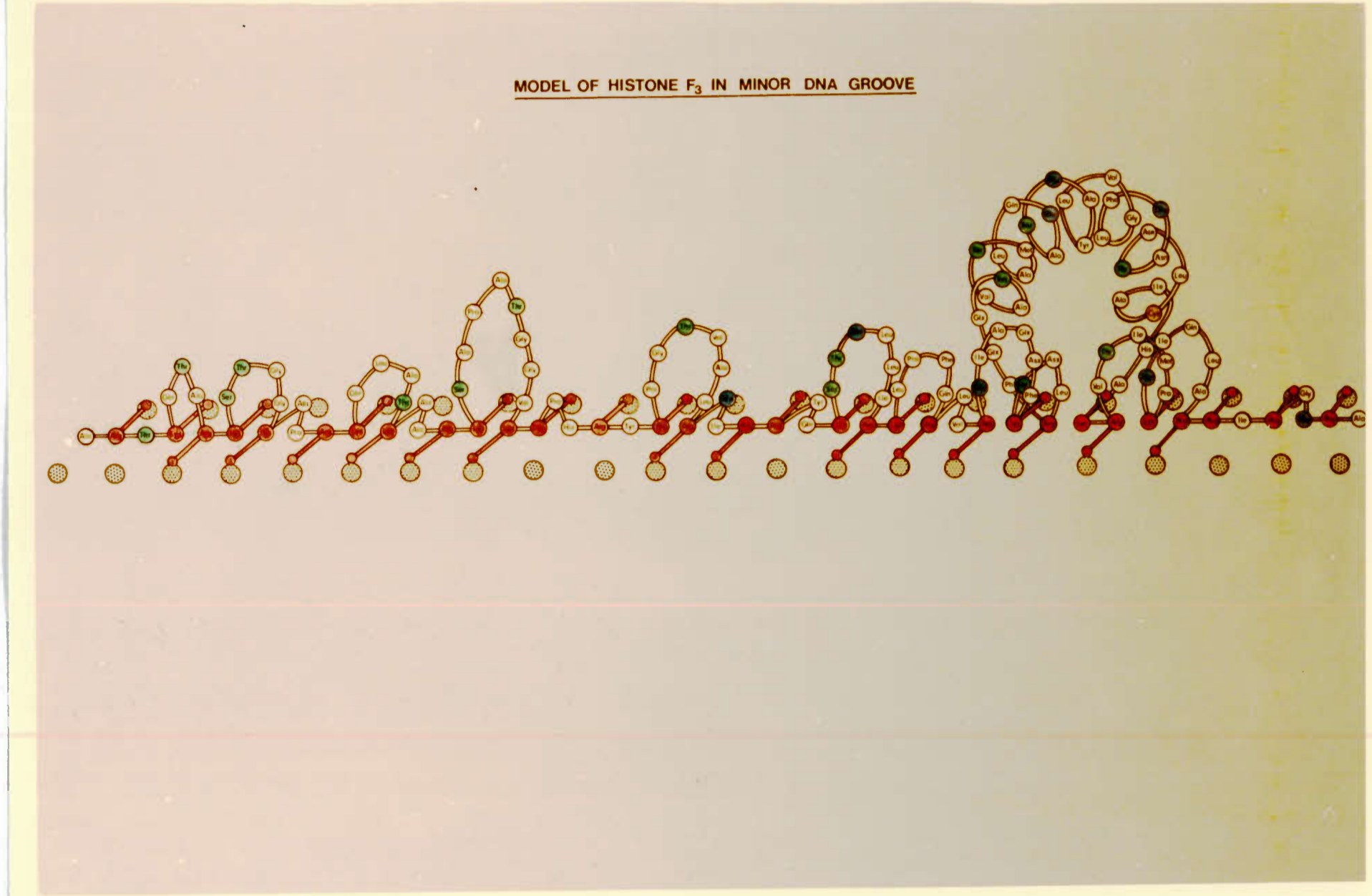


Fig. 3.2 : Fit of histone F3 into the extended groove of double stranded DNA. Hatched spheres represent PO_3 groups, red groups Lys or Arg residues, green spheres hydroxy amino acids and blue spheres acidic residues.

The DNA in the dense part of chromatin or in chromosomes is very tightly packed. It is conceivable that such structures between the individual supercoil strands are achieved by histone-histone or histone-chromosomal protein interaction.

This cross linking could involve either direct binding of the ends of a given histone molecule to different segments of the DNA supercoil, or alternatively, both ends of a histone molecule could be bound to a segment of DNA, and the looped out portion could form a bridge through hydrophobic interaction with histones on an adjacent segment of DNA. A particular structure formed could be stabilized and made more permanent by the formation of disulphide bridges.

3.4.3 Biological significance of the deoxyribonucleoprotein complex

Electrostatic repulsive forces among the phosphates and the stacking forces of the bases provide the rigidity of DNA. Thus the interaction between histone and DNA may change the conformation of the naked DNA by altering the constraints on this molecule and causing it to form a supercoil. If relaxation of the supercoil is a prerequisite for transcription (Richards & Pardon, 1970), the influence of histones on the structure of DNA could be considered as a second order mechanism for gene regulation. This relaxation of the DNH complex, however, does not necessarily require the detachment of histones, but can be achieved by modification of histones through acetylation of lysine residues or loss of α-helical content in the loops not attached to DNA.

The supercoiled conformation of DNA may not be a suitable substrate for binding of RNA polymerase. This assumption is consistent with results obtained after exposing DNH to trypsin (Simpson, 1972). After some of the basic residues of all five histone fractions had been cleaved with trypsin a loss in supercoiling and increase in template activity followed. However, the basic stretches of the histones remain attached to the DNA. This implies that the entire molecule of the histone is required to induce the chromatin to supercoil. From this one can conclude that histones influence the template activity through the modification of DNA conformation.

The histones could also be responsible for packing the DNA beyond the supercoil more closely together in order to achieve the super-supercoil (Bahr, 1970). This form of chromatin may represent the more permanent repressed DNA in the cell nucleus.

The structure of the DNA-histone complex in turn can be altered by a change in the charge distribution of the molecule causing a local detachment of the histone from the DNA. Such modification like methylation and acetylation of lysine residues has been observed (Allfrey & Mirsky, 1963).

Furthermore, the regions of the histone not bound by DNA could interact with other proteins through salt linkages or hydrophobic interactions and thus alter the structure of the complex. In accordance with this is the observation that the removal of some of the histone (mainly F1) from chromatin with solutions of increasing NaCl concentration causes nucleoprotein particles to form a less compact structure of DNA-loops which probably possesses a supercoiled structure (Sonnenbichler, 1969).

The supercoiled loops of chromatin may in turn represent part of the genome to be expressed in a cell of particular differentiation by relaxation of the supercoil.

Despite extensive experimentation, however, the role of histones in gene repression remains obscure. It is generally admitted that <u>in vitro</u> experiments are imperfect replicas of the cellular mechanism.

It seems likely that other primary control mechanisms are involved in differentiation and growth which must involve interactions between DNA and other components of the chromosome at a molecular level making RNA synthesis at a particular genetic locus possible or impossible.

However, extensive work has to be done on the DNH complex in order to gain more insight into the structure and function of (chromosomes) the genetic material in the cell nucleus of higher organisms. No doubt the knowledge of the amino acid sequence of the histones will aid in the elucidation of the structure and function of the DNH in the cell nucleus.

PART 4

MATERIALS AND METHODS

4.1 ISOLATION OF HISTONES

4.1.1 Preparation of calf thymus nucleoprotein

Calf thymi were collected from the local abattoir and transported on ice to the laboratory. All subsequent operations were carried out at 4°C. The thymus was trimmed and minced. Crude nucleoprotein was isolated by the method of Hnilica and Bess (1965), modified by adding bisulphite to the saline-citrate wash medium to prevent proteolysis of histone (Panyim et al., 1968). The minced thymus was homogenized in 5 volumes of 0.14 M NaCl, 0.01 M trisodium citrate and 5 mM sodium bisulphite in a Waring Blendor at high speed for 2 minutes, filtered through 4 layers of cheesecloth, and centrifuged at 5000 x g for 10 minutes in a refrigerated centrifuge. The supernatant was discarded. The sediment was rehomogenized in about 5 volumes of wash medium by blending at low speed for 30 seconds. The homogenate was centrifuged and the sediment washed at least 4 times with the saline-citrate-bisulphite solution by repeated blending and centrifugation.

The final pellet is referred to as crude nucleoprotein.

4.1.2 Preparation of chicken erythrocyte nucleoprotein

4.1.2.1 Isolation of nuclei

Blood of white Leghorn chickens was collected in 10% (w/v) ice cold trisodium citrate (100 ml/l of blood) with stirring to prevent clotting. The blood was transported on ice to the laboratory. All subsequent

operations were performed at 4°C.

In a typical experiment 3 l blood (40 chickens) were processed. The blood was strained through two thicknesses of muslin. Cells were sedimented at 600 x g for 15 minutes in a refrigerated centrifuge. The plasma was discarded. The cells were washed three times by suspending them in 5 - 8 volumes of 0.14 M NaCl - 0.01 M trisodium citrate followed by centrifugation.

All the erythrocytes were pooled and suspended in two volumes of 0.14 M NaCl - 0.01 M trisodium citrate - 0.01 M EDTA - 5 mM sodium bisulphite (medium A). (Murray et al., 1968, Panyim et al., 1968). Digitonin was dissolved in medium A (0.5 mg/20 ml). Erythrocytes were lysed by slowly adding 20 ml of the digitonin solution, with vigorous stirring, per liter of blood initially used.

After an hour of stirring the red viscous solution was centrifuged at 10,000 x g for 10 minutes (Neelin et al., 1966). The supernatant was discarded. The sediment, consisting of intact cells and nuclei, was again suspended in medium A and treated with the same amount of digitonin.

After centrifugation a homogeneous yellowish coloured sediment was obtained. The sediment was washed for at least 5 times by repeatedly suspending it in 5 - 8 volumes of medium A followed by centrifugation at 10,000 x g for 10 minutes. At this stage the supernatant was colourless. All the stages were monitored by phase contrast microscopy which revealed that the final sediment was a homogeneous preparation of morphologically intact erythrocyte nuclei. At this stage nuclei were sometimes stored at -20°C.

4.1.2.2 Preparation of nucleoprotein

The nuclei were homogenized in 1 - 2 volumes of medium A in a Waring Blendor at low speed for two minutes

The homogenate was suspended in 5 volumes of the wash medium followed by centrifugation at 10,000 x g for 10 minutes. The clear supernatant was discarded. This process was repeated four times on the sediment obtained. The yellowish coloured sediment was twice homogenized in 2 volumes of 95% (v/v) ethanol (analytical grade) in the Waring Blendor and the crude nucleoprotein each time recovered by centrifugation at 10,000 x g for 10 minutes (Johns 1964). The yellowish supernatant was discarded (it exhibited a typical carotenoid spectrum). The cream coloured, leathery pellet obtained is referred to as crude nucleoprotein and was either processed further or stored in 95% (v/v) ethanol at -20°C (Johns 1964).

Crude nucleoprotein was also prepared by omitting the sodium bisulphite or substituting it with β-mercaptoethanol. No significant differences in the yields and properties of subsequently prepared crude histone F3 were observed.

4.1.3 Preparation of total histones

Total histones were prepared by extracting the crude nucleoprotein with 20 volumes of 0.25 M HCl at 4°C. The nucleoprotein, from which the ethanol had been removed by centrifugation, was homogenized in the acid at full speed for 30 seconds in a Waring Blendor (Mauritzen et al., 1967). The suspension was centrifuged at 10,000 x g for 20 minutes. The sediment was re-extracted with half of the original amount of acid. The supernatants were separately filtered through a no. 4 sintered glass funnel. Histones were recovered from the extract by dialysing it against distilled water followed by freeze drying. Alternatively, histones were precipitated from the clarified extract by slowly adding 6 volumes of ice cold acetone (analytical grade). Histones were recovered by centrifugation at 10,000 x g for 10 minutes followed by washing the white sediment twice with acetone (Johns 1967).

The final pellet was broken up into small lumps with the aid of a glass rod followed by removing the acetone in a vacuum.

Typical yields of total histones isolated from erythrocyte nucleoprotein obtained from 1 litre of blood was 3 g for the first and 0.6 for the second extraction.

4.1.4 Preparation of crude F3 histone from chicken erythrocyte nucleoprotein

Bulk fractionation of histones have been carried out by the method of Johns (1964). Histones are selectively extracted when ethanol is added to the extracting medium. The histones in the extract can be selectively precipitated by adding organic solvents. The following schematic diagram illustrates the procedure which had been applied to the fractionation of calf thymus histones (Johns, 1964).

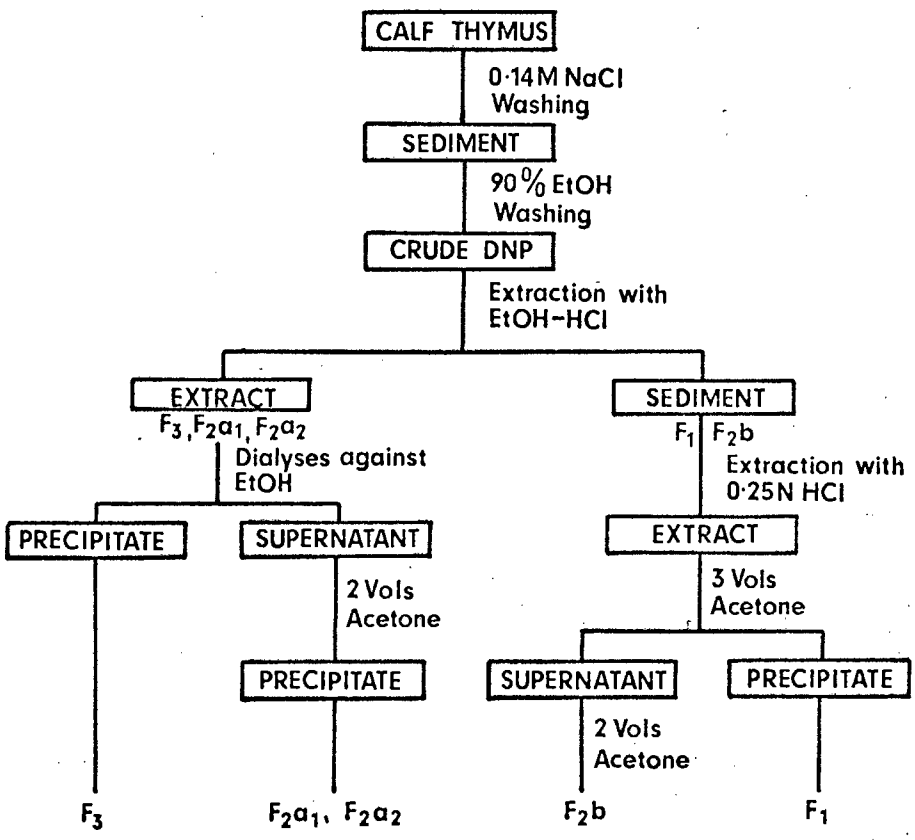


Fig. 4.1 : A method for the large-scale preparation of the main histone fractions (Johns, 1964).

All operations were performed at 4°C. Chicken erythrocyte nucleoprotein obtained from e.g. 3 l of blood was suspended in 300 ml 80% (v/v) ethanol in 0.25 M HCl. The suspension was transferred to a 1 litre polyethylene bottle containing 20 (0.5 cm) glass balls. The bottle was slowly rolled on a shaker for 18 hours, after which the suspension was centrifuged at 10,000 x g for 20 minutes. The sediment was re-extracted with ethanol-HCl. The yellowish-brown, but clear, supernatants were filtered through a no. 4 sintered glass funnel. The filtrates were dialysed against 2 volumes of absolute ethanol for 4 hours followed by another period of 4 and 18 hours each against the same volume of fresh absolute ethanol (Johns 1964). In most cases the precipitate formed after the two 4 hour dialysis periods (1st precipitate), and the 18 hour period (2nd precipitate), were recovered separately by centrifugation. The white sediments were resuspended in 95% (v/v) ethanol and sedimented out again at 10,000 x g. The pellet was then washed twice in acetone followed by drying it in a vacuum.

Typical yields of crude histone F3 isolated per litre of blood were : first extract first precipitate 0.3 g and the second precipitate 0.2 g. The second extract yielded 0.1 g (Fig. 1.3).

The white powdery substance was stored at -20°C. The remaining histone fractions were isolated as indicated in Fig. 4.1.

4.2 PURIFICATION OF HISTONE F3

4.2.1 Gel filtration

Columns of various dimensions were packed as recommended by the manufacturers with various grades of Sephadex. The eluent used was 0.01 M HCl. Initially

0.02% (w/v) sodium azide was added to inhibit bacterial growth. Since it contributed to the absorbancy between 270 and 220 nm it was later omitted. This also simplified freeze drying, since the dialysis step could be omitted. Azide was, however, added when columns were stored.

Protein samples were generally dissolved in 6 M cyanate-free urea - 0.01 M HCl (the final sample volume did not exceed 2% of the column volume), allowed to stand at 4°C for a few hours, and then layered on top of the gel or applied via a flow adaptor. The flow rate was regulated by a constant hydrostatic head maintained by a mariotte flask. The fraction size was controlled by a drop counter. Protein concentrations were monitored with a Beckman DBG spectrophotometer. Selected fractions were pooled, if the molecular weight was large enough they were dialysed before freeze drying. Additional details of each column run are given in the subscripts to the figures in the text.

Pooled fractions from the 14.5 x 100 cm column were first concentrated by ultrafiltration and washed on the filter with H_2O before freeze drying.

4.2.2 Preparation of cyanate-free urea

It was observed that if protein samples were dissolved in old solutions of urea, or allowed to stand for longer times in urea, the electrophoretic bands became progressively more diffuse. Urea is in equilibrium with ammonium cyanate (Wöhler 1828) which is known to carbamylate amino and sulfhydryl groups in proteins (Stark et al., 1960).

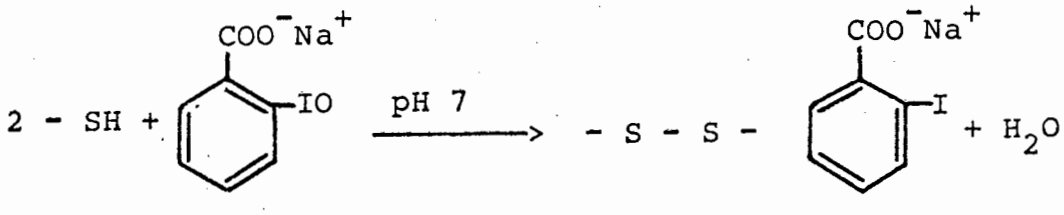
Cyanate concentration can easily be measured by the method of Werner (1923). Freshly made urea solutions were generally found to have low cyanate concentrations.

In crucial experiments 6 - 8 M urea solutions were adjusted to pH 2.5 with HCl and allowed to stand for an hour.

This decomposes cyanate to NH_3 and CO_2 (Stark et al., 1960). This solution was either directly used or the urea recrystallised by cooling the solution.

4.2.3 Dimerization of F3 histone with o-iodosobenzoate

Hellerman et al. (1943) used o-iodosobenzoate to study the role of SH groups in the enzyme urease. This reagent selectively oxidised sulfhydryl groups to disulfide bonds.



At neutral pH no other amino acids were modified to any noticeable extent, although their destruction was evident at low pH values (Hellerman et al., 1943).

In a typical experiment 600 mg (approximately 40 µmoles) of crude F3 histone (prepared by method 4.1.4) was dissolved in 20 ml 6 M urea previously acidified to pH 2.5 (4.2.2). All subsequent operations were carried out at 4°C. The protein solution was allowed to stand for a few hours to disaggregate. Ten ml of 0.5 M Tris-buffer was added and the solution carefully adjusted with 0.1 N NaOH to pH 7. o-Iodosobenzoic acid (40 µmoles) was dissolved in an equivalent amount of 0.1 NaOH and 5 ml H_2O and was slowly added to the protein solution with stirring. The oxidation was allowed to proceed for an hour. The mixture was dialysed for 18 hours against three changes of

distilled H_2O to remove unreacted o-iodosobenzoate. An equal volume of 8 M cyanate-free urea (pH 3) was added and the solution allowed to stand for 12 hours. This solution was finally filtered and applied to a 14.5 x 100 cm Sephadex G-100 column (1.4.2).

4.2.4 Ultrafiltration

Protein solutions were concentrated in an Amicon model 202 ultrafiltration cell using an UM-10 membrane which retains proteins larger than 10,000 daltons. The pressure applied to the cell was approximately 5 kg/cm^2.

4.3 CHARACTERISATION OF HISTONE F3

4.3.1 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis was carried out by a slightly modified method of Panyim and Chalkley (1969). The following solutions were prepared :

A. 60% (w/v) Acrylamide - 0.4% (w/v) bisacrylamide
B. 40% (v/v) Glacial acetic acid - 0.4% (w/v) N,N,N'N'-tetramethylethylenediamine (TEMED)
C. 0.2% (w/v) $(NH_4)S_2O_3$ - 36% (w/v) urea

All solutions were stored at $4^{\circ}C$. The persulphate-urea solution (C) was prepared fresh every week.

Gels were prepared by mixing solution A, B and C in the ratio 2:1:5 (15% acrylamide). This solution was degassed in a vacuum for 1 minute, poured into 9 x 0.5 cm glass tubes and covered with a layer of 2.5 M urea. After 2 hours the gels were rinsed with electrode solution (0.8 M acetic acid) and pre-electrophoresed for 4.5 hours at 2 mA/gel with 0.8 M acetic acid as electrode solution. The final constant voltage was near 150 V. After the electrode solution had been replaced with fresh 0.8 M acetic acid, salt-free samples were dissolved in fresh 6 M urea and carefully layered on the top surface of the gel. Loads of

5 - 10 µg per fraction in 20 µl were generally used. The upper buffer reservoir contained the anode and the lower reservoir the cathode. The current applied was 2 mA/gel.

In some cases gels were prepared in the late afternoon and allowed to polymerize. All subsequent operations were carried out at 4°C. Gels were automatically pre-electrophoresed during the very early morning using time clock operated electrophoresis power supply. The complete electrophoretic analysis could now be carried out in a single working day.

Gels were removed from the tubes by cooling them in ice followed by loosening them with cold H_2O forced between the gel and the glass tube with the aid of a syringe.

Staining and de-staining of gels : Gels were stained for 0.5 hours with 0.2% (w/v) amido black in 7% (v/v) acetic acid, 25% (v/v) ethanol and water. The stained gels were transferred to destaining solution (25% (v/v) ethanol and 7% (v/v) acetic acid in H_2O) and allowed to stand for a few hours. Gels were destained in a transverse electrophoretic destainer filled with the destaining solution.

Gels were kept in the ethanol-acetic acid solution at 4°C until photographed.

Photography : Gels were placed in test tubes, filled completely with destaining solution and stoppered. They were then photographed against an evenly illuminated background with a Nikon FTN camera through a Macro Nikkor lens (1:3.5), fitted with a red filter (R.60). A number of exposures at different speeds were made on an Ilford FP-4 film (200 ASA) which was developed in Kodak microphen diluted 1:3 and using a developing time of 12 minutes. The film was fixed with Unifix. The negative with the highest contrast was printed on a hard paper.

4.3.2 Amino acid analysis

Proteins and peptides were hydrolysed in constant boiling HCl. To prevent oxidative destruction of various amino acids, anti-oxidants like phenol and thioglycollic acid were added (Sanger & Thompson, 1963; Matsurbara & Sasaki, 1969). Destruction of the various amino acids was determined experimentally by studying the recoveries versus hydrolysis time and extrapolating to zero time (Fig. 4.2). Decomposition values of 3% for Thr, 8% for Ser, 10% for Tyr and 6% for cystine were generally found for hydrolyses at 110°C for 24 hours. Corrections to subsequent amino acid determination were applied.

Twice distilled constant boiling HCl (5.7 M) was prepared by collecting the fraction distilling at a constant temperature (near 108°C) from an all-glass still (Moore & Stein, 1963).

Protein samples (0.5 - 1.0 mg) were dissolved in 1 ml 5.7 M HCl containing 1% (w/v) phenol or thioglycollic acid in rimless neutral glass tubes. The tubes were necked out, their contents frozen and evacuated below 0.02 torr (Pyrani guage). The sample was allowed to thaw completely while boiling was prevented by spreading the solution as a thin film on the glass wall of the tube with the aid of a whirly mixer. The top half of the tube was warmed to prevent protein foam from escaping. With the aid of a two way stopcock the tube was flushed with N_2. After freezing the contents of the tube again the evacuation was repeated. The hydrolysis tube was then flame sealed and the samples hydrolysed at 110°C in a thermostatically controlled oven for 24 hours.

After the hydrolysates had been cooled the tubes were opened and the HCl evaporated over NaOH in an evacuated desiccator. The dry hydrolysates were dissolved in pH 2.2 buffer and the analysis performed on a Beckman Model 116 amino acid analyser using the method of Spackman et al. (1958). Basic amino acids were eluted from a 5.5 cm column of PA-35

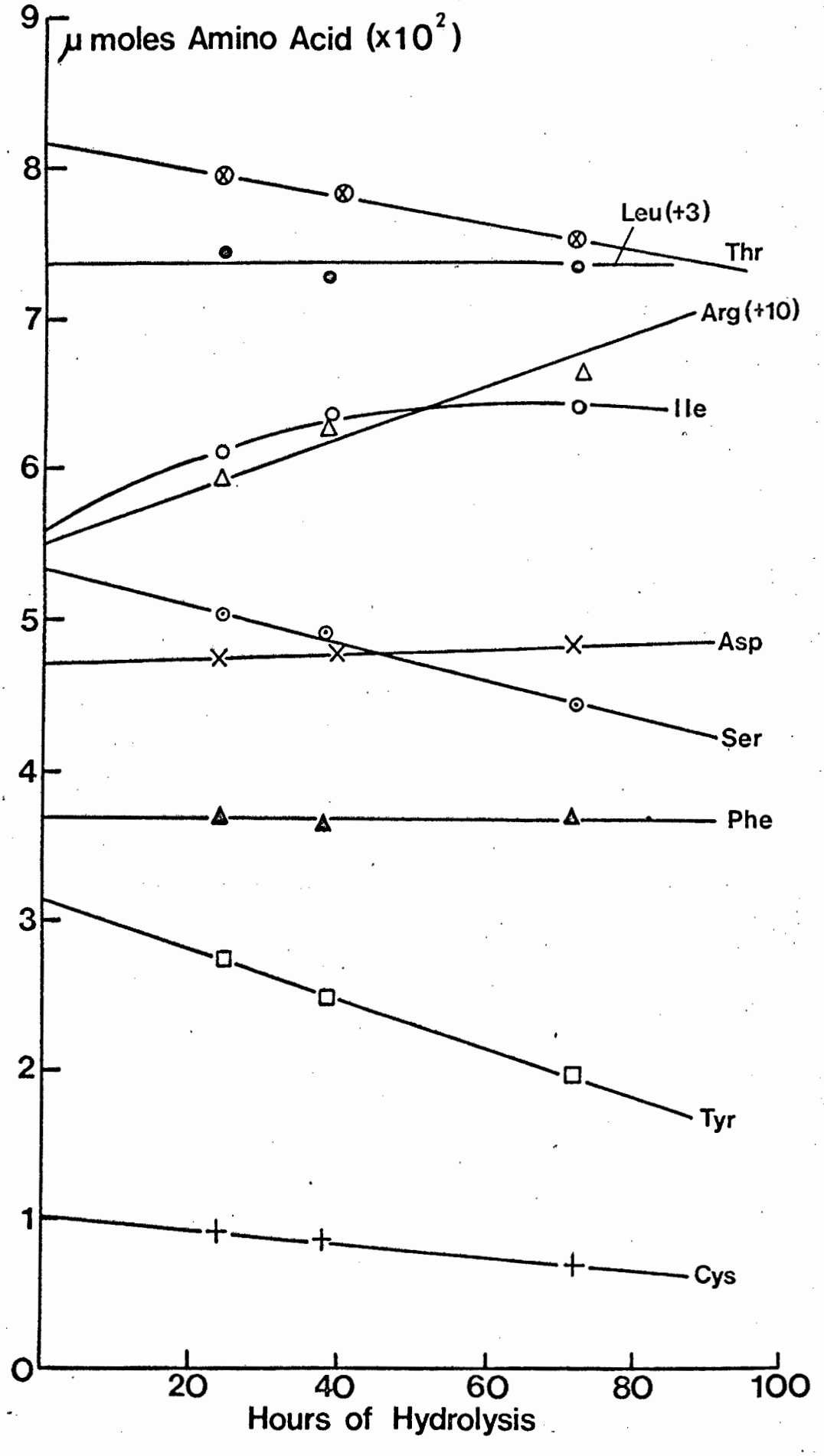


Fig. 4.2 : Recoveries of some amino acids from HCl hydrolysis of histone F3 as a function of time.

resin using 0.35 M citrate buffer at pH 5.25. Acid and neutral amino acids were eluted from a 56 cm column of UR-30 resin using 0.2 M citrate buffer at pH 3.25 and pH 4.30. Peaks were integrated by the "half-height width times height method". The amino acid analyser was frequently restandardised with a calibration mixture. To correct for the ageing of the ninhydrin solution, these calibration runs were used for computations in amino acid determination of analytical runs separated by not more than one week.

4.3.3 Molecular weight determination by electrophoretic methods

Gel electrophoresis has been extensively used as a technique for the separation and characterization of macromolecules. This technique requires very little sample, is simply performed and has remarkable high resolving power. In the case of proteins the separation achieved is the resultant of size and charge. Differences due to charge can be excluded by forming complexes between protein and an anionic detergent, like sodium dodecyl sulfate (SDS). The mobility in the gel will then mainly be determined by size, and possibly shape. The size in turn is related to the molecular weight. A few "anomalous" proteins, however, have been described. The reason for such behaviour is not understood (Dunker & Rueckert, 1966).

Alternatively, by choosing a certain set of conditions and treatment of the electrophoretic data, the effect of the charge on the mobility can be eliminated.

The electrophoretic mobility (m) of a particle is related to the nett charge (Q) and the effective radius (f) by the following expression :

$$m = XQ/_f$$ (Parish & Marchalonis, 1970)

X is a dimensionless factor which changes the effective charge and the size due to ionic environment. If one considers the mobility of a protein in two gels which differ only in their pore size, the ionic environment would remain

identical. Therefore, in gel 1 the mobility of a protein would be $m_1 = QX/f_1$ and in gel 2 $m_2 = QX/f_2$

$$\frac{m_1}{m_2} = \frac{f_2}{f_1}$$

The frictional coefficient (m_1/m_2) for various proteins was found to be directly related to their molecular weight (Parish & Marchalonis, 1970).

4.3.3.1 Electrophoresis of the sodium dodecyl sulfate histone complex

Gels were prepared by mixing 15 ml of 0.2 M sodium phosphate buffer pH 7.2 containing 0.2% (w/v) SDS (buffer A) with 13.5 ml 22.2% (w/v) acrylamide - 0.6% (w/v) methylene bisacrylamide, 1.5 ml 1.5% (w/v) ammonium persulphate and 0.045 ml TEMED. This solution was evacuated for 1 minute, poured into 9 x 0.5 cm glass tubes, overlaid with buffer A and allowed to set for 2 hours. Gels were rinsed with electrode buffer (buffer A diluted 1:1).

Samples were dissolved (1 mg/ml) in a solution containing 30% (w/v) urea, 1% (w/v) SDS and 1% (v/v) β-mercaptoethanol and incubated at 50°C for 60 minutes. Aliquots (20 µl) were carefully layered on the top surface of the gel. Electrophoresis was carried out at 8 mA per tube for 4 hours with the negative electrode in the top tray.

Gels were removed, stained and destained as described in 4.3.1. The mobilities of the various proteins were plotted against the log. of the molecular weight (Fig. 1.7a).

4.3.3.2 Frictional coefficient on gels

Gels were prepared and the samples run as described under polyacrylamide gel electrophoresis (4.3.1). The acrylamide-bisacrylamide solution was prepared from stock solution by diluting it with H_2O to give the desired percentage gels (10 - 15% acrylamide).

Total histones from chicken erythrocytes were applied to the various pore sized gels. The mobility of the various

histones were measured. Since the molecular weight for histone F2a1, F2b and F1 is known, these served as molecular weight standards (Fig. 1.7b).

4.3.4 Estimation of sulfhydryl groups of spectrophotometric titration with p-chloromercuribenzoate

Sulfhydryl groups of proteins react with a variety of metals to form mercaptides. However, p-chloromercuribenzoate (PCMB) is thus far the only mercurial giving rise to an analytical adequate spectral change due to mercaptide formation. PCMB exhibits an absorption maximum at 233 nm ($\varepsilon = 1.69 \times 10^4$). When PCMB reacts with a SH group a maximal difference of absorption between PCMB and the mercaptide occurs in the region 250 - 255 nm. This increase is proportional to the amount of mercaptide formed (Benesch & Benesch, 1962).

Approximately 100 nmoles of protein were dissolved in 0.3 ml 6 M urea in a stoppered 1 x 1 cm quartz cuvette. To this 2.7 ml 0.05 M sodium phosphate buffer pH 7 was added. The blank consisted of the same solutions except protein. A 1.75 mM PCMB solution was prepared by dissolving p-chloromercuribenzoic acid in a sufficient amount of 0.1 N NaOH to bring the solution just above pH 7 (Boyer, 1954).

The molarity calculated from the amount taken and the absorbancy of the solution at 233 nm was in very close agreement. Ten µl aliquots of the PCMB reagent were added to both cuvettes and, after mixing, the difference in absorbancy of the blank and protein solution measured. This was repeated until no further change in absorbance could be detected (Fig. 1.10).

4.3.5 Performic acid oxidation

4.3.5.1 Removal of halide ions

In a typical experiment 100 mg histone F3 dimer was dissolved in 3 ml 0.1 M acetic acid and passed over a column

containing 20 ml Dowex-2 (x 10) resin which had previously been washed and converted to the acetate form. The protein was eluted with 0.1 M acetic acid. The eluent containing the protein was freeze dried (Hirs, 1967).

4.3.5.2 Performic acid oxidation

Performic acid was prepared by mixing 19 ml formic acid and 1 ml 35% (v/v) H_2O_2 and allowing it to stand for 2 hours. The dehalogenated protein was dissolved in a mixture of 3 ml 98% (v/v) formic acid and 0.5 ml methanol and cooled to $-8^{\circ}C$. Performic acid was added in 0.2 ml aliquots at intervals until a total of 0.6 ml had been added. The oxidation was allowed to proceed for 2 hours (Hirs, 1967). The protein solution was diluted in several volumes of cold H_2O, frozen in liquid N_2 and freeze dried.

4.3.6 N-terminal group determination

4.3.6.1 Dansylation of proteins and peptides

The method used for labelling N-terminal amino acids of proteins and peptides is that of Gros and Labouesse (1969).

Proteins : Protein (1 - 10 nmoles) was dissolved in 500 µl 0.4 M disodium phosphate buffer pH 8.2, 250 µl dimethylformamide and 100 µl 0.2 M dansyl chloride (1-dimethylaminonaphththalene-5-sulfonyl chloride) in acetonitrile was added. After an hour the protein was precipitated by adding trichloroacetic acid (TCA) to a final concentration of 10% (w/v). The residue after centrifugation was washed with 1 N HCl.

Peptides : To 1 - 10 nmoles of peptide in 10 µl H_2O and 20 µl 50 mM bicarbonate buffer (pH 8.3) 30 µl 10 mM dansyl chloride in acetone was added. The coupling was allowed to proceed for an hour. Excess dansyl chloride

was hydrolysed by adding 10 µl 0.1 M NaOH (change from a yellow to colourless solution). The mixture was lypholized. This method can also be used for proteins.

4.3.6.2 Hydrolysis of dansyl proteins and peptides :

To the solid residue in a hydrolysis tube 0.5 ml of 5.7 N HCl was added, the tube evacuated and sealed (see amino acid analysis 4.3.2). After hydrolysis at 110°C for 4 hours the HCl was evaporated in a desiccator filled with solid NaOH.

4.3.6.3 Identification of dansylated amino acids :

Polyamide layer chromatography was chosen for the identification of the DNS-amino acids (dansyl) since all common amino acids are readily separated and up to 10^{-11} mole per derivative can be detected. Since the particular type of polyamide used by Woods and Wang (1967) was not obtainable, Merck polyamide 11 F_{254} on aluminium sheets were used. The original solvent systems had to be modified.

A. Benzene : acetic acid 9:1 (v/v)
B. Methanol: H_2O: acetic acid: formic acid 50:50:1.5:1.5 (v/v)
C. Butanol : acetic acid 9:1 (v/v)

Although the dry residue obtained after the hydrolysis of the DNS-peptide or protein can be directly dissolved in acetone-acetic acid and applied to the polyamide sheet, better separations were achieved by fractionating the DNS-amino acid into two groups.

The dry residue after hydrolysis was dissolved in 0.5 ml 0.5 mM HCl followed by two 1 ml ethyl acetate extractions. The organic layer contained all DNS-amino acids except His, Arg, $CySO_3H$, DNS-O-Tyr and DNS-ε-Lys together with DNS-OH and free amino acids and peptides which remain in the aqueous layer. The ethyl acetate and the aqueous layers were evaporated under a stream of N_2 and dissolved in

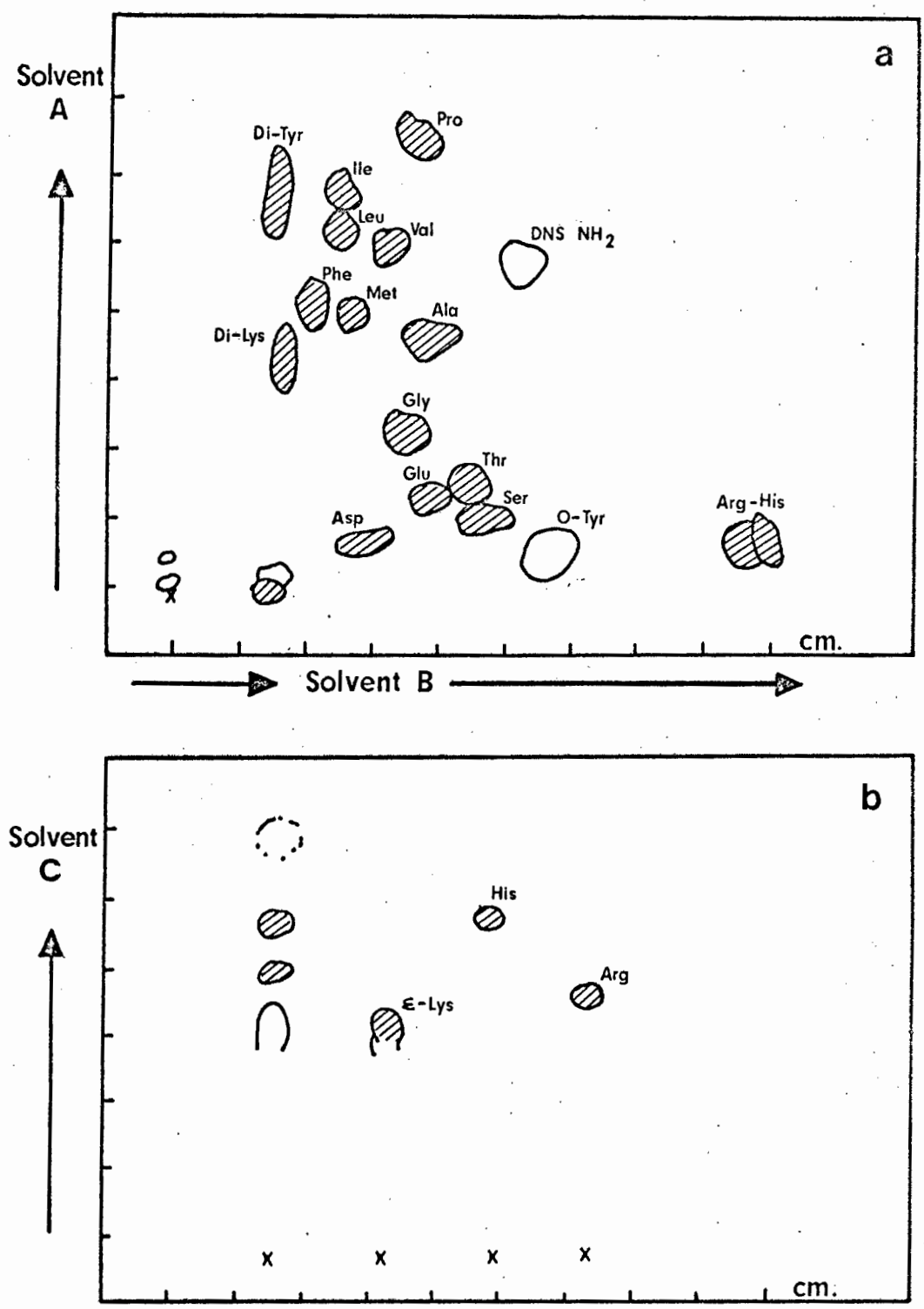


Fig. 4.3 : a) Two dimensional thin layer chromatograms of a mixture of DNS-amino acids. The derivatives were separated on a polyamide layer using solvent systems A and B.

b) Separation of the basic DNS-derivatives on polyamide layer using solvent C in a single dimension.

50 µl acetone-acetic acid (3:2 v/v).

Samples and standard solutions of DNS-amino acids were applied as a spot not larger than 3 mm in the corner of a 10 x 10 cm polyamide sheet. The chromatogram was developed in solvent system A followed by drying and then redeveloped in solvent system B in the other direction (Fig. 4.3a). Separation could be improved by multiple application of the solvent systems. DNS-ε-Lys, DNS-Arg, DNS-His and DNS-$CySO_3H$ were separated by developing the polyamide layer at least twice in solvent system C (Fig. 4.3b).

DNS-amino acids were detected under ultraviolet light where they appeared as greenish spots on the polyamide layer.

4.3.7 C-terminal group determination by hydrazinolysis

The method used is that of Nui and Fraenkel-Conrat (1955) adapted to histones by Phillips and Simson (1969). About 1 mg of protein was dissolved in 0.5 ml anhydrous hydrazine and the tube frozen in liquid N_2, evacuated and sealed. After heating the tube for 5 hours at 110°C the hydrazine was evaporated in a desiccator. The residue was dissolved in 1 ml 0.1 M HCl followed by three extractions with 1 ml benzaldehyde. The layers were separated by centrifugation. The lower yellow layer containing amino acid hydrazides was discarded. The aqueous layer was washed three times with benzene and then freeze dried. The residue was dissolved in 40 µl 0.05 M bicarbonate buffer followed by dansylation and identification of the DNS-amino acids as described under N-terminal group determination (4.3.6).

4.3.8 Determination of the acetyl content in histone F3

Ward and Coffey (1964) described a quantitative procedure for the determination of acetyl groups in proteins. Ethyl acetate had, however, to be substituted by ether in the extraction of acetic acid since the ethyl acetate contained appreciable amounts of acetic acid. Since the extraction of acetic acid from the protein hydrolysate with ether is

incomplete the partition of acetic acid was determined experimentally using radioactive acetic acid.

Two samples of 10 mg F3 histone were hydrolysed for 16 hours in 5.7 M HCl at 110°C in an evacuated and sealed tube. After hydrolysis the tube was cooled in ice, opened and immediately adjusted to pH 3 with saturated NaOH. The total volume was 2 ml. The hydrolysate was extracted with 0.5 ml ether and dried over anhydrous Na_2SO_4.

The extraction was repeated at 0.1 ml of C^{14} acetate (specific activity 40 mCi/mM), had been added. The amount of radioactive acetate in 0.1 ml of the aqueous phase and all of the ether phase was determined by scintillation counting. The cocktail consisted of 0.5 ml ether, 0.1 ml H_2O or aqueous phase, 0.1 ml biosolve and 10 ml toluene-PPO. The total counts, due to the radioactive acetate in the ether extract, were 40,000 counts/min while 360,000 counts/min remained in the aqueous layer indicating that only 10% of the acetic acid had been extracted.

Acetic acid was identified in the ether extract on a 4 mm x 4 ft glass column packed with 10% LAC-446 diatoport S. The column was conditioned at 200°C for 18 hours and finally operated at 110°C using a flow rate of 50 ml N_2/minute (Fig. 1.8).

The area under the peak, resulting from 43 nmoles acetic acid at the 1 x 10^{-9} amp range in the gas chromatographic identification was 5.5. The area under the acetic acid peaks corresponding to 1.6% of the ether extract at the 3 x 10^{-11} amp range was 1.9 and 1.54.

Therefore, 288 and 227 nmoles of acetic acid have been initially present in the two hydrolysates.

Since 10 mg F3 histone (590 nmoles) had been hydrolysed, one mole of histone F3 contains approximately 0.4 moles of acetyl groups.

4.3.9 Enzymatic deacetylation of histone F3

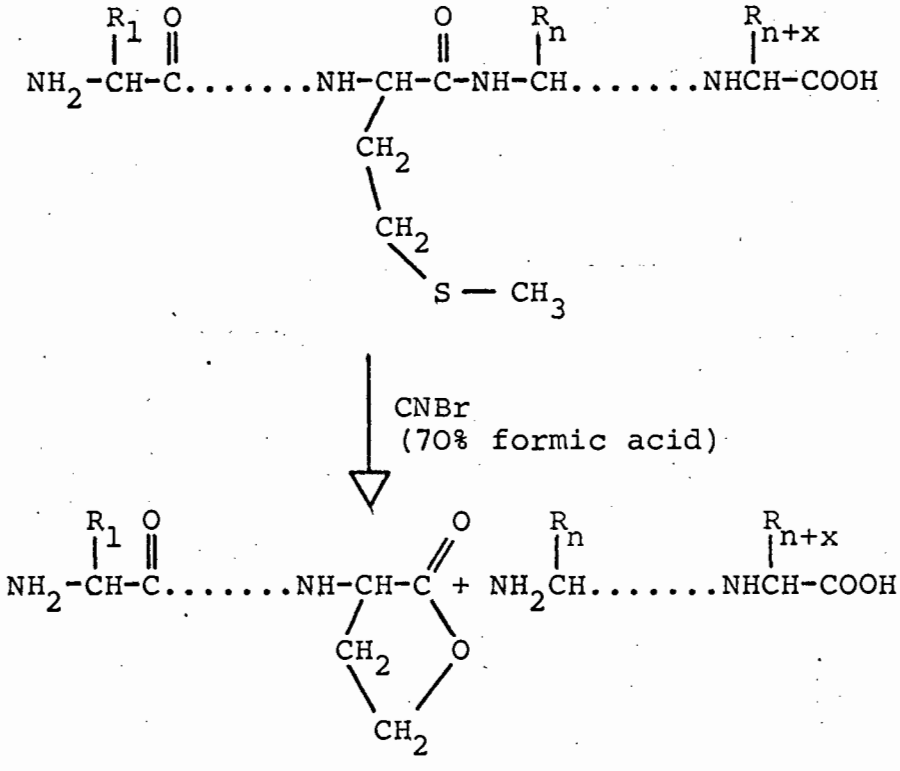
Inoue et al. (1969) reported that a 0.15 M NaCl extract of calf thymus nuclei possessed histone deacetylation actively. Hasler (1970) in this laboratory succeeded in

partially purifying the enzyme(s) by gel filtration. An unidentified dialysable co-factor was required for the deacetylation. To deacetylate histone F3 it was dissolved in 0.025 M pH 7.2 Tris-HCl buffer (1 mg/ml). To 0.3 ml aliquots of this protein solution 0.1 ml co-factor and 0.2 ml partially purified histone deacetylase were added. A white precipitate formed. The mixture was incubated for various times at 37°C with shaking. The reaction was stopped by adding 250 mg solid urea after which the solution became clear again. These solutions (10 - 20 µl) were layered on polyacrylamide gel and electrophoresed (4.3.1) (Fig. 1.9).

4.4 CHEMICAL FRAGMENTATION OF HISTONE F3

4.4.1 Cyanogen Bromide cleavage at methionine

Gross and Witkop (1961) have shown that cyanogen bromide (CNBr) reacts with methionine under mild conditions and that the reagent may be applied to nonenzymatic cleavage of methionyl peptide bonds. The conditions for cleavage are mild and specific and yields are generally high making this the most useful chemical cleavage known to date (Gross, 1967).



Except for cysteine, which is slowly oxidized to cysteic acid, no other amino acid undergoes modification. After the cleavage all peptides, except the C-terminal of the uncleaved protein or polypeptide, have a homoserine or its lactone as C-terminal group. The homoserine lactone is easily detected during amino acid analysis since it is eluted just after the NH_3 peak (Gross, 1967).

In a typical experiment 200 mg F3 dimer were dissolved in 30 ml 70% (v/v) formic acid to which 200 mg CNBr was added with stirring. The solution was kept under N_2 at room temperature (20°C). After 4 and 24 hours 100 mg of CNBr was again added. The solution was diluted after 30 hours with 1 volume of H_2O, frozen in liquid N_2 and freeze dried.

The dried material was dissolved in 6 M urea at pH 3 and applied in two separate runs to a Sephadex G-100 2.5 x 100 cm column (Fig. 2.1).

4.4.2 N-Bromosuccinimide cleavage

Peptides and proteins containing unsaturated amino acids with the double bond in the γ,δ position relative to the carboxamide group can be cleaved with N-bromosuccinimide (Witkop, 1961; Spande et al., 1970).

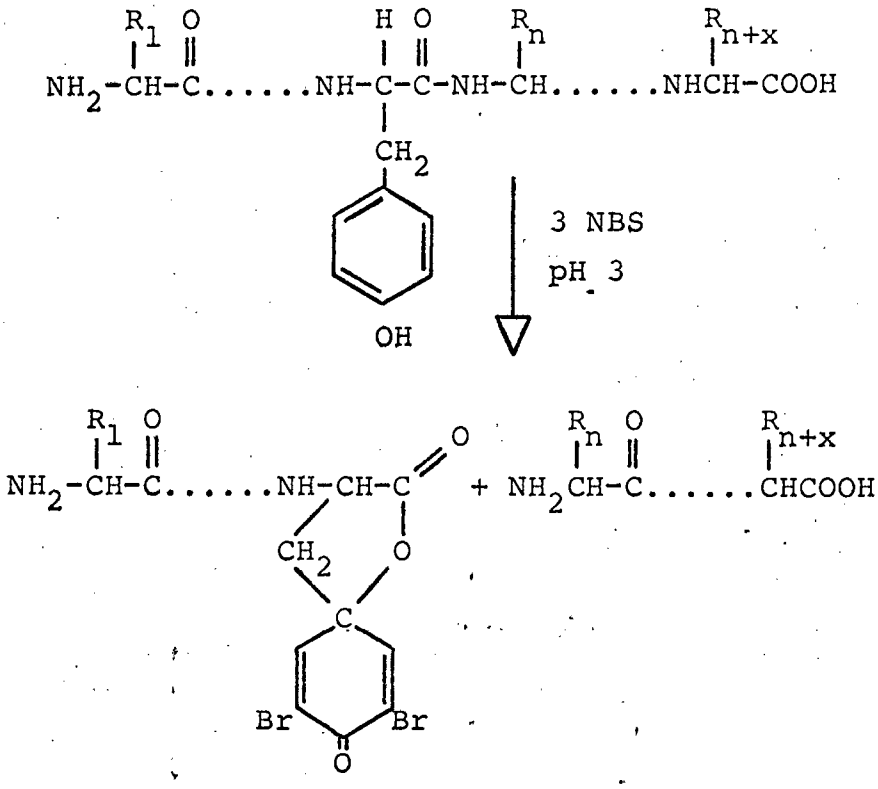
4.4.2.1 Cleavage at tyrosine residues

Between pH 0 - 5 tyrosine reacts with 3 moles of N-bromosuccinimide (NBS) resulting in the cleavage of the tyrosyl-peptide bond, formation of dibromodienone spirolactone and the release of a peptide with a new N-terminal group.

The reaction may be followed spectrophotometrically at 260 nm at which the dienone spirolactone shows characteristic absorption (ε = 10,000 - 11,000).

The tyrosyl-peptide bond will be cleaved selectively if tryptophan is absent. Histidine reacts slower with NBS and requires more drastic conditions for the cleavage. Yields of tyrosyl-peptide bonds cleaved in model peptides vary from 30% to 60%.

The proposed reaction is (Spande et al., 1970) :



The cleavage was performed in the form of a spectrophotometric titration instead of adding an experimentally determined or calculated amount of NBS (Wilson & Cohen, 1963; Ramachandran & Witkop, 1967).

The oxidation was also performed on a standard amino acid mixture. An amino acid analysis was performed and it was found that methionine was oxidized to the sulphone while Cys, Tyr and His disappeared completely from the chromatogram. However, the presence of the spirolactone could be revealed by the ultraviolet spectrum of the solution while the brominated histidyl residues still gave a positive Pauly reaction (Easley, 1965).

In a typical experiment 67 mg fragment CN-1 were dissolved in 30 ml 50% (v/v) acetic acid. NBS freshly recrystallised from 90% (v/v) acetic acid was dissolved in 50% acetic acid (20 mg/ml) and added in 100 µl aliquots to the peptide solution at 20°C with stirring. The ultraviolet spectrum was recorded at intervals and further additions made when no further change occurred during 5 minutes. Spectra

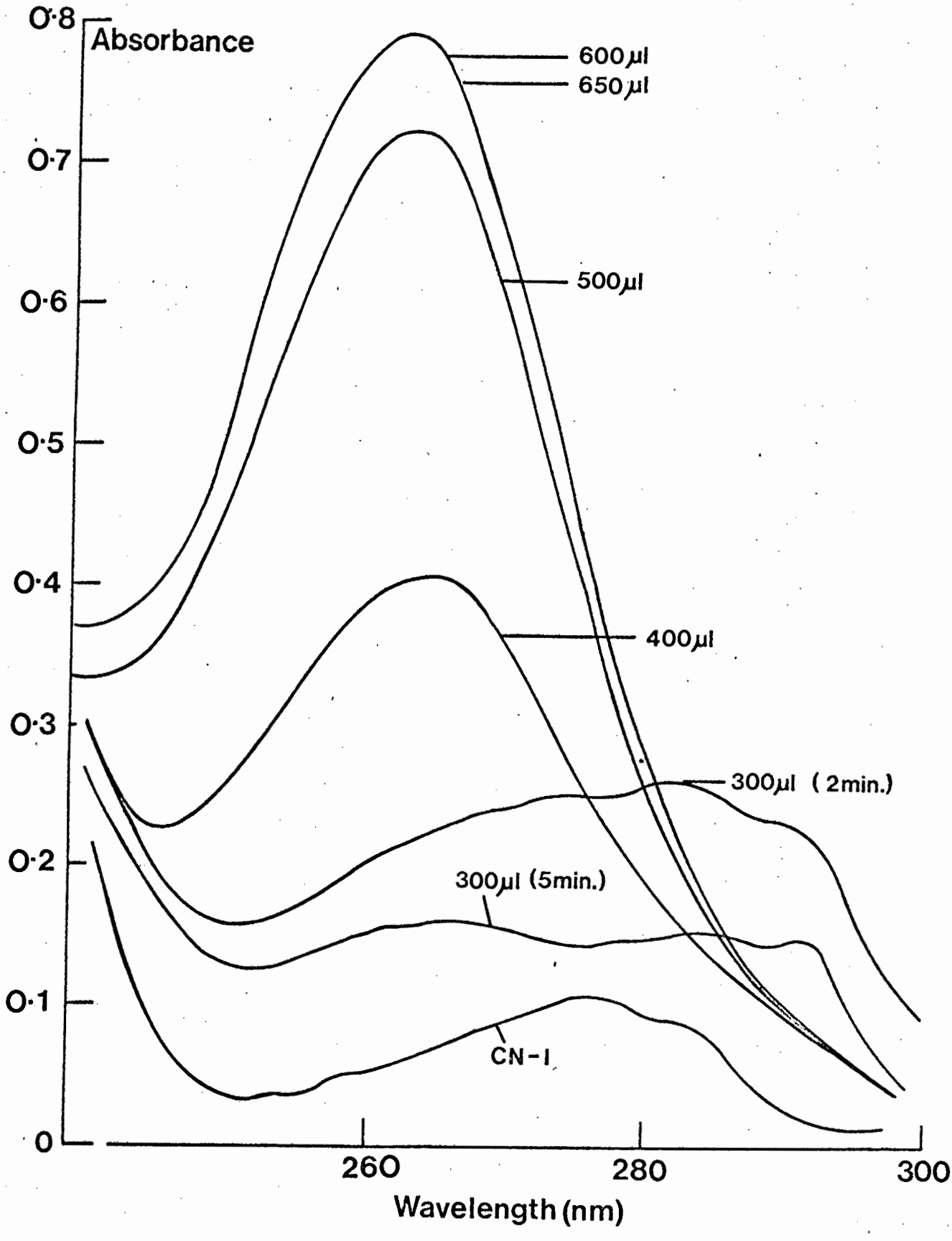


Fig. 4.4 : Typical ultraviolet spectra of a tyrosine containing peptide to which progressively more NBS had been added. In this case fragment CN-1 in 50% acetic acid (6 7 mg/30 ml) was titrated with 0.11 M NBS. The calculated cleavage yield is over 90%. The optical pathlength of the difference spectra was 2 mm.

were obtained in quartz cells of 0.2 cm path length and using as reference 50% (v/v) acetic acid that had been treated in the same way as the sample. Near the end point the solution turned yellow which faded initially, but on further additions, persisted. Addition of NBS was stopped when no further increase in the 260 nm absorption could be observed (Fig. 4.4). The solution was then diluted with an equal amount of H_2O, frozen in liquid N_2 and freeze dried.

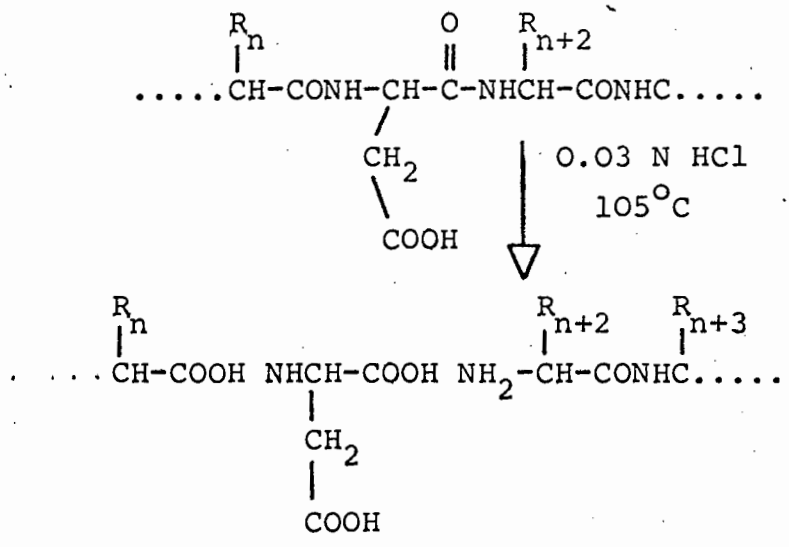
Performic acid oxidized histone F3 and fragment CN-2 were cleaved in a similar fashion.

4.4.2.2 Cleavage at histidine residues

Approximately 1 mg of fragment CN-2 NB-1 was dissolved in pyridine-acetic acid-H_2O = 1:10:19 (v/v). Assuming a molecular weight of 2,400 for this peptide, 3 moles NBS per mole of peptide were added and left standing for 1 hour. Excess NBS was destroyed by adding excess solid imidazole. The tube was sealed and heated in a boiling water bath for 1 hour (Shaltiel & Patchornik, 1963). The peptide solution was lyophilized and redissolved in 0.01 N HCl and applied to Sephadex G-50 1.5 x 90 cm column (2.3.3).

4.4.3 Cleavage of fragment CN-1 NB1(1) at aspartic acid with dilute HCl

Partridge and Davis (1950) observed that refluxing a protein with dilute acetic acid resulted in the almost selective release of free aspartic acid to the exclusion of other free amino acids. It was established that when acids are used at a concentration just sufficient to maintain the β-carboxyl group of aspartic acid residues in a peptide in the undissociated state the aspartic acid peptide bonds are cleaved at a rate at least 100 times greater than any other peptide bond (Schultz, 1967).



It has been postulated that the β-carboxyl group of Asp takes part in a trans-peptidation reaction (Witkop, 1961). asparagine is hydrolysed slightly slower since the amide group has first to be hydrolysed.

The optimum cleavage time was established by dissolving 1.5 mg CN-1 NB-1(1) in 4 ml 0.03 N HCl. One ml aliquots were placed in hydrolysis tubes and one of them was adjusted to a 5.7 N HCl concentration. All tubes were evacuated and sealed as described (4.3.2). Hydrolysis was performed for various times at 105°C. At the end of hydrolysis the content of the tubes was dried (4.3.2) and subjected to amino acid analysis (4.5.3.2). The difference in the amount of aspartic acid between the tubes reflects the degree of cleavage (Fig. 2.1.3).

4.4.4 Ion exchange chromatography on CMC-cellulose

CMC-cellulose (Sigma, 0.7 meq./mg) was allowed to soak in H_2O-EDTA. The CMC-cellulose was retained in a Buchner funnel on a fine meshed nylon gauze and washed successively with 0.5 N HCl, H_2O, 0.5 N NaOH, H_2O, 0.5 N HCl H_2O, 0.5 N NaOH, H_2O and 0.05 M acetate buffer pH 4.4. The water washes were continued each time until neutrality. Finally the cellulose was suspended in the same acetate buffer and packed in 1.5 cm column to 10 cm bed height. The column was fitted with a flow adaptor and the flow rate

controlled with the aid of peristaltic pump. The column was flushed with several volumes of 0.05 M acetate buffer pH 4.4. Peptide mixtures (25 mg per run) were dissolved in the acetate buffer containing 3 M urea and applied to the column. After about two volumes of the starting buffer had been pumped through the column peptides were eluted using a linear gradient of NaCl.

The gradient former contained 150 ml 0.05 M acetate buffer pH 4.4 and the second vessel 150 ml 0.05 M acetate - 0.7 M NaCl pH 4.2. Finally 0.2 N HCl was pumped through the column.

The peptide concentration in the fractions was monitored at 236 nm. Selected fractions were pooled and freeze dried.

Peptides were desalted on a 2.5 x 30 cm Sephadex G-25 column using 0.01 N HCl as eluent. Peptide containing fractions were again pooled and freeze dried. The dry peptides were kept at -20°C.

4.5 SEQUENCE ANALYSIS

4.5.1 Manual Edman degradation

The Edman degradation used is a modification of Fraenkel-Conrat's filter strip method (1954).

Approximately 0.5 µmoles of peptide or protein in H_2O was applied to two 1 x 7 cm Whatman No. 1 filter paper strips. The dried strips were wetted with 0.2 ml 20% (w/v) phenylisothiocyanate in dioxane, suspended on a rack and placed in a 300 ml screw-capped jar containing 20 ml of a mixture of equal volumes of pyridine, dioxane and H_2O. The tightly closed jar was heated at 40°C for 3 hours. The strips were briefly dried and extracted three times in benzene for a total of 2 hours. The dry strips were placed in a desiccator containing one beaker each of glacial acetic acid and 6 N HCl. The pressure was reduced to 100 torr and the cleavage allowed to proceed for 7 hours. The strips were dried in another desiccator and the PTH-amino acids

extracted with acetone. After drying the strips they were ready for the next degradation.

Several samples were degraded simultaneously.

Identification of PTH-amino acids : The acetone containing the PTH-amino acids was evaporated on a rotary evaporator and the residue redissolved in 50 µl acetone.

PTH-amino acids were separated and identified on silica thin layers (Eastman silica gel (F254) on plastic sheets) by consecutively applying solvents systems A, B and C. The amino acids derivatives were detected in ultraviolet light.

Solvent systems (Brenner et al., 1965) :

A. Chloroform
B. Chloroform : methanol 9:1 (v/v)
C. Heptane : ethylene chloride : formic acid: propionic acid 90:30:21:18 (v/v), upper phase

4.5.2 Automatic Edman degradation

4.5.2.1 The Protein-Peptide Sequencer

After Edman's (1950) discovery of the sequential degradation of proteins and peptides with phenylisothiocyanate, Edman and others contributed to the better understanding of the basic chemistry involved (Edman, 1957, 1960; Sjöquist et al., 1960; Blombäck et al., 1966). It was found that separation of the cleavage from the conversion reaction, avoidance of oxygen and careful purification of the reagents improved the efficiency of the degradation dramatically.

This degradation procedure has been automated by Edman and Begg (1967) who described the design and operation of a sequenator that was able to sequentially remove the first 60 amino acid residues from myoglobin.

The operation of the instrument is based on the principle that solutions of reactants and extracting solvents are spread in a thin film inside a rotating cylindrical glass cup. The large surface area of the spinning film is ideally

suited for rapid reaction, drying and extraction manoeuvres. Extractions are carried out by allowing an immiscible solvent to slide over the surface of the protein film to the top of the cup where the liquid is scooped off and collected.

The sequencer performs only the 'coupling' and 'cleavage' operation of the Edman degradation. The 'conversion' of the extracted amino acid thiazolinones to the phenylthiohydantoins is carried out by heating the dried fraction in aqueous acid (Ilse & Edman, 1963) (4.5.2.4). The amino acid derivatives have then to be identified by one or more of the chromatographic methods.

The design of the sequencer required modification of certain reagents and solvents as compared to the manual procedure. The relative large volume of the bell jar necessitates the replacement of more volatile reagents with reagents of lower vapour pressure in order to reduce losses due to evaporation. Quadrol (N,N,N'N'-tetrakis (2-hydroxypropyl)-ethylenediamine) and n-heptafluorobutyric acid have been substituted for N-dimethyllylamine and trifluoroacetic acid respectively used in the manual degradation (Edman & Begg, 1967).

In the Beckman Model 890 sequencer used in our laboratory, evaporation is further minimized by a space reducer plug introduced into the glass cup.

4.5.2.2 The automatic degradation cycle

4.5.2.2.1 The Beckman Model 890 sequencer : The Beckman protein-peptide sequencer controls have both automatic and manual modes - the manual modes being used during development of the program. When the automatic modes are in use, the instrument is under command of a 42-channel programmer that utilizes punched mylar-backed paper tape.

After the optimum conditions for sequencing a given sample have been established, the various steps for each cycle are punched into the tape which is inserted into the programmer. Sequencing then proceeds automatically. There is no necessity of reprogramming at the end of a cycle since the tape is circular and is thus repeated indefinitely. Table 4.1 shows the programmed equivalents of the chemical and mechanical operation involved in the Edman degradation.

4.5.2.2.2 Sample application : Proteins and peptides are applied in a suitable solvent to the sequencer cup. Throughout this investigation a known amount of protein or peptide (uncorrected for moisture or salt content) was dissolved in H_2O and transferred to the rotating sequencer cup (1500 rev./min). The amount of H_2O used was just sufficient to cover the lower half of the spinning cup. The aqueous layer was then freeze dried by applying restricted, rough and fine vacuum in fairly rapid succession, resulting in the formation of an even, white layer of protein or peptide.

4.5.2.2.3 The protein program : The main features in the program described by Edman and Begg (1967) are the use of the non-volatile buffer Quadrol and a 'double cleavage' with heptafluorobutyric acid. This was introduced because Edman (1970) established that the cleavage reaction is an equilibrium rather than a quantitative reaction.

It was found in this laboratory that Quadrol contaminated the extract containing the amino acid thiazolinones contributing substantially to the general background in the gas chromatographic identification.

Secondly, it is known that Quadrol is difficult to purify and still contains aldehydes after extensive purification (Edman & Begg, 1967).

The protein program was thus modified by replacing the Quadrol buffer with 3-dimethylamino-1-propyne buffer. This buffer is easily purified (Braunitzer & Schrank, 1970) and found to contain no aldehydes (4.5.2.3). To compensate for the higher volatility of this buffer a few small extra additions of the buffer were made during the coupling reaction.

After the 'coupling' step most of the buffer could be volatilized in the vacuum steps. Therefore, the time for the ethyl acetate extraction could be reduced. The complete program is given in Table 4.2.

TABLE 4.1

The Beckman model 890 sequencer program statements and codes.

Beckman®

SEQUENCER

PROGRAM _____

DATE _____

NAME _____

TEMP. 55°C SETTING 700

REAGENTS

	TYPE	PRESS.
N_2 Cell		40
R_1	PITC/heptane	110
R_2	HFBA	180
R_3	–	160
R_4	DMAP-buffer	170
R_5		
S_1	benzene	134
S_2	ethyl acetate	
S_3	butyl chloride	

PROGRAM STATEMENTS		STEP TIME	SPEED
Stop slew	Mechanical or		
Blank	chemical equivalent		
Delay			
N_2 Cell			
R or S vent.	deliver chemicals,		
R or S press.	wash or extract		
R or S deliver	sample		
S deliver and collect			
S deliver and waste			
Reaction	coupling or cleavage		
Vac restricted			
N_2 cell + waste	dry protein or		
Vac. rough	peptide film		
Vac. fine			
Fraction collector			
(F/C)vent.			
Fraction collector vac. - dry sample			
Fraction collector step			
High speed (1500 r.p.m.)			H
Low speed (1000 r.p.m.)			L
Start slew			
Conditional stop			

PROGRAM STEP

PROGRAM CONTROL: START SLEW, STOP SLEW, PROGRAM

DRIVE SPEED: HIGH

STEP TIME IN SECONDS: HUNDREDS (800, 400, 200, 100), TENS (80, 40, 20, 10), UNITS (8, 4, 2, 1)

REAGENT TRANSFER: REAGENT 1, REAGENT 2, REAGENT 3, REAGENT 4, REAGENT 6

SOLVENT TRANSFER: SOLVENT 1, SOLVENT 2, SOLVENT 3

REACTION CELL: PRESSURIZE, RESTRICTED VAC, ROUGH VACUUM, FINE VACUUM

FRACTION COLLECTOR: NITROGEN, VACUUM, STEP

EFFLUENT VALVE: WASTE, COLLECT

PRESSURIZE: R & S BOTTLES

PROGRAM STAGE: STAGE

TABLE 4.2

PROTEIN PROGRAM

The conditions and chemicals are listed in Table 4.1

Step	Program Statement	Step Time in sec	Cup Speed	Step	Program Statement	Step Time in sec	Cup Speed
1	Stop slew	2	L	20	Reaction	300	L
2	Delay	6	L	21	R5 deliver	120	L
3	Blank	2	L	22	Blank	2	L
4	R4 vent.	14	L	23	Vac. restricted	60	L
5	R4 press.	14	L	24	Delay	4	L
6	R1 vent.	14	H	25	N_2 cell + waste	300	L
7	R1 press.	14	H	26	Vac. restricted	100	H
8	R1 deliver	6	H	27	Vac. rough	200	H
9	Blank	2	H	28	Vac. fine	400	H
10	Vac. restricted	30	H	29	Blank	6	H
11	Delay	6	H	30	S1 vent.	30	H
12	N_2 cell	60	H	31	S1 press.	30	H
13	R4 deliver	28*	H	32	S1 deliver + waste	200	H
14	Reaction	300	H	33	N_2 cell + waste	60	H
15	R4 deliver	4	L	34	Vac. restricted	30	H
16	Reaction	300	L	35	Vac. rough	140	H
17	R4 deliver	4	L	36	Delay	3	H
18	Reaction	300	L	37	S2 vent.	30	H
19	R5 deliver	10	L	38	S2 press.	30	H

TABLE 4.2 cont'd....

Step	Program Statement	Step Time in sec	Cup Speed	Step	Program Statement	Step Time in sec	Cup Speed
39	S2 deliver + waste	300	H	59	Vac. fine	60	H
40	N_2 cell + waste	82	H	60	Delay	4	H
41	Vac. restricted	60	H	61	R3 vent.	14	L
42	Vac. rough	40	H	62	R3 press.	14	L
43	Vac. fine + F/C step	300	H	63	R3 deliver	24*	H
44	Delay	4	L	64	Reaction	120	H
45	R3 vent. + F/C vent.	20	L	65	Vac. restricted	60	L
46	R3 press.	14	L	66	Vac. rough	60	L
47	R3 deliver	24*	H	67	Vac. fine	140	L
48	Reaction	120	H	68	Delay	4	L
49	Vac. restricted	60	L	69	S3 vent.	30	L
50	Vac. rough	60	L	70	S3 press.	30	L
51	Vac. fine	140	L	71	S3 deliver + waste	160	H
52	Delay	4	L	72	Delay	40	H
53	S3 vent.	30	L	73	Vac. restricted	60	H
54	S3 press.	30	L	74	Vac. rough	40	H
55	S3 deliver + collect	200	H	75	Vac. fine	600	H
56	N_2 cell + collect	40	H	76	Delay	4	L
57	Vac. restricted	60	H	77	Start slew	2	L
58	Vac. rough	60	H	78	Conditional stop	-	L

* These values vary from sequencer to sequencer
The cup speed L and H correspond to 1000 and 1500 r.p.m. respectively

4.5.2.2.4 The peptide program : Although the instrument design and the use of non-volatile reagent system devised by Edman and Begg (1967) work excellently for proteins, problems appear when attempts are made to degrade peptides. The prolonged solvent extractions necessary to remove the Quadrol after the 'coupling' step cause large losses of peptide material from the cup. Shorter and more hydrophobic peptides are almost completely lost during the ethyl acetate extraction.

This problem has largely been solved by replacing Quadrol with volatile N-dimethylallylamine buffer (Niall et al., 1969). Since most of the buffer can be removed by evaporation prolonged solvent extractions are unnecessary, minimizing peptide losses. Finally, a single 'cleavage' with heptafluorobutyric acid is used to reduce the number of extractions.

In the peptide program used N-dimethylallylamine was replaced with 3-dimethylamino-1-propyne buffer (Braunitzer & Schrank, 1970). To insure that the peptide was properly insolubilized a precipitation step before every extraction manoeuvre was included in the program. This consisted of just covering the peptide film with solvent followed by a drying cycle. These precipitation steps were performed before the benzene wash and the butyl chloride extraction, resulting in a decrease of losses at these steps. The complete program is given in table 4.3.

4.5.2.3 Reagents and solvents

The following chemicals used in the automatic sequence analysis were purchased from Beckman Instruments (sequencer grade) :

Phenylisothiocyanate
Heptafluorobutyric acid
Dimethylallylamine buffer (N-allyl-N,N-dimethylamine-trifluoroacetic acid in pyridine-H_2O)
Quadrol-trifluoroacetic acid buffer
Heptane
1-chlorobutane
Benzene and ethyl acetate

TABLE 4.3

PEPTIDE PROGRAM

The conditions and chemicals are listed in Table 4.1

Step	Program Statement	Step Time in sec	Cup Speed	Step	Program Statement	Step Time in sec	Cup Speed
1	Stop slew	2	L	20	R5 deliver	10	L
2	Delay	6	L	21	Reaction	300	L
3	Blank	2	L	22	R5 deliver	120	L
4	R4 vent.	14	L	23	Blank	2	L
5	Delay	2	L	24	Vac. restricted	60	L
6	R4 press.	14	L	25	Delay	6	L
7	R1 vent.	14	H	26	N_2 cell + waste	400	L
8	R1 press.	14	H	27	Vac. restricted	100	H
9	R1 deliver	6	H	28	Vac. rough	200	H
10	Blank	2	H	29	Vac. fine	400	H
11	Vac. restricted	30	H	30	Blank	2	H
12	Delay	6	H	31	Delay + F/C step	6	H
13	N_2 cell + waste	60	H	32	S1 vent.	30	H
14	R4 deliver	26*	H	33	S1 press. + F/C vent	30	H
15	Reaction	300	H	34	S1 deliver + waste	22*	H
16	R4 deliver	4	L	35	Vac. restricted	60	H
17	Reaction	300	L	36	N_2 cell + waste	120	H
18	R4 deliver	4	L	37	Vac. rough	60	H
19	Reaction	300	L	38	Vac. fine	60	H

TABLE 4.3 cont'd.....

Step	Program Statement	Step Time in sec	Cup Speed	Step	Program Statement	Step Time in sec	Cup Speed
39	Delay	6	H	58	Vac. restricted	20	L
40	S1 deliver + collect	120	H	59	Vac. rough	30	L
41	N_2 cell + waste	30	H	60	Delay	6	L
42	Blank	2	H	61	S3 deliver	22*	L
43	Vac. restricted	30	H	62	Vac. restricted	200	L
44	Delay	6	H	63	N_2 cell + waste	60	L
45	N_2 cell + waste	200	H	64	Vac. rough	60	L
46	Vac. restricted	100	H	65	Vac. fine	60	L
47	Vac. rough	300	H	66	Blank	2	L
48	Vac. fine	300	H	67	Delay	6	L
49	Blank	2	H	68	S3 deliver + collect	120	H
50	Delay	6	H	69	N_2 cell + collect	30	H
51	S3 vent.	30	H	70	Blank	2	H
52	S3 press + F/C step	30	H	71	Vac. restricted	60	H
53	R3 vent. + F/C vent.	14	H	72	N_2 cell + waste	200	H
54	R3 press.	14	H	73	Vac. rough	300	H
55	R3 deliver	33*	H	74	Vac. fine	800	H
56	Reaction	80	H	75	Start slew	-	L
57	N_2 cell + waste	40	L	76	Conditional stop	-	L

* These values vary from sequencer to sequencer
The cup speed L and H correspond to 1000 and 1500 r.p.m. respectively.

3-dimethylamino-1-propyne (90% in H_2O) (DMAP) was purchased from Fluka. The product was twice distilled. Propanol and trifluoracetic acid were purified as recommended by Edman and Begg (1967). n-Propanol (analytical grade) was refluxed for 48 hours with powdered Zn (5 g/l) and saturated NaOH (5 ml/l) and then dried over $CaSO_4$. The dried propanol was then fractionated by distillation on a 30 cm Vigreux column. Various fractions were taken but only those used which gave a negative Tollen's reaction.

Trifluoroacetic acid was refluxed for 24 hours with CrO_3, distilled off, dried over $CaSO_4$ and redistilled. The fraction boiling at 119°-120° was kept. All chemicals prepared gave negative Tollen's reaction (Edman & Begg, 1967).

The buffer was prepared from a 1 M DMAP solution in twice distilled H_2O which was adjusted to pH 9 with trifluoroacetic acid and subsequently diluted with propanol in the ratio 3:4 (Braunitzer & Schrank, 1970). The buffer was always prepared fresh just before use.

4.5.2.4 Conversion of amino acid thiazolinones

The conversion of the thiazolinones into the PTH-derivatives was carried out as follows :

To the dried fraction 0.2 ml 1 N HCl containing 1% (v/v) ethanethiol was added, the tube flushed with N_2, stoppered with a silicone stopper and heated at 80°C for 10 minutes in a heating block. If glycine was expected the heating was prolonged to 15 minutes while for Ser and Thr it was reduced. After the solution had cooled it was twice extracted with 1 ml peroxide free ethyl acetate (sequencer grade ethyl acetate to which solid ascorbic acid had been added) with the aid of a vibration mixer followed by centrifugation to separate the two layers. The organic layer was evaporated in a separate tube with a stream of N_2 at 50°C. The residue was dissolved in 50 µl peroxide-free ethyl acetate. Aliquots (2 - 8 µl) of this solution were used for gas chromatographic identification of the PTH-amino acids. The aqueous phase was evaporated in a stream of N_2

also at 50°C and the residue transferred with methanol to a hydrolysis tube. To the same tube 25 µl of the ethyl acetate sample was added and evaporated to dryness. It was found that samples could be kept for months at -20°C without deterioration until they were hydrolysed to the free amino acids.

4.5.3 Identification and quantitation of PTH-amino acids

4.5.3.1 By gas chromatography

4.5.3.1.1 Column preparation : 25 g gas chrom P (100/120) mesh were placed into a 0.5 M $NaCO_3$ solution followed by exhaustively degassing of the suspension. The next day the support was repeatedly suspended in one litre of distilled H_2O, allowed to settle for 2 minutes. The unsettled fines were decanted. This was repeated for at least ten times. The de-fined support was filtered and dried. The support was then suspended in 250 ml conc. HCl, degassed and allowed to stand overnight. After decanting the acid it was 3 times washed with the same volume of conc. HCl. Fines were again removed as described above. The support was washed with acetone, dried at 80°C until free-flowing, transferred to a 5% (v/v) dichlorodimethyl silane - toluene solution, degassed and allowed to stand for 5 minutes. The support was then washed 4 times with toluene followed by several washings with dry methanol. The material was dried at 80°C until it was again free-flowing.

10 g of this silylated support were suspended in 75 ml 10% (w/v) DC 560 in acetone and degassed. The suspension was allowed to stand for 5 minutes followed by filtration through a coarse sintered disc funnel. Excess liquid was removed by gentle suction. The wet support was fluidized in a sintered glass funnel by a warm stream of air.

Glass wool and 2 mm x 4 ft glass columns were cleaned with HCl and silylated as described. The outlet end of the column was plugged with glass wool and the support was packed into the column with the aid of suction and gentle tapping.

4.5.3.1.2 Column conditioning and operation : A Beckman GC-45 and a Packard gas chromatograph, both fitted with a flame ionisation detector, were used.

To condition the column the N_2 carrier gas flow rate was set to 120 ml/min. The oven temperature was kept at 50°C for 30 minutes and then increased to 325°C over a 320 minute period. The final temperature was held for at least 14 hours.

PTH-amino acids were separated using a temperature program consisting of a two minute isothermal period at 165°C followed by a 110°C rise over a 16 minute period (Pisano & Bronzert, 1969).

4.5.3.1.3 Identification and quantitation : Standard solution of PTH-amino acids were prepared by dissolving them in ethyl acetate (5 nmoles/µl). Basic amino acids and acid amide derivatives of Glu and Asp were dissolved in methanol. PTH-amino acids were silylated by evaporating the solvent containing the derivatives (10 µl) in microfuge tubes followed by adding 10 µl of bisilylacetamide. The tube was heated for 2 minutes at 80°C followed by gas chromatographic analysis.

PTH-amino acids were quantitated from the chromatogram by their peak height (Table 4.4).

The gas chromatographic behaviour of the PTH-amino acids can be divided into three groups (Pisano & Bronzert, 1969) (Table 4.4).

Group 1. These are the most volatile derivatives and thus give symmetrical peaks, are easily identified and fairly reproducibly quantitated. Ile is identified after silylation since it racemizes to allo-Ile giving rise to a double peak. Relative amounts of Leu and Ile could, however, not be determined.

Group 2. These derivatives are either labile or not very volatile. The recovery of Ser is very low due to dehydroalanine formation. Thr always gives two characteristic peaks. Generally silylation improves their gas chromatographic behaviour. They cannot be very reproducibly quantitated.

Group 3. These PTH-amino acids have to be silylated. Lys and His were found to give variable results and could often not be detected at all.

Arg cannot be identified by gas chromatography. In some cases it was identified by a spot test using the phenanthrenequinone method of Yamada and Itano (1966).

TABLE 4.4

GAS CHROMATOGRAPHIC BEHAVIOUR OF PTH-AMINO ACIDS (5 nmoles)

Results were obtained using the Beckman GC-45 gas chromatograph and a 10% DC-560 gas chrom P column (2 mm x 2 ft). Peak heights are given in 1/10th inch pen deflection of 5 nmoles at attenuation of 1600.

Group	1		2		3	
Elution time in min	PTH-amino acid	Peak height	PTH-amino acid	Peak height	PTH-amino acid	Peak height
4	Ala	62				
			Ser	16	TMS-$CySO_3H$	30
	Gly	55	Thr	20		
	Val	60				
	Pro	60	ΔThr	10		
	Ile-Leu	75				
					TMS-Asp	70
	Met	62				
			Asn	27	TMS-Glu	74
	Phe	85	(TMS-Asn)	34		
			Gln	10	TMS-Lys	—
			(TMS-Gln)	40		
16	Tyr	60	(TMS-His)	—		

4.5.3.2 By amino acid analysis after hydrolysis

4.5.3.2.1 Hydrolysis : PTH-amino acids can be converted to the free amino acids by acid or base hydrolysis (Van Orden & Carpenter, 1964; (Africa & Carpenter, 1966).

In a preliminary investigation it was found that higher yields and more reproducible results were obtained using acid hydrolysis.

Half the ethyl acetate extract (25 µl) containing the PTH-amino acids derived from an automatic Edman degradation cycle was placed in a hydrolysis tube. The dry residue of the aqueous layer from the same degradation was transferred with methanol (4.5.2.4).

The organic solvents were evaporated and the hydrolysis performed in 0.5 ml 5.7 N HCl containing 1% (v/v) thioglycollic acid for 24 hours at 130^{o}C in evacuated and sealed tubes (4.3.2).

Standard PTH-amino acids were treated in the same way (Table 4.5).

4.5.3.2.2 Identification and quantitation : The sensitivity of a Beckman 116 amino acid analyser was increased by inserting an extended range card, (4-5 mV full scale) into the recorder using 6 mm pathlength cuvettes and an accelerated elution procedure for the amino acids.

Basic amino acids were eluted from a 5.5 cm column of PA-35 resin with 0.35 M citrate buffer pH 5.36. Acidic and neutral amino acids were eluted from a 56 cm column of UR 30 with 0.2 M citrate buffer at pH 3.49 and pH 4.4.

Less than 5 nmoles amino acid per fraction could be detected. The time for a complete anlaysis is 2.5 hours.

Table 4.5 gives the recoveries of amino acid after acid hydrolysis of the PTH-amino acids.

TABLE 4.5

RECOVERIES OF AMINO ACIDS AFTER HYDROLYSIS OF PTH-DERIVATIVES

Hydrolysis was performed in 5.7 N HCl containing 1% (v/v) thioglycollic acid at 130°C for 24 hours

Amino acid	Amount PTH-amino acid hydrolysed	Amount amino acid recovered	Recovery
	nmoles	nmoles	(%)
Lys	80	31	39
His	80	17	20
Arg	80	33	41
Asp	40	35	87
Thr	40	2	5
Ser	80	1	1
Glu	40	37	92
Pro	40	21	53
Gly	80	60	75
Ala	40	34	85
Val	40	26	65
Met	40	35	88
alloIle	0	12	60
Ile	40	12	
Leu	40	24	60
Tyr	40	23	58
Phe	40	26	65

REFERENCES

AFRICA, B. and CARPENTER, F.H. (1966) Biochem. Biophys. Res. Comm. 24, 112.

ALLFREY, V.G. (1971a) In Histones and Nucleohistones, p. 241. Ed. D.M.P. Phillips. Plenum Press, London.

ALLFREY, V.G. (1971b) Ibid. p. 281.

ALLFREY, V.G. (1971c) Ibid. p. 253.

ALLFREY, V.G. (1971d) Ibid. p. 260.

ALLFREY, V.G. (1970) Federation Proc. 29, 1447.

ALLFREY, V.G., FAULKNER, R. and MIRSKY, A.E. (1964) Proc. Natl. Acad. Sci. U.S. 51, 786.

ALLFREY, V.G., LITTAU, V.C. and MIRSKY, A.E. (1963) Proc. Natl. Acad. Sci. U.S. 49, 414.

ALLFREY, V.G. and MIRSKY, A.E. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 247.

BAHR, C.F. (1970) Exptl. Cell Res. 62, 39.

BARTLEY, J.A. and CHALKLEY, R. (1972) J. Biol. Chem. 247, 3647.

BAUER, K. (1971) Die Naturwissenschaften, 7, 364.

BAUER, R.D. and JOHANNSON, R. (1966) Biochim. Biophys. Acta, 119, 418.

BELLAIR, J.M. and MAURITZEN, C.M. (1965) Aust. J. Biol. Sci. 18, 160.

BELLAIR, J.T. and MAURITZEN, C.M. (1967) Biochim. Biophys. Acta, 133, 263.

BENESCH, R. and BENESCH, R.E. (1962) In Methods in Biochem. Anal. 10, 43.

BENSON, A.M., MOWER, H.F. and YASUNOBU, K.T. (1967) Arch. Biochem. Biophys. 121, 563.

BIJVOET, P. (1957) Biochim. Biophys. Acta, 25, 502.

BLACK, J.A. and DIXON, G.H. (1967) Nature, 216, 152.

BLOMBÄCK, B., BLOMBÄCK, M., EDMAN, P. and HESSEL, B. (1966) Biochim. Biophys. Acta, 115, 371.

BONNER, J., DAHMUS, M.E. and FAMBROUGH, D. (1968) Science, 159, 47.

BOUBLIK, M., BRADBURY, E.M. and CRANE-ROBINSON, C. (1970) Eur. J. Biochem. 14, 486.

BOYER, P.D. (1954) J. Am. Chem. Soc. 76, 4331.

BRADBURY, E.M. (1969) Biochem. J. 114, 49P.

BRADBURY, E.M., CRANE-ROBINSON, C., GOLDMAN, H., RATTLE, H.WE. and STEPHENS, R.M. (1967) J. Mol. Biol. 29, 507.

BRADBURY, E.M. and CRANE-ROBINSON, C. (1971a) In Histones and Nucleohistones, p. 96. Ed. D.M.P. Phillips. Plenum Press, London.

BRADBURY, E.M. and CRANE-ROBINSON, C. (1971b) Ibid. p.85.

BRAM, S. and RIS, H. (1971) J. Mol. Biol. 55, 325.

BRANDT, W.F. and VON HOLT, C. (1972) FEBS Letters, 23, 357.

BRAUNITZER, G. and SCHRANK, B. (1970) Hoppe-Seyler's Z.Physiol. Chem. 351, 417.

BRAUNITZER, G., SCHRANK, B. and RUHFUS, A. (1970) Hoppe-Seyler's Z. Physiol. Chem. 351, 1589.

BRAUNITZER, G., SCHRANK, B., RUHFUS, A., PETERSEN, S. and PETERSEN, U. (1971) Hoppe-Seyler's Z. Physiol. Chem. 352, 1730.

BRENNER, M., NIEDERWIESER, A. and PATAKI, G. (1965) In Thin-Layer Chromatography, p. 430. Ed. E. Stahl. Springer Verlag, Berlin.

BUSCH, H. (1965) In Histones and other Nucleoproteins, p. 121. Ed. H. Busch. Academic Press, New York.

BUTLER, J.A.V., JOHNS, E.W. and PHILLIPS, D.M.P. (1968) Progr. Biophys. Mol. Biol. 18, 209.

BUTLER, J.A.V., POWER, D.F. and PALAU, J. (1967) Biochem. J. 102, 539.

CHALKLEY, R. (1970) Private communications.

CLARK, R.J. and FELSEFELD, G. (1971) Nature New Biology, 229, 101.

CRUFT, H.J., MAURITZEN, C.M. and STEDMAN, E. (1957) Nature, 180, 1107.

CRUFT, H.J., MAURITZEN, C.M. and STEDMAN, E. (1958) Proc. Roy. Soc. (London), Ser. B. 149, 21.

DALY, M.M. and MIRSKY, A.E. (1955) J. Gen. Physiol. 38, 405.

DAVISON, P.F. (1957) Biochem. J. 66, 703.

DELANGE, R.J., FAMBROUGH, D.M., SMITH, E.L. and BONNER, J. (1969a) J. Biol. Chem. 244, 5669.

DELANGE, R.J., GLAZER, A.N. and SMITH, E.L. (1969b) J.Biol. Chem. 244, 1385.

DELANGE, R.J., HOOPER, J.A. and SMITH, E.L. (1972) Proc. Nat. Acad. Sci. U.S. 69, 882.

DICKERSON, R.E. (1972) Sci. Amer. 226, 58.

DIXON, G.H. (1966) In Essays in Biochemistry, Vol. 2, p. 147. Ed. P.N. Campbell and G.D. Greville. Academic Press, New York.

DOUNCE, A.L., CHANDA, S.K., ICKOWICS, R., VOLKMAN, D. and TURK, R. (1972) In Gene Transcription in Reproductive Tissue, p. 91. Ed. E. Diczfalusy. Bogtrykkeriet Forum, Copenhagen.

DUNKER, A.K. and RUECKERT, R.R. (1966) J. Biol. Chem. 244, 5074.

DUS, K., SLETTEN, K. and KAMEN, M. (1968) J. Biol. Chem. 243, 5507.

EASLEY, C.W. (1965) Biochim. Biophys. Acta, 107, 386.

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

EDMAN, P. (1957) Proc. Roy. Australian Chem. Inst. 24, 434.

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

EDMAN, P. (1970) In Protein Sequence Determination. p. 211. Ed. S.B. Needleman. Chapman and Hall Limited, London.

EDMAN, P. and BIGG, G. (1967) European J. Biochem. 1, 80.

FAMBROUGH, D.M. and BONNER, J. (1966) Biochemistry, 5,2563.

FAMBROUGH, D.M. and BONNER, J. (1968) J. Biol. Chem. 243 4434.

FAMBROUGH, D.M., FUJIMURA, F and BONNER, J. (1968) Biochemistry, 7, 575.

FEUGHELMAN, M., LANGRIDGE, R. and SEEDS, W.E. (1955) Nature, 175, 834.

FRAENKEL-CONRAT, H. (1954) J. Am. Chem. Soc. 76, 3606.

FRENSTER, J.H. (1965) Nature, 206, 680.

GALL, J.G. and CALLAN, H.G. (1962) Proc. Nat. Acad. Sci. U.S. 48, 562.

GARRET, R.A. (1968) J. Mol. Biol. 38, 249.

GERSHEY, E.L., HASLETT, G.W., VIDALI, G. and ALLFREY, V.G. (1969) J. Biol. Chem. 244, 4871.

GOLDSTONE, A.D. and NEEDLEMAN, S.B. (1970) In Protein Sequence Determination. p. 256. Ed. S.B. Needleman. Chapman and Hall Limited, London.

GRAY, W.R. (1968) Nature, 220, 1005.

GROSS, E. (1967) In Methods in Enzymology, 11, 294.

GROSS, G. and LABOUESSE, B. (1969) Eur. J. Biochem. 7, 463.

GROSS, E. and WITKOP, B. (1961) J. Am. Chem. Soc. 83, 1510.

HARRIS, J.I., FRAENKEL-CONRAT, H. and LEVY, A.L. (1965) In Methods of Biochemical Analysis, 2, 359.

HASLER, J. (1970) Hons. Project.

HELLERMAN, L., CHINARD, F.P. and DEITZ, V.R. (1941) J. Biol. Chem. 140, 57.

HELLERMAN, L., CHINARD, F.P. and DEITZ, V.R. (1943) J. Biol. Chem. 147, 443.

HIRS, C.H.W. (1956) J. Biol. Chem. 219, 611.

HIRS, C.H.W. (1967) In Methods in Enzymology, 11, 199.

HIRS, C.H.W., MOORE, S. and STEIN, W.H. (1960) J. Biol.Chem. 235, 633.

HNILICA, L.S. (1964) Experientia, 20, 13.

HNILICA, L.S. (1967) Progr. Nucleic Acid Res. Molec. Biol. 7, 25.

HNILICA, L.S. and BESS, L.G. (1965) Anal. Biochem. 12, 421.

HNILICA, L.S. and BUSCH, H. (1963) J. Biol. Chem. 238, 918.

HNILICA, L.S., KAPPLER, H.A. and JORDAN, J.J. (1970)Experientia, 26, 353.

HUANG, R.C. and BONNER, S. (1962) Proc. Natl. Acad. Sci.U.S. 48, 1216.

HUANG, R.C., BONNER, J. and MURRAY, K. (1964) J. Mol. Biol. 8, 54.

HUXLEY, H.E. and ZUBAY, G. (1961) J. Biophys. Biochem.Cytol. 11, 273.

ILSE, D. and EDMAN, P. (1963) Australian J. Chem. 16, 411.

INMAN, J.K., HANNON, J.E. and APPELLA, E. (1972) Biochem. Biophys. Res. Commun. 46, 2075.

INOUE, A. and FUJIMOTO, D. (1969) Biochem. Biophys. Res. Comm. 36, 146.

INOUE, A. and FUJIMOTO, D. (1970) Biochim. Biophys. Acta, 220, 307.

IWAI, K. and ANDO, T. (1967) In Methods in Enzymology, 11, 263.

IWAI, K., ISHIKAWA, K. and HAYASHI, H. (1970) Nature, 226, 1056.

JOHNS, E.W. (1964) Biochem. J. 92, 55.

JOHNS, E.W. (1967a) Biochem. J. 104, 78.

JOHNS, E.W. (1967b) Biochem. J. 105, 611.

JOHNS, E.W. (1968) J. Chromatog. 33, 563.

JOHNS, E.W. (1971a) In Histones and Nucleohistones, p. 34. Ed. D.M.P. Phillips. Plenum Press, New York.

JOHNS, E.W. (1971b) Ibid. p. 6.

JOHNS, E.W. and BUTLER, J.A.W. (1964) Nature, 204, 853.

JOHNS, E.W. and FORRESTER, S. (1969) Eur. J. Biochem. 8, 547.

JOHNS, E.W. and FORRESTER, S. (1970) Biochim. Biophys. Acta, 209, 54.

JOHNS, E.W., PHILLIPS, D.M.P., SIMSON, P. and BUTLER, J.A.V. (1960) Biochem. J. 77, 631.

JOHNS, E.W., PHILLIPS, D.M.P., SIMSON, P. and BUTLER, J.A.V. (1961) Biochem. J. 80, 189.

JUKES, T.H. (1966) In Molecules and Evolution. Columbia University Press, New York.

KASPER, C.B. (1970a) In Protein Sequence Determination, p. 137. Ed. S.B. Needleman. Chapman and Hill Limited, London.

KASPER, C.B. (1970b) Ibid. p. 150.

KLEINSMITH, L.J., ALLFREY, V.G. and MIRSKY, A.E. (1966) Proc. Nat. Acad. Sci. U.S. 55, 1182.

KOSSEL, A. (1884) Physiol. Chem. 8, 511.

KURASHINA, Y., OHBA, Y. and MIZUNO, D. (1970) J. Biochemistry, 67, 661.

LEWIN, S. and JACOBSON, R.E. (1967) Biochem. J. 103, 66P.

LI, H.J., CHANG, C and WEISKOPF, M. (1972) Biochem. Biophys. Res. Comm. 47, 883.

LIEW, C.C., HASLETT, G.W. and ALLFREY, V.G. (1970) Nature, 226, 414.

LIGHT, A. (1967) In Methods in Enzymology, 11, 417.

LITTAU, V.C., BURDICK, C.J., ALLFREY, V.G. and MIRSKY, A.E. (1965) Proc. Nat. Acad. Sci. U.S. 54, 1240.

LUCK, J.M. (1964) In The Nucleohistones, p. 3. Ed. J. Bonner and P. Ts'O. Holden-Day Inc., San Francisco, London.

LUCK, J.M., RASMUSSEN, P.S., SATAKE, K. and TSVETIKOV, A.N. (1958) J. Biol. Chem. 233, 1407.

MARUSHIGE, K. and BONNER, J. (1966) J. Mol. Biol. 15, 160.

MARZLUFF, W.F. and McCARTY, K.S. (1972) Biochemistry, 11, 2672.

MAURITZEN, C.M., STARBUCK, W.C., SAROJA, I.S., TAYLOR, C.W. and BUSCH, H. (1967) J. Biol. Chem. 242, 2240.

MATSUBARA, K. and SASAKI, Y. (1969) Biochem. Biophys. Res. Comm. 35, 175.

McLACHLAN, A.D. (1971) J. Mol. Biol. 61, 409.

McLACHLAN, A.D. (1972) Ibid. 64, 417.

MILSTEIN, C. and PINK, J.R.L. (1970) In Progress in Biophysics and Molecular Biology, 21, 209.

MIRSKY, A.E. and RIS, H. (1950) J. Gen. Physiol. 34, 475.

MOORE, S. and STEIN, W.H. (1963) In Methods in Enzymology, 6, 818.

MURRAY, K. (1966) J. Mol. Biol. 15, 409.

MURRAY, K. (1969) J. Mol. Biol. 39, 125.

MURRAY, K. VIDALI, G. and NEELIN, J.M. (1968) Biochem. J. 107, 208.

NEELIN, J.M. and CONNELL, G.E. (1959) Biochim. Biophys.Acta, 31, 539.

NEELIN, J.M. and NEELIN, E.M. (1960) Can. J. Biochem. Physiol. 38, 354.

NEELIN, J.M. and PURKAIYASTHA, R. (1966) Biochim. Biophys. Acta, 127, 468.

NIALL, H.D., PENHASI, H., GILBERT, P., MYERS, R.C., WILLIAMS, F.G. and POTTS, J.T. (1969) Federation Proc. 28, 661.

NUI, C.I. and FRAENKEL-CONRAT, H. (1955) J. Am. Chem. Soc. 77, 5882.

OGAWA, Y., QUAGLIANOTTE, J., JORDAN, J., TAYLOR, C.W., STARBUCK, W.C. and BUSCH, H. (1969) J. Biol. Chem. 244, 4387.

PALLOTTO, D., BERLOWITZ, L. and RODRIQUES, L. (1970) Exp. Cell Res. 60, 474.

PANYIM, S., BILEK, D. and CHALKLEY, R. (1971) J. Biol. Chem. 246, 4206.

PANYIM, S. and CHALKLEY, R. (1969a) Arch. Biochem. Biophys. 130, 337.

PANYIM, S. and CHALKLEY, R. (1969b) Biochem. 8, 3972.

PANYIM, S., CHALKLEY, R., SPIKER, S. and OLIVER, D. (1970) Biochim. Biophys. Acta, 214, 216.

PANYIM, S., JENSEN, R.H. and CHALKLEY, R. (1968) Biochim. Biophys. Acta, 160, 252.

PARISH, G.R. and MARCHALONIS, J.J. (1970) Anal. Biochem. 34, 436.

PARTRIDGE, S.M. and DAVIS, H.F. (1950) Nature, 165, 62.

PAUL, J. and GILMOUR, R.S. (1966) J. Mol. Biol. 16, 242.

PELLING, C. (1964) Chromosoma, 15, 71.

PHILLIPS, D.M.P. (1958) Biochem. J. 68, 35.

PHILLIPS, D.M.P. (1963) Biochem. J. 87, 258.

PHILLIPS, D.M.P. (1965) Biochem. J. 97, 669.

PHILLIPS, D.M.P. (1967) Biochem. J. 105, 46.

PHILLIPS, D.M.P. (1971a) In Histones and Nucleohistones, p. 47. Ed. D.M.P. Phillips. Plenum Press, New York.

PHILLIPS, D.M.P. (1971b) Ibid. p. 67.

PHILLIPS, D.M.P. (1971c) Ibid. p. 78.

PHILLIPS, D.M.P. and JOHNS, E.W. (1959) Biochem. J. 72, 538.

PHILLIPS, D.M.P. and SIMSON, P. (1969) Biochim. Biophys. Acta, 181, 154.

PISANO, J.J. and BRONZERT, T.J. (1969) J. Biol. Chem. 24, 5597.

RAMACHANDRAN, L.K. and WITKOP, B. (1967) In Methods in Enzymology, 11, 283.

RICHARDS, B.M. and PARDON, J.F. (1970) Exptl. Cell Res.62, 184.

SADGOPAL, A. and BONNER, J. (1970a) Biochim. Biophys. Acta, 207, 227.

SADGOPAL, A. and BONNER, J. (1970b) Ibid. 207, 206.

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

SANGER, F. and THOMPSON, E.O.P. (1963) Biochim. Biophys. Acta, 71, 468.

SCHIFFER, M. and EDMUNDSON, A.B. (1967) Biophys. J. 7, 121.

SCHULTZ, J. (1967) In Methods in Enzymology, 11, 255.

SHALTIEL, S. and PARTCHORNIK, J. (1963) J. Am. Chem. Soc. 85, 2799.

SIMPSON, R.T. (1972) Biochemistry, 11, 2003.

SJÖQUIST, J., BLÖMBÄCK, B. and WALLEN, P. (1960) Arkiv Kemi, 16, 425.

SMYTH, D.G. (1967a) In Methods in Enzymology, 11, 214.

SMYTH, D.G. (1967b) Ibid. 11, 421.

SONNENBICHLER, J. (1969) Hoppe-Seyler's Z. Physiol. Chem. 350, 761.

SPACKMAN, D.H., STEIN, W.H. and MOORE, S. (1958) Anal.Chem. 30, 1190.

SCHERAGA, H.A. (1972) Private communications.

SPANDE, T.F., WITKOP, B., DE GANI, Y. and PATCHORNIK, J. (1970) In Advances in Protein Chemistry, p. 97. Ed. C.B. Anfinsen, J.T. Edsall and F.M. Richards. Academic Press, London.

STARBUCK, W.C., MAURITZEN, C.M., TAYLOR, C.W., SAROJA, I.S. and BUSCH, H. (1968) J. Biol. Chem. 243, 2038.

STARK, G.R., STEIN, W.H. and MOORE, S. (1960) J. Biol.Chem. 235, 3177.

STEDMAN, E. and STEDMAN, E. (1943) Nature, 152, 556.

STEDMAN, E. and STEDMAN, E. (1950) Nature, 166, 780.

STELLWAGEN, R.H. and COLE, R.D. (1969) Ann. Rev. Biochem. 38, 951.

STOCKEN, L.A. (1966) In Histones, p. 62. Eds. A.V.S. De Reuck and J. Knight. J.A. Churchill Ltd., London.

SWIFT, H. (1964) In The Nucleohistones, p. 169. Ed. J.Bonner and P.O.P. Ts'O. Holden-Day Inc., San Francisco.

TUAN, D.V.H. and BONNER, J. (1969) J. Mol. Biol. 45, 59.

VAN ORDEN, H. and CARPENTER, F.H. (1964) Biochem. Biophys. Res. Comm. 14, 399.

VIDALI, G. and NEELIN, J.M. (1968) Eur. J. Biochem. 5, 330.

VON WILM, M. (1970) Angew. Chem. internat. Edit. 9, 267.

WALKER, I.O. (1965) J. Mol. Biol. 14, 381.

WARD, N.D. and COFFEY, J.A. (1964) Biochemistry, 3, 1577.

WERNER, E.A. (1923) J. Chem. Soc. 123, 2577.

WILHELM, X. and CHAMPAGNE, M. (1969) Eur. J. Biochem. 10, 102.

WILKINS, M.H.F. (1959) Rapp. Cons. Chim. Solvay, 11, 45.

WILLIAMS, F.G. and POTT, J.T. (1969) Federation Proc. 28, 661.

WILSON, J.G. and COHEN, L.A. (1963) J. Am. Chem. Soc. 85, 564.

WITKOP, B. (1961) Advan. Protein Chem. 16, 221.

WÖHLER, F. (1828) Pogg. Ann. 12, 253.

WOODS, K.R. and WANG, K.T. (1967) Biochim. Biophys. Acta, 133, 369.

YAMADA, S. and ITANO, H.A. (1966) Biochim. Biophys. Acta, 130, 538.

YOKOTSUKA, K., KIKUCHI, A. and SHIMURA, K. (1972) J. Biochem. 71, 133

YOKOTSUKA, K. and SHIMURA, K. (1969) In Histones and Gene Function, p. 20. (K. Iwai ed.). "The Research Group on Histones and Gene Function in Japan" .

ZUBAY, G. and WILKINS, M.H.F. (1962) J. Mol. Biol. 4, 444.