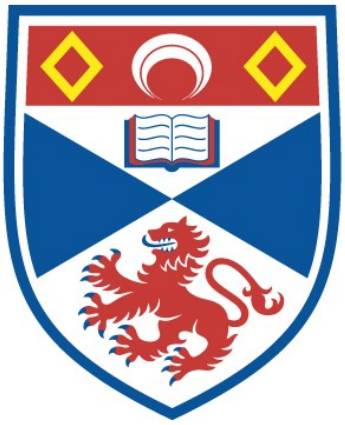
ENZYMATIC INVESTIGATIONS INTO EXTRA-HELICAL AND TERMINAL STRUCTURES OF COLLAGEN

David C. Steer

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1966

Enzymatic Investigations into Extra-Helical and Terminal Structures of Collagen

By

David C. Steer.

A thesis, presented to the University of St.Andrews for the Degree of Doctor of Philosophy.

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

DECLARATION

I hereby declare that the following thesis is based on work carried out by me, that the thesis is my own composition, and that no part of it has been presented previously for a higher degree.

The research was carried out in the Department of Biochemistry in the United College of St. Salvator and St. Leonard, St. Andrews, under the direction of Professor G. R. Tristram.

CERTIFICATE

I hereby certify that David Steer has spent nine terms engaged in research work under my direction, and that he has fulfilled the conditions of ordinance No.16 (St.Andrews), and that he is qualified to submit the accompanying thesis for the degree of Doctor of Philosophy.

ACADEMIC RECORD

I matriculated at the University of St. Andrews in October 1959, and graduated with the degree of Bachelor of Science, Second Class Honours in Biochemistry, in June 1963. My subsidiary subjects were: Botany, Chemistry, Physics at General Level, and Chemistry at Special Level. In October 1963 I matriculated as a research student in the Department of Biochemistry, St. Andrews University.

ACKNOWLEDGMENT

I should like to thank my supervisor Professor G. R. Tristram for his advice and criticism during this work. I should also like to thank Mrs. A. Serafini-Fracassini for performing the amino acid analyses, and Mr. W. J. Blyth for the photographs.

CONTENTS

INTRODUCTION

NOMENCLATURE

The name collagen covers a group of proteins widespread throughout the animal kingdom. The classification is one based on secondary structure, collagen having a unique triple helical polypeptide "backbone" quite different from the α- helix which is a feature of most proteins (Rich and Crick 1961). The amino acid composition of different collagens varies considerably, with the reservation that there is always an amount of glycine, proline and hydroxyproline which is compatible with the existence of the collagen fold in which these amino acids play an essential role.

In most collagenous tissues the majority of the collagen is highly insoluble. Usually some of the protein can be extracted by procedures involving relatively mild solvents; the amount depends on the tissue concerned and the solvent used. Thus with calf-skin and rat-skin collagen, extraction with dilute salt solution produces a fraction which has been termed neutral salt soluble collagen. Acetic acid and citric acid solutions dissolve still more viz. acid soluble collagen or procollagen. The residual

protein which can only be brought into solution by methods which probably involve the rupture of covalent bonds, is regarded as the mature form, and in most tissues represents the bulk of the collagen present (> 90%).

The term tropocollagen has been used to designate the single collagen molecule. It is believed that mature collagen, which is always laid down in a fibrous form, consists of tropocollagen units polymerised by intermolecular crosslinking. Thus the limit for solubility will be one of size of aggregate, some of the polymers still being sufficiently small to pass into solution. However, it is thought that neutral salt soluble and acid soluble collagen solutions contain mostly tropocollagen monomers, the former approximating most closely to the homogeneous condition. Jackson and Fessler (1957) reported that purified extracts of rabbit skin collagen in solution behave as rigid rods of length 4000 - 6000°A and width 100 - 200°A; this would correspond to an average aggregate size of twelve tropocollagen units.

Neutral salt soluble collagen and acid soluble collagen have been proposed as precursors of mature collagen, however the role of acid soluble collagen in this connection has been questioned by Harkness et.al. (1958). They observed a higher rate of incorporation

of C^{14} glycine into neutral salt soluble collagen than into acid soluble collagen, and took this as evidence against a mechanism of formation of tissue collagen of the form:- neutral salt soluble collagen → acid soluble collagen → mature collagen. However the distinction between the different forms of soluble collagen is probably not a hard and fast one (Jackson and Bentley 1960). Thus there probably exists a continuous spectrum of aggregation states, from the most recently formed and most easily solubilised molecules, to the older, partially polymerised ones, which are dissolved with most difficulty. Viewed in this light, the observations of Harkness can be explained by the probability that acid soluble collagen will contain a proportion of "older" molecules which have escaped polymerisation.

Aggregation States, the fibrous form and electron microscopy.

At the morphological level, collagen fibres are organised in a number of ways. One remarkable form exists in cornea where sheets of parallel collagen fibres are laid one on the other with their grains at right angles in a laminated structure (Gross 1961). When suitably stained and examined with the electron microscope, collagen fibres exhibit a characteristic banding pattern. The technique of positive staining attaches heavy metal ions to polar regions of the molecule. Phospho tungstic acid (P.T.A.) and cationic uranium are the most commonly used reagents, showing the distribution of basic and acidic side chains respectively. In both cases an axial period of about 640°A is observed in which 12 or 13 bands are normally present. The most prominent of these has been called the d. band.

Two other types of banding pattern can be obtained by precipitating collagen from solution under specific conditions. These have been termed "segment long-spacing" (S.L.S.) and "fibrous long-spacing" (F.L.S.); the former precipitates as a 'bobbin'-shaped crystallite of length 2,800°A and the latter as fibres similar to native collagen but with a periodicity of approx. 2,500°A.

The most widely accepted interpretation of these data is that whilst in native collagen, the tropocollagen molecules overlap by 640°A; in S.L.S. aggregates the molecules are joined only side to side, and are in register. In the F.L.S. forms the molecules are attached both side to side and end to end, but there is no overlap the individual molecules being laterally in register. Thus the length of the S.L.S. crystallite is equal to the length of the tropocollagen molecule, and the periodicity of the F.L.S. is four times that of the native fibre. The discrepancy between the periodicity of F.L.S. and the total length of the molecule, can be accounted for by the fact that co-axial molecules, overlap or "interlock" at their ends by approx. 300°A. (Schmitt et. al. 1953, Hall & Doty 1953, Hodge & Schmitt 1960). It has been suggested (Boedtker & Doty 1956 and Hodge and Schmitt 1953) that terminal peptide appendages are concerned in this overlap region.

By examination of dimorphic forms in which native-type fibrils were used as nucleation sites for subsequent growth of S.L.S. crystallites, Hodge and Schmitt 1960 demonstrated the exact correspondence of S.L.S. bands with those of native fibrils with respect to axial location, but not with respect to intensity. This was taken to show that "all bands in the native

type pattern arise by summation of sets of equivalent bands which contribute to the staining intensity by virtue of their lateral apposition in the staggered array." If it is accepted that the bands are due solely to distinct polar regions in the molecule (Kühn 1960) then this hypothesis requires that the tropocollagen macromolecule can be divided into four regions of equal length, each containing a total of approx. 12 polar zones whose axial positions are identical in each of the four regions. In each molecule there will be a total of approx. 50 polar locations, possibly a few more if the terminal peptide overlaps contribute to the banding at all. From amino acid analyses of collagen there are about 600 Acidic and basic side chains; i.e. if all of the polar residues are involved in the formation of the polar zones these would contain an average of 12 polar residues each (as well as any other amino acids present). Only one half of this average of 12 residues will be responsible for the band intensity at any one time depending upon whether staining is for acidic or basic residues (collagen contains approx. equal numbers of acidic and basic residues).

No information has been offered regarding the

contribution of each of the three constituent polypeptide chains of tropocollagen, to the band forming zones. There is good evidence that these three chains are chemically distinct (Piez 1965), so their primary structure must be such that when assembled as a triple-helix, the four equivalent sets of bands result. If all three chains are involved in the formation of each band, the number of polar residues per band from each chain will average four. If only one chain is involved then there is the question of the distribution of the polar regions between the three of them. The former possibility is probably closest to the true state of affairs, for alignment of the constituent chains of tropocollagen could occur initially by attraction between acidic and basic areas of the polypeptides. This would account for the identity of the bands resulting from the two methods of positive staining viz. for acidic and basic groups.

In the final construction of fibres from the tropocollagen molecules, it is necessary to postulate a two stage process. The molecules are first polymerised by end-to-end linkage through interaction of the terminal peptides to form proto fibrils; and these are subsequently displaced by 0.25 of their length relative to one another (Schmitt and Hodge 1960).

Obviously to explain the electron micrographs of collagen fibres in terms of a quarter-staggered arrangement and an orderly distribution of polar locations, necessitates a knowledge of the organisation and structure of the molecule at several levels. Current theories of protein biosynthesis can account for the most specific requirements of primary structure; species differences do occur, but always those residues in vital positions for the correct function or stereochemical arrangement of a protein, are the same. It is unlikely though that proteins are synthesised as units of molecular weight greater than about 60,000; the molecular weight of each of the three constituent chains of tropocollagen is approx. 100,000. Thus it seems likely that collagen is synthesised from a number of subunits; these cannot be identical because of the dissimilarity of the three chains of tropocollagen, and also because the four equivalent sets of bands of the native fibre correspond to the bands in the S.L.S. form in position but not in intensity. Nevertheless, the subunits must possess some relationship to one another in respect to the spacing of polar residues, if these equivalent bands are to exist. Does this stretch concepts of biochemical evolution too far?

There is no other example of a protein assembled

from subunits which although different, bear a mathematical relationship to one another in their primary structure; nor of a single protein which has repeating zones of this nature. These problems seem hard to explain but biology has probably had greater surprises for the biochemist in the past.

An attempt to explain electron micrographs of collagen on somewhat simpler lines has been made by Grant et. al. (1965). Pointing out that a completely quarter staggered arrangement has been demonstrated to be theoretically impossible in a three dimensional system where more than two molecules can be mutually in contact (Smith 1965), they suggest an essentially random aggregation based on the alignment of so called "bonding regions", five of which are present in the length of each tropocollagen molecule.

This theory can not only account for the 640 Å periodicity of positively stained collagen, it also explains the pattern obtained with negative staining which has been something of an embarrassment to protagonists of the "quarter stagger" system. The idea of Grant et. al. is that negative staining involves the retention of heavy metal in a neutral form in the regions of loose packing of the material. Negatively stained collagen shows up as a series of light bands

(a-bands) alternating with dark bands (b-bands). In a distance corresponding to the length of the molecule there are five a-bands and four b-bands. The light zones correspond to bonding regions with a high content of polar amino acids which by polar and hydrogen bonding, hold these parts of the molecule tightly together. In these bands there will be no room for the entrance of molecules of stain. Conversely the dark bands are more loosely packed non-bonding zones where there has been ready access for the negative stain.

Support for this concept of negative staining comes from examinations performed by Grant on glutaraldehyde treated collagen. In this material the size and density of the a-bands is increased whilst the b-bands becomes less distinct. This is thought to be due to additional intermolecular crosslinks being introduced by the glutaraldehyde. Thus negative staining reflects the molecular packing of collagen fibres and bears little or no relationship to the variations in crystallinity within individual molecules.

The length of the a-bands is approx. 265°A and of the b-bands 375°A. When the collagen molecules are aggregated into fibres there is a random choice as to which bonding region of one molecule crosslinks with one on another molecule. As the combined length of a and b

bands is in the region of 640°A, a periodicity of this distance will result.

An interesting point raised was that the collagen molecules in negatively stained preparations appear to be quite flexible, and a good deal of crossing over occurs in both a and b bands.

The advantage of a hypothesis of bonding regions giving an essentially random mode of aggregation, is that there is no need to postulate a two stage process of end to end linkage into protofibrils followed by quarter staggering. However the model of Grant et. al. does no more to help explain the 12 - 13 bands in the 640°A period of positively stained collagen. The only mention is that these characteristic spacings can be considered to be determined by a basically random process of aggregation of tropocollagen units. It would appear to be again necessary to assume the presence of four equivalent sets of bands and in this case they would need to be compatible with the concept of the simultaneous presence of five bonding zones.

The Primary Structure of Collagen

Amino acid analyses of soluble collagens give quite reproducible results, whilst mature collagen and derived gelatins tend to give inconsistent values, due no doubt to the difficulties in purification of insoluble proteins. Thus reliable comparisons have been possible only for the amino acid analyses of soluble forms of collagen, notwithstanding this, some interesting species, differences have come to light.

The earliest amino acid analyses were achieved by macro-chemical methods and although those techniques have been outdated for some time, the analyses of Chibnall and coworkers (Chibnall 1946) and Bowes and Kenten (1948) have certainly stood the test of time. Ion exchange chromatography as developed by Moore and Stein (1951) has revolutionised methods of amino acid analysis and the modern automatic analysers based on this principle have improved the speed and accuracy of analyses enormously. Eastoe (1955) using the original Moore and Stein technique, and published extensive data on the amino acid composition of collagen and gelatins. A large number of analyses have been published in recent years, probably as a result of the advent of the automatic amino acid analyser; a few of these are given in table 1.

As amide nitrogen masks in the region of 45 of the anionic side chains of aspastic acid and glutamic acid, there is a slight excess of cationic groupings in collagen which is therefore a slightly basic protein. Most values for the iso electric point of collagen lie between pH7.5 - 8 which would be in agreement with this concept. Other factors must also be considered in evaluating iso electric points, including non-covalent interactions involving ionisablegroups. These interactions may affect the dissociation of the groups concerned. Another factor which must be considered is the binding of ions.

Grettie (1965) produces evidence that the true iso electric point of collagen is only just above pH7. This could be accounted for by anion binding, or by interactions involving basic groups in excess of acidic groups. The anomalously steep titration curve of collagen in the region pH4 - pH6 raises the possibility of masked groups. Only part of the lysine appears to be titrated, so the possibility that some of these are masked must be considered (Steinhardt and Zaiser 1955).

The most interesting feature of the amino acid composition of collagen is the high proportion of glycine and of the imino acids. Hydroxyproline and hydroxylysine are unusual amino acids; elastin contains some

hydroxyproline, but otherwise they are found in no other commonly occuring proteins. It is generally accepted that collagen contains no tryptophan or cystaine. The tyrosine content is very low and it has been suggested that this amino acid is confined to terminal regions of the molecule or to covalently attached peptide appendages (Hodge et. al. 1960). Arginine and Glutamic acid give probably the most consistent values for different analyses, even with collagens obtained from quite different sources their amount seems to be fairly constant.

Although the content of glycine, proline and hydroxyproline varies to a certain extent, their amounts are always unusually high when compared with the composition of other proteins. (For collagen glycine usually represents approx. $\frac{1}{3}$ and proline + hydroxyproline $\frac{1}{5}$ of the total number of amino acid residues). This feature is a consequence of the unusual configuration of the polypeptide chains in collagen. Glycine and imino acids tend to distort the normal arrangement of a chain of amino acids, the α-helix. If their amount is as high as it is in collagen then the unique configuration known as the collagen-fold in which the 'pitch' of the helix is much wider than in the α helix, is probably a necessary result.

TABLE I. AMINO ACID ANALYSES OF SOME SOLUBLE COLLAGENS

Expressed as moles amino acid /1000 residues recovered (I,II,III,IV) & moles 10^5 g.protein (V, VI).

AMINO ACID.	SOURCE					
	Calf Skin Tropocollagen I	Calf Skin Tropocollagen II	Calf Skin Procollagen III	Calf Skin Total Acid Soluble Collagen IV	Human Tendon Acid Extract V	Cod Skin Acid Soluble Collagen VI
Hydroxyproline	98.0	92.0	82.3	86.0	96.2	62.4
Aspartic Acid	44.7	45.1	47.6	45.05	50.5	50.7
Threonine	17.0	18.2	18.8	17.8	19.3	27.3
Serine	33.6	36.8	29.2	37.2	38.5	61.1
Glutamic Acid	72.2	71.7	75.8	71.3	75.5	54.5
Proline	123.0	120.0	131.8	134.7	132.0	104.0
Glycine	330.0	333.0	327.3	325.2	338	333.0
Alanine	111.0	115.0	107.8	111.7	115.6	111.8
Valine	22.2	21.6	22.6	22.4	26.5	22.0
Methionine	6.0*	5.0*	4.5	6.4	6.0	13.0
Isoleucine	11.3	11.2	12.2	11.3	11.6	15.6
Leucine	27.4	25.5	26.6	24.6	27.2	28.6
Tyrosine	3.5	3.8	3.9	3.0	3.8	n.d.
Phenyl Alanine	12.8	13.4	14.7	13.3	14.8	n.d.
Hydroxylysine	7.5	7.4	5.1	5.0	9.3	5.2
Lysine	26.0	25.5	28.8	26.6	22.5	28.6
Histidine	4.4	4.3	6.5	5.0	5.6	7.3
Arginine	50.2	50.6	54.5	51.2	51.2	54.2
TOTAL	1000.8	1000.1	1000	1000	1044.1	979.4
Anide N.	30.0+	30.0+	--	44.0	46.0	65.0

* Value for Methionine and Methionine sulphone. + Ammonia.

I and II. Samples from different skins prepared in an identical manner. Rubin et.al.1965.

III Grassmann et.al. 1960. IV. Piez et.al. 1960. V. Eastoe 1955.

VI Young and Lorimer 1960. Analyses IV and VI obtained from Tristram & Smith 1963.

The hydroxyproline content of mammalian collagen is approx. 14% whilst that of reptiles lies between 10.2 - 9.3% and fish 5.8 - 7.9%. Takahashi and Tamaka (1953) have shown that the threshold temperature for hydrothermal stability of collagens also varies, from 60 - 70°C for mammalian collagens to 38 - 45°C for cold water fish. They propose that the two factors are related. Gustavson (1954) supported these views, suggesting that the hydroxyl group of hydroxyproline is involved in hydrogen bonding, giving rise to the extra thermal stability of collagens with a high content of this amino acid. More recent studies by Piez and Gross (1960), including accurate analyses of collagens from a wide range of animals have however, lead them to conclude that it is the total imino acid content rather than the hydroxyproline content which determines the degree of stability of the protein. The determining factor would therefore appear to be one of configuration rather than amount of hydrogen bonding.

Astbury (1940), basing his conclusions on chemical data and X-ray diffraction studies proposed that the collagen molecule is composed of repetitions of the fundamental sequence - P - G - R - (P, proline; G, glycine and R one of the remaining residues). Schrohenloher et. al. (1959) suggested - Glycine - Proline - Hydroxyproline -

to be the most important single tripeptide sequence, occuring frequently throughout the length of the molecule. By attacking heat denatured collagen with trypsin and analysing the resulting 114 peptides, Grassmann et. al. (1960) deduced that as with only one exception the peptides all contained approx. ⅓ of their amino acids as glycine, this amino acid is distributed evenly through the molecule.

These observations sum up the most important aspects of collagen primary structure, but that they hold true for the entire molecule is open to doubt, as has been shown by other work. Schroeder et. al. (1953 and 1954) and Kroner et. al. (1953 and 1955) described more than 60 peptides obtained by controlled acid and basic hydrolysis of collagen, the composition of some of these is shown in table 2.

TABLE 2. Peptides isolated from partially hydrolysed collagen and gelatin. From Hannig and Nordwig (1965).

Ala Gly	Pro Ser	Gly Pro Hypro
Ser Gly	Pro Thr	Gly Pro Ala
Thr Gly	Gly Gly	Gly Pro Glu
Val Gly	Gly Ala	
Glu Gly		Gly Pro Gly
		Ala Hypro Gly
Ala Ala	Gly Pro	Glu Hypro Gly

It is evident that glycine is often bound via its amino group to amino acids other than proline; the

dipeptide - Gly - Pro - in fact appears to be a very frequently occuring sequence. The carboxyl group of proline is bound to amino acids other than glycine in several peptides, and one of the dipeptides isolated had the structure - gly - gly -. These observations taken together show that the sequence - G - P - R - of Astbury cannot be the rule. This sequence does however account for 33 - 35% at least of the primary structure according to the results of Grassmann et. al. (1961), and must therefore be an important structural feature of the molecule.

Peptide sequences of four or five residues containing no glycine have been found (Hannig and Nordwig 1965) so the assumptions that glycine repeats in every third position, and is distributed evenly through the molecule can be no more than partially true.

Much indirect evidence for the existence of polar and apolar regions in collagen has come from electron microscope studies, and Kühn (1960) showed conclusively that the dark staining bands giving the characteristic cross striations of positively stained collagen fibres contain an accumulation of acidic and basic amino acids. The first chemical evidence came when Grassmann et. al. (1957) isolated six peptides of chain length ranging from 29 - 79 amino acid residues in which it was possible to

show that on the whole areas which were rich in proline and hydroxyproline were devoid of polar amino acids and vice versa. It has been thought for some time that the collagen triple-helix is composed of regions containing a preponderance of non-polar amino acids (crystalline regions) alternating with possibly non-helical or only slightly helical regions with a high proportion of polar side chains (amorphous zones). The crystalline regions are thought to consist largely of glycine and the imino acids, and here probably the sequence - G - P - R - assumes its major importance.

Grassmann pointed out that areas of some thirty amino acids which are completely devoid of polar groups may represent macro repeating units which could be associated with a single turn of the complete collagen helix (as opposed to the individual helices of three constituent polypeptide chains). Further chemical evidence appeared when Manahan and Mandl. (1961) isolated two peptides of the composition:-

Ala - Pro (Gly_3 Ala_2 Glu_2 Asp_2 Phe_2 Thr) Pro - Ala

and Gly - Pro (Gly_6 Ala_3 Val_2 Leu_2 Lys_2 Glu_2 Asp. Phe Ser) Pro-Phe.

The areas within the brackets agree with the idea of polar sequences devoid of imino acids.

Some of the most detailed peptide analyses yet achieved for collagen were produced by Grassmann et.al.(1960).

Using highly purified, chymotrypsin - free trypsin, they digested 62gms. procollagen which had been denatured at 90° for a short period. The digestion was performed at 37° and was followed titrimetrically. Only C-terminal arginine and lysine was found, and from the titration data it appeared that 60% of the bonds involving these amino acids had been ruptured. On this basis approx. 160 peptides of average chain length should be obtained assuming a molecular weight of 360,000 for collagen. The peptides were separated by continuous preparative electrophoresis, ion-exchange chromatography, molecular-seive chromatography and paper chromatography, resulting in the isolation of 114 peptides. Of these peptides, 55 fulfilled all criteria for homogeneity, 51 of which were submitted to quantitative amino acid analysis, 18 to end group analysis, and 10 to sequence studies. The results of the end group and sequential analyses of some of the peptides is shown in table 3.

In their interpretation of these results, Grassmann et. al. extend the hypothesis of alternating polar and apolar regions, adding that the polar zones can be further subdivided into acidic and basic sub-regions. A large amount of their evidence must come from the percentage figures for imino acids and polar amino acids in each of the 51 peptides analysed, for actual sequences are given for only short lengths of a few peptides. The

TABLE 3. from Grassmann et. al. (1960) and Hannig and Nordwig 1965

1. Polar Peptides without proline

B_1: H-Gly-His-Arg-OH.

S_{86}: H-Gly-Asp-Glu-Gly-HyLys- 3 Gly,Thr,2 Glu,Asp - Arg-OH.

---

2. Peptides containing proline + hydroxyproline.

S_1: H-Gly-( ) [7 Pro 8 Hypro 15 Gly, -Arg-OH
H-Gly-(Gly,Ser,2 Asp,Glu)-[2 Ala, Glu, 13 Asp, -Arg-OH
H-Gly-( ) [Arg. -Arg-OH

S_f: H-Gly-Ser-Ala-(Thr.Leu.Gly.)-[10 Pro 5 Hypro 12 Gly 2 Ala,Ser, 2 Tyr, Glu]-Phe-Glu-Leu-Asp-Asp-OH.
polar. apolar.

N_{IV1}. H-Gly. H-Gly.(Gly,Ser,Glu,Ala,)-[Pro,2 Hypro,7 Gly,Ala Val,2 Ser,2 Asp,2 Glu] -Lys OH -Lys OH -Arg OH
H-Gly.

N_{IV2}. H-Gly. H-Gly.(4 Gly,4 Ala,2 Glu- 2Asp) polar [10 Pro,2 HyPro,13 Gly 2 Ala,Val,Ser,Asp, 4 Glu 2 Lys, Arg.] -Lys OH -Lys OH -Arg OH
H-Gly.

N_{IV3}. H-Gly. H-Gly.(Gly,Ala,Ser, Thr,3 Glu,Asp.) - [5 Pro,Hypro,8 Gly,Ala Ser,2 Val,Asp,2 Glu.] -Lys OH -Arg OH -Arg OH
H-Gly. polar

N_{VI1}. H-Gly. H-Gly.(Gly,Ala,2 Glu,Asp)-[14 Pro,12 Hypro,29 Gly, 9 Ala,2 Val,4 Leu,2 Phe,4 Ser,9 Glu,2 Asp, Lys.] - Lys OH - Arg OH - Arg OH
H-Gly. polar

N_{VIv}. H-Gly. H-Gly.(Gly,Glu,Ala,Ser)-[5Pro,4 Hypro,18 Gly,7 Ala 2 Val,Leu,2 Phe,3 Ser, 2 Thr,4 Asp,5 Glu,2 Lys.] -Lys OH -Lys OH -Arg OH
H-Gly.

N_{VIW} H-Gly.(Ala,Ser,Glu) - [3 Pro,4 Hypro,14 Gly,5 Ala 3 Val, 2 Leu,Phe, 7 Ser, 6 Thr, Tyr, Met, 3 Asp, 3 Lys, 3 Arg, 2 Glucosamine] -Lys OH

---

longest peptides which were free of imino acids were of chain length 21 and 22 residues; peptide S_{36} of 13 amino acids and a sequence of 13 amino acids in peptide N_{IV2} as well as two tripeptides form the remainder of this category. The imino acid free zones represented by the four long peptides account for only 2.7% of the protein molecule. No peptides were isolated from this digest which were free of polar amino acids.

Of the peptides which were scrutinised for amino acid sequence (see Table 3) most contained regions to which can be ascribed either a polar or apolar character. No information is available as to the way in which these peptides fit together or about the order of the amino acids in the remainder of the peptides (the major portion in most cases). Most of the polar areas of peptides shown in Table 3 contain acidic amino acids rather than basic ones and the question arises as to the number of these which are present as asparagine or glutamine; some of these areas may not be truly polar, rather free of imino acids. That polar regions do exist cannot be disputed, but the chemical evidence to date is not sufficient to give any information as to their frequency or distribution, so it is difficult to draw any conclusions concerning the theories for the banding patterns obtained with electron micrographs.

The enzyme collagenase has been used to probe the primary structure of collagen. Numerous workers have

shown that the enzyme cleaves the bond -P-X↓G-P- where P = proline, G = Glycine and X one other residue which in collagen is usually hydroxyproline or alanine (Michaels et. al. 1958, Nagai and Noda 1959, Grassmann et. al. 1959, Schrohenloher et. al. 1959, Gallop and Seifter 1962). Thus collagenase digestion results in the formation of a large number of di and tri peptides containing glycine and proline, originating from the apolar areas of the molecule, as well as larger peptides derived from the polar regions. Franzblau et. al. (1964) isolated these larger peptides by dialysis and by chromatography on Sephadex G.25. Amino acid analysis of these non-dialysable peptides obtained from different collagens revealed basic similarities and a pattern of amino acids quite different from that of intact collagen. Thus the non-dialysable fraction (See table 4) has increased glutamic acid, aspartic acid and lysine, contains the majority of carbohydrate, aldehyde and tyrosine; but has reduced amounts of imino acids arginine and leucine.

It would be expected that collagenase only attacks non-polar regions of the molecule where the highest proportions of glycine and proline occur, and that the non-dialysable fraction as a result should be composed largely of peptides from the residual polar areas. This cannot be entirely true however, because the reputedly polar regions in electron micrographs of collagen fibres

stain with phospho tungstic acid, a reagent specific for arginine and listidine, being readily washed off the lysine residues. Thus the polar regions of the electron microscopist are rich in arginine whilst the non-dialysable fraction of collagen digests has relatively little, indicating that there can be no simple identity between the two fractions.

Table 4 from Seifter et. al. 1965.

Composition of non-dialysable fraction of various collagens. (As residues/1000 residues).

	Icthyocol	Rabbit Skin	Rat tail	Calf Skin
Asp.	72.3	73.8	83.8	82.9
Thr.	32.2	16.2	20.3	16.8
Ser.	37.2	38.4	50.5	40.5
Glu.	104.7	103.3	142.5	110.1
Pro.	72.9	73.6	82.2	83.7
Gly.	333.0	332.8	332.2	332.2
Ala.	123.0	95.0	99.4	103.4
Val.	20.5	25.6	29.4	22.0
Isoleu.	3.7	6.9	5.8	10.0
Leu.	14.0	16.9	18.6	23.3
Tyr.	6.0	7.3	8.3	12.6
Phe.	11.7	13.0	9.8	14.4
Lys.	27.0	29.8	28.8	28.1
His.	3.0	4.7	3.8	5.6
Arg.	36.8	33.2	31.2	36.1
OH.Pro.	63.3	83.2	95.5	97.7
OH.Lys.	10.5	5.6	6.0	6.8
NH_3	54.8	41.6	54.3	48.2

The non dialysable fraction has a high content of reactive side chains, especially aspartic acid and glutamic acid. That hexose may be bound to one of these residues in some way is a possibility as the majority of the hexose of collagen is found in this fraction. The chemical character

of this peptide portion, is also compatible with the theory that inter and intra molecular crosslinking and any other covalent modifications of the molecule will originate from it.

Butler and Cunningham (1965) have isoleted a glycopeptide from guinea pig skin collagen the amino acid composition of which (Table 5) indicates that it is of a highly polar nature. Their results suggest that the hexose is bound via an -O-glycosidic bond to hydroxylysine.

Table 5. Amino Acid Composition of Glycopeptide. From Butler and Cunningham 1965.

Amino Acid	Preparation 1. *	Preparation 2.
Hydroxylysine	1.0	1.0
Lysine	0.28	0.30
Histidine	1.09	1.23
Ammonia	0.79	0.95
Arginine	1.87	1.86
Glutamic Acid	0.39	0.42
Glycine	3.33	3.46
Methionine	0.93	1.07
Hexose	1.98	2.34

* Results expressed as ratio to hydroxylysine.

Peptides liberated from ichthyocol by collagenase have been investigated by Greenberg et. al. (1964) using the Edman degradation technique on the entire enzymic digest. By following the course of the enzymic hydrolysis with a pH stat it was estimated that 220 peptides were liberated per 1000 amino acid residues.

Obviously a large number of the smaller peptides will have the same structure. Seifter et. al. (1959) in a peptide map of the products of complete collagenase digestion of ichthyocol observed only 30 definite spots. The results of the Edman degradation showed that glycine is most abundant in positions 1 and 4., proline in position 2 and hydroxyproline in position 3. See table 6.

Table 6 Position of Various Amino Acids in Peptides. From Greenberg et. al. 1964
(As residues per 1000 in digest)

Amino Acid	Position I	Position II	Position III	Position IV
Hypro			55	
Asp.		6.2	13.7	3
Thr.		8.9	9.6	3
Ser.			11	7.6
Glu.		6.2	20.6	
Pro.		1.7	1	8
Gly.	200	10	36	53
Ala.	16.8	21	38	20
Val.		16		
IsoLeu.		6.2	2.7	
Leu.		1.6		3
Phe.		11		2
His.		3.7		27
Arg.			31.5	-3
TOTAL	217	193	219	100

There is a certain amount of glycine in position III evidence for peptides of the form -G-X-G-. However, on the whole these results are in agreement with the popular concept that the sequence -G-P-R- forms the most common tripeptide. The fact that such a high proportion of

hydroxyproline occurs in position III must be regarded as showing a definite characteristic of the amino acid sequence of collagen. Although alanine is more abundant than hydroxyproline in collagen, and is a common residue in the non-polar regions, it is less frequent than hydroxyproline in position III.

Thus two undisputable features of the primary structure of collagen have come to light. 1. The frequency of the dipeptide - glycine - proline -, if the observation of Grassmann 1961 that -G-P-R- accounts for 35% of the molecule is accepted, then 90% of the proline of collagen is in this arrangement, and there is no question of a random distribution of this amino acid. 2. The high proportion of hydroxyproline in the third position of the sequence -G-P-R-.

Apart from these two observations, the picture of the primary structure of collagen emerging from information obtained to date, is one of considerable heterogeneity. The polypeptide chain can be divided into distinct polar and crystalline regions but within the former areas the amino acid sequence appears to be fairly random. The results of Grassmann et. al. (1960) show that within the polar regions the polar amino acids themselves are arranged in no particular ordered fashion. These workers interpret their results as indicating polar

zones to correspond with the banding pattern of electron micrographs of collagen. This is difficult to explain, considering the spread of polar residues in the polar regions, and the somewhat limited number of polar amino acids to give rise to the large number of bands observed. One possibility which does not seem to have been examined, is that the banding pattern of positively stained collagen is in fact a representation at the level of one amino acid residue. If for example, collagen stained with uranyl acetate is considered: there are approximately 75 side chains which will bind this stain in each of the constituent chains of tropocollagen (1000 amino acids per chain; 75 glutamic acid, 45 aspartic acid, less 45 as glutamine and asparagine). The banding pattern requires the presence of something like 50 positions at which stain is bound in the length of the molecule. Thus 3 x 75 = 225 residues must give rise to 50 bands. With a proportion of bands arising from two or more acidic or basic groups at a time, as a result of either: (i) close proximity of polar residues in the polypeptide giving no resolution of bands corresponding to each of them; and (ii) the presence of polar residues in corresponding positions on the three constituent chains of tropocollagen; also non-staining of some of the polar side chains due to involvement in inter and intra-molecular ionic, cross-linking, there is still the likelihood that a

proportion of the banding pattern will be due to single polar groups. This concept allows for an essentially random distribution of acidic and basic residues and involves a lower degree of organisation of the so-called polar regions, this is perhaps more in keeping with the present state of knowledge of the primary structure of collagen.

The effect of proteolytic enzymes on collagen.

Collagen is remarkable for its resistance to proteolytic attack, only the highly specific enzyme collagenase brings about any large scale destruction of the molecule. There has been a great deal of conjecture as to whether or not proteinases in general have any activity with collagen. Kühn et. al. (1961) maintained that trypsin degrades only tyrosine-containing impurities and that the collagen molecules remain unchanged. Conversely Hodge et. al. (1960) reported that when soluble tropocollagen is treated with proteolytic enzymes, extra-helical peptide appendages are released which they term telopeptides. Concurrent with the release of telopeptides, the interactions of the molecule are modified; thus fibrous-long-spacing aggregates can no longer be formed but the ability to form segment-long-spacing crystallites is unimpaired.

Rubin et. al. (1963) demonstrated that pepsin liberates terminal or near terminal covalently bonded-peptides the amino acid composition of which is quite different from that of the residual major portion of the molecule. They observed that pepsin converts most of the β chains to α chains, and take this to indicate that the inter chain link is external to the body of the macromolecule.

Grant and Alburn (1960) showed that rat tail tendon collagen could be solubilised at pH 7.4 in the presence of calcium salts or salicilates by a variety of enzymes including trypsin and chymotrypsin. Other chemicals which could be present in in-vivo conditions e.g. arginine and creatinine, enhanced this solubilisation. The significance of such findings is difficult to assess, for rat-tail tendon collagen is an unusually soluble form. However collagen is normally quite insoluble at neutral pHs, so the solubilisation observed in this work could be due to telopeptide liberation and subsequent separation into -chains at the temperature of the experiment (38°C) i.e. a combination of proteolysis and thermal denaturation.

A picture of the tropocollagen molecule as a "hairy-rod", i.e. carrying a number of protruding peptide chains has been put forward by Rosmus et. al. (1966). They suggest that there are at least 15 places where peptide chains of low molecular weight are sticking out from the molecule, and that these chains are often almost identical in amino acid sequence.

Thus the majority of information is in favour of the concept that proteolytic enzymes find points of attack in peptide appendages of collagen; whilst the main body of the triple-helix is resistant to proteases, being degraded only by collagenase. It is reasonable to

suggest that the peptide appendages might be important in connection with crosslinking and end to end polymerisation interactions of tropocollagen molecules.

Pretreatment with the enzyme α-amylase at pH5.4 has been used as a method for the solubilisation of ox-hide collagen (Nishihara 1963). Steven (1964) has used the technique to extract collagen from human connective tissue and suggests that the enzyme destroys covalent linkages which stabilise connective tissue collagen, while leaving the actual telopeptides intact.

N-terminal residues of collagen.

Bowes and Moss (1953) and Grassmann and Hörmann (1953) could detect no N-terminal amino acid in native insoluble collagen by Sanger's dinitrophenyl technique. With procollagen Bowes and Moss found small amounts of dinitrophenyl (D.N.P.) aspartic acid and D.N.P. alanine which they thought must arise from extraneous matter, not representing true terminal groups. Other workers have found significant amounts of N-terminal residues in soluble forms of collagen. Hörmann et. al. (1965) detected 0.17 moles/1000 mole amino acid of D.N.P. Aspartic Acid and D.N.P. Glycine of which D.N.P. Glycine represented 81% using the D.N.P. technique with procollagen. Chandrarajan and Bose (1965), using the phenyl isothiocyanate method, found 0.16 moles aspartic acid and 0.23 moles glycine per 1000 moles amino acid in citrate soluble collagen. Hörmann et. al. also discovered similar amounts of the same amino acids in insoluble collagen.

Steven and Tristram (1962), by hydrolysing the entire reaction mixture after dinitrophenylation of acid soluble collagen, obtained D.N.P. derivatives of several amino acids (See Table 7), in small amounts. This procedure would detect the N-terminal residues of any peptides or amino acids present in the collagen preparation

which would otherwise be removed, by washing of the D.N.P. protein after dinitrophenylation. It was suggested that these N-terminal derivatives originate from a collagen non-protein nitrogen fraction which is not removed from the material by normal purification procedures. It was found that the non-protein nitrogen could be more or less completely removed by acetone precipitation or dialysis at low pH, and it was postulated that the fraction was important in connection with the fibre forming interactions of collagen.

Table 7. N-terminal residues of Soluble Collagen.
(From Steven and Tristram 1962)
Values expressed as moles/10^9 protein

D.N.P. Amino Acid	Amount
Gly.	0.13
Glu.	0.17
Asp.	0.17
Thr.	0.19
Ala.	0.10
Val.	0.10
Leu.)	0.19
IsoLeu.)	
Phe.	0.03
Ser.	0.37
Lys.	0.03
Ty.	Trace
Pro.	Trace
Hypro.	Trace
TOTAL	1.48

Just how many N-terminal residues are present in collagen is as much of a question now as it was ten years ago. The non-protein-nitrogen fraction of Steven and Tristram could account for all of the residues so far

observed by the numerous workers who have investigated this problem. The fairly constant amounts of D.N.P.-glycine and D.N.P. aspartic acid observed by Hörmann et.al. could easily belong to bound peptides which are impossible to remove from the parent protein, no covalent linkages being involved.

Hörmann et. al. do not comment on the significance of their 0.1 mol. of free α amino groups per 1000 mol. of amino acid residues, beyond saying that it is only one tenth of that expected if one amino end group occured in each of the three peptide chains of collagen. On the basis of determinations of acetyl groups present in collagen in fact, they propose that the peptide chains of collagen consist of an average of six subunits, whose amino end groups are acetylated.

Chandrarajan and Bose on the other hand found that they could detect no end groups in insoluble collagen, but alkali treatment liberated amounts similar to those they had estimated in soluble collagen. They took this to indicate that the N-terminal residues of insoluble collagen are masked by hexose, and actually isolated and characterised a glycopeptide from insoluble collagen containing galactose and glucose. Again there was no attempt to rationalise the low yield of N-terminal amino acids from soluble collagen in terms of number of residues

per molecular weight. Possibly the only logical explanation, if it is assumed that glycine and aspartic acid are present as terminal residues in the amounts stated, and that the theories for the molecular weight and subunit composition of collagen are accurate; is that the N-terminal residues of collagen are masked at some stage subsequent to its biosynthesis, possibly by an enzymatic process, and that this marking is either not always complete or occurs over a relatively long period of time so that in any preparation there are bound to be some molecules with free α-amino groups.

The Availability of ε-lysyl groups for substitution.

A fundamental problem in the chemistry of collagen has been the possibility that the ε-amino groups of lysine are not all equivalent; that some of them are not free to react with substituting reagents, while the majority can. This question has come to light as a result of studies using the dinitrophenyl technique of Sanger, mostly in conjunction with investigations of N-terminal residues as described above. Bowes and Moss (1953) in their experiments reported that 40% of the ε lysyl residues appeared to not be available to substitution by 1.fluoro 2.4 dinitro benzene (F.D.N.B.) the reagent used in the D.N.P. technique of Sanger (Sanger 1945). The reaction medium used in this case was 70% ethanol plus 4% saturated sodium bicarbonate, at room temperature. Several forms of collagen and gelatin were investigated but in no case did the degree of substitution exceed 70%. These workers pointed out that in the case of gelatin and formic acid treated collagen, the D.N.P. derivatives were soluble, so the low degree of substitution observed cannot be due to a low rate of reaction with an insoluble substrate. However they were only able to detect a small amount of lysine in their hydrolysates (4 moles/1000 amino acids) leaving approx. 10 moles/1000 amino acid residues unaccounted for. It

was put forward that ε-D.N.P. lysine is much less stable to acid hydrolysis when combined in collagen than when present as the free amino acid or in another D.N.P. protein.

Other workers to report incomplete substitution of ε-lysyl groups were Solomons and Irving (1958) who obtained 80% recovery of ε-D.N.P. lysine and Hallsworth (1964) who quoted percentage availabilities of 40-70% for different types of collagen aggregate, using corrections based on hydrolytic recoveries of ε-D.N.P. lysine of 88.4 - 96.9% as determined by control experiments. Hallsworth was also able to show that under his conditions of reaction viz. aqueous medium pH 7.4 and 37°C, the degree of substitution was proportional to ionic-strength.

Mechanic and Levy (1959) isolated the tripeptide LL N^{ε} - (glycyl-α-glutamyl) - lysine from bovine achilles tendon. They postulated that the 40% of lysine reported to be unassailable to F.D.N.B. by Bowes and Moss 1953 may be present as ε lysyl peptides of this nature. However it is a known fact that under certain conditions of hydrolysis peptide synthesis and rearrangement can occur, so it must be considered unlikely that this peptide occurs in such large amounts if at all.

Hörmann et. al. (1965) report that effectively all

of the lysine of collagen is substituted by F.D.N.B. when the reaction is carried out in a denaturing medium 2.5M with respect to sodium perchlorate. The degree of substitution was determined by separation of the basic amino acids from the hydrolysate of the D.N.P. protein on a column of amberlite IRC50. The results are given in Table 8.

Table 8. From Hörmann et. al. 1965.

Analysis of basic amino acids from normal and dinitrophenylated soluble and insoluble Collagen. Yields in mol./100 mol. amino acid.

	His.	Lys.	HyLys.	Arg.
Soluble Collagen.				
intact.	0.65	2.88	0.50	5.45
dinitrophenylated undenatured.	0.23	1.26	0.22	5.45
dinitrophenylated denatured 2.5M $NaClO_4$	0.03	0.07	0.03	5.45
Insoluble Collagen.				
intact.	0.60	2.70	0.55	5.35
dinitrophenylated undenatured.	0.06	0.27	0.11	5.35
dinitrophenylated denatured 2.5M $NaClO_4$	0.15	0.10	0.12	5.35

Heyns and Wolff (1956) also suggest that with excess bicarbonate and denaturants ε-lysyl residues of collagen become more or less completely substituted. Many other workers have recorded a whole range of availabilities to a variety of reagents, for ε lysyl residues. Leach (1966) gives values for substitution by potassium cyanate (carbamoylated residues) and benzensulphonylation ranging

from 40% - 99% and 45% - 98% respectively. Harding (1966) in a very complete survey of the subject gives data on reactivity of ε-lysyl groups to many reagents including:- acetylation, benzoylation, benzenesulphonylation, succinylation, guanidination, sodium bromoacetate, pp difluoro m m dinitrophenyl sulphone, 1, 5 difluoro, 2, 4 dinitro benzene, nitrous acid, nitrosyl chloride, ninhydrin and trypsin. The figures vary enormously but some are very high including 100% for acetylation.

In the face of this wealth of information, one must conclude that the ε-lysyl groups of collagen are in the region of 100% free for substitution reactions. Possibly in native collagen, some steric hindrance or ionic bonding prevents complete reaction with some reagents, but in most cases this can be overcome by denaturation or increasing the molarity of the solution. The fact that ionic strength has such an effect on reaction of ε-lysyl residues with F.D.N.B., could indicate that this form of masking is in fact due to ionic bonding of these residues rather than a purely steric effect; and that the observations in connection with denaturing agents could also arise from a simple increase in ionic strength causing the enhanced availability.

The experiments used in estimating the reactivity of ε-lysyl groups, in general are not sufficiently sensitive to preclude the participation of a very few (say one or

two per molecule) ε-lysyl groups in covalent linkages. These, whilst few in number, could be highly important features of the molecule. Thus the tripeptide of Mechanic and Levy cannot be completely ruled out. Franzblau (1962) produced evidence in favour of the idea, when he found that 12% of the lysine of a collagenase digest of ichthyocol, is not free to react with F.D.N.B.; masking in such small peptides is unlikely to be anything but covalent, although here again ionic linkages could be important. The most contradictory evidence against this theory came from the work of Hörmann et. al. when they actually estimated the free lysine in hydrolysates of D.N.P. collagen. Their figure of only 0.07 mol./100 mol. amino acid (see Table 3) corresponds to 2.1 lysine residues unsubstituted per molecule (molecular wt. 300,000), so the hypothesis is not completely ruled out, although of course this tiny amount of free lysine could have originated from breakdown of ε D.N.P. lysine.

Lockhart and Abraham (1956), reported that ε-lysyl peptides are extremely stable to acid hydrolysis, having found that on 43 hrs. hydrolysis with 11N. HCl at 80°C there was no release of lysine or aspartic acid from an ε-lysyl peptide of the two amino acids. This could be a special case; the proximity of two free carboxyl groups to the peptide bonds would tend to repel protons and confer a stability on the molecule which would not be

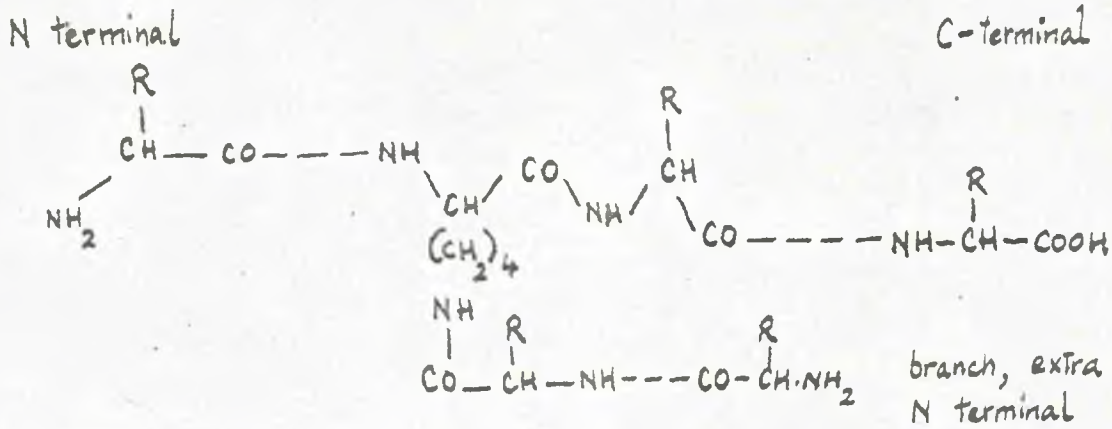
characteristic of all ε-lysyl peptide bonds. However, if this stability is a property of all ε-lysyl bonds, estimations of free lysine in hydrolysates of D.N.P. proteins would be subject to considerable errors, and the results of experiments such as those of Hörmann et. al. would be misleading. This is only a remote possibility, nevertheless.

The possibilities for covalent linkages involving lysine are numerous. Bensusan 1965 put forward a hypothesis for crosslinking between the polypeptide chains of collagen, by N-glycosyl linked carbohydrate residues taking on the form of a Schiff-base, and thought that the ε amino group of lysine would form the amino group donors in this system.

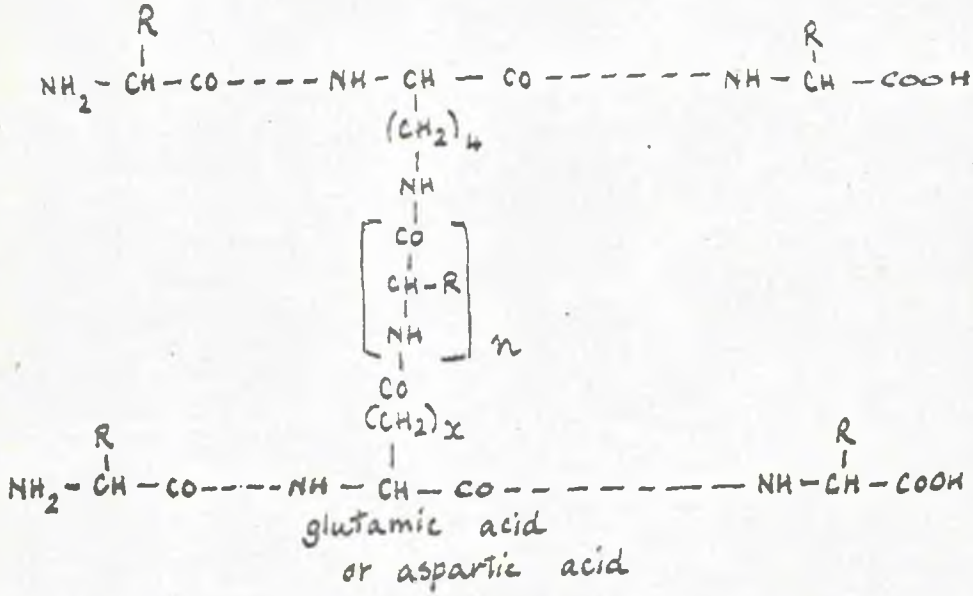
Harding (1966) enumerated the possible peptide structures in which the ε amino groups of lysine could participate; see figure I. He pointed out that in the latter case (straight chain ii) the lysine residue involved would appear as N-terminal lysine in investigations into the N-terminal amino acids of collagen. This is not completely true: with Sanger's technique the normal N-terminal derivative of lysine is α, ε Di D.N.P. lysine, which is ether soluble. The derivative of lysine combined in an ε-lysyl bond would be the unusual derivative α - mono D.N.P. lysine, which is water soluble and would not be extracted with ether. α - mono D.N.P.

FIGURE I. Possible linkages involving ε-lysyl amino groups. (From Harding 1965.)

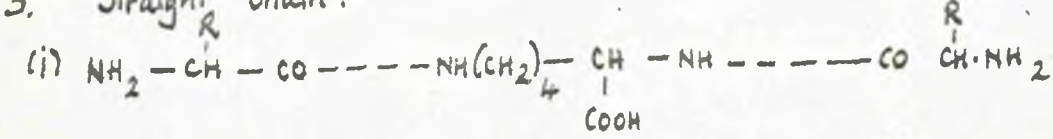
1. Branching Point



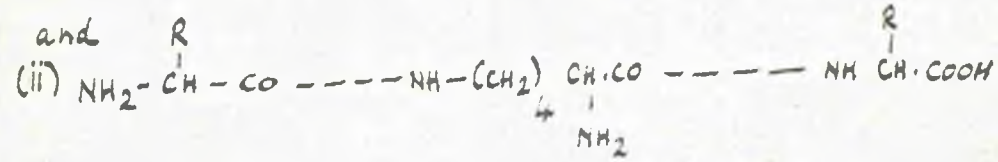
2. Cross linking to a carboxyl residue.



3. Straight chain.

(i) 

and

(ii) 

lysine would almost certainly escape detection in the aqueous phase of an analysis of a D.N.P. protein hydrolysate, due to the large excess of ε D.N.P. lysine which would also be present.

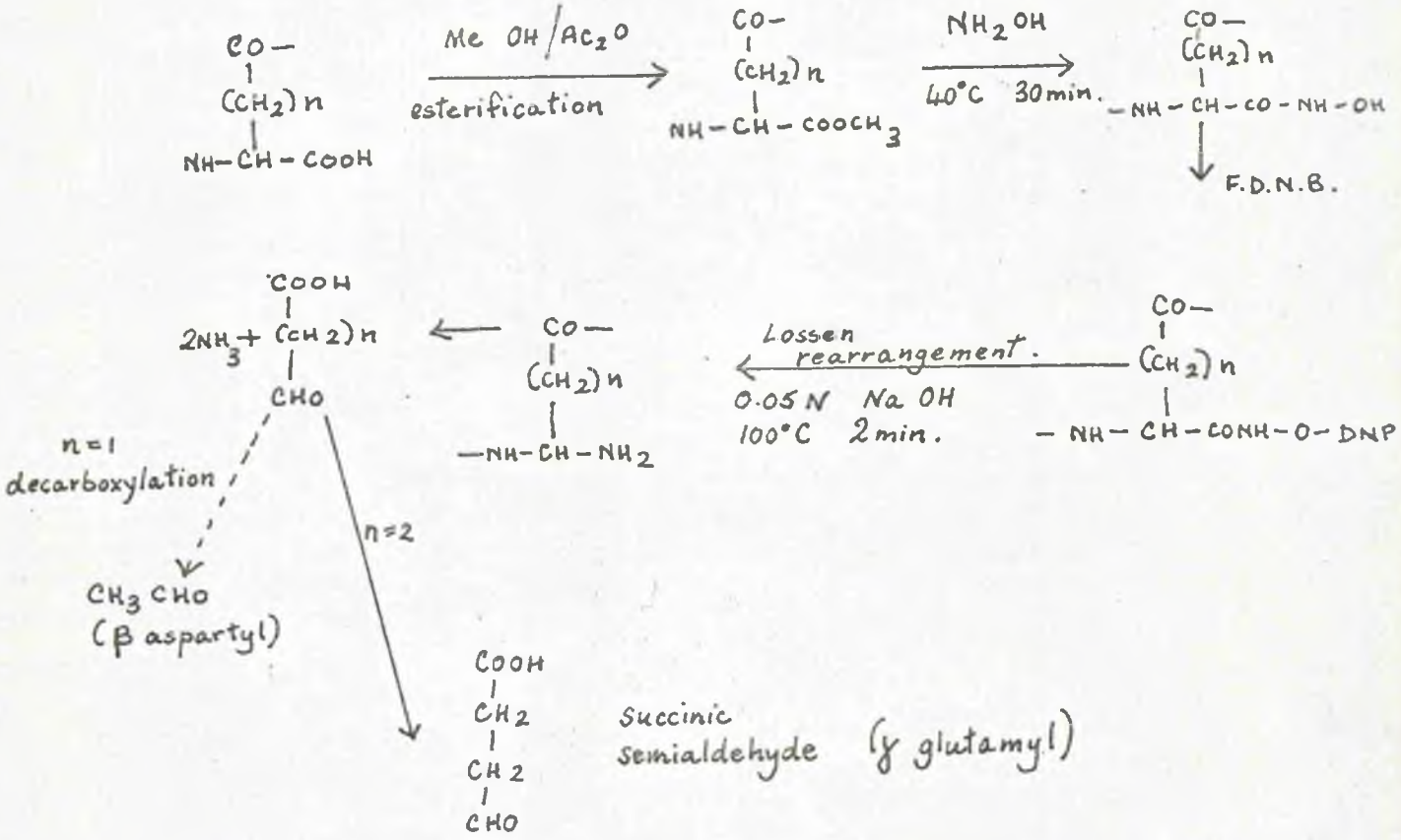
γ - Glutamyl and β aspartyl linkages.

The presence of a considerable number of γ - glutamyl linkages in collagen has been suggested by Haurowitz et. al. (1957) as a result of experiments using the thiohydantoin method. The investigations were not quantitative but were taken to indicate the presence of γ - glutamyl linkages rather than β aspartyl linkages. Gallop et. al. (1960) used the Lossen - Rearrangement to obtain a quantitative estimation of the amount of the γ - glutamyl linkages in collagen; see figure 2.

Under these conditions, α linked glutamic acid + aspartic acid give α γ diamino butyric acid and α β diamino proprionic acids respectively. In the gelatin obtained from ichtyocol, the formation of these compounds was noted as well as a decrease in glutamic acid and aspartic acid, and the formation of succinic semialdehyde and ammonia. A criticism of this work was that the anhydrous medium used could promote ring closure by the γ carboxyl groups, with reopening to form γ-links. Franzblau (1962) and Franzblau et. al. (1963) developed a method for the formation of

FIGURE 2

Identification of γ glutamyl linkages.
(From Gallop et al. 1960.)

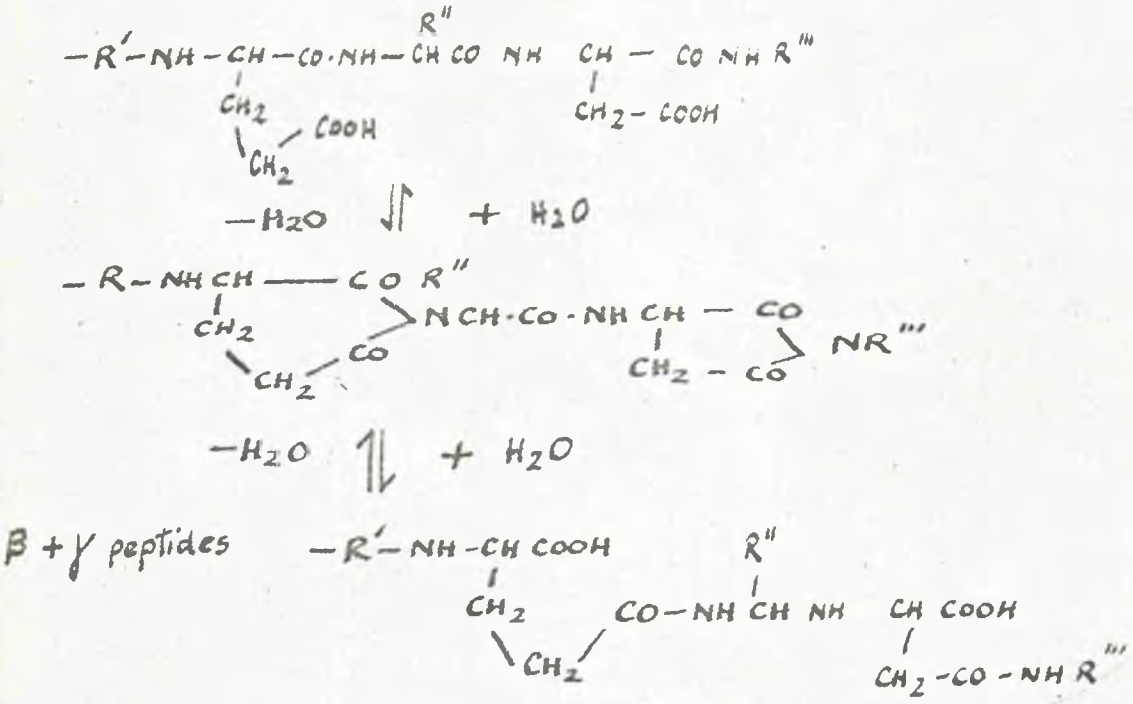


hydroxamic acids from unmodified proteins in aqueous conditions. Using water soluble 1 - cyclohexyl - 3 [2 - morpholinyl - (4) ethyl] - carbodiimide metho p-toluene sulphonate at pH4 and 25°C in an aqueous medium, they suggested that there would be no possibility of α - γ conversion. Their results indicated 30% glutamic acid of ichthyocol and 13% of acid soluble calf skin collagen is present in γ peptide linkage.

Tristram (1964) considered that the possibility of α - γ conversion should not be ruled out. With a polypeptide containing normal α - glutamyl and α - aspartyl linkages, the reactions and equilibria in figure 3 might be derived under certain conditions. Curtis and Spikes (1962) have shown that carboimides catalyse the formation of peptide linkages without themselves being involved, and so the glutarimide ring system could be formed in this fashion. Tristram also referred to work by Kovacs et. al. (1953) on the Hofmann degradation of polyanhydro aspartic acid, which showed that this polymer opens to give both α and β forms of polyaspartic acid.

If γ - glutamyl linkages are present in collagen in substantial numbers, as Franzblau suggests, then the two methylene groups introduced into the polypeptide "backbone" would have a marked effect on its configuration. Franzblau states in this respect that the γ - glutamyl linkages may be largely responsible for the disorder found

FIGURE 3. Equilibria involving Ring closure of glutamyl and aspartyl functions.
(From Tristram 1965.)



in the polar regions of tropocollagen, which contain the majority of the glutamic acid residues.

Free α - carboxyl groups of proteins tend to have lower pk values than β and γ carboxyl residues. It would be expected that the free α - carboxyl groups present as a consequence of γ - glutamyl linkages would have been detected in some of the many titration studies performed on collagen. An appreciable titration of groups over a lower pH range than that normally observed for β and γ carboxyl groups has not been reported for collagen as yet. However, this problem is complicated by the fact that pk values for groups sometimes differ between proteins.

Subunit composition of Collagen.

Orekhovich and Shpikiter (1955) first observed two peaks when soluble collagen at 40°C (parent gelatin) was examined in the ultracentrifuge. They assigned molecular weights of 80,000 (α component) and 160,000 (β component) to the two fractions, and proposed that the collagen molecule consisted of two α chains and one β chain. Doty and Nishihara (1958) reported molecular weights of 120,000 (α) and 230,000 (β), and suggested that tropocollagen was formed from only one of each subunit, the triple helix arising from the doubling-over of the

chain.

The β components were found to be relatively labile, mild alkali treatment or prolonged heating at 40°C or more resulting in breakdown into α - chains, or molecules of the same size as chains. This β → α conversion could not be achieved using hydrogen - bond breaking agents and it was concluded that the β component is held together by labile covalent linkages with an energy of activation of approximately 24k. cals.

Rubin et. al. (1962) suggested that the alkali labile bonds of β components are located in tyrosine containing pepsin-vulnerable end regions of the polypeptides, since after pepsin treatment they observed a markedly increased α : β ratio in soluble calf-skin collagen.

Piez et. al. (1960) fractionated solutions of parent gelatin by column chromatography at 40°C, using carboxymethyl cellulose adsorbent. They discovered two fractions which were thought to correspond to the α and β components observed in the ultracentrifuge. In a later paper, Piez (1963) reported that with improved chromatographic technique, four components were shown: these have been called α1, α2, β1 and β2. Both α components have a molecular weight of approximately 100,000 while the β components have a molecular weight of about 200,000. Thus one α and one β together would form the tropocollagen molecule. The α1 and α2 chains differ in their amino

acid composition, and β 1 corresponds to the combination $\alpha 1 - \alpha 2$, whilst $\beta 2 = \alpha 1 - \alpha 1$. It was suggested that every collagen molecule basically consists of two $\alpha 1$ chains with one $\alpha 2$ chain, and crosslinking occurs in two ways to give either β 1 or β 2 components. The possibility that all three chains can become crosslinked was recognised, and the term γ- component used for this molecule.

Table 9. Subunit distribution in Rat collagens.
(From Piez et. al. 1963)

	Collagen subunits				Collagen structures*		
	$\alpha 1$	$\alpha 2$	$\beta 1$	$\beta 2$	$(\alpha 1)_2\alpha 2$	$\alpha 1\,\beta 1$	$\alpha 2\,\beta 2$
Ratskin:-							
Salt extracted	55	27	13	5	72	20	8
Acid extracted	35	15	37	13	24	56	20
Rat tail tendon	30	10	45	15	9	68	23

* Calculated from subunit distribution.

$(\alpha 1)_2 \alpha 2$ represents excess α components.

Further fractionation studies by Piez et. al. (1964), and Bornstein et. al. (1964), provided more insight into the system of crosslinking in collagen. Pointing out that the majority of investigations into the components of collagen have been carried out on acid soluble forms which usually represent a minor and possibly atypical fraction of tissue collagen, they presented results of experiments using collagen remaining after salt and acid extraction, but solubilised by the fairly rigorous method of extraction

with 5m. guanidine. This collagen contained more total β - component (75 - 80%) than the theoretical maximum of 67% that could arise from intramolecular crosslinking alone. A new β component was identified which was shown to be the dimer of α 2 by amino acid composition and sedimentation properties. Since each collagen molecule contains only one α 2 chain, this dimer could only arise from an intermolecular crosslink.

Piez et. al. propose that both inter and intramolecular crosslinking take place by what is probably a single continuous process. The rate of this crosslinking process varies between collagens from different tissues; but, with an increasing amount of intermolecular bonding, the condition of high insolubility found in many collagens can be explained.

As intermolecular crosslinking gives rise to such a complex pattern of subunits, the nomenclature whereby β and γ components are labelled according to which chains they are formed from, has been introduced.
Thus $\beta_{12} = \alpha 1 - \alpha 2$ and $\gamma_{122} = \alpha 1 - \alpha 2 - \alpha 2$ etc.

Figure 4 (from Piez et. al. 1964) shows some of the possible molecules formed during the early stages of crosslinking. All of these will be extractable by 5M guanidine, but it is likely that larger aggregates are impossible to extract by solvents, unless some degradation is involved.

From Piez et al 1965

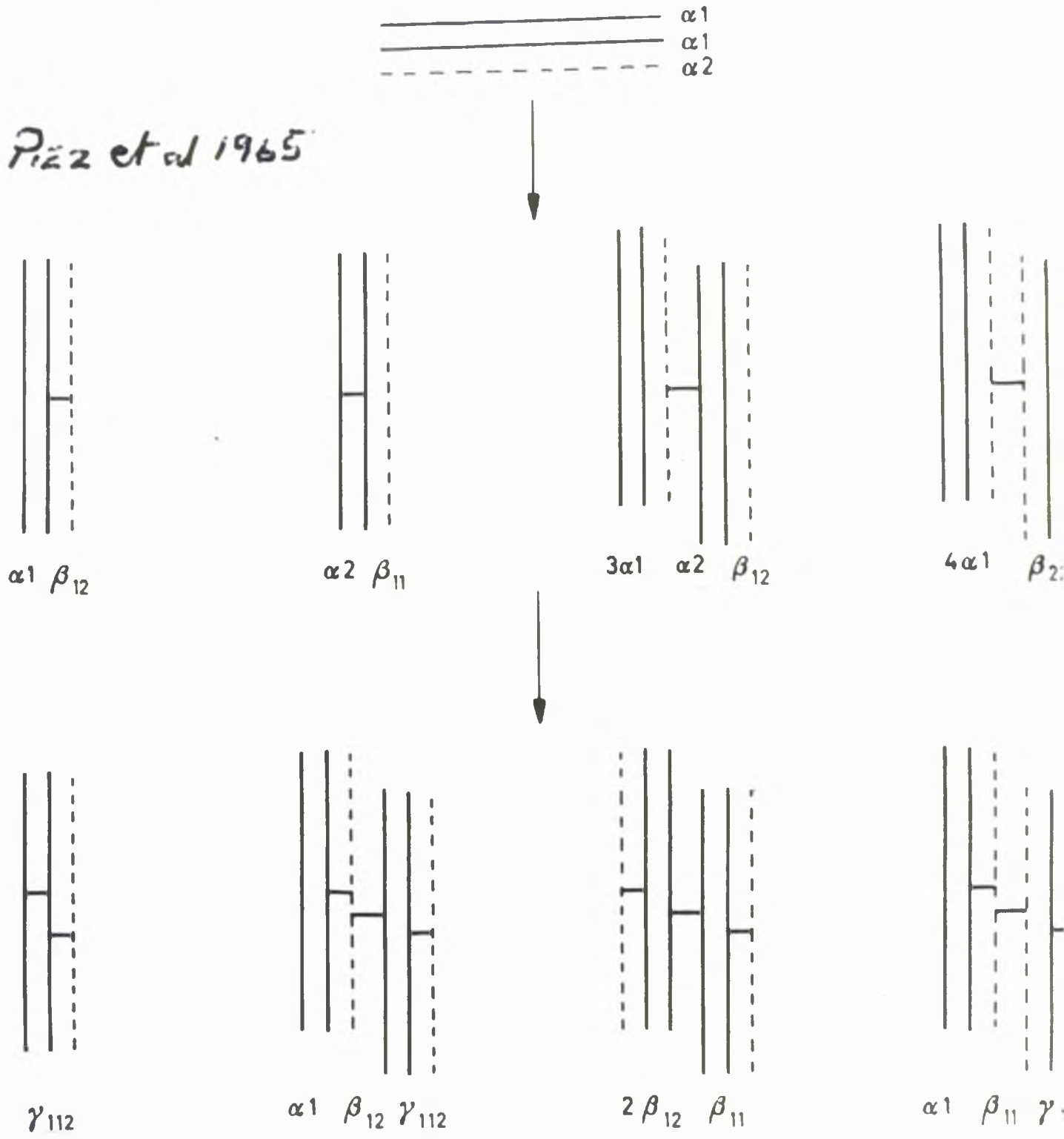


FIG. 4. A diagrammatic representation of the early stages of collage crosslinking. The triple-chain molecule is initially not crosslinked (top One of the three chains (α2) is different from the other two (α1). On crosslink produces β-components (β_{11}, β_{12}, and β_{22}) which can be isolate from the denatured collagen (middle). Equivalent products can aris from either inter- or intramolecular crosslinking. Additional cros linking produces larger aggregates including γ-component (bottom Crosslinking may continue giving rise to a continuous three dimensiona structure

Piez (1965) and Heidrich and Wynston (1965) have further amended the structure of collagen by postulating the existence of a third α component. Piez, working with codfish skin collagen, and using molecular sieve chromatography and carboxymethyl cellulose chromatography, obtained a third α component ($\alpha 3$) with an amino acid composition very similar to $\alpha 1$. (See Table 10). He also speculated that the $\alpha 3$ chain may contain one tryptophan residue on the basis of absorption-spectra of the three α - chains. Heidrich and Wynston demonstrated three α components in calf-skin and rat-skin tropocollagens. The new component, which they also called α 3, had a similar amino acid composition to α 1 and was difficult to separate from $\alpha 1$ except by starch gel electrophoresis.

Nänto et. al. (1965) report the presence of six bands during the electrophoresis of guinea-pig skin, neutral salt soluble collagen, at pH 4.8 38°C on starch gel. Four of the bonds are equated with $\alpha 1$, $\alpha 2$, β 12 and β 11; the other two, designated X_1 and X_2, move more slowly and are thought to be larger aggregates than β components.

Table 10. Amino Acid Composition of Codfish Skin Collagen and its constituent α chains.
(From Piez 1965)

Residues/1000 total residues.

Amino Acid	Collagen	α1	α2	α3
4-Hydroxyproline	56	55	52	58
Aspartic Acid	52	50	54	51
Threonine	24.5	23.5	26.9	24.5
Serine	73	70	73	73
Glutamic Acid	73	76	62	78
Proline	97	93	97	96
Glycine	339	339	348	347
Alanine	109	119	107	101
Valine	17.4	15.5	19.5	20.1
Methionine	13.2	16.7	13.3	16.8
IsoLeucine	9.3	10.8	9.2	8.9
Leucine	20.9	18.8	24.4	17.5
Tyrosine	4.2	1.8	4.7	2.6
Phenylalanine	10.8	13.2	9.1	11.0
Hydroxylysine	6.5	5.5	9.5	5.3
Lysine	27.9	31.3	20.6	30.3
Histidine	8.5	5.2	11.5	7.0
Tryptophan	-	0	0	1.0
Arginine	53	51	54	51
Amide Nitrogen	(46)	(43)	(46)	(54)

Further evidence for soluble aggregates with a greater molecular weight than β components has come from the ultracentrifuge work of Veis et. al. 1960, Altgelt et. al. 1961, and Grassmann et. al. 1961. Schleyer (1962) obtained a chromatographic fraction of collagen larger than β. Kulonen et. al. (1962), fractionating gelatins by chromatography, described a component which is eluted with 0.1N sodium hydroxide and which in their opinion resembles the parent collagen. Tristram et. al. (1965), by continuing the gradient of ionic strength to 0.5, and

finally eluting with alkali, obtained two fractions additional to the normal α and β pattern produced by carboxymethyl cellulose chromatography. The first of these fractions (component 3) appeared in greatly reduced amounts when the temperature of the fractionation was increased, and is probably a polymerised form of α, held together by hydrogen bonding. The fraction eluted with sodium hydroxide (component 4) is probably similar to the one obtained by Kulonen et. al. from gelatin, consisting of covalently linked high molecular weight aggregates of the size of γ - components, or larger.

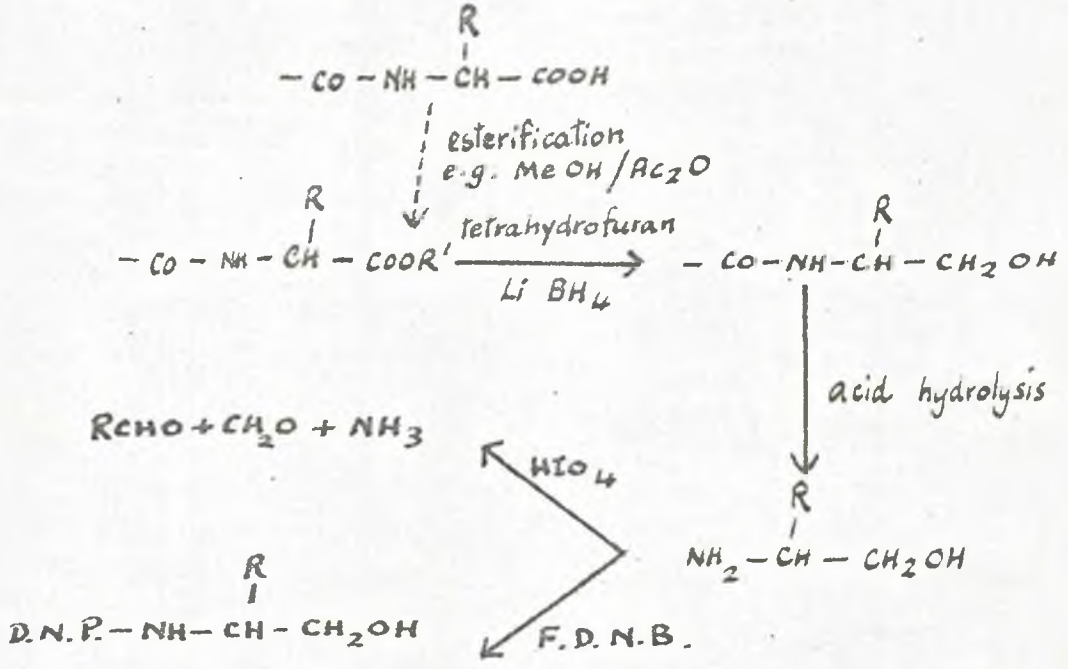
In a recent publication, Levy and Fishman (1966), report that by using a system of automated stepwise chromatography on carboxymethyl cellulose, but with similar overall conditions of pH, ionic strength and temperature to those used by Piez et. al. (1963), they have been able to resolve denatured ichthyocol into a number of fractions considerably larger than has been noticed previously. Whether these fractions represent true subunits of collagen, or are the products of degradation of covalent linkages, these authors are not prepared to commit themselves as yet, but it does seem that there might be some factors which have not been fully elucidated, in connection with the subunit structure of tropocollagen.

Ester-like linkages.

Evidence that the α chains of collagen are built up of a series of smaller subunits has come to light as a result of investigations into ester-like bonds in collagen. Grassmann et. al. (1954), used the method of reduction with Li BH_4. With and without a preliminary esterification stage the same result was obtained, viz. 40 groups per 1000, which presumably reacted as esters. Thus, for no increase to be detected with preliminary esterification, any aspartic acid or glutamic acid participating must be esterified at both carboxyl functions. The reactions involved in reduction of esterlinkages are shown in figure 5.

In this way, a large number of bonds were detected, the acyl group donors of which were estimated as amino alcohols. However, this work has been criticised by Crawhill and Elliot 1955, and Chibnall and Rees 1953, on the grounds that there was a strong likelihood that some of the amino alcohol present was a product of peptide fission. Grassmann et. al. themselves pointed out that some of the carboxyl groups of aspartic and glutamic acid are also reduced by Li BH_4. The value of this work is difficult to assess, but further indications of the presence of ester linkages have come from the work of Gallop et. al. (1959) using specific cleavage of the

FIGURE 5. Reduction of Ester bonds.
(From Harding 1965.)



bonds with hydroxylamine and hydrazine.

Gallop et. al. identified the carboxyl donor using the Lossen Rearrangement resulting from the thermal decomposition of hydroxamic acids. See figure 6. Their results indicated the involvement of α and β carboxyl functions of aspartic acid only. (Table 11.)

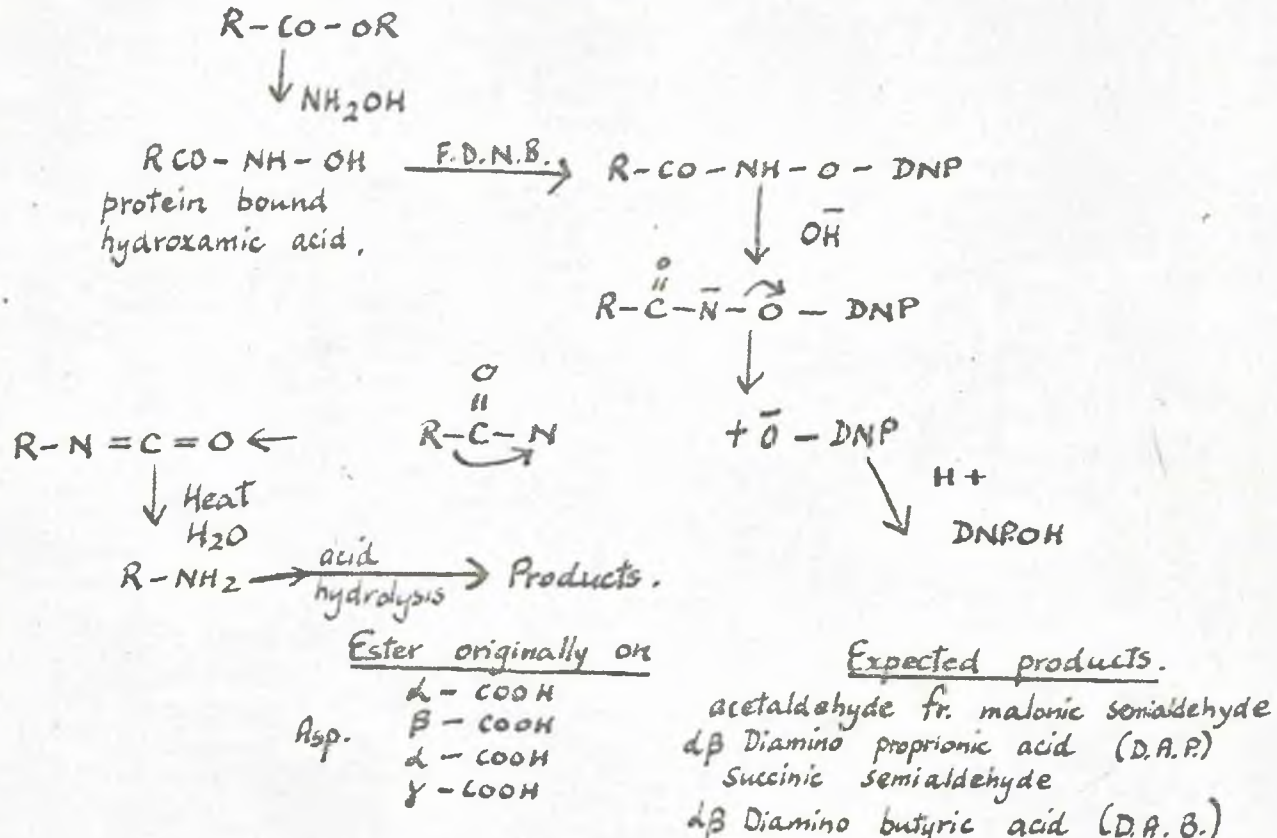
Table 11. Analysis of Material following Lossen Rearrangement of D.N.P. hydroxamate, derivative of gelatin from ichthyocol, as compared with the analysis of a suitable control gelatin. (Blumenfeld and Gallop 1962).

	Control A.	Treated gelatin B.	Diff. B-A (residues/1000)	
Hydroxamic Acid	0	4.9	+ 4.9	
Glu.	71.0	71.4	+ 0.4	
D.A.B.	0	0	0	
Asp.	46.0	41.5	- 4.5	
D.A.P.	0	2.6	+ 2.6)	4.4
Aldehyde	4.9	6.7	+ 1.3)	

The amounts of hydroxamic acid estimated correspond to a cleavage of 6 bonds per α chain of 1000 residues. It appeared that these bonds occurred in groups of two and that the α chains were built up of an average of 4 segments held together by 3 pairs of bonds. Blumenfeld 1965 showed that the subunits were of three types, represented by molecular weights of 34,000, 18,000 and 8,000; within each of these molecular weight groups there are probably several components which differ chemically.

Blumenfeld and Gallop (1962) treated collagen with

FIGURE 6. Identification of Ester bonds with Lossen rearrangement. (From Gallop et al. 1960.)



hydrazine and got results similar to those with hydroxylamine. They also digested hydrazine treated collagen with collagenase, and isolated the peptides containing hydrazone functions. They found that each peptide had 2 aspartic acid and 2 hydrazone groups, as α and β carboxyls are both involved in the ester linkages, and no β aspartyl peptide bonds occur in collagen, they concluded that ester linkages involving C - terminal aspartic acid must be present.

Hodge et. al. (1964) correlate the findings of Gallop, with the data obtained for S.L.S. aggregates of collagen, by analysis of electron micrographs, and proposed a subunit structure for tropocollagen in which the 2 α1 chains consist of 5 identical polypeptide subunits, and the α2 chain of 7 identical subunits. This theory would necessitate molecular weights of 14,000 and 20,000 for the two types of subunit, which is not in keeping with the results of Blumenfeld (1965).

Gallop (1966) has also postulated a subunit hypothesis for the α chains of tropocollagen. Each α chain consists of 6 subunits of molecular weight 17,000 of three compositional types A, B and C. Thus tropocollagen contains 6A + 6B + 6C in the form:-

$\alpha 1 = 3A + 2B + C$

$\alpha 2 = 2C + 2A + B$

$\alpha 3 = 3B + 2C + A$

This theory was tested by comparison with Piez's analysis of the three α components of codfish skin collagen (see table 10). Gallop remarked on the frequency of the number 6 or multiples thereof, and examined the amino acid composition with respect to Tyrosine, Hydroxylysine, Histidine, Phenylalanine, Methionine, Valine, Threonine, Lysine, Glutamic Acid. He found that the predicted composition of α chains on the basis of his subunits corresponded well to observed data. The subunits of molecular weight 17,000 can be equated to the 18,000 molecular weight fraction of Blumenfeld et. al. (1965), with the extension that the 8,000 molecular weight fraction represents a breakdown product, and the 34,000 molecular weight group, a dimer form of the basic subunit. Such a subunit structure has the advantage that it leaves considerable scope for variations between 640°A repeat periods of the molecule, and therefore does not contradict any of the criteria imposed by electron microscope studies. However, it is based on the assumption that the chromatographic experiments of Piez give definite separations of the three α components and that these fractions are indeed distinct species, with the new slant imposed by the work of Levy and Fishman (1966) this premise can no longer be regarded as unequivocal.

The exact significance of the subunits linked by

"ester-like" linkages, from the point of view of biosynthesis of collagen, is uncertain, until further information has been presented, it seems equally likely that (i) the ester linkages are synthesised while the polypeptide chain is attached to ribosomal material, by the action of a specific enzyme, or (ii) that the ester linkages are synthesised subsequent to removal of the polypeptide from the ribosomes, in a process of polymerisation of subunits by a mechanism (again, probably enzymatic) about which as yet little is known. If possibility (i) is true then the fractions identified by Blumenfeld et. al. will not be true subunits from a biosynthetic point of view. In some ways, (i) appears most likely, for it does not seem an economical system that for the extra-ribosomal polymerisation necessary for (ii) the incorporation of a very small subunit of only a few amino acid residues should be necessary. This ought to be the case, when the fact that the ester linkages occur in groups of two, is considered. Possibly the true mechanism is a combination of (i) and (ii). Hodge et. al. and Gallop in their hypotheses for the subunit structure of tropocollagen, seem to assume that (ii) is the likely mechanism.

<u>The cross-linking of collagen</u>.

Fractionation studies have shown that the inter - and intra - molecular crosslinking reactions which occur

in collagen, are effected by what is probably a single continuous process involving the same functional groups acting competitively (Piez et. al. 1964 and Veis and Anesey 1965). The linkages are covalent, and their formation occurs at some stage after biosynthesis of the molecule, the majority after incorporation into the fibrous tissue. Thus it seems likely that the mechanism of formation of crosslinks is either a spontaneous process or an enzymatic process. The possibility that extra groups, probably amino acids, carbohydrates or difunctional aldehydes, are included into the polymer during cross-linking, cannot be excluded. The nature of the crosslinks in collagen however, is still a matter for conjecture. Numerous theories have been proposed, but only a few are supported with good chemical evidence.

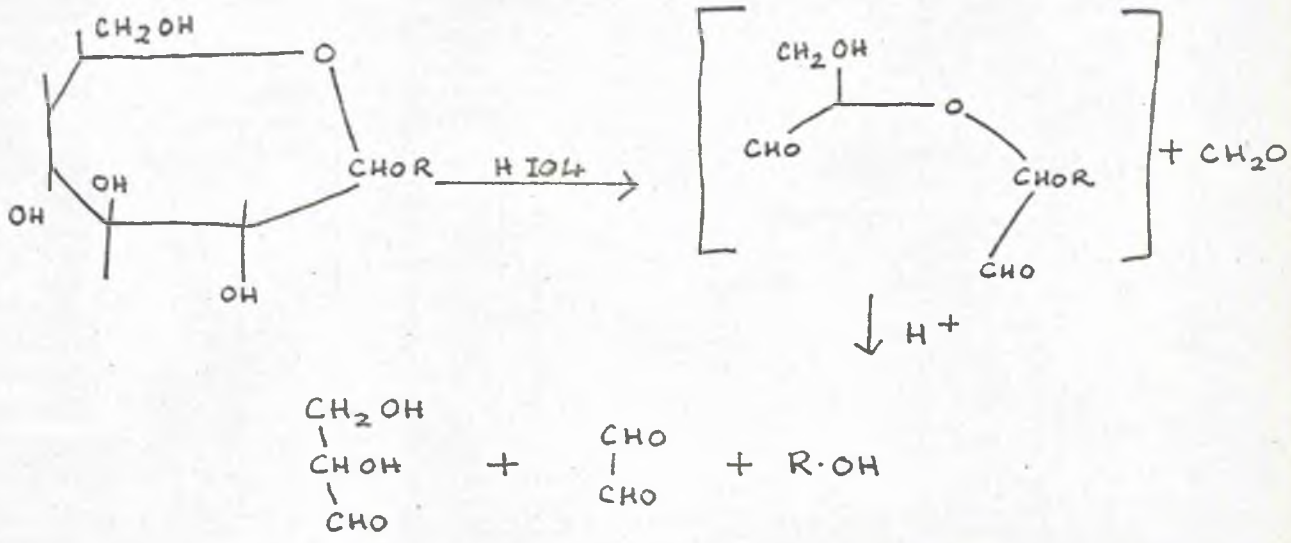
The possibility of ε-lysyl peptide bonds has been considered previously in connection with the availability of ε-lysyl residues to substitution reactions. That ε-lysyl residues may be concerned in linkages involving carbohydrate residues was also noted, and in this connection it is also worth mentioning a recent communication by Bensusan et. al. 1966. They further define their hypothesis for the role of hexose; thus an N-glycosidic linkage between the sugar molecule and the ε amino group of collagen is stabilised by the Amadori

Rearrangement, with the formation of 5-hydroxy methyl furfural. This can condense with the imidazole ring of histidine to give a crosslink with the formula:- 1256 tetra hydro pyrido - 2 - (5 hydroxy methyl - 2 - furyl) - 3, 4 - iminazole - 6 - carboxylic acid. This structure was found to be quite feasible from experiments using model compounds, but has not been isolated and can only be regarded as hypothetical.

That carbohydrates are involved in crosslinking is not a new idea, Grassmann and Schleich (1935) being probably the first to suggest it. The problem is fraught with difficulties because of the difficulties of purification of collagen from all mucopolysaccharide. Notwithstanding, Hörmann 1960 showed that mature insoluble collagen can be dissolved completely by periodate and a hydrogen bond breaking agent to give a high molecular weight product. This seems good evidence that periodic acid breaks the crosslinks, which must therefore contain carbohydrate.

Blumenfeld et. al. identified the products of periodate oxidation of the hexose of ichthyocol as 2:4 D.N.P. osazones. See figure 7. Glyoxal and glyceraldehyde were found, demonstrating that the hexose is bound via the C1 position, with C, 2, 3 and 4, free. It was also shown that the C6 position (not determined by

FIGURE 7. Periodate Oxidation of Hexose in Icthyocol.
(From Blumenfeld. et al. 1963.)



$$\begin{array}{l} CH_2OH \\ | \\ CHOH \\ | \\ CHO \end{array} + \begin{array}{l} CHO \\ | \\ CHO \end{array} + R\cdot OH$$

periodate oxidation) was free in galatose residues, by using the enzyme galatose oxidase. Thus the only carbohydrate functions free to occur in covalent linkage are: C1 of glucose and galatose and possibly C6 of glucose. However these experiments were performed with ichthyocol which has no intermolecular and few intramolecular crosslinks, thus intermolecular crosslinking could involve residues other than the above. The participation of galatose intramolecular crosslinks is more or less ruled out. The participation of hexoses in the subunit attachment sites of the ester-linked subunits (Blumenfeld et. al. 1963) is also virtually ruled out on the grounds of stoichiometry - α 1 chains contain only three hexose residues, but six ester linkages.

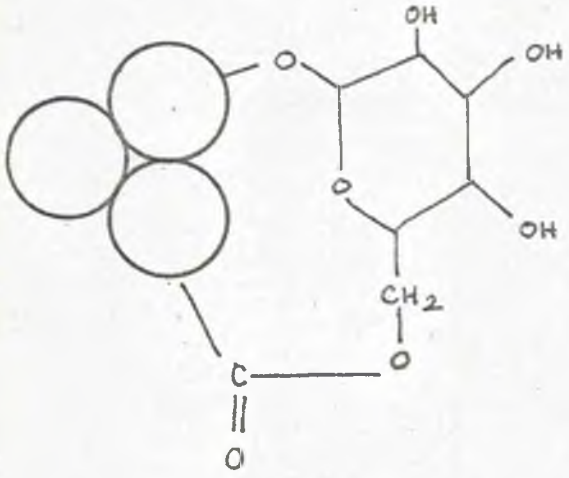
Hörmann 1960 showed that, although mature collagen and procollagen have similar amounts of hexose, insoluble collagen has a number of ester-like links more than procollagen which is approximately equal to the hexose content of either. He postulates that the extra ester linkages are involved in intermolecular crosslinks (see figure 3), and that intramolecular crosslinking also involves an ester-linkage.

Butler and Cunningham 1965 (see section on primary structure) isolated a glycopeptide from collagen in which

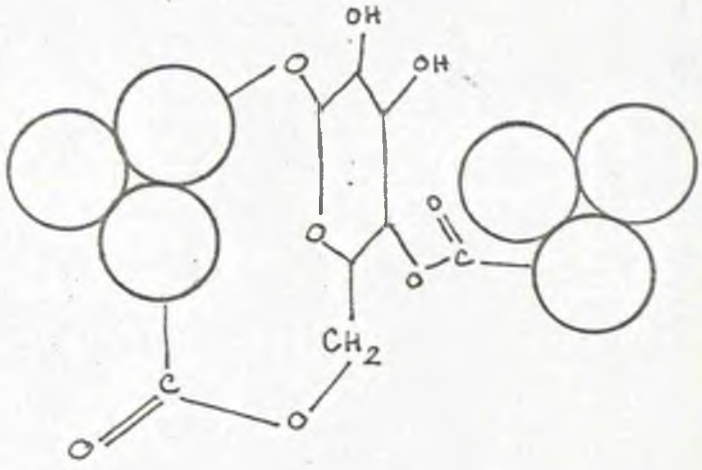
FIGURE 8.

Ester Crosslinking.

(From Hörmann 1960.)



procollagen
inter and intramolecular bonds.



collagen
inter and intramolecular bonds.

the hexose was bound to hydroxylysine by an O-glycosidic bond. This could account for the way in which hexose is attached to collagen. The ester-linkages forming attachments for intra and intermolecular crosslinking would then probably involve one of the free hydroxyl groups, and a carboxyl group of aspartic and glutamic acid. Hormann et. al. (1959) examined the periodate consumption of mature collagen, and assumed that the hexose must be linked either 1 - 2 or 1 - 4. No evidence is available as to the carboxyl donor in these ester linkages; investigations would be difficult, due to interference from the 'subunit' ester bonds of Gallop et. al. (1959).

Harding (1966) points out that if hexoses are involved in crosslinking, then the observation that insoluble collagen can be dissolved by proteolytic enzymes which presumably cleave intermolecular crosslinks, must be explained. Thus crosslinks must either contain a peptide chain or be situated at a position close to the enzyme - sensitive bonds of the polypeptide chain.

A further type of crosslink, present in rat skin collagen, has been proposed by Bornstein et. al. (1966). By cleavage of $\alpha 1$ and $\alpha 2$ chains and $\beta 12$ double chains with CNBr at methionine bonds, they obtained peptides in some of which the presence of lysyl - derived aldehydes could be demonstrated. From $\beta 12$ they isolated a peptide

corresponding to the sum of two peptides from α 1 and α2. It was postulated that the two peptides were linked in β 12 by an aldol-type condensation between two aldehyde functions. By experiments utilising lysine - C^{14} incorporation into the collagen, it was found that although all of the peptides participating in the crosslink possessed activity, none of them contained lysine when subjected to amino acid analysis. No amino acid other than lysine in the hydrolysates possessed activity. Thus it was concluded: a specific lysyl residue located in a non-helical region near the N - terminus of each α chain, participates in the formation of the intramolecular inter chain crosslink in collagen. The δ - semialdehyde of α - amino adipic acid is a probable intermediate.

Other data given by Bornstein et. al. (1966) to support this theory of crosslinking, concerns the use of the reagent β - amino proprionitrile (B.A.P.N.) as a lathyritic agent. Collagen from lathyritic animals yielded peptides which had the same structure as those from normal rats, except that one lysine residue was present. In the condition of lathyrism crosslinking is either prevented or destroyed. Bornstein et. al. suggest that lathyrism is due to the inhibition of an enzymatic process by which lysine in peptide linkage is

converted to the aldehyde. Recently Fessler and Bailey (1966) incubated gelatin obtained by mild denaturation of collagen with B.A.P.N. at neutral PH and 37°C. They observed the cleavage of β into α units by an apparently first order reaction mechanism, and concluded that only one crosslink per molecule was involved, or that, if there were several links, one was more stable than the others. It was also stated that these results are compatible with the linkages existing as either a Schiff's base or an aldol-like condensation product.

Rojkind et. al. (1966) localised an aldehyde component from tropocollagen which they believe may be involved in the maturation and crosslinking of soluble collagen to form insoluble tissue collagen. The aldehydic component was present in a heptapeptide having no free amino groups; this locates the aldehyde component within 7 residues of a blocked aminoterminus of a tropocollagen chain. It occurs in an amount of approximately 2eq./300,000 units of molecular weight, but little is known as to its structure and mode of attachment to the protein. Whether or not this aldehyde bears any relationship to the crosslinking agent of Bornstein et. al. (1966) is an interesting matter for conjecture. The possibility is not ruled out by amino acid analyses of the peptides concerned, see table 12.

When one looks at the vast complexity of data concerning the crosslinking and subunit structure of collagen, one is forced into the opinion that the chemistry of the molecule is extremely varied. Perhaps with time some of the possibilities which have been explored, will be excluded, for it does seem that a number of the hypotheses which have been proposed with respect to collagen, are not supported by very strong evidence.

Table 12. Amino acid composition of crosslinking peptide. (Bornstein et. al. 1966) and Aldehyde containing peptide. (Kefkind et. al. 1966)

	Crosslinking peptide ex. α 1.	Aldehyde containing peptide
Aspartic Acid	0.9	1.2
Serine	2.1	0.5
HomoSerine	0.9	-
Glutamic Acid	1.1	1.1
Proline	1.9	1.1
Glycine	3.1	2.9
Alamine	1.0	1.3
Valine	2.0	
Tyrosine	1.1	
Aldehyde	- *	present.

* present, but in this analysis represented by homoserine.

EXPERIMENTAL

Preperation of Collagen

1. Acid soluble collagen.

Acid soluble collagen was prepared according to the method of Steven and Tristram (1962 a), with some minor variations. A fresh calf-skin was, as far as possible, shaved free of hair and fatty tissue, using a scalpel. The skin was then cut into small pieces and minced, first coarsely and then finely. At this stage, fat was removed by performing three 3-hour extractions of the tissue with chloroform/methanol (70:30 by volume), with constant stirring.

The skin was homogenised in a large Waring Blendor in 0.2M $Na_2H PO_4$ solution. This procedure was repeated four times, approximately 50gm. wet weight of tissue being used with every litre of extracting solution. The homogenised material was isolated at each stage by centrifugation.

Extraction of the acid soluble collagen was achieved by stirring overnight at 5°C in 0.1M acetic acid. This process was repeated three times, and the collagen in the pooled supernatant liquids precipitated by the addition of 30% Na Cl solution to a final concentration of 5%. The collagen was collected by centrifugation, redissolved

in 0.1M acetic acid solution, and carefully salted out again by the gradual addition of 30% Na Cl with continual stirring. This purification by salt precipitation was repeated twice more. Finally, a solution of the collagen was dialysed against tap water in the presence of chloroform, for several days. The resultant precipitate was spun down and stored as such in a refrigerator.

2. Insoluble Collagen.

For this preparation, a sample of the insoluble residue after the extraction of acid soluble collagen, was taken. The material was lyophilised, and treated with liquid air. The frozen tissue could then be reduced to a fine powder, by rapidly pounding it in a pestle-and-mortar. (This method is described by P. Taylor, 1965 - Honours dissertation - University of St. Andrews).

The powdered material contained a proportion of hard granular pieces which were discarded, the majority of the tissue was in a very fine fibrous condition and was quite soft in texture.

Two extractions of the powdered insoluble collagen with 0.1M acetic acid, were performed, followed by an extraction with 0.2M Na_2HPO_4. The collagen was then suspended in distilled water and dialysed against 0.05M acetic acid for several days at 5°C. The dialysed material was finally centrifuged down, and freeze dried for storage.

(1) Investigations into the N-terminal residues of soluble collagen.

Dinitrophenylation

Dinitrophenylation of collagen was carried out in 50% alcoholic medium with excess sodium bicarbonate, and potassium chloride added to a total ionic strength of 0.3. Collagen was used as a solution of approximately 4mg./ml., and the total final volume was adjusted to approximately 500 mls. per gm. of protein. 1 fluoro 2: 4 dinitro benzene, (F.D.N.B.) 0.5 gms. (in solution in 20mls. ethanol) per gm. of collagen was added, and the reaction vessel shaken for 48 hours in the dark.

Dinitrophenyl collagen (D.N.P. - collagen) was isolated by centrifugation of the reaction products (using an international High Speed centrifuge - model HR-I). The derivative was washed three times with distilled water and five times with acetone IN HCl (99 : 1), dried in vacuo and weighed. This sample is referred to as DNP tropocollagen. Further samples were prepared in which the whole reaction mixture was acidified and taken to dryness by rotary evaporation, prior to hydrolysis according to the method of Steven and Tristram (1962 b.). This preparation is referred to as D.N.P. - total collagen.

Hydrolysis.

(i) Hydrolysis of the D.N.P. tropocollagen was

effected by transferring to a sealed tube with 25mls. of 5.7N HCl (redistilled over $SnCl_2$) per 500mg. derivative. The sealed tube was placed in an oven at 120°C for 10 hours.

(ii) Hydrolysis of D.N.P. - total collagen was performed by refluxing with 5.7N HCl according to the method of Steven and Tristram (1962 b.).

(iii) Samples of D.N.P. tropocollagen were also hydrolysed by refluxing in the presence of a 10-fold weight excess of Dowex - 50 resin (H^+ form) (Method of Steven 1962 a). Usually 250 mg. of D.N.P. collagen were refluxed with 50 mls. of resin-slurry for 48 hours.

Isolation of Ether-soluble D.N.P. amino acids.

The acid hydrolysates after cooling were diluted to approximately 1N HCl. and extracted three times with equal volumes of ether. The ether extracts were washed twice with a small volume of distilled water, and taken to dryness.

The resin hydrolysates were washed five times with hot water on a sintered-glass filter funnel. A few drops of concentrated HCl. were added to the washings which were then extracted with ether and taken to dryness as above.

The ether soluble D.N.P. amino acids from hydrolyses of D.N.P. tropocollagen, were chromatographed as such.

Before chromatographing the ether soluble fraction from hydrolyses of D.N.P. total collagen, however, it was necessary to remove the large amount of dinitrophenol and dinitroaniline formed as a by-product of the dinitrophenylation procedure. This was accomplished by the chromatographic technique of Steven (1962 b.): 5g. silicic acid (Mallinckrodt 100 mesh) is ground with 2.5 mls. 0.67M. Na_2HPO_4 and then made into a slurry with chloroform. The slurry is used to pack a 1 x 30 cm. column. The ether soluble D.N.P. amino acids are applied to the top of the column in solution in chloroform, and first of all elution is performed with chloroform equilibrated with 0.67 M. Na_2HPO_4. This elution carries off the artefacts leaving the D.N.P. amino acids adsorbed to the column; these are subsequently eluted with chloroform/glacial acetic acid (99 : I) and brought to dryness for chromatography.

Chromatography.

A two-dimensional system, using the toluene, 2 chloroethanol, pyridine, aqueous ammonia system of Biserte and Osteux (1951) for the first dimension, and 1.5M. phosphate buffer pH 6 in the second, was employed. Chromatograms were performed in the dark, care being taken always to equilibrate the papers in the presence of fresh 0.3 N ammonia, prior to running the first

dimension; and to ensure complete removal of the first solvent by heating to 50°C in a current of air in a darkened cupboard for 1 - 2 hours before the second development. When these conditions were observed, and so long as there was no overloading of the spot, complete separations of all D.N.P. - amino acids, except D.N.P. - aspartic acid and D.N.P. glutamic acid, and the D.N.P. leucines, were obtained.

The D.N.P. amino acids were estimated quantitatively by elution of the spots with 1% $NaHCO_3$ solution, and the optical density of the resultant solutions read at 360 m.u. in a spectrophotometer (Unicam S.P. 600). If the D.N.P. amino acids of any hydrolysate were not transferred quantitatively to a paper, the residue was taken up in a suitable volume of 1% Na HCO_3 solution and also read at 360 m.u., the value obtained being used to increase proportionately the results from the chromatogram to account for the whole hydrolysate. Molar extinction values for D.N.P. amino acids were obtained from Rao and Sober (1954).

Control Experiments.

For control hydrolyses a solution containing approximately 100 ug/ml. of each of certain D.N.P. amino acids (pure samples - Mann Assayed Chemicals) was made up. 1 ml. samples of this solution were included in

normal acid hydrolyses of 250 mg. samples of D.N.P. tropocollagen. Triplicate hydrolyses were performed; these were taken through and chromatographed as described above. To determine the overall recovery, the final amounts were reduced by the amounts obtained for hydrolyses of D.N.P. tropocollagen alone.

Estimation of Protein Quantities.

The weight of collagen represented by a certain weight of D.N.P. derivative was estimated by a correlation of values for total nitrogen and amide nitrogen of normal collagen, with amide nitrogen of D.N.P. tropocollagen and weights of freeze dried samples of the two materials.

Estimation of amide nitrogen was accomplished by hydrolysing approximately 70 mg. samples of protein in 2N HCl for 2 hours at 120°C in a sealed tube. The ammonia was liberated in a Micro Kjeldahl apparatus, Na OH solution being added until the hydrolysate was alkaline to thymol phthalein. Distillation was performed in the presence of octan - 2 - ol to prevent undue frothing of the solution.

In the case of the D.N.P. - total collagen, the initial quantity of collagen used gave the amount represented by the final derivative, for no preparative losses were involved, as necessarily occurred during the washing procedures for the D.N.P. tropocollagen.

Enzyme Treatment.

500 mg. samples of D.N.P. tropocollagen were brought to pH 1.9 with dilute HCl. in a total volume of 100 mls., 5 mg. crystalline pepsin (Armour Co.) was added, and the mixtures were stirred in the dark at 25°C overnight. The D.N.P. collagen following this treatment was washed thoroughly with distilled water. Initially, the washings were visibly yellow, because of the presence of soluble pepsin - liberated peptides, but the final washings contained no trace of such peptides. One sample was hydrolysed as such, and the other redinitrophenylated prior to hydrolysis. The ether soluble D.N.P. amino acids from these preparations were examined in the usual way.

Results.

Estimation of Collagen.

1. Intact collagen.

Total nitrogen: using a value of 18% as the nitrogen content of collagen it was estimated from triplicate determinations that the sample of freeze dried collagen contained 92% protein.

Amide nitrogen: freeze dried collagen = 0.56% amide N

$$\therefore \text{collagen} = 0.56 \times \frac{100}{92}$$

$$= 0.61\% \text{ amide nitrogen.}$$

2. D.N.P. Tropocollagen.

Amide nitrogen: D.N.P. tropocollagen = 0.57% amide nitrogen.

Thus to convert weights of freeze dried D.N.P. collagen to their equivalent weights of collagen a factor of 57/61 = 0.95 was used.

On the basis that collagen contains approximately 41 residues/100,000gms. which will bind one D.N.P. residue each viz. 26 lysine, 7 hydroxylysine, 5 histidine, and 3 tyrosine, a theoretical correction factor of 0.94 is obtained, as opposed to the observed figure of 0.95. It would be expected that the freeze-dried D.N.P. derivative would be slightly heavier than the value calculated from increase in molecular weight of the anhydrous protein, due to any ash-content and the presence of water not removable by the freeze-drying procedure. Perhaps these two sources of error are not so marked in the altered protein, water and ions being bound less strongly than in the native protein. However the correction factor itself can only be fairly approximate, for the figure of 18% for the nitrogen content of collagen (an average value from reliable amino acid analyses) may not be quite accurate for this preparation.

N-terminal residues of collagen derivatives.

Table 13. N-terminal amino acids of D.N.P. tropocollagen, expressed as moles/100,000 gms. collagen.

	D.N.P.Tropo-collagen uncorrected	% overall recovery control	D.N.P.Tropo-collagen corrected	D.N.P.Tropo-collagen - resin hydrolysis *
Glycine	0.031	30%	0.103	0.12
Aspartic acid	0.0085)	60%	0.016	0.022
Glutamic acid	trace)		trace	-
Alamine	0.0050	70%	0.0071	0.013
Serine	0.0005	65%	0.0007	0.0015
Threonine	trace	65%	trace	0.0005
			0.13	0.16

* The recovery figures of Steven (1962 b) were used for correction.

Table 14. N-terminal residues of D.N.P. - total collagen. (Corrected Values) (as moles/100,000 gms. collagen).

Glycine	0.120
Aspartic acid	0.030
Glutamic acid	0.0005
Alamine	0.019
Serine	0.025
Threonine	0.0008
Leucine + isoleucine	trace
Phenylalanine	0.001
Valine	trace
Tyrosine	trace
Proline	trace
Hydroxyproline	trace
Lysine	0.001
	0.20

Table 15. N-terminal residues of pepsin treated D.N.P. tropocollagen and redinitrophenylated pepsin treated D.N.P. tropocollagen. (Corrected Values). (as moles/100,000 gms. collagen).

	pepsin treated	pepsin treated re-dinitrophenylated	D.N.P. Tropocollagen. (Control)
Glycine	0.099	0.220	0.103
Aspartic acid	0.016	0.025	0.016
Glutamic acid	-	0.0005	trace
Alanine	0.005	0.008	0.0071
Serine	0.001	0.001	0.0007
Threonine	trace	trace	trace
Leucine	-	trace	-
	0.12	0.25	0.13

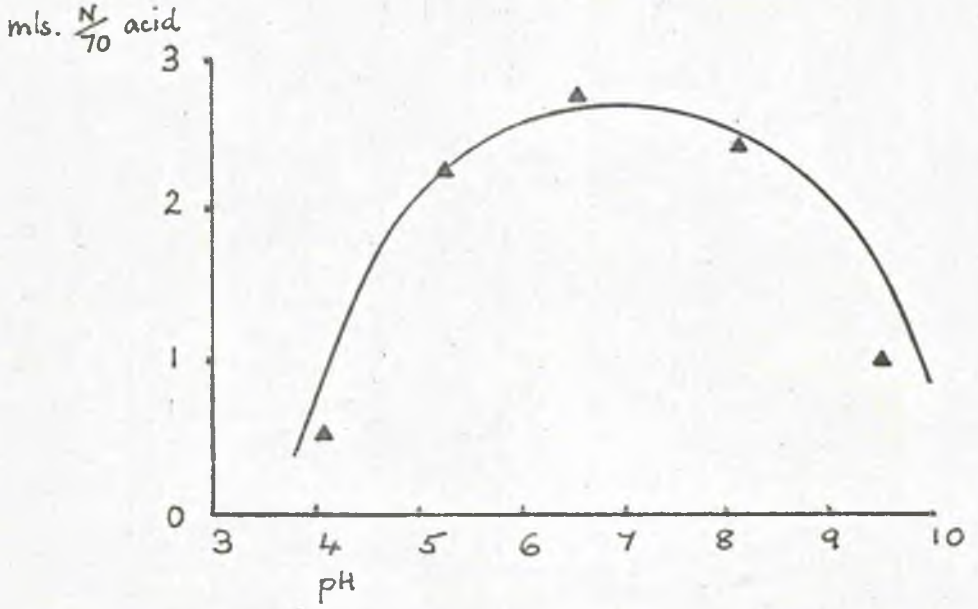
2. Investigations into the activity of α-amylase on collagen.

Solubilisation of mature collagen by treatment with α-amylase.

Optimum pH.

100 mg. samples of powdered insoluble collagen were suspended in buffers in the pH range 4-9. Phosphate buffers were used with the addition of HCl to reduce the pH below 5, and of Na OH to increase the pH to above 8. Incubation with 1 mg. α-amylase (hog pancreas) per 100 mg. collagen, at 25°C with shaking for one hour, was carried out. The collagen was then centrifuged down, and the pH of the supernatant liquid noted. The precipitates were washed once with distilled water and then suspended in 50 mls. 0.1 N acetic acid and allowed to shake overnight at room temperature. The acid extracts were centrifuged and 5 ml. portions of the supernatant liquids removed for total nitrogen determinations by the Micro Kjeldahl technique. A graph was plotted showing the variation in solubility, with the pH of the initial amylase pretreatment of the samples of insoluble collagen. The estimate of solubility used was simply the number of mls. of N/70 H Cl. required for titration during the total nitrogen determination performed on the supernatant after extraction with dilute acetic acid. The pH values used were those obtained for the supernatant liquids after amylase pretreatment of each sample. See figure 9.

FIGURE 9. Optimum pH for collagen solubilisation by α amylase.



Preparation of soluble collagen by amylase treatment of insoluble collagen.

Insoluble collagen was digested with α - amylase (hog pancreas), at pH 6.3, and an enzyme; substrate ratio of 1: 500. The reaction mixture was stirred overnight at room-temperature, in the presence of a few drops of merthiolate. The collagen was collected by centrifugation. After washing the collagen three times with distilled water, it was extracted with 0.1 M acetic acid, during which treatment the majority of the protein went into solution. The solution was centrifuged, and after discarding the sediment the supernatant liquid was precipitated, and purified, by salt precipitation and dialysis as described in the preparation of normal acid soluble collagen. The material was stored in freeze-dried condition and will be referred to as "amylase solubilised collagen".

Carbohydrate content of amylase solubilised and normal collagen.

Carbohydrate was assayed by a modification of the method described by R.M. Howell et. al. (1964). Approximately 10 mg. samples of finely divided, freeze-dried material were suspended in 1 ml. of water in 50 ml. test tubes. 4 mls. of 0.15% anthrone solution in 75% sulphuric acid were pipetted carefully into each tube, while the tubes were being cooled in an ice bath. The two layers

were mixed, the tubes being kept as cool as possible during the process by maintaining contact with the cooling medium. Once all the tubes were mixed they were immersed in a water bath at 90°C for 10 minutes, cooled, and the solutions read at 620 m.u. against a blank prepared by taking 1 ml. of water through the same procedure. Standard solutions of glucose, at concentrations 25, 50 and 100 ug/ml. were taken through, and a standardisation graph constructed. Triplicate samples of acid soluble, insoluble and amylase solubilised collagen were assayed.

Results.

Table 18. Hexose content of collagens.

Protein	Total Hexoses %
Acid soluble	0.62
insoluble	0.70
Amylase solubilised	0.36

Assay for the proteolytic activity of the enzyme. (Method of Davis & Smith (1955)).

Materials.

Freeze dried acid soluble calf skin collagen;

Freeze dried powdered insoluble calf skin collagen;

Crude bacterial α-amylase (Cambrian chemicals);

α-amylase extract of hog pancreas, freeze dried enzyme cited "free of proteolytic enzymes." (B.D.H.);

Haemoglobin, purified (B.D.H.)

Phenol reagent of Folin and Ciocalteu (B.D.H.)

Citrate-phosphate buffer pH. 6-8.

1. Crude bacterial α-amylase.

A 1% solution of Haemoglobin in the buffer was prepared. An approximately 1% suspension of acid soluble collagen was made up, by allowing 500 mg. of the collagen to swell in 50 mls. of the buffer, and then homogenising thoroughly. 5 ml. samples of these preparations were pipetted into 50 ml. tubes. For the collagen suspension, a broken pipette with a wide tip was used, and the suspension was shaken before each withdrawal.

1 ml. samples of a solution of 1 gm. of the crude enzyme in 25 mls. of the buffer were used. The solution was centrifuged prior to use, to sediment the considerable proportion of insoluble material. The supernatant solution was brown in colour.

The tubes containing protein were equilibrated at 35°C in a water bath, and 1 ml. of the enzyme solution was added to them all. At appropriate intervals of time, pairs of tubes, one containing haemoglobin and the other collagen, were removed, and 10 mls. of 0.3 M. trichloroacetic acid added with thorough mixing and warming. The solutions were filtered through Whatmann 3 mm. filter paper.

10 mls. of 0.5 M. NaOH and 3 mls. of diluted phenol reagent (1 vol. reagent with 2 vols. H_2O) were added to 5 mls. of the trichloroacetic acid filtrate. After 10 minutes the solutions were read against a standard solution

of tyrosine reacted with Folin's reagent in the same way, at 750 m.u.

A blank determination was performed in which the enzyme sample was added to 10 mls. of trichloroacetic acid solution, followed by mixing with the protein. The extinction coefficient of the blank was subtracted from that of the unknown solutions.

Results.

Table 16. Proteolytic activity of bacterial α-amylase. (See figure 10).

	optical density 750 m.u.	
	Haemoglobin	Soluble Collagen
2 min.	0.04	0.03
5 min.	0.17	0.05
25 min.	0.35	0.13
120 min.	0.71	0.26

2. α-amylase ex. hog pancreas.

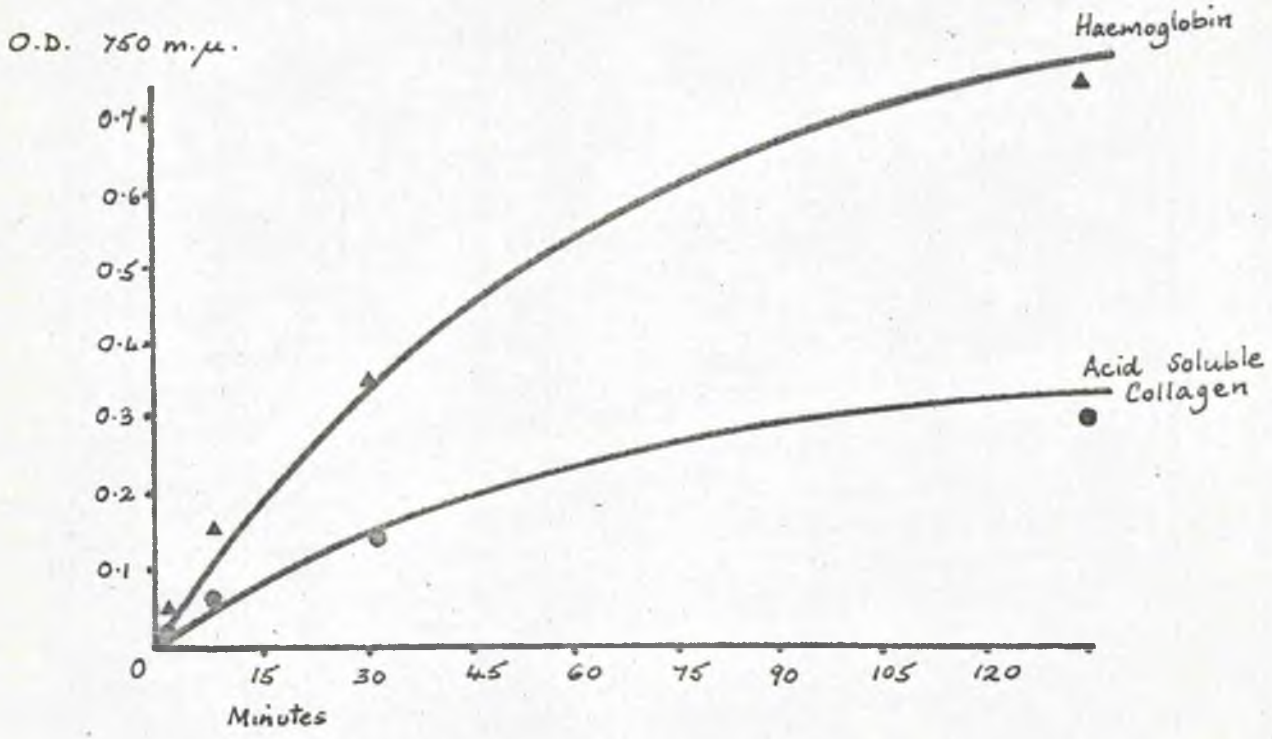
This assay was performed as above except that an enzyme concentration of 1.5 mg./ml. was used. Also, another series of tubes was taken through, containing 50 mg. samples of insoluble collagen, prepared as a suspension as described for acid soluble collagen.

Results.

Table 17. Proteolytic activity of pancreatic α-amylase optical density at 750 m.u.

	soluble collagen	insoluble collagen	Haemoglobin
5 min.	0.03	0.13	0.08
25 min.	0.06	0.48	0.18
1 hr.	0.09	0.63	0.25
2 hr.	0.09	0.78	0.35

FIGURE 10. Action of crude α amylase on Haemoglobin and Collagen.



See figure 11.

A standard solution of tyrosine containing 8×10^{-4} milli equivalents per 5 mls. and taken through the same procedure as the enzyme assays, gave an optical density of 0.18. On this basis, from the initial slopes of the graphs:-

(a) With experimental conditions used, crude bacterial amylase liberated colour equivalent to:

from Haemoglobin, 1.59×10^{-4} millimoles tyrosine/minute/40 mg. enzyme

and from acid soluble collagen,

3.18×10^{-5} millimoles tyrosine/minute/40 mg. enzyme.

(b) With experimental conditions used, amylase from hog pancreas liberates colour equivalent to;

from Haemoglobin,

4.78×10^{-5} millimoles tyrosine/minute/1.5 mg. enzyme.

from acid soluble collagen,

1.52×10^{-5} millimoles tyrosine/minute/1.5 mg. enzyme.

and from insoluble collagen,

1.32×10^{-4} millimoles tyrosine/minute/1.5 mg. enzyme.

<u>Action of pepsin on amylase solubilised collagen</u>.

Suspensions containing approximately 6 mg. amylase-solubilised and normal acid soluble collagen per ml. were prepared by homogenising the freeze-dried material at neutral pH in the absence of acetate or citrate. 4 ml.

FIGURE 11 Action of α amylase ex. hog-pancreas.

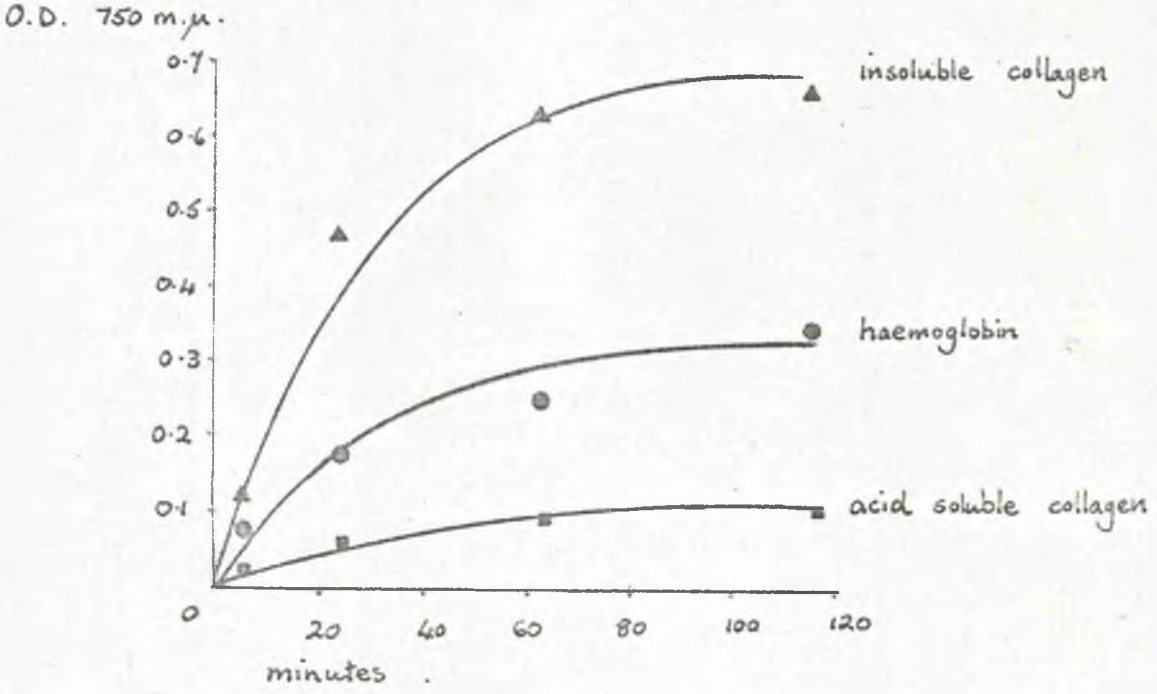
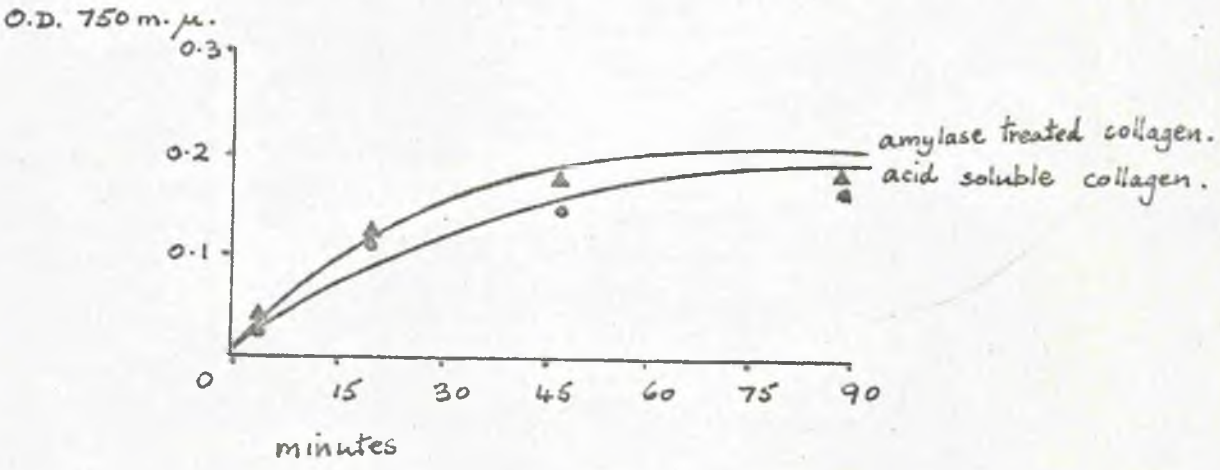


FIGURE 12. Activity of pepsin on amylase treated collagen and acid soluble collagen.



samples of the suspensions were pipetted into test tubes and the pH reduced to pH 1.8 - 2 by the addition of 1 ml. of 0.05 N HCl. 1 ml. of a solution containing approximately 0.2 mg./ml. of crystalline pepsin (Armour Co.) was added to each of the tubes which were incubated at 35°C and assayed for peptide release as described for the estimation of proteolytic activity of ∝-amylase.

Results.

Table 18. Action of Pepsin. (See figure 12).

	amylase solubilised	acid soluble
5 min.	0.035	0.02
25 min.	0.13	0.12
1 hr.	0.17	0.14
2 hrs.	0.19	0.17

Column-fractionation of amylase solubilised collagen.

Chromatography was carried out on a 1.5 x 12" column packed with carboxymethyl cellulose (Bio-rad; Cellex-C.M.) and fitted with a water jacket to maintain the column at 40°C. Gradient elution was performed, using two chambers of an "autograd" (Technicon Ltd.), the first of which contained 0.08 M sodium acetate buffer pH 4.3, and the second, the same buffer, but brought to a total ionic strength of 0.2 by the addition of sodium chloride. Thus a gradient was obtained from ionic strength of 0.08 to

0.20.

20 mls. of a solution of collagen, denatured at 40°C for 30 minutes, and containing approximately 30 mg./ml. were applied to the surface of the column. The protein solution was "pumped-on" to the column before gradient elution was commenced. A flow rate of 200 mls. was chosen by suitable adjustment of the pump (D.C.L. micro-pump). 8 ml. samples were taken, using an automatic fraction collector.

After elution with the gradient (600 mls.), 200 mls. of 0.5 M sodium acetate buffer pH 4.8 were pumped through, followed by 200 mls. of 0.1 M NaOH, containing NaCl to a total ionic strength of 0.5.

The effluent fractions were analysed for protein, using a modification of the Folin-Lowry method, employing an autoanalyser (Technicon Instruments Ltd.).

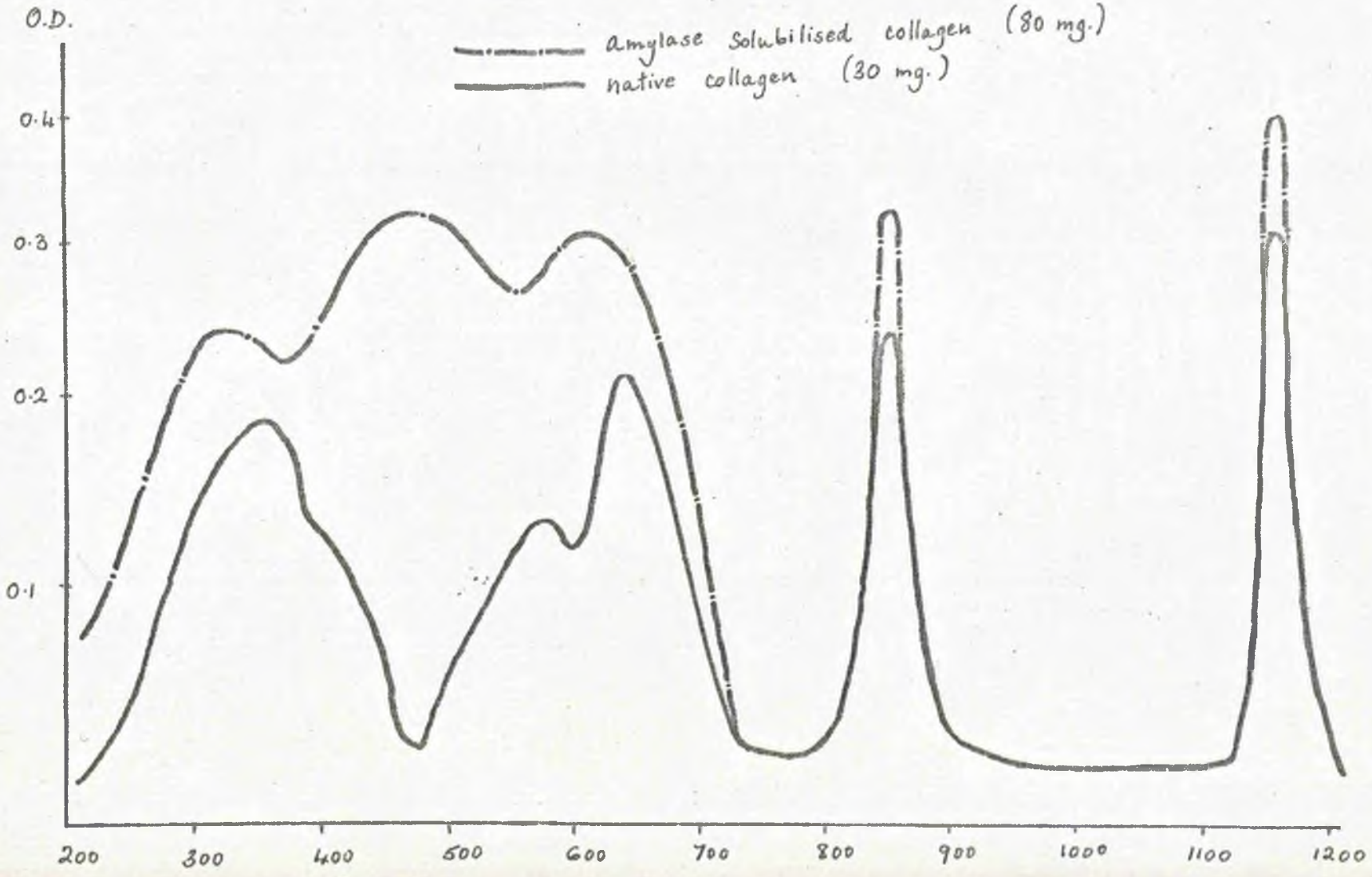
A diagram of typical traces obtained with normal acid soluble collagen, and amylase extracted collagen is shown in figure 13.

N-terminal residues of amylase solubilised collagen.

Amylase solubilised collagen was dinitrophenylated in the same way as described previously for acid soluble collagen. The arylated derivative was washed with water and acetone/N Cl to give a produce corresponding to D.N.P. - tropocollagen. Hydrolysis was effected in a sealed tube with 5.7N H Cl. Ether extraction and chromatography were carried out as in Section 1.

FIGURE 13.

Chromatography of collagen on carboxy methyl cellulose at 40° C.



Results.

Table 19. N-terminal residues of amylase solubilised collagen.
(Corrected values) (as moles/100,000 gms. collagen).

glycine	0.063
aspartic acid	0.016
glutamic acid	0.001
alanine	0.012
serine	0.0045
threonine	0.0005
leucine + isoleucine	0.0084
phenylalamine	trace
valine	0.008
lysine	trace
	0.13

Ultracentrifugation of solutions of amylase solubilised collagen.

Ultracentrifuge work was performed using a Spinco Model E analytical ultracentrifuge, and using the sedimentation velocity method. Solutions of amylase solubilised collagen in 0.25 M. sodium acetate buffer pH 4.3, at different protein concentrations, were examined, and the runs were performed at 40°C to show the presence of the different subunits of heat denatured collagen. The values for sedimentation coefficient calculated for different solutions were plotted against concentration, and the graph extrapolated to give the sedimentation coefficient at infinite dilution. This observed sedimentation coefficient was corrected to the value which would be theoretically obtained in a solvent

with the density and viscosity of water at 20°C, by use of the equation:-

$$S_{20w} = S_{obs}\left(\frac{\eta_t}{\eta_{20}}\right)\left(\frac{\eta}{\eta_0}\right)\left(\frac{1-\bar{v}.\rho_{20w}}{1-\bar{v}.\rho_t}\right).$$

where $\frac{\eta t}{\eta 20}$ = viscosity of water at t° relative to that at 20°C

$\frac{\eta}{\eta o}$ = relative viscosity of the solvent to that of water

P20W and Pt are the densities of water at 20° and the solvents at t° respectively.

$\bar{v}$ is the partial specific volume of collagen.

Viscosity measurements were carried out, using a capillary viscometer, in a water bath the temperature of which was controlled to ± 0.005°C.

A value for the partial specific volume of collagen was obtained from the literature, the figure of 0.655 ml./g. of Davison and Drake (1966) being used.

Results.

A single uniform peak was obtained, which did not split at any stage.

From graph:- Observed sedimentation coefficient at infinite dilution = 4.62 S

Corrected coefficient S_{20w} = 3.25 S

Investigation of the low molecular weight products of amylase digestion of collagen.

3 gms. of insoluble collagen were stirred at room temperature with 10 mgs. of the pancreatic α-amylase at pH 6.8 overnight. The entire digest was transferred to a dialysis sac, and dialysed against three changes of 1.5 litres of distilled water in the cold for several days. The dialysate was concentrated by rotary evaporation, and a total nitrogen determination performed on a portion of the concentrate. Separate control dialyses of 10 mg. amylase and 1 gm. of insoluble collagen were carried out, and the dialysates obtained, examined for total nitrogen in the same way.

The dialysed material from the amylase digest was desalted using a column of ion-exchange resin. The solution was applied to the surface of a 2 x 10 cm. column of Dowex - 50 resin in the H^+ form, the column having previously been washed free of excess acid. Elution was commenced with 200 mls. distilled water: the first 30 mls. of effluent from the column being retained and concentrated to dryness at room temperature (Fraction A.). 10% aqueous ammonia (A.R. reagent) was then passed through and the first 20 mls. of alkaline effluent was collected. This was taken to dryness by rotary evaporation and washed once with distilled water (Fraction B.).

Fraction A was examined by thin layer chromatography

in the system, Butanol : acetic acid : water (30 : 20 : 20). Acid present in this fraction (anions are not bound by Dowex - 50 which is a cation exchange resin), does not interfere very seriously with developments in solvent systems of the type Butanol : acetic acid : water, because of the high proportion of acid already present. The chromatogram was sprayed with aniline phthalate reagent, and examined for spots corresponding to carbohydrates.

Fraction B was subjected to 2-dimensional thin layer chromatography in the systems -

(1) Butanol : acetic acid : water (30 : 20 : 20)

(2) Phenol : water (75 : 25)

Visualisation of spots was effected by spraying with ninhydrin reagent. A portion of Fraction B was hydrolised with 6N. HCl in a sealed tube at 120°C for 18 hrs., and chromatographed in the same way.

Results.

Total Nitrogen:-

50% of the dialysed material from the amylase digest was taken, this gave a titration of 0.89 mls. N/70 acid.

∴ Nitrogen liberated by amylase from 3 gms. of insoluble collagen = 0.356 mg. N.

Control amylase dialysate gave 0.25 mls. = 0.05 mg. N.

Control collagen dialysate gave 0.05 mls./gm. insoluble collagen = 0.03 mg. N/3gms.

∴ Nitrogen liberated after correction = 0.286 mg N/ 3gms. collagen.

Assuming a value of 18% for the nitrogen content of the amino acids and peptides represented by this value of 0.286 mg.N/3gms. collagen, the weight of amino acid + peptide liberated = 1.59 mg., corresponding to approximately 2 residues per tropocollagen molecule.

Chromatography.

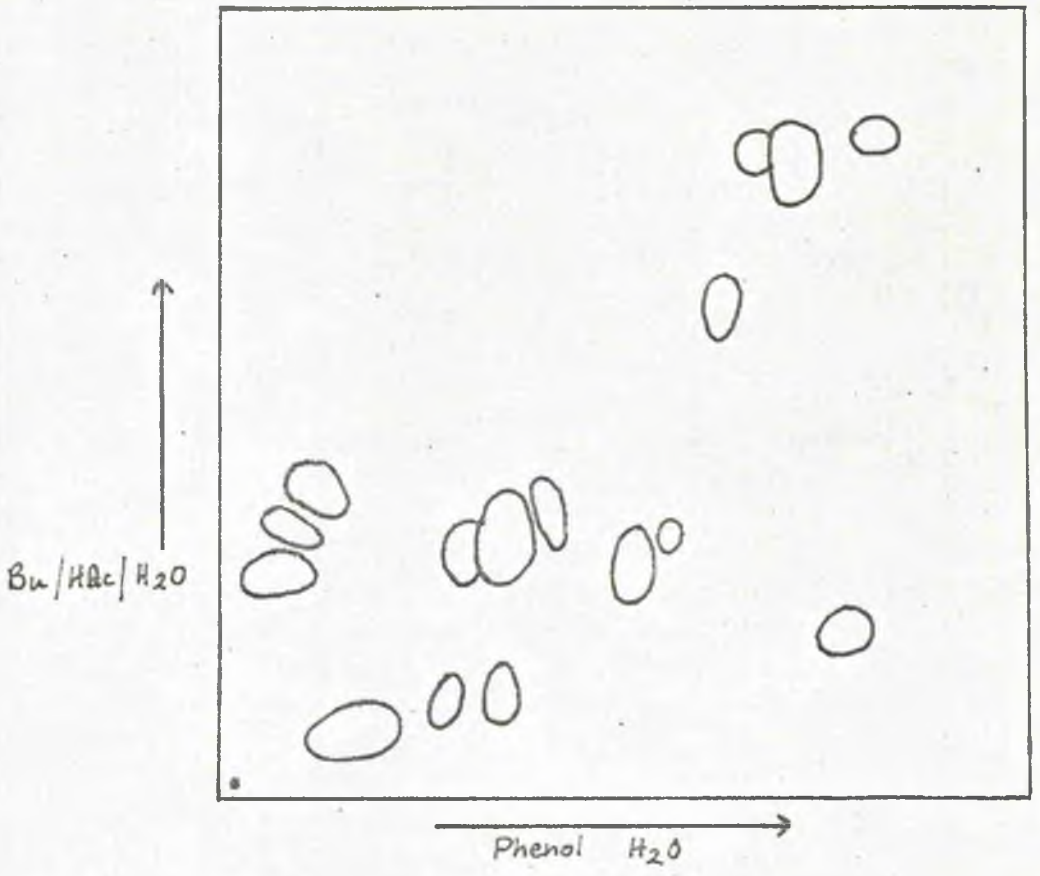
Fraction A: Two spots were observed upon staining with aniline phthalate reagent, the Rf values of which corresponded to glucose and galactose.

Fraction B: A complex pattern of spots was observed upon staining with ninhydrin, the majority of which corresponded in position to amino acids. (See figure 14). The hydrolysate of Fraction B showed the following amino acids when subjected to 2 dimensional chrometography in the same way:-

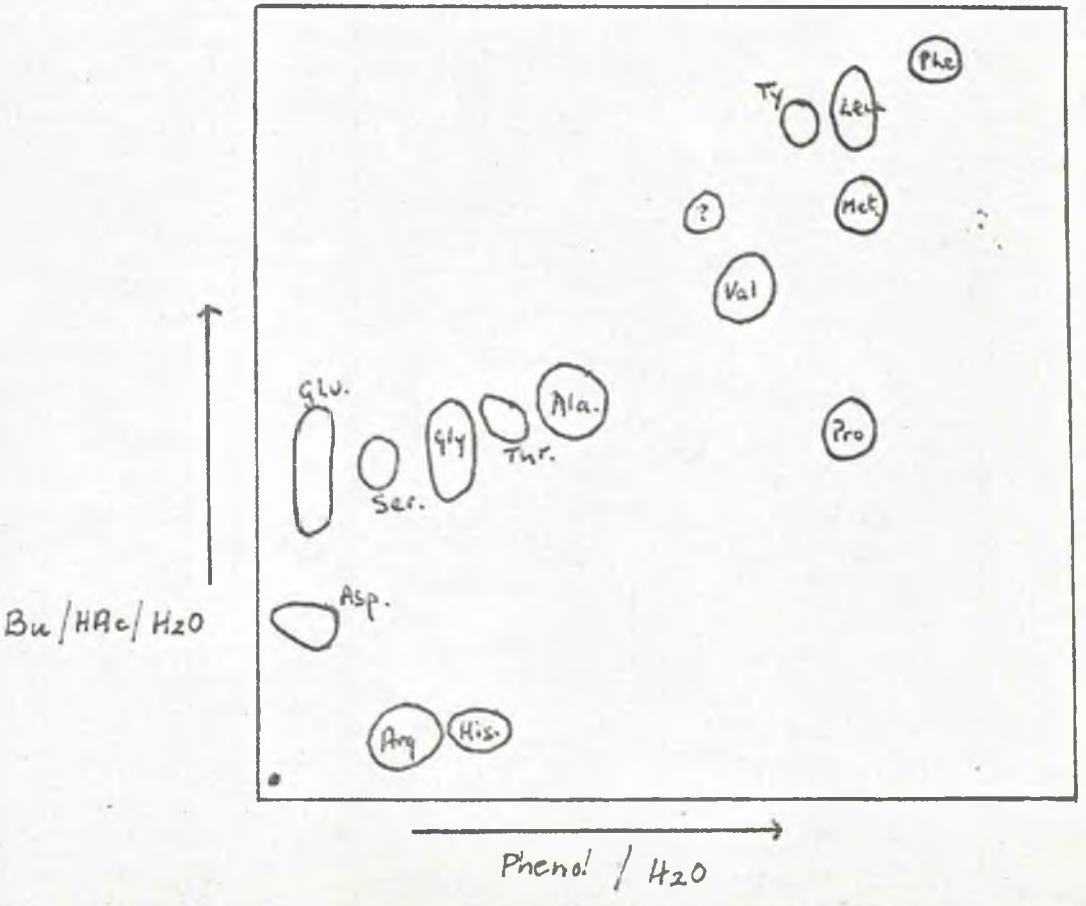
Leucine	Phenylalanine	)
Tyrosine	Methionine	) faint
Valine	Threonine	)
Alanine		
Glycine		
Serine		
Glutamic acid		
Aspartic acid		
Proline		
Arginine		

One unidentified spot appeared which travelled close to Valine.

FIGURE 14. Chromatography of amylase liberated peptides.



Amino acids from hydrolysate of the peptides.



3. The availability of ε-lysyl residues to dinitrophenylation.

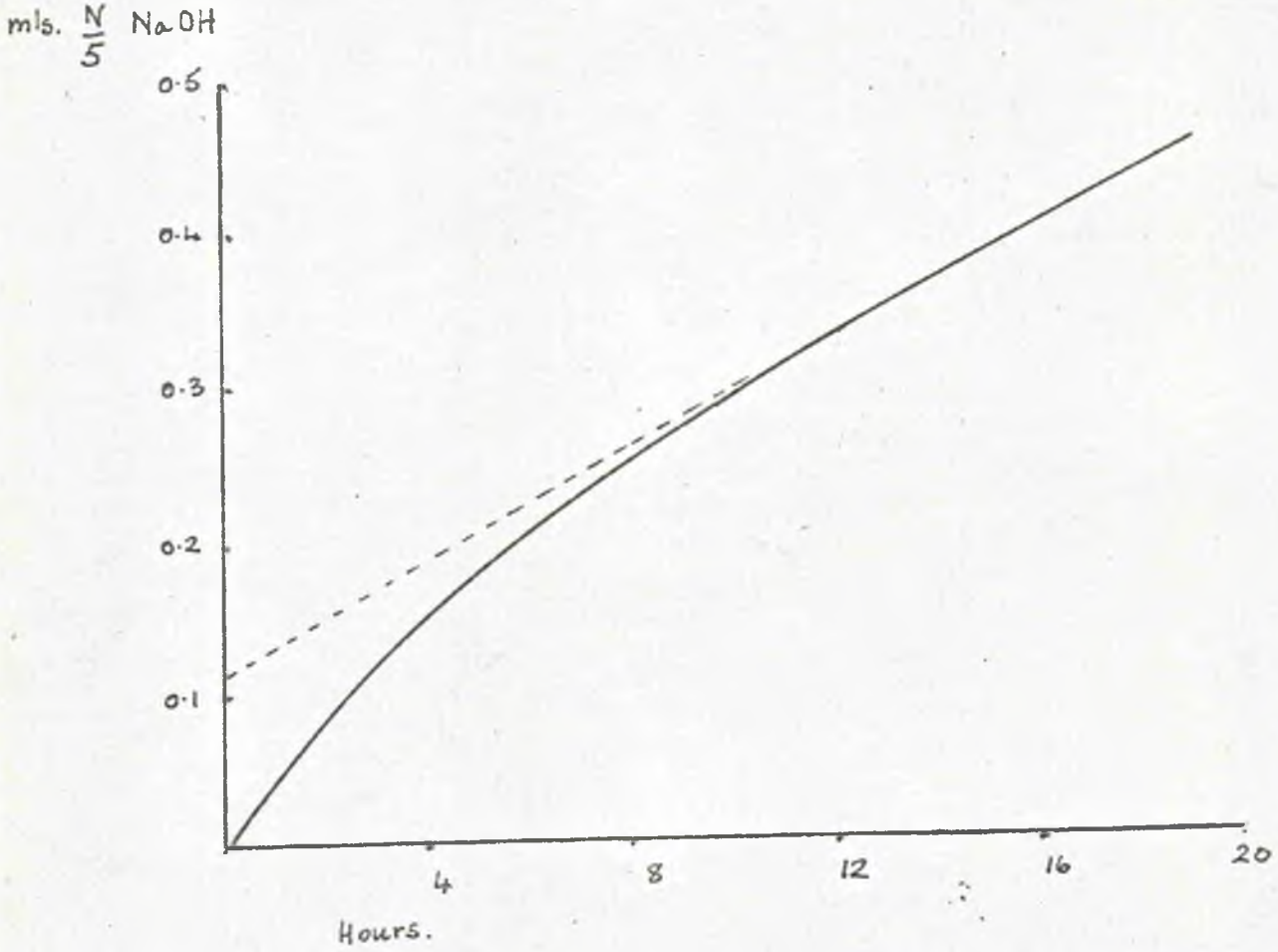
Dinitrophenylation of heat-denatured collagen using a pH-stat.

50 mg. of acid soluble calf skin collagen was dinitrophenylated at pH 7.5 and at 40°C in an aqueous medium ionic strength 0.3. The reaction was performed using a pH-stat (Radiometer), the electrode vessel being fitted with a water jacket to produce the elevated temperature. A gentle stream of nitrogen was introduced at the surface of the reaction mixture, to eliminate CO_2. A graph was obtained of consumption of alkali against time; the final gradient of the graph gave the reaction due to hydrolysis of the 1-fluoro 2:4 dinitrobenzene reagent; extrapolation of this slope to zero time gave the consumption of alkali due to the H.F. liberated from the reaction between the F.D.N.B. and protein. (See figure 15).

The D.N.P. derivative obtained was dialysed versus distilled H_2O for 2 days, and then freeze dried. The sample was weighed and then hydrolysed for 12hrs. with 5mls. 5.7 N redistilled H Cl in a sealed tube under nitrogen at 120°C. The hydrolysis vessel used in this experiment and in all subsequent experiments unless otherwise stated, was a pyrex test tube with a screw-on cap, fitted with a teflon-disc. The lip of the pyrex

FIGURE 15.

Titration of the reaction of 50 mg. Collagen with excess F. D. N. B.



tube was rubbed on a block of carborundum to give a very flat surface which, when in contact with the teflon-disc, provided an efficient seal. The hydrolysate was made up to 500mls. with distilled water, and a 50 ml. sample taken. This was extracted twice with ether, and its optical density at 360 m.u. measured, after blowing off the dissolved ether with a current of air for 1 minute. The amount of ε-D.N.P. lysine and ε-D.N.P. hydroxy-lysine was assessed by comparison with a standardisation graph, constructed by measuring the optical density of solutions of ε-D.N.P. lysine in 0.1 N. H Cl. For these purposes the molar extinction of ε-D.N.P. hydroxylysine and ε-D.N.P. lysine were assumed to be the same.

Matheson (1966) has reported that imidazole D.N.P. histidine is virtually colourless in the spectral range where the absorption of the other D.N.P. - amino acids is measured. The same author points out that this derivative is probably rather unstable to acid hydrolysis. O-D.N.P. tyrosine has considerably reduced absorption at 360 m.u. (Porter 1957). No correction was made for O-D.N.P. tyrosine and imidazole D.N.P. histidine in this experiment but an allowance was made in the next one for the estimation of the total amount of arylated ε amino groups. (See results for that experiment).

Hydrolytic release of ε-D.N.P. lysine from D.N.P. collagen: the total amount of arylated ε amino groups and their destruction during acid hydrolysis.

1 gm. of acid soluble collagen was dinitrophenylated in aqueous medium at 40°C, ionic strength 0.3 and in the presence of excess $NaHCO_3$. The reaction was carried out in a flat-bottomed 1-litre flask which was supported in a water bath at 40°C. The contents of the flask were stirred continuously, and light was excluded by means of a black cloth. Reaction was allowed to proceed for 15 hours.

There was a considerable proportion of soluble D.N.P. protein in the final reaction mixture; this was precipitated by cooling and reducing the pH to 3 -4 with H Cl. After standing in a refrigerator for 1 day a fine yellow precipitate was found which could be centrifuged down along with the initially insoluble D.N.P. protein. The precipitate was washed with distilled water and ether, and freeze dried.

Approximately 50 mg. samples of the D.N.P. derivative were accurately weighed into hydrolysis tubes. 5 mls. of 5.7 N H Cl redistilled were added to each, and the tubes were transferred to an oven at 120°C. The samples were exposed to hydrolysis for periods of from 5 hrs. to 14 hrs., and the final hydrolysate assayed for ε-D.N.P.-lysine and ε D.N.P. hydroxylysine, as before.

A control, in which a known weight of ε D.N.P. lysine was hydrolysed in the presence of 50 mg. of D.N.P. collagen, was performed. 1 ml. of a solution containing 1.5 mg. ε D.N.P. lysine per ml. of 0.1 N H Cl, was pipetted into a hydrolysis tube, and the solution was taken to dryness in vacuo overnight. 50 mg. of D.N.P. collagen was added, and 5 mls. of 5.7 N H Cl, and the mixture was hydrolysed for 10 hours. The optical density of the hydrolysate was measured, and the recovery value estimated by deducting the value for a corresponding hydrolysate of 50 mg. D.N.P. collagen alone.

The hydrolysates after measurement of optical density were taken to dryness in a rotary evaporator, and washed free of H Cl. The concentrates were examined for the presence of D.N.P. - peptides by high voltage electrophoresis at pH 1.9, pH 3.5 and pH 8.6; and by thin layer chromatography in the system: chloroform, tertiary amyl alcohol, acetic acid (80 : 20 : 1). Control samples of ε D.N.P. lysine (Mann Research Laboratories Inc. New York) and ε D.N.P. hydroxylysine (synthesised from hydroxylysine - Sigma Chemicals Co. Ltd., method of preparation below) were used. Electrophoresis runs were performed in a Pherograph High Voltage Electrophoresis apparatus, at 2000 volts for pH 1.9 and pH 3.5, and at 1,500 volts for pH 8.6. The duration of the runs was usually 30 minutes. The buffers employed were:-

pH 1.9 150 mls. glacial acetic acid) in 1 litre
50 mls. formic acid)

pH 3.5 100 mls. glacial acetic acid) in 1 litre
10 mls. pyridine)

pH 8.6 12 g. Veronal were dissolved in 23 mls. hot 2N Na OH, the pH was adjusted with 2N NaOH after cooling, and the solution was made up to 1 litre.

In another experiment, the relative amounts of the soluble and insoluble fractions of the D.N.P. parent gelatin obtained by dinitrophenylating collagen at 40°C, was estimated, by separating the precipitate and supernatant after reaction, dialysing against distilled water, freeze drying and weighing.

Preparation of ε-D.N.P. hydroxylysine.

The method of Porter and Sanger (1948) for the preparation of ε-D.N.P. lysine was adapted to prepare ε-D.N.P. hydroxylysine. 100 mg. of hydroxylysine (the sample was a mixture, containing some allo-hydroxylysine) was dissolved in 5 mls. of H_2O and $CuCo_3$ added to the boiling solution until no more dissolved. The excess $CuCo_3$ was filtered from the solution and washed with 2 mls. water. Excess $NaH\ CO_3$ was added and 0.5 g. F.D.N.B. dissolved in 10 mls. ethanol. The mixture was

shaken for 2 hours at room temperature and the precipitate filtered off, washed with water, ethanol and ether. Copper ions were removed by suspending in 5 mls. H_2O, adding enough H Cl to produce a clear solution, and after cooling in ice, bubbling H_2S through for 2 minutes. Charcoal was added and the mixture filtered. The filtrate was taken to dryness and recrystallised from water.

The spectrum of a solution of the derivative was recorded using a Unicam S.P. 800 Ultra Violet Spectrophotometer, and this was compared with Spectra of E-D.N.P. lysine, Di D.N.P. lysine, and D.N.P. threonine. The major peak of the derivative occurred at 360 m.u. as compared with 360 m.u. for E D.N.P. lysine, 355 m.u. for Di-D.N.P. lysine, and 350 m.u. for D.N.P. threonine (solutions made up in 0.1 N H Cl). This would indicate that the sample of E-D.N.P. hydroxylysine was authentic.

The E-D.N.P. hydroxylysine preparation was subjected to thin layer chromatography in the system: Butanol: acetic acid: water (80 : 20 : 20). Two overlapping spots appeared; the faster-moving of the two had an Rf. value of approximately 0.6. These were assumed to be E-D.N.P. hydroxylysine and E-D.N.P. allo hydroxylysine. A yellow artefact was present; this moved as a faint spot just behind the solvent front.

Dinitrophenylation of Hippuryl-lysine.

A 20 mg. sample of Hippuryl-lysine (Benzoyl - glycyl -

lysine - Mann Research Laborat ries Inc. New York) was dinitrophenylated in aqueous medium in a pH-stat. It was found that at pH 7 little or no reaction occurred between F.D.N.B. and the peptide, whilst at pH 9.3 reaction was quantitative.

The D.N.P. hippuryl-lysine obtained was in the form of an insoluble precipitate; this was centrifuged down and washed with water and ether.

A sample of the derivative was hydrolysed for 48 hours in 5.7 N H Cl. at 120°C in a sealed tube. The hydrolysate was taken to dryness, washed, and subjected to thin layer chromatography in the system: Butanol: acetic acid: water (80 : 20 : 20), alongside spots of the unhydrolysed material, and a control mixture of glycine, ε-D.N.P. lysine and lysine.

It was found that the original material ran as a single yellow spot which did not stain with ninhydrin. The hydrolysed sample gave two strong spots corresponding to ε-D.N.P. lysine and glycine (both of which stain with ninhydrin), and a faint spot just behind ε-D.N.P. lysine, which appeared on staining with ninhydrin. There was no trace of lysine in the hydrolysate.

Titration of D.N.P. collagen.

A sample of collagen which had been dinitrophenylated at 40°C in aqueous bicarbonate medium, as described

previously, was freeze dried, and reduced to a finely divided state, by grinding in a pestle and mortar. The material was transferred to a dialysis-sac as a suspension, and dialysed against three changes of approximately 0.01 M Citrate/phosphate buffer pH 6.8, for several days. After dialysis the D.N.P. collagen was centrifuged down, and approximately 500 mg. resuspended in the buffer against which it had been finally dialysed, to a total volume of 60.5 mls.

This suspension was titrated against 1.00N H Cl using an automatic-titrator (Radiometer) to a final pH of 2.0. The time for titration selected was 1 hour, and the pH was noted and recorded on the diagram at regular intervals as an additional check to the pH scale obtained from the automatic recorder. To obtain the corresponding buffer titration, 60 mls. of the buffer solution against which the protein had been dialysed, was titrated against 1N H Cl in the same way.

Similar up-scale titrations were performed for another 500 mg. sample of the D.N.P. collagen and the buffer, against 1.00 N NaOH to a final pH of 11.3. The graphs corresponding to the up-scale and down-scale titrations of the D.N.P. protein and buffer, were put together (see figure 16) and the difference, corresponding to the titration of the protein alone, plotted on a separate graph as m. moles acid or base per gm. of collagen

FIGURE 16.

Titration of 590 mg. D.N.P. collagen in citrate / phosphate buffer.

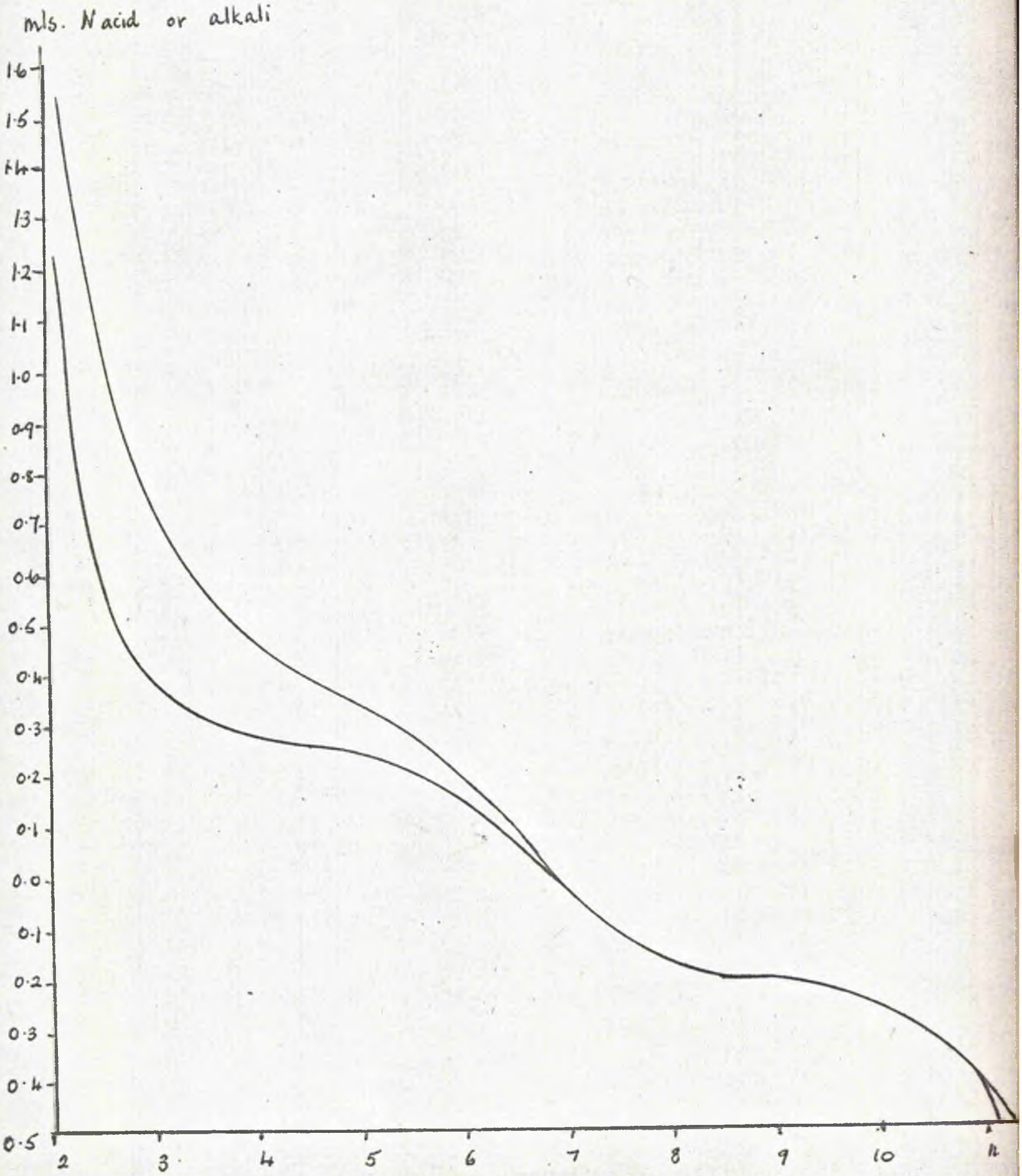
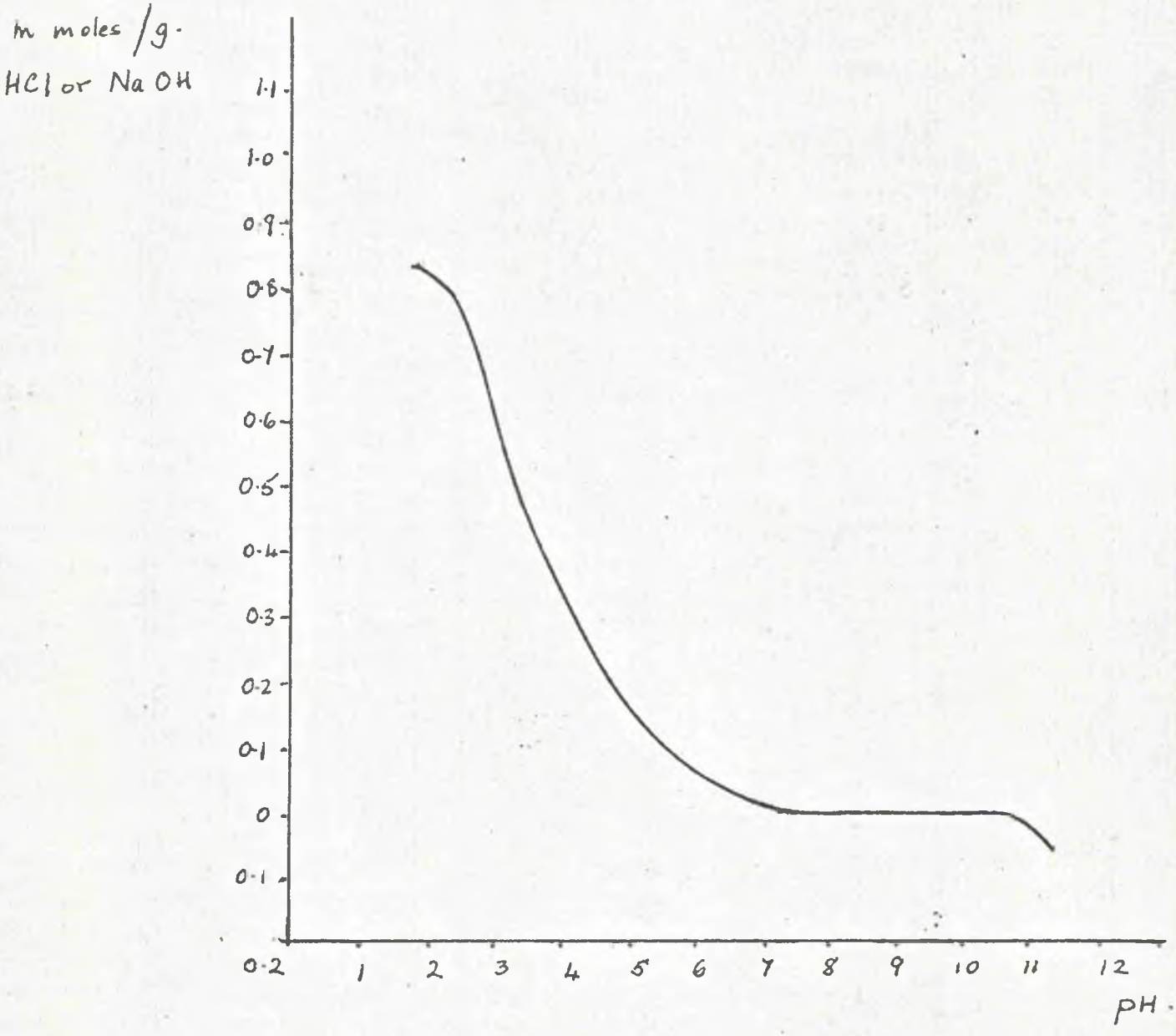


FIGURE 16A.

Titration of D.N.P. Collagen.



against pH (see figure 16A.).

It will be seen that no titration occurred from pH 7 - 11 and the pk of the acid-titrating group was at approximately pH 3.6. A small titration occurred at pH 6 - 7, corresponding to approximately 5 moles per 100,000 g. collagen. This could be due to histidine as all the carboxyl groups should have titrated at below this pH and any amino groups unsubstituted would titrate at a higher pH.

Results.

Collagen dinitrophenylated at pH 7.5 employing a pH-Stat.

1. Consumption of Na OH due to reaction between protein and F.D.N.B. (from figure 15) = 42 equivalents per 100,000 gms. At pH 7.5 the majority of the ε amino groups will be in the $- NH_3^+$ condition. Thus for the reactions:

(a) $R - NH_3^+ \longrightarrow R - NH_2 + H^+$

(b) $R - NH_2$ + F.D.N.B. $\longrightarrow$ R - NH - D.N.P. + HF.

two equivalents of Na OH will be consumed per mole of reacting lysine. Thus the titration observed is considerably lower than the theoretical figure, assuming that all of the ε amino groups of lysine and hydroxylysine react viz. approximately 68 equivalents per 100,000 gms. It was assumed that at the elevated temperature of 40°C. the equilibrium:

$$R - NH_3^+ \longrightarrow R - NH_2 + H^+$$

is displaced to the right.

From the graph it was observed that the slope became a straight line after 14 hours. Thus reaction between protein and F.D.N.B. was complete under these experimental conditions in 14 hours.

2. By calculation from the optical density of the hydrolysed material, it was found that approximately 33 ε amino radicals per 1,000 were dinitrophenylated. This corroborates the fact that the titration value obtained above is low.

Table 20. Collagen dinitrophenylated with excess $NaHCO_3$ at 40°C.

Liberation of ε D.N.P. lysine and hydroxylysine.

Duration of hydrolysis	Wt. D.N.P. collagen	O.D. 360*	equiv./† 100,000g.	average‡
6 hours	51.0	0.460	32.9)	
6 "	50.5	0.459	33.2)	31.35
6 "	49.7	0.460	33.8)	
8 "	51.0	0.485	34.7)	
8 "	50.6	0.466	33.6)	31.95
8 "	54.15	0.493	33.3)	
10 "	50.8	0.470	33.7)	31.5
10 "	51.1	0.470	33.6)	
12 "	54.7	0.494	32.8)	
12 "	50.9	0.453	32.6)	30.35
12 "	55.6	0.480	31.6)	
14 "	59.8	0.502	30.6)	28.75
14 "	54.8	0.452	30.2)	
10 hr. control (+ 1.5 mg. ε D.N.P. lysine)	53.2	0.62	-	-

* made up to 500 mls. i.e. in 0.057 N HCl.

† As equivalents/100,000 g. D.N.P. collagen. Assuming a molar extinction coefficient of ε D.N.P. lysine (and ε D.N.P. hydroxylysine) of 1.45×10^4, a value obtained by experiment, using standard ε D.N.P. lysine in solutions in 0.1N H Cl.

‡ Corrected to equivalents of reacting ε-amino groups/100,000 g. collagen. Also including a deduction of 2.25 equivalents, namely, 0.25 for tyrosine, and 2.0 for histidine calculated as follows:-

1. Assuming 4 residues/1000 for tyrosine and an approximate molar extinction coefficient of 9×10^2 at 360 m.u. for O-D.N.P. tyrosine in 0.1N H Cl. (computed from data of Porter, 1957).
2. Assuming a molar extinction of 0.7×10^4 for imidazole D.N.P. histidine at 360 m.u., and 4 residues/1000 for histidine in collagen. No information seems to be available concerning the extinction of imidazole D.N.P. histidine, so the value used is an approximation, from consideration that the molar extinction for di D.N.P. histidine in $Na\,HCO_3$ is 2.1×10^4.

From the control 10 hour hydrolysis it was calculated that 1.49 mg. of ε-D.N.P. lysine were recovered.

Thus 10 hours hydrolysis in 5.7 N HCl at 120°C in presence of D.N.P. collagen gives 99.4 ± 3% recovery of ε - D.N.P. lysine.

The maximum value of 31.95 equivalents for the 8-hour hydrolysate can be taken to indicate that 32 ± 2 equivalents ε-amino groups per 100,000 gms. collagen are reactive towards F.D.N.B. under the employed reaction conditions.

Examination of hydrolysates for peptides.

The electrophoresis experiments and chromatography demonstrated that, whilst in the 6-hour hydrolysates there was still evidence of D.N.P. peptides, in the 8-hour hydrolysates these were present in tiny amounts which could only be detected by heavy loading, and in 10-hour hydrolysates none were detectable. The peptides in the 6-hour hydrolysates were particularly noticeable when electrophoresis was performed at pH 3.5. Three faint bands of acidic D.N.P. peptides appeared, and one of basic D.N.P. peptides, the majority of the yellow material remained at the origin, as did the control samples of α-D.N.P. lysine and ε-D.N.P. hydroxylysine. These control samples did not separate under any of the conditions of electrophoresis, and always corresponded in position exactly to the predominating band in the electrophoresis pattern of the hydrolysates.

Thin layer chromatography gave slight separation of the ε D.N.P. lysine and ε D.N.P. hydroxylysine, the former travelling slightly further. Single applications of the

ε-D.N.P. hydroxylysine showed a degree of separation into spots corresponding to the allo and normal forms, and in the hydrolyses there was evidence that both these forms existed. The D.N.P. peptides of the 6-hour hydrolysate travelled faster than the major spots of the amino acid derivatives, and showed up as a faint yellow streak.

Estimation of amount of soluble D.N.P. protein after dinitrophenylation of collagen at 40°C.

From the weights of the soluble and insoluble derivatives it was calculated that the soluble D.N.P. protein represented 54% of the original collagen.

4. Experiments with the peptides liberated from soluble and insoluble collagen by the enzyme collagenase.

Preparation of Peptides.

Enzyme digestion was performed at pH 7 - 7.5 in the presence Ca^{+} ions; the total ionic strength was kept very low. Crystalline collagenase (Sigma Chemical Co. Ltd.) was added in an enzyme: substrate ratio of 1:500, and the mixture stirred at room temperature for 12 hours. 2 g. samples of soluble and insoluble collagen were digested in this way, and the soluble peptides were collected by concentrating the resultant solution in a rotary evaporator and finally freeze drying the concentrate. Undigested or insoluble material at the end of the digestion was discarded; there was very little of this in the soluble collagen digest, but a definite insoluble residue was found after collagenase treatment of insoluble collagen.

Dinitrophenylation of peptides.

50 mg. samples of the peptides from soluble and insoluble collagen were dinitrophenylated in a 0.5% trimethylamine, 50% alcoholic medium with a 20-fold weight excess of F.D.N.B. to peptide. Reaction was carried out for 2 hours at room temperature in the dark. A considerable amount of the D.N.P. peptide obtained was insoluble. A few drops of 1% trimethylamine were added to the final

reaction mixture, to ensure alkaline conditions, and three ether extractions were performed. The D.N.P. peptides were taken to dryness, and 10 mg. samples hydrolysed in 5 mls. 5.7N HCl in a sealed tube for 12 hours. The hydrolysate was diluted to approximately 1N with water, and the ether soluble D.N.P. amino acids extracted; these were separated and estimated quantitatively by two-dimensional paper chromatography as described in Section I. The aqueous phase was taken to dryness, and a sample taken for amino acid analysis; the residue was examined by two-dimensional thin layer chromatography in the systems: Butanol: acetic acid: water (80 : 20 : 20)

and: phenol: water (75 : 25). E-D.N.P. lysine, O-D.N.P. tyrosine (Mann Research Laboratories Inc. New York) E-D.N.P. hydroxylysine (prepared sample) and chromatographically pure amino acid samples (B.D.H.) were used as standards.

Amino Acid Analyses.

Amino acid analyses were obtained using a Technicon Automatic amino acid analyser.

Samples of hydrolysed acid - soluble, and insoluble collagen, and dinitrophenylated peptides obtained from both materials by the action of collagenase, were analysed.

The D.N.P. amino acids present in the hydrolyses of D.N.P. peptides viz. E-D.N.P. lysine, E-D.N.P. hydroxylysine, O-D.N.P. tyrosine, and probably imidazole

D.N.P. histidine, were not estimated. These derivatives remained attached to the column, while the free amino acids were being eluted, and could be finally washed off the column with alkali.

Results.

Table 21. I: N-terminal amino acids of collagenase liberated peptides.

1. Peptides ex. soluble collagen.

D.N.P. amino acid	moles/100,000g. peptide.
alamine	1.2
valine	1.8
leucine	3.9
phenylalanine	1.6
glycine	108
serine	trace
threonine	trace
aspartic acid	trace

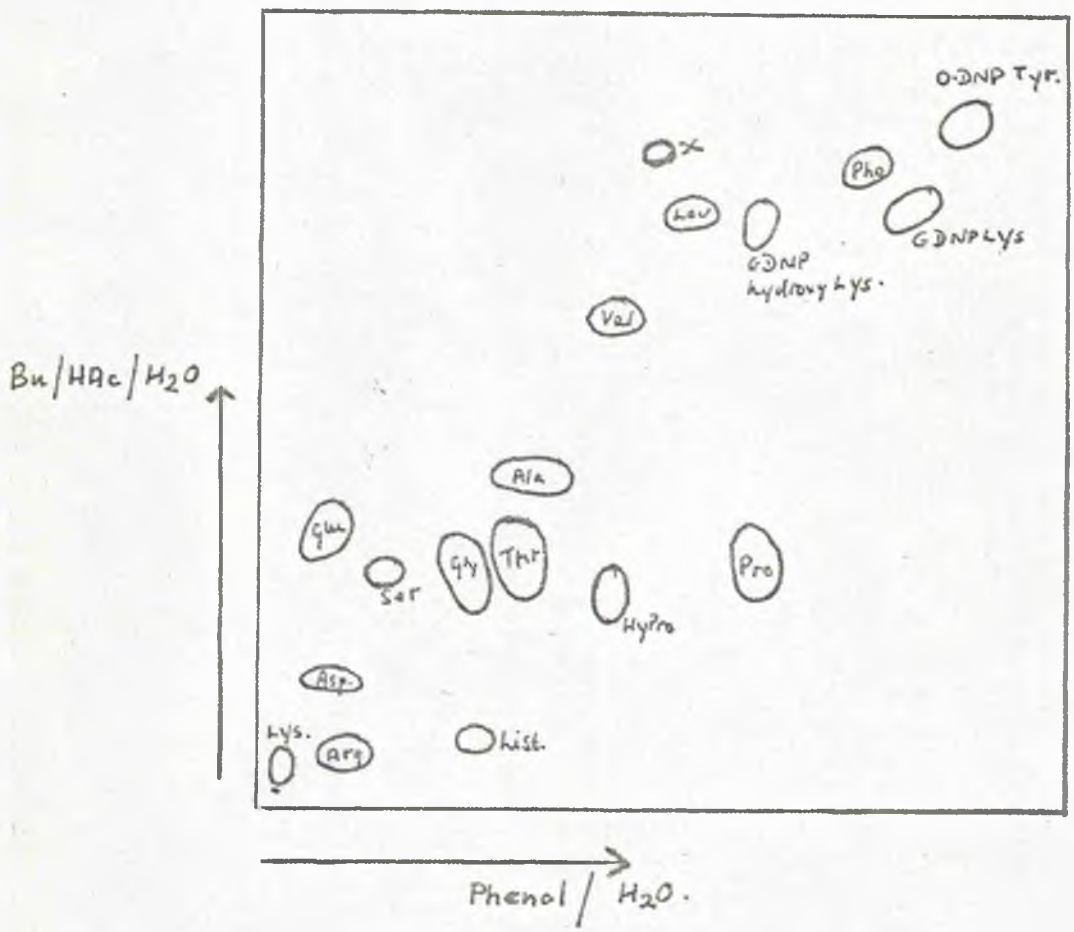
2. Peptides ex. insoluble collagen.

Quantitative determination was not performed but the pattern of N-terminal residues obtained was identical. viz. alamine, glycine, valine, leucine, phenylalanine, (serine, threonine, aspartic acid traces).

II: Thin layer chromatography of Aqueous Phase.

A copy of a typical separation using the system: Butanol: acetic acid: water (80 : 20 : 20), phenol: water (75 : 25) is shown in figure 17. spots corresponding to

FIGURE 14. Chromatogram of amino acids and D.N.P. amino acids from aqueous phase of hydrolysates of D.N.P. collagen.



ε D.N.P. lysine and ε D.N.P. hydroxylysine were visible before staining with ninhydrin; after staining their colour changed from yellow to brown. There was no qualitative difference between chromatograms of material from soluble collagen to those of material from insoluble collagen.

Free lysine and histidine were apparent, but no free tyrosine or hydroxylysine. O-D.N.P. tyrosine was identified with certainty. Imidazole D.N.P. histidine was not identified but may correspond to the spot marked x. The positions of the spots corresponding to the D.N.P. amino acids and the free amino acids which travelled farthest in the "Butanol" Solvent was somewhat variable. It is possible that the presence of D.N.P. amino acids affects the Rf. values for the free amino acids. The colour given by O-D.N.P. tyrosine with ninhydrin was similar to that given by free tyrosine viz. purple; this colour is quite characteristic.

III: Amino Acid Analyses.

Table 22./ (See p.105.)

Table 22. Amino acid composition of acid soluble collagen and insoluble collagen. As residues per 1,000 total residues.

Amino Acid	Soluble Collagen*	Insoluble Collagen*	Ox-Hide† Collagen
Hydroxyproline	84.8	79.4	99.8
Aspartic Acid	49.1	51.4	44.2
Threonine	16.4	18.3	18.0
Serine	29.1	31.6	29.9
Glutamic acid	75.8	78.3	71.8
Proline	134.1	129.0	122.6
Glycine	319.0	302.8	339.0
Alanine	115.1	110.7	99.8
Valine	23.2	29.2	27.1
Methionine	3.4	4.8	5.1
Isoleucine	14.3	17.0	)
Leucine	26.8	30.7	) 39.9
Tyrosine	4.1	5.1	5.1
Phenylalanine	14.2	14.4	14.2
Hydroxylysine	8.3	8.9	6.4
Lysine	28.1	31.3	23.7
Histidine	4.5	6.1	4.5
Arginine	50.6	50.8	46.2
	1000	1000	1002
Ammonia	44.1	29.7	
% age Nitrogen	18.6%	18.75%	

* For comparison with literature values see Table I.

† Literature values for comparison with values for insoluble Calf-skin collagen - From Tristram 1953.

Table 23. Amino Acid composition of D.N.P. peptides.

As residues per 84.8 residues of hydroxyproline in soluble collagen and residues per 79.4 residues of hydroxyproline in insoluble collagen.

Amino Acid	Ex. acid sol. collagen	Ex. insol. collagen	control†
Hydroxyproline	84.8	79.4	79.4
Aspartic Acid	32.4	35.1	32.4
Threonine	14.9	14.6	15.0
Serine	32.6	32.4	28.8
Glutamic acid	65.5	72.4	63.1
Proline	130.2	124.2	120.3
Glycine*	137.1	149.0	144.0
Alanine	97.7	98.0	101.2
Valine	13.8	14.9	17.0
Methionine	6.8	2.4	1.5
Isoleucine	6.6	9.7	9.2
Leucine	18.1	21.1	20.0
Tyrosine	0.0	0.0	0.0
Phenylalanine	11.8	9.7	8.2
Hydroxylysine	0.0	0.0	0.0
Lysine	1.9	3.6	3.7
Histidine	1.4	2.3	2.1
Arginine	54.8	47.6	46.0
NH_3	33.0	20.1	24.3

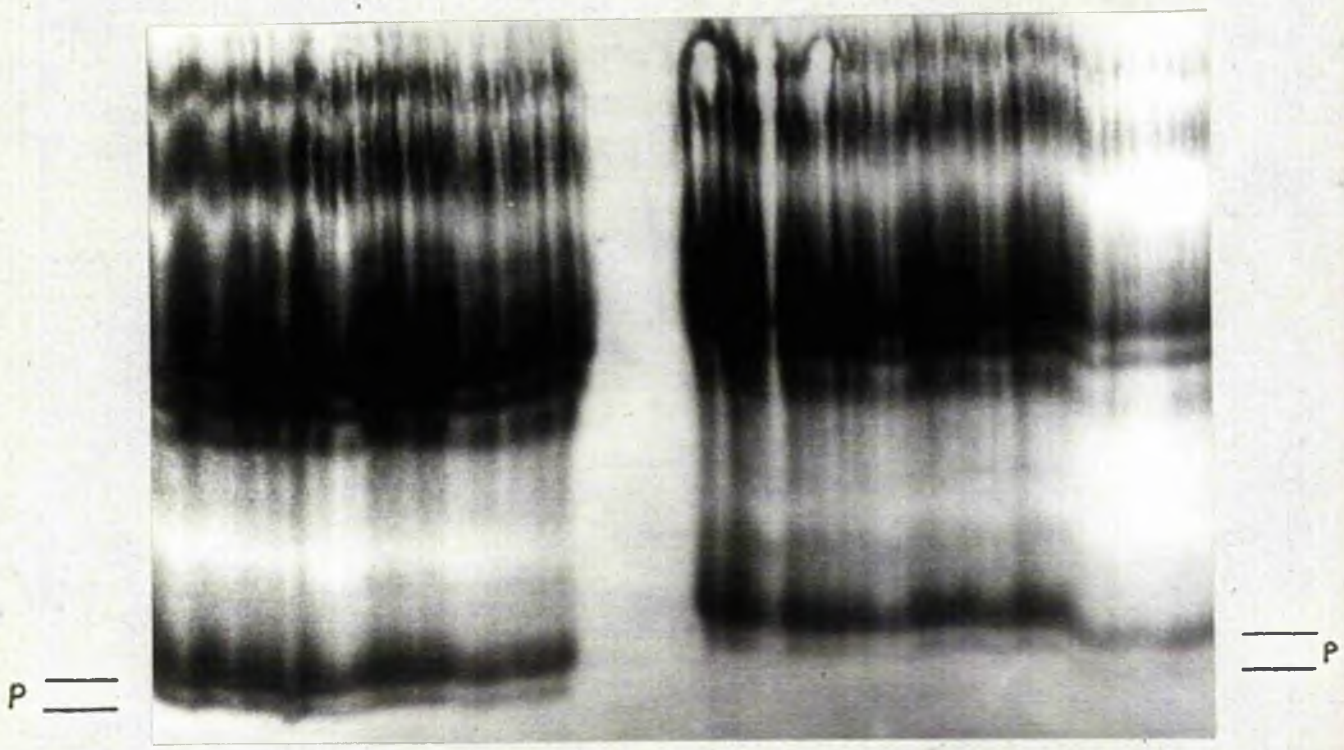
† A sample obtained from the hydrolysis of 10 mgs. of the D.N.P. peptides from insoluble collagen + 0.5 mg. ε D.N.P. lysine + 0.5 mg. O D.N.P. tyrosine.

* The glycine values appear low, due to the large amount of N-terminal glycine in the peptides being removed as D.N.P. glycine.

Separation of free collagenase-liberated peptides.

Attempts to separate the peptides were made by a variety of techniques, including high voltage electrophoresis and paper chromatography, and combinations of both. Although "fingerprinting" with electrophoresis at pH 3.5 followed by chromatography in the system : pyridine: isoamylalcohol : water (117 : 117 : 100 - Baglioni et. al. 1961) gave probably the most complete overall separation (c.2[illegible] spots), it was decided to use electrophoresis only, as it was then possible to load heavily across the whole width of the paper, and to elute the resultant zones for subsequent experiments. It was considered that for investigations into the possible existence of trifunctionally involved lysine residues, polar peptides would be the most likely source to consider. The majority of hypotheses concerning the site of cross-linking in collagen, indicate the likelihood that polar regions are involved. If any lysine residues were present as an ε-lysyl peptide, peptides containing this grouping would have an extra free α-amino group due to the branch, compensating for the loss of one α-amino group. Thus basic peptides would probably be best investigated in this connection. Electrophoresis at pH 1.9 gave good separation of peptides with a high proportion of basic functions, and was used to give the initial fractionations.

A photograph of the separation obtained is shown



insoluble

Soluble.

FIGURE 18. Peptides from collagenase digestion of soluble and insoluble collagen, separated by High Voltage electrophoresis. pH 1.9.

in figure 18. The conditions used were : buffer, pH 1.9 (acetic acid: formic acid: water - 150 : 50 : 800), 2000 v. 50 m.A.

The paper was soaked in the buffer, and excess was carefully pressed out, using a hand-roller, before placing the paper on the glass platform of the apparatus. The peptide sample (5 - 10 mg.) in solution was applied to the paper near the +ve end, using a micro-pipette. The voltage was applied gradually, and usually a period of 1½ hours was allowed for the separation.

After drying the paper, the edges were stained using an ethanolic solution of ninhydrin. The strip from the centre unstained portion of the paper, corresponding to the zone marked P. in figure 18 was cut out, and the peptide material collected by elution.

The method used for elution was to place the paper strips between two glass plates, with both ends protruding slightly. One end of the paper strip was dipped into water, which then travelled along the paper carrying the peptide material into the other tip of the paper which was exposed to the atmosphere. The peptide material from zone P of several electrophoresis runs of material from both soluble and insoluble collagen, was collected in this way.

Dinitrophenylation of peptide fraction P.

All stages were carried out in a 50 ml. roundbottomed quickfit flask. The peptides were dissolved in 1% trimethylamine and excess of an alcoholic solution of F.D.N.B. added. After 2 hours contact more trimethylamine was added and the reaction mixture extracted thoroughly with ether (the ether layer being removed by suction). The D.N.P. peptides so obtained were dried in vacuo and then fractionated crudely into acetone soluble, water soluble and water insoluble portions. It was found that the peptides obtained from soluble collagen contained little or none of the water insoluble fraction whilst the insoluble collagen peptides had a substantial amount. The three D.N.P. peptide fractions were investigated separately:-

1. Acetone soluble fraction.

This was strip loaded onto a thin layer plate and run in the 'butanol' system, using an S-chamber. (All chromatograms of D.N.P. derivatives were performed in the dark). Dinitrophenol was present in this fraction; it ran with the solvent front. Approximately 8 yellow peptide - bands were seen; the three with the highest Rf. values were the strongest. There was some qualitative difference between the fractions from soluble and insoluble collagen. This is indicated in table 24.

Table 24.

Peptide bands.	approx. Rf.	insoluble collagen peptides.	soluble collagen peptides.
1.	0.855	+	trace
2.	0.810	++	++
3.	0.740	+++	+++
4.	0.665	+	trace
5.	0.57	trace	none
6.	0.48	trace	trace
7.	0.45	trace	trace
8.	0.40	trace	trace

"Peptide band3." from a chromatogram of material from insoluble collagen was scraped off the plate and brushed into a small centrifuge tube. The peptide material was eluted with three portions of acetone: acetic acid (99 : 1), taken to dryness and hydrolysed for 16 hours with 5.7 N H Cl. The hydrolysate was examined for N-terminal amino acids, and free amino acids (2 dimensional T.L.C. in the systems: chloroform : methanol : 17% ammonia 2 : 2 : 1 and phenol : water 75 : 25). A portion of the hydrolysate was dinitrophenylated and normal quantitative D.N.P. amino acid chromatography performed. For estimation of D.N.P. arginine and Di D.N.P. histidine (not extracted by ether), the residual material after ether extraction was taken to dryness, dissolved in acetone applied to Whatman paper No.1, and run in the "Toluene" solvent which gave good separation of these derivatives.

Results.

Table 25. N-terminal amino acids of "peptide band 3."

Amino Acid.	u moles x 10^3 in total hydrolysate
Valine	0.93
Serine	3.8
Leucine	45.0
Aspartic Acid	7.9
Glycine	125.0

Table 26. Amino Acids of "peptide band 3."
(estimated as dinitrophenyl derivatives).

Amino Acid.	u moles x 10^3 in total hydrolysate.
Lysine	49
Glutamic acid	60
Glycine	136
Alanine	40
Proline	144
Hydroxyproline	32
Valine	20
Serine	24
Threonine	16
Histidine	30
Arginine	54

Amino Acids and derivatives in aqueous phase.

Thin layer chromatography revealed the presence of the following substances:-

ε-D.N.P. lysine	Proline	Hydroxyproline	Glycine
Glutamic acid	Alanine	Arginine	Serine
Threonine	Histidine	Valine.	

There was no trace of free lysine, but a spot corresponding to free histidine was present.

2. Water soluble fraction.

Only the fraction obtained from insoluble collagen was examined. It was also found to be complex on chromatography in the system : chloroform : tertiary amyl alcohol: acetic acid (60 : 40 : 5). The entire material was hydrolysed and qualitatively examined for N-terminal amino acids, and free amino acids and derivatives of the aqueous phase.

Results.

1. N-terminal amino acid = glycine only.

2. The aqueous phase contained:

- D.N.P. lysine
- D.N.P. hydroxylysine
- lysine (trace)
- Phenylalanine
- Proline
- Glutamic acid
- Glycine
- Arginine
- Alanine
- Valine
- Leucine

There was no hydroxyproline in this peptide fraction.

3. Acetone and Water insoluble fraction.

The unfractionated material was hydrolysed an

examined in the same way as "peptide band 3" of the acetone soluble fraction. The optical density at 360 m.u. of the hydrolysate was also measured to determine the ε D.N.P. lysine and ε D.N.P. hydroxylysine, before conversion to the Di D.N.P. derivatives.

Results.

Table 27. N-terminal amino acids of insoluble peptide fraction.

Amino acid	u. moles x 10^3 in total hydrolysate.
Glycine	30
Leucine	trace

Table 28. Amino acids of insoluble peptide fraction. (estimated as Dinitrophenyl derivatives).

Amino Acid	u. moles x 10^3 in total hydrolysate.
Arginine	24
Serine	20
Glycine	101
Alanine	37
Glutamic acid	46
Valine	25
Proline	57
Hydroxyproline	13
Histidine	19
Lysine + Hydroxylysine	32

From the optical density of the hydrolysate at 360 m.u. it was calculated that the total amount of lysine and hydroxylysine present as ε D.N.P. derivatives was: 15×10^3 u. moles.

∴ Approximate free lysine content of total hydrolysate

$= 17 \times 10^3$ u. moles.

Two dimensional thin layer chromatography revealed the presence of the following compounds:-

ε D.N.P. lysine

ε D.N.P. hydroxylysine

Proline

Glutamic acid

Serine

Glycine

Alanine

Arginine

Hydroxyproline

Valine

* Histidine

Lysine

* A spot which may have been imidazole D.N.P. histidine was present near valine. Free histidine was not confirmed, although its presence as a spot overlapping with that of alanine was indicated.

Free lysine was present as a very strong, well-defined spot.

DISCUSSION

1. N-terminal residues of collagen.

The value of 0.13 moles/100,000 gms. for N-terminal residues of D.N.P. tropocollagen compares with the figures:

0.10 moles aspartic acid + alanine per 100,000 g. procollagen (Bowes and Moss 1953);

0.39 moles glycine + aspartic acid per 100,000 g. citrate soluble collagen (Chandrarajan and Bose 1965);

and 0.17 moles glycine + aspartic acid per 1000 moles of amino acids of procollagen (Hörmann et. al. 1965).

Thus glycine and aspartic acid seem to be the most important N-terminal residues. That Bowes and Moss reported no N-terminal glycine could be due to destruction of the highly labile derivative, D.N.P. glycine, under their conditions of hydrolysis (refluxing with 6N H Cl for 12 hours, and sealed-tube hydrolysis with 10N H Cl for 16 hours; as opposed to the conditions of Hörmann et. al. viz. reflux with 2N H Cl for 4 hours).

Apart from glycine and aspartic acid, traces of other N-terminal amino acids were found with D.N.P. tropocollagen. This could be due to inefficient washing of the D.N.P. protein, for Hörmann et. al. and Bowes and Moss both used Soxhlet - extraction with ethanol to remove soluble material whilst washing with acetone: H Cl

was employed here.

The value of 0.20 moles/100,000 g. collagen for the N-terminal residues of D.N.P. - total collagen compares with 0.148 moles/100,000 g. reported by Steven and Tristram (1962). These values are in keeping with the concept that collagenous tissue contains a proportion of free amino acids and peptides which become bound to the protein and can only be incompletely removed. The difference between the value 0.20 moles/100,000 g. for D.N.P. - total collagen and 0.13 moles/100,000 g. for D.N.P. tropocollagen i.e. 0.07 moles/100,000 g. probably represents the majority of this non-protein nitrogen fraction. The true amount, however, may be higher due to material not removed from the D.N.P. tropocollagen by washing, and also because of unpredictable hydrolytic losses during the hydrolysis of D.N.P. total collagen. During this hydrolysis procedure it was impossible to avoid the presence of a considerable amount of salt which is likely to have accentuated the destruction of D.N.P. derivatives. The alternative of carrying out the dinitrophenylation at a low ionic strength would have led to an incomplete degree of substitution of the amino groups of the protein.

Hörmann et. al. carried out their dinitrophenylations in the presence of 2.5M Na ClO_4 to cause complete

denaturation of the protein. This could be responsible for the slightly higher value for N-terminal residues quoted by these workers, than that found in the above experiments. However, the difference (23%) may not be significant when it is considered that the accuracy of this type of experiment is rather limited. Hallsworth (1964) reported that ionic strength of the medium has a marked effect on the degree of substitution of ε amino groups. It was considered that under the conditions of dinitrophenylation used for these experiments viz. 50% alcohol, ionic strength 0.3; substitution of the α - amino groups would be complete.

To sum up, therefore, it must be concluded that collagen contains approximately 0.4 - 0.6 equivalents of free - amino groups per molecule (molecular weight approximately 300,000). These groups mainly arise from glycine and aspartic acid, although some N-terminal alanine, serine, threonine and glutamic acid may occur. It seems likely that the actual amount of N-terminal residues varies between different collagen preparations.

The results obtained by pepsin digestion of D.N.P. collagen indicate that although the enzyme has some activity on the derivative, yellow peptides being liberated, the N-terminal content is probably unaffected. Thus the N-terminal residues cannot terminate peptide

appendages which are susceptible to pepsin.

On redinitrophenylation the presence of an increased number of N-terminal residues was revealed, although only in small amount. Thus the pepsin treatment must either have liberated peptides with masked N-terminal residues, or ruptured peptide linkages without actually releasing peptides from the protein. The former case is most likely, as the presence of soluble yellow material was observed in the enzymic digest. The liberation of peptides with C-terminal residues would not have been noticed in this experiment, thus the enzyme activity could have been higher than the extremely low value indicated by the formation of only 0.12 moles new N-terminal residues per 100,000 g. protein. On the other hand, pepsin could have a reduced effect on the D.N.P. protein as opposed to native collagen.

To explain the small but definite content of N-terminal amino acids in collagen in a way which is compatible with the above findings, three theories can be proposed:-

1. That collagen possesses a larger number of N-terminal residues, but that these are masked at some stage during the process of biosynthesis and assembly of subunits. The masking in some cases is not complete - for example, it could be because of incorporation of the molecule into a fibre before completion of an enzymatic

process, thus giving rise to a non-stoichiometric number of free amino groups in the tissue as a whole.

The fact that pepsin treatment releases peptides from D.N.P. tropocollagen with the formation of new N-terminal residues is evidence for masking of α-amino groups in collagen. For the D.N.P. tropocollagen after pepsin treatment showed no decrease in substituted α-amino groups, and the released peptides must therefore have possessed no terminal free amino groups.

2. Biosynthesis of the collagen molecule is a multi-stage process, with the possibility that some incompletely "assembled" molecules bearing free α-amino groups "escape", and are secreted from the fibroblast and incorporated into the collagenous tissue as such.

3. That, although collagen has an extremely long half-life, some tissue catabolism of the molecule is bound to occur. If the proteolysis involved were a fairly gradual process, at the time of laboratory-preparation of the collagen sample from its tissue, some partially degraded molecules would be present. Possibly some of the molecules could have been attacked at their "telopeptide" appendages by a "non-collagenolytic" enzyme, leaving a "stub" bearing a free α-amino group. The results obtained with pepsin digestion of D.N.P.

tropocollagen support this theory, for the N-terminal residues present are not removed, possibly having been produced originally by an enzymatic process similar to that being effected by the pepsin on other intact molecules.

Again the presence of masked α- amino groups in collagen seems likely if this theory is at all valid. If telopeptides are to leave a free α-amino group after proteolysis, when there is no evidence for such groups in the intact molecule, then these peptides must be terminated by masked α-amino groups.

The possibility that all the telopeptides bear only free C-terminal amino acids cannot be overlooked. In that case, the observations on the action of pepsin must be explained by two processes.

(a) attack on telopeptides - not forming new N-terminal groups on the macromolecule but liberating peptide material.

(b) rupture of bonds without the release of peptide material.

In this case possibility (3) can only be explained in terms of the latter process, although the proteolysis of C-terminal telopeptides could occur at the same time.

Recently Bouchilloux and Cheftel (1966) have demonstrated that the incorporation of carbohydrate

residues into Thyroglobulin does not occur while the protein is still attached to the polysomes, but at some later stage, probably, while the polypeptide chains are still within the confines of the endoplasmic membranes. This could be taken as support for the first possibility, when it is remembered that Chandrarajan and Bose (1965) demonstrated a masking of amino groups of collagen by hexose.

The evidence of Hörmann et. al. (1965) for the presence of acetylated α-amino groups in collagen is very convincing. Masking of these groups in this manner does seem more likely than masking with hexose by a presumably N-glycosidic linkage, for the latter groups (unlike O-glycosidic bonds) form osazones without prior hydrolysis and such a reaction has never been demonstrated for collagen. (Grassmann et. al. 1957).

2. The action of ∝-amylase on collagen.

The property of the enzyme to convert insoluble collagen into soluble collagen was observed over a wide pH range. This is a similarity to the fundamental action of ∝-amylase as a carbohydrase, for this activity also occurs over a wide pH range.

The fact that amylase solubilised collagen contains less carbohydrate than the insoluble protein, and that free glucose and galactose were detected in the dialysate of the enzyme digest, also supports the idea that rupture of linkages involving carbohydrate residues occurs during the digestion of insoluble collagen by ∝-amylase.

Using, however, the method of Davis and Smith (1955), it was demonstrated quite conclusively that the two preparations of ∝-amylase used both had definite proteolytic activity, not only for collagen, but also for haemoglobin. It was interesting to note that pancreatic amylase had more activity with insoluble collagen than with haemoglobin. For the majority of proteolytic enzymes the situation would probably be reversed. This might be an indication that some of the peptide release with insoluble collagen occurs as a result of carbohydrase-like action.

Pepsin had approximately equal action on both acid soluble collagen and amylase solubilised collagen,

indicating that the pepsin vulnerable bonds present in the former were also present in the latter.

The pattern of N-terminal amino acids of amylase - solubilised collagen was slightly more complex than that of acid soluble collagen, although the total amount was approximately the same. The significance of this is difficult to interpret, for obviously there is no indication of liberation of N-terminal residues in anything like the amounts corresponding to one residue per protein molecule.

Ultracentrifugation of the collagen extracted by pretreatment of insoluble collagen with α-amylase, and fractionation on carboxy methyl cellulose of the same material, appeared to give conflicting results. Only one peak appeared in the Ultracentrifuge; this had a sedimentation coefficient of 3.255, which compares quite well with the value obtained by Piez et. al. (1960) of 3.35 for α-components. There was no indication of β-components in the ultracentrifuge studies of amylase - solubilised collagen.

Fractionation on carboxy methyl cellulose on the other hand, gave three initial peaks, which must be taken to show the existence of components in the material, unless the presence of more than 2 forms of α-component in calf-skin collagen is accepted. The two final peaks of the column fractionation correspond to higher molecular

weight aggregates; these would have sedimented very rapidly in the Ultracentrifuge, and escaped detection; however, their existence in amylase solubilised collagen is difficult to explain when it is considered that the enzyme treatment appears to degrade all β components to α, as well as degrading most of the more highly cross-linked aggregates to α. Why then should there be some high molecular weight aggregates remaining, but no β components?

Investigation of the products of amylase treatment of insoluble collagen, as already mentioned, demonstrated the presence of glucose and galactose. According to the assays of carbohydrate content by the anthrone method, 0.34% of the total protein weight is lost as carbohydrate, during such enzyme treatment. This corresponds to 5.6 hexose residues per tropocollagen molecule - rather more than the two amino acid residues calculated to be released concurrently. Fraction A when examined by chromatography was shown to contain the two hexose residues; but a very approximate calculation, on the basis that the spots revealed by spraying with aniline phthalate reagent represent in the region of 5 u.g. carbohydrate, indicated that fraction A only contained a small portion of the carbohydrate released. The losses which must therefore have occurred may be due to carbohydrate being bound by

the Dowex-50 column to a certain extent. Any hexose present as glyco peptides in the supernatant after enzyme treatment would certainly bind to the ion exchange column, and not be collected with fraction A.

The amino acids in the hydrolysate of the material released by amylase from collagen were very numerous. This indicates that the enxymatic process is not a very specific one, or that only a minor fraction (in amount) of the collagenous tissue (one not present as a unit in numbers equal to those of the tropocollagen molecules themselves) is being degraded; otherwise, only a limited number of amino acids would be released.

The fact that no hydroxyproline appeared in the hydrolysate shows that the main body of the triple-helix of the tropocollagen molecules is probably not attacked by the enzyme.

Thus the process by which the enzyme α-amylase causes the solubilisation of insoluble collagen, may involve a combination of proteolytic and carbohydrase-like activities. It is possible that in fact the process is analogous to the normal activity of amylase, the enzyme perhaps recognising certain hexose-protein linkages or even certain peptide bonds, in the same way as the α 1 — 4 glycosidic bond of its accepted substrate.

As amylase solubilised collagen contains considerably

reduced amounts of carbohydrate, but similar amounts of N-terminal amino acids to acid-soluble collagen, it is certain that the carbohydrate residues released by amylase treatment are not bound to free α-amino groups of normal collagen.

3. Availability of ε-lysyl residues to dinitrophenylation.

Experiments using the pH-stat show that, even at the elevated temperature of 40°C, complete reaction of the groups which can be substituted by F.D.N.B. took at least 15 hours. The titration study suggested that all of the histidine of collagen does not react under the conditions used. The normal method of preparation of α-mono D.N.P. histidine is to react the amino acid with 1 equivalent of F.D.N.B. Thus the imidazole group probably reacts only very slowly, and may not have been fully substituted in these experiments.

The value of 31.95 for the maximum number of equivalents of substituted ε lysyl and ε hydroxylysyl residues detected by measurement of the optical density of the hydrolysates at 360 m.u. may in fact be slightly low, because of overcorrection for histidine, for the above reasons. However, a value of 32 ± 2 residues per 100,000 g. corresponds quite well with the figure of 8.3 + 28.1 = 36.4 residues per 1,000 total residues (i.e. approximately 35.1 residues per 100,000 g.) From the amino acid analysis of acid soluble collagen in Section 4, the value is slightly low and therefore does not discount the possibility that 1 or 2 residues of lysine per 1,000 residues may be involved in covalent linkages via their ε amino groups.

The titration data showed no titration of free ε-lysine amino groups in D.N.P. collagen. This does not rule out the possibility that some free ε amino groups are present, being non-covalently masked, in a manner which prevents their titrating; but it does seem unlikely when the above facts are taken into consideration. Probably all the free ε amino groups of collagen react with F.D.N.B.

Prolonged hydrolysis of ε D.N.P. hippuryl-lysine yielded no free lysine, but traces of another ninhydrin positive compound other than glycine and ε D.N.P. lysine. (This compound had an Rf. value similar to that of nor-leucine in the solvent system used). Thus the hydrolytic destruction of the ε D.N.P. lysine resulting from hydrolysis of D.N.P. proteins probably does not give rise to free lysine.

Investigation of the hydrolysates of D.N.P. collagen by electrophoresis and thin layer chromatography, demonstrated that ε D.N.P. lysine and ε D.N.P. hydroxy-lysine are liberated from the protein by acid hydrolysis very rapidly. After 6 hours of hydrolysis at 120°C with 5.7N H Cl, in the region of 99% of these residues are present in an uncombined form. After 10 hours hydrolysis there was no trace of peptides containing dinitrophenylated ε lysyl or ε hydroxylysyl residues.

4. Experiments with collagenase-liberated peptides.

The sample of collagenase used must have contained some traces of proteolytic enzyme impurities, for N-terminal residues other than glycine (notably leucine) were present whilst the proved specificity of collagenase only allows for the liberation of N-terminal glycine.

The detection of free lysine in hydrolysates of the D.N.P. peptides by both thin layer chromatography and normal quantitative ion exchange analysis, using an automatic analyser, is strong evidence for the presence of covalently masked ε lysyl residues; for it is unlikely that steric hindrance would be a factor involved in the reaction of small peptides with substituting reagents such as F.D.N.B. A similar experiment has been performed by Franzblau (1962); with collagenase peptides from icthyocol, he found that 12% of the ε lysyl amino groups did not react. In these experiments the values obtained were 5.2% in acid soluble collagen and 9.0% for insoluble collagen.

Free histidine was also noticed; this may have arisen from incomplete substitution of free imidazole groups as explained for Section 3. No free tyrosine or free hydroxylysine was detected, thus these amino acids must be completely free for substitution in calf skin collagen.

Free lysine cannot be produced from hydrolytic destruction of ε D.N.P. lysine if the results of the experiments with hippuryl-lysine in Section 3 are considered to be relevant to the situation existing during the hydrolysis of D.N.P. proteins. Thus the only explanation for the existence of free lysine in the hydrolysates of D.N.P. collagenase liberated peptides, is that, either some of the lysine residues are covalently linked through their ε amino groups, or, that Steric factors are involved, which are manifest even in relatively small peptides.

The fact that there is more free lysine in hydrolysates of peptides derived from insoluble collagen than in hydrolysates of material from soluble collagen, could be taken to indicate that these lysine residues are involved in intermolecular crosslinkages. It does seem unlikely that if a Steric factor is concerned which prevents ε lysyl residues in some positions from reaction with F.D.N.B. that this should be more prevalent in insoluble collagen.

Further evidence in support of the concept of covalently linked ε-lysyl residues is found in the results of experiments with peptides obtained by fractionation of the collagenase digests by high voltage electrophoresis. Although it was found to be impossible to obtain completely pure peptides by the techniques

used, a peptide fraction was found (acetone and water insoluble peptide fraction) from insoluble collagen containing approximately 4.4% of its total residues as lysine, the ε amino groups of which could not be substituted by F.D.N.B.

The zone selected for investigation was highly complex; for example, the acetone soluble fraction obtained, contained more than eight peptides, the average chain length of which was approximately 4 residues. The insoluble fraction containing the free lysine, however, may have been less complex; from the stoichiometry of the molar proportions obtained it seems that the average chain length would be about 12 residues, with a total number of 2-3 peptides.

The water soluble fraction could be the most rewarding to investigate. It was not so complex as the acetone soluble one, and being soluble would be more easily fractionated than the peptides which were insoluble in both water and acetone. No hydroxyproline was detected in the water soluble fraction but some free lysine was found. Thus it seems likely that a peptide is also present in this fraction containing lysine, the ε-amino group of which is involved in covalent linkage in the original protein.

The nature of the covalent bond in which the ε-amino groups of lysine, detected in the above experiments, are

involved, is uncertain. However the linkage must be a different one to that discovered by Bornstein et. al. (1966) for rat skin collagen. In their experiments they demonstrated the involvement of lysine, and postulated that the final cross-link took the form of an aldol-type condensation of aldehyde functions derived from the ε-amino groups of lysine. Linkages of this type upon hydrolysis would not yield free lysine.

SUMMARY

The presence of 0.1 - 0.2 moles of free N-terminal groups per 100,000 g. of soluble collagen was demonstrated, and it was proposed that this arises as a result of either tissue catabolism or an incomplete masking of N-terminal residues. Liberation of carbohydrate from collagen by the action of α-amylase, without a concomitant release of significant amounts of extra N-terminal residues, precluded the involvement of hexose as a masking factor for α - amino groups; although linkage of carbohydrate to ε amino groups or histidine residues, is still a possibility.

All of the ε lysyl amino groups of collagen were demonstrated to be effectively free for substitution by the reagent 1 Fluoro 2:4 dinitrobenzene, with the reservation that one or two residues may be covalently bonded at the ε amino group and so not available to F.D.N.B. The presence of free lysine in hydrolyses of D.N.P. peptides obtained by collagenase digestion of collagen, and subsequent dinitrophenylation, was shown to be probably due to this. As more free lysine was found in analyses of material from insoluble collagen than of material from soluble collagen, it was concluded that the free lysine probably had been involved mainly in intermolecular crosslinking.

Further evidence for covalently linked ε-lysyl groups came from investigations on polar peptides obtained by electrophoretic separation of the collagenase-liberated peptides. It was suggested that these linkages may be involved in intermolecular crosslinking, but were not involved in intramolecular crosslinking of the type proposed by Bornstein et. al. (1965).

BIBLIOGRAPHY

Altgelt k. Hodge A.J. Schmitt F.O. (1961) Proc. Natl. Acad. Sci Wash 47 1914.

Astbury W.T. and Bell F.O. (1940) Nature 145 421.

Baglioni C. Ingram V.M. Sullivan E. (1961) Nature 183 467.

Bensusan H.B. (1965) Proceedings of M.A.T.O. Conference - Structure and Function of Connective and Skeletal Tissue. University of St. Andrews. Butterworths.

Bensusan H.B. Mcknight S. Naidu M.S.R. (1966) Fed. Proc. 25 No. 2 part 1. p.716.

Biserte G. and Osteux R. (1951) Bull. Soc. Chim. biol 33 50.

Blumenfeld O.O. Paz M.A. Gallop P.M. Seifter S. (1963) J. Biol. Chem. 238 3835.

Blumenfeld O.O. and Gallop P.M. (1962) Biochemistry 1 947.

Boedtker H. and Doty P. (1956) J.A.M. Chem. Soc. 78 4267.

Bornstein P. Kang A.H. Piez K.A. Proc. Nat. Acad. Sci. (1966) 55 417.

Bornstein P. Martin G.R. Piez K.A. (1964) Science 144 1220.

Bouchilloux S. and Cheftel C. (1966) Bioch. and Biophys. Res. Comms. 23 305.

Bowes J.H. and Kenten R.H. 1948 (a) Bioch. J. 43 358.

Bowes J.H. and Kenten R.H. 1948 (b) Bioch. J. 43 365.

Bowes J.H. and Moss J.A. (1953) Bioch. J. 55 735.

Butler W.T. and Cunningham L.W. (1965) J. Biol. Chem. 240 3449.

Chandrarajan J. and Bose S.M. (1965) Indian J. Biochem. 2 127.

Chibnall A.C. (1946) J. Soc. Leather Trades Chem. 30 1.

Chibnall A.C. and Rees M.W. (1953) Bioch. J. 63 105.

Cole S.W. (1933) Practical Physiological Chemistry p.327.

Crawhill J.C. and Elliott D.F. (1955) Bioch. J. 61 264.

Curtis D.H. and Spikes J.D. (1962) Proc. Utah Acad. Sci. Arts and Letters 39 189.

Davis N.C. and Smith E.C. (1955) Methods of Biochemical Analysis Vol. 2 Interscience.

Davison P.F. Drake M.P. (1966) Biochemistry 5 313.

Deane K.R. and Truter E.V. (1955) Bioch. Biophys. Acta. 18 436.

Doty P. and Nishihara T. (1958) Recent Advances in Gelatin and Glue Research Pergammon p.92.

Eastoe J.E. Bioch. J. (1955) 65 589.

Fessler J.H. and Bailey A.J. (1966) Fed. Proc. 25 No. 3 part 1. 1022.

Franzblau C. Gallop P.M. Seifter S. (1963) Biopolymers 1 79.

Franzblau C. (1962) Ph.D. Dissertation. Yeshiva University New York.

Gallop P.M. and Seifter S. (1962) Collagen. Ed. Ramanathan N. Interscience p.249.

Gallop P.M. Seifter S. Lukiu M. Meilman E. (1960) J. Biol. Chem. 235 2619.

Gallop P.M. Seifter S. Meilman E. (1959) Nature 183 1659.

Gallop P.M. (1966) Nature 209 73.

Grant R.A. Horne R.W. Cox R.W. (1965) Nature 207 822.

Grant N.H. & Alburn H.E. (1960) Arch. Bioch. and Biophys. 89 262.

Grassman W. Hannig K Schleger M. (1960) Z. Physiol. Chem. 322 71.

Grassmann W Nordwig A. Hörmann H. (1961) Z. Physiol. Chem. 323 48.

Grassman W. Hörmann H. Nordwig A. Wunsch E. (1959) Z. Physiol. Chem. 316 287.

Grassman W. and Schleich H. (1935) Bioch. Z. 277 320.

Grassman W. Hannig K. Endres H. Reidel A. (1957) Connective Tissue Ed. Tunbridge R.E. Blackwell.

Grassman W. Hannig K. Engel J. (1961) Z. Physiol. Chem. 322 71.

Grassman W. Endres H Steber A. (1954) Z. Naturforsch 96 513.

Grassman W. and Hörmann H. (1953) Z. Physiol. Chem. 292 24.

Greenberg J. Fishman L. Levy M. (1964) Biochemistry 3 1826.

Grettie D.P. (1965) J.AM. Leather Chem. Assoc. LX No.10. 572.

Gross J. (1961) Scientific American May 1961 p.33.

Gustavson K.H. (1954) Acta. Chem. Scand. 8 1298.

Hall C.E. and Doty P. (1958) J.AM. Chem. Soc. 80 1269.

Hallsworth A.S. (1964) Bioch. J. 93 255.

Hannig K. and Nordwig A. (1965) Proceedings of N.A.T.O. Conference - Structure and Function of Connective and Skeletal Tissue. University of St. Andrews. Butterworths.

Harding J. (1966) Advances in Protein Chemistry 20 Academic Press.

Harkness R.D. Marko A.M. Muir H.M. Neuberger A. (1954) Bioch. J. 56 558.

Heidrich H.G. and Wynston L.K. (1965) Z. Physiol. Chem. 342 166.

Hodge A.J. Highberger J.H. Deffner G.G.J. Schmitt F.O. (1960) Proc. Natl. Acad. Sci. U.S. 46 197.

Hodge A.J. Petruska J.A. Bailey A.J. (1965) Proceedings of N.A.T.O. Conference - Structure and Function of Connective and Skeletal Tissue. University of St. Andrews. Butterworths.

Hodge A.J. and Schmitt F.O. (1958) Proc. Natl. Acad. Sci. Wash. 44 418.

Hodge A.J. and Schmitt F.O. (1960) Proc. Natl. Acad. Sci. Wash. 46 186.

Hodge A.J. and Petruska J.A. (1963) Aspects of Protein Structure Ed. Ramanathan Academic Press.

Hörmann H. Joseph K. Wilm M. (1965) Z. Physiol. Chem. 341 284.

Hörmann H. (1960) Beitr. Silikose-Forsh Sonderband 4 205.

Hörmann H. Hannig K. Pries G. (1958) Z. Physiol. Chem. 315 109.

Howell R.M. Cook M.A. Scott G.B.D. (1964) Nature 204 1089.

Jackson D.S. and Fessler J.H. (1957) Connective Tissue Symposium. Ed. Tunbridge R.E. Blackwell p.59.

Jackson D.S. and Bentley J.P. (1960) J. Biophys. and Bioch. Cytol. 7 37.

Kovacs J. Konyes I. Pusztai A. (1953) Experimentia 9 459.

Kroner Th. D. Tabtoff W. McGarr J.J. (1953) J. Am. Chem. Soc. 75 4084.

Kroner Th.D. Tabroff W. McGarr J.J. (1955) J.Am. Chem. Soc. 77 3356.

Kühn K. (1960) Leder 11 110.

Kühn K. Kühn J. Hannig K. (1961) Z. Physiol. Chem. 326 50.

Kulonen E. Virtanen U.K. Viljanto J. Seppala P. Acta Chem. Scand. (1962) 16 695.

Leach A.A. (1966) Bioch. J. 98 506.

Levy M. and Fishman L. (1966) Fed. Proc. 25 No.2 792.

Lockhart I.M. and Abraham E.P. (1956) Bioch. J. 62 645.

Manahan J. and Mandl I. (1961) Bioch. Biophys. Res. Comms. 4 368.

Mechanic G. and Levy M. (1959) J. Am. Chem. Soc. 81 1889.

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

Michaels S. Gallop P.M. Seifter S. Meilman E. (1953) Bioch. Biophys. Acta. 29 450.

Matheson N.A. (1966) Bioch. J. 100 389.

Nagai Y. and Noda H. (1960) Bioch. Biophys. Acta. 34 298.

Näutö V. Pikkarainen J. Kuloneu E. (1965) J.Am. Leather Chem. Assoc. 60 63.

Nishihara T. (1963) Japanese Patent 9295/63.

Orekhovich V.N. and Schpikiter V.O. (1955) Doklady. Akad. Nauk. S.S.S.R. 101 529.

Piez. K.A. Gross J. (1960) J. Biol. Chem. 235 995.

Piez K.A. Weiss E. Lewis M.S. (1960) J. Biol. Chem. 235 1987.

Piez K.A. (1965) Biochemistry 4 2590.

Piez K.A. Eigner E.A. Lewis M.S. (1963) Biochemistry 2 58.

Piez K.A. Bornstein P. Lewis M.S. Martin G.R. (1965) Proceedings of N.A.T.O. Conference - Structure and Function of Connective and Skeletal Tissue. University of St. Andrews. Butterworths.

Porter R.R. (1957) Methods in Enzymology IV p.234. Academic Press.

Porter R.R. and Sanger F. (1948) Bioch. J. 42 287.

Randall J.T. (1954) J. Soc. Leather Trades Chemists 38 362.

Rao K.R. and Sober H.A. (1954) J. Am. Chem. Soc. 76 1328.

Rich A. and Crick F.H.C. (1961) J. Mol. Biol. 3 483.

Rojkind M. and Gallop P.M. (1963) Meetings Am. Chem. Soc. (Sept.) New York. (In Collagen Currents 4 No.542).

Rojkind M. Blumenfeld O.O. Gallop P.M. (1966) J. Biol. Chem. 241 1530.

Rubin A.L. Drake M.P. Davison P.F. Pfahl D. Speakman P.T. Schmitt F.O. (1965) Biochemistry 4 181.

Sanger F. (1945) Bioch. J. 39 507.

Schleyer M. (1962) Z. Physiol. Chem. 329 97.

Schmitt F.O. Gross J. Highberger J.H. (1953) Proc. Natl. Acad. Sci. Wash. 39 459.

Schmitt F.O. and Hodge A.J. (1960) J. Soc. Leather Trades Chem. 44 217.

Schroeder W.A. Kay L.M. Legette J. Honnen L. Green F.C. (1954) J. Am. Chem. Soc. 76 3556.

Schroeder W.A. Honnen L. Green F.C. (1953) Proc. Natl. Acad. Sci. Wash. 39 23.

Schrohenloher R.E. Ogle J.D. Logan M.A. (1959) J. Biol. Chem. 234 59.

Seifter S. Gallop P.M. Klein L. Meilman E. (1959) J. Biol. Chem. 234 285.

Smith J.W. (1965) Nature 205 356.

Solomons C.C. and Irving J.T. (1953) Bioch. J. 68 499.

Steinhardt J. and Zaiser E.M. (1955) Advances in Protein Chemistry X. Academic Press.

Steven F.S. and Tristram G.R. (1962a) Bioch. J. 83 240.

Steven F.S. and Tristram G.R. (1962b) Bioch. J. 83 245.

Steven F.S. (1962a) Analytical Biochemistry 4 316.

Steven F.S. (1962b) J. Chromatog. 8 417.

Steven F.S. (1964) Ann. rheum. Dis. 23 300.

Takahashi T. and Tanaka T. (1953) Bull. Japan Soc. Sci. Fisheries 19 603.

Taylor P. (1965) Honours Dissertation. University of St. Andrews.

Tristram G.R. (1953) The Proteins. Vol. 1A. p.181. Academic Press. New York.

Tristram G.R. (1964) Symposium - Calcified Tissues University of Liege p.215.

Tristram G.R. and Smith R.H. (1963) Advances in Protein Chemistry 18 Academic Press.

Veis A. Anesey J. Cohen J. (1960) J.Am. Leather Chem. Assoc. 55 548.

Veis A. and Anesey J. (1965) J. Biol. Chem. 240 3899.

Young E.G. Lorimer J.W. (1960) Arch. Bioch. Biophys. 88 373.