U. S. DEPARTMENT OF COMMERCE
W. AVERELL HARRIMAN, Secretary
NATIONAL BUREAU OF STANDARDS
E. U. CONDON, Director

CIRCULAR OF THE NATIONAL BUREAU OF STANDARDS C458

CHEMISTRY OF COLLAGEN

BY JOSEPH R. KANAGY

[Issued May 29, 1947]

UNITED STATES
GOVERNMENT PRINTING OFFICE
WASHINGTON : 1947

For sale by the Superintendent of Documents, U. S. Government Printing Office
Washington 25, D. C. - Price 10 cents

PREFACE

One of the most important industrial proteins is collagen. On treatment with tannins or certain inorganic salts, it is converted into leather. This process is known as tanning. By prolonged heating with water, collagen is changed into glue, which has a large number of commercial applications, such as the making of paper, photographic plates, and adhesives. This Circular gives a general outline of the physical and chemical properties of collagen, and was prepared for the large number of chemists employed in industries in which this material or its transition product, glue, is processed.

E. U. Condon, *Director.*

CHEMISTRY OF COLLAGEN

By Joseph Kanagy

ABSTRACT

One of the most important industrial proteins is collagen. On treatment with tannins or certain inorganic salts, it is converted into leather. This process is known as tanning. By prolonged heating with water, collagen is changed into glue which has a large number of commerical applications, such as the making of paper, photographic plates and adhesives. This Circular gives a general outline of the physical and chemical properties of collagen and was prepared for the large number of chemists employed in industries where this material or its transition product, glue, is processed.

CONTENTS

I. INTRODUCTION

Collagen occurs in the skin, connective tissues, and bones of animals. It is the principal constituent of tendons. As it forms about 35 percent of the coria of fresh cattle hide, it is the most important fibrillar material in leather. It is insoluble in organic solvents, in water, and in dilute solutions of acids and alkalis at ordinary temperatures.

A distinguishing and highly important property of collagen is its very large swelling capacity in aqueous acid and alkaline systems in the absence of high-salt concentrations. The prolonged action of hot water converts collagen into gelatin. This reaction, which is of great technical importance, may be considered as the most important single characteristic feature of the protein, aside from its leather-forming and swelling powers. It is responsible for the name collagen, which is derived from the Greek and means "glue former."

II. SHRINKAGE TEMPERATURE OF COLLAGEN

Collagen exhibits a characteristic shrinkage temperature. When heated in water, it contracts suddenly at a temperature which varies somewhat with the environment and previous treatment, but which is approximately 60° to 65° C for fibers from fresh untreated skin.

This so-called shrinkage temperature is a transition point, or fusion, similar to the melting point of other more definitely crystalline materials. The fibers contract to about one-third their original length, increase considerably in thickness, and acquire rubber-like elasticity in the wet condition. The material has apparently changed from a fibrous form to a rubberlike one. On further heating it loses its elasticity and, in approaching the form of gelatin, it assumes properties closely resembling plastic flow.

On warming, tendons, which consist of collagen, behave in a remarkable way. They contract in the direction of the fiber axis and become elastic like rubber. At the same time the X-ray diagram disappears. The behavior of unwarmed tendon is similar to that of stretched frozen rubber, that of the contracted tendon is similar to amorphous rubber obtained from stretched and frozen rubber by warming.

According to Meyer and Ferri [1][1], the restoring force of stretched tendon in the rubber-like condition has the same origin as that of stretched rubber. The chains that have been induced to abandon probable twisted configurations for the less probable stretched formation tend to resume the statistically favored contracted positions and configurations under the influence of thermal agitation. If fully stretched and at a sufficiently low temperature, the chains snap into position in the crystal lattice with the evolution of heat of crystallization. Hollemann [44] and Kuntzel [45] have determined calorimetrically the heat of fusion of collagen, which amounts to about 12.5 cal/g.

Collagen chains provide an example that protein chains may exist in the stretched conditions under certain circumstances, while under other circumstances they are contracted. Wohlisch [2] supposed at one time that the contraction of tendon like that of stretched, frozen rubber upon heating was due to the reversible conversion of one form into another. Later he suggested that the contraction represented the first state of the irreversible conversion into gelatin. In the light of more recent work, the first interpretation must be accepted and the contraction regarded as a fusion similar to the fusion of rubber.

It is more precise, however, to speak of a two-dimensional solution of the crystallites in the water of imbibition as the melting point of water-free tendon lies at about 140° C, and is lowered in the presence of water to about 60° C. A system consisting of a crystalline substance and solvent possesses a definite saturation concentration for each temperature; conversely at each concentration there is a definite saturation temperature above which crystallites disappear and below which crystallites are formed. In all chain polymers, the temperature at which the product is sparingly soluble is separated by only a very small interval from that at which it dissolves in considerable quantities; as much as 20 percent, for example. For this reason the solution of the crystallites then appears like fusion.

[1] Figures in brackets indicate the references given at the end of this paper.

The greater the solubility of the collagen chain in the solvent, the lower is the temperature of fusion. Salts that increase the solubility of peptides diminish the fusion temperature, whereas salts that lead to "salting out" raise this point. If the fibers are soaked in formamide, which has a high solvent effect on proteins, contraction takes place at about 35° C, whereas collagen tanned with vegetable tanning materials contracts at 75° to 80° C and that tanned with chromium salts at over 100° C.

III. FORMATION OF GELATIN AND GLUE

On prolonged treatment with hot water, collagen fibers are converted into water-soluble gelatin. The nature of this reaction has been the subject of considerable discussion, and it is believed related to the transformation of native collagen into the contracted, amorphous form. A possible means of detecting even slight chemical changes is provided, according to reports [3], by the behavior in the presence of trypsin; native fibers are unattacked but gelatin is hydrolyzed.

On the other hand, collagen that has been converted into the contracted condition at 35° C in the presence of formamide is not attacked by trypsin. The contraction can also take place without the fibers losing their resistance to the trypsin attack if they are dipped for a few seconds into water at pH 8 and at a temperature of 55° C. On prolonged heating, the specimen soon shows plastic flow and goes into solution.

Gelatin formation may occur without fusion as a result of treatment for considerable time with cold alkali. The fibers retain their original length, and the crystalline regions remain unmelted. Swelling takes place at right angles to the fiber axis only. On washing out the alkali, the swelling persists. In this way an oriented gelatin gel is formed, which is rapidly attacked by trypsin and which, on warming, passes into solution.

Two superimposed processes seem to be involved in the formation of gelatin: (1) reversible fusion in the water of imbibition, and (2) irreversible hydrolysis of linkages, including possibly a few peptide linkages leading to formation of soluble products consisting of chain fragments. In figure 1 is shown the increase in amino nitrogen and decrease in the amount of protein precipitated with phosphotungstic acid as the hydrogen-ion concentration of a gelatin solution is increased. These results indicate that the formation of water-soluble gelatin from collagen may occur at high pH values without increasing appreciably the amino-nitrogen content. Therefore, few peptide linkages are broken in the process.

Prolonged heating of soluble gelatin in water or in acid solution leads to progressive diminution of chain lengths and hence to reduction in viscosity. Osmotically active particles of molecular weights from 20,000 to 70,000 [4] have been found in commercial samples of gelatin. Hot dilute water solutions of gelatin display no marked structural viscosity, whereas, at room temperature, they display a number of properties which can be explained only by the simultaneous presence of single molecules, molecular aggregates, and larger micells. Such a system reaches equilibrium slowly.

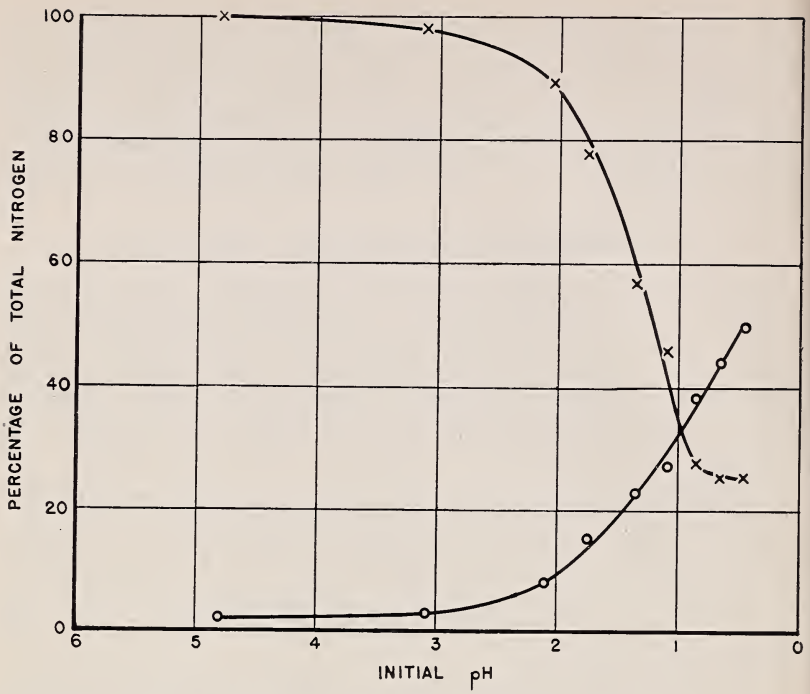


FIGURE 1.—*Change with* pH *of the amino nitrogen and nitrogen precipitated with phosphotungstic acid for solutions of hide substance (gelatin) hydrolyzed with sulfuric acid.*

x, precipitated with phosphotungstic acid; *o*, determined as amino nitrogen.

Some work with the ultracentrifuge on the physicochemica behavior of solutions of degraded gelatin has recently been done by Scatchard, Oncley, Williams, and Brown [5]. They assume that the bonds between the long chains of polypeptide residues are usually hydrolyzed at about the same rate, but that there are a few bonds equally spaced along the chain that hydrolyze very much more rapidly. In the preparation of gelatin, the more reactive bonds are hydrolyzed. The degradation of gelatin then consists in the hydrolysis of the less reactive peptide linkages. There is thus an ideal parent undegraded gelatin molecule, which is the length of chain between two reactive bonds. Gelatins thus consist of a mixture of such molecules with the products of their degradation, which include every possible peptide from single amino acids to chains containing only one less residue than the parent molecule. They obtained data with undegraded gelatin and four samples of gelatin in different degrees of degradation. The results are given in table 1.

The quantity $Mz/Mn-1$ is a good measure of the size spread. If all molecules are of the same size, it is 0, and for random splitting of an infinitely long chain, it becomes 2. The ratio $Mw/Mn-1$, which increases from 0 to 1, is also an indication of the size spread. As the fraction of parent molecules decreases the size spread increases to

TABLE 1.—*Physicochemical characteristics of degraded gelatins*

Sample [1]	α [2]	C_p	Mn	$Mw/Mn-1$	$Mz/Mn-1$	H_0 at 55° C
A	0.00120	0.244	45,700	0.61	0.92	0.47
B	.00175	.129	36,000	.73	1.17	.36
C	.00227	.071	30,000	.81	1.38	.31
D	.00409	.008	19,000	.92	1.59	.18
E	.00540	.002	15,000	.96	1.82	.14

[1] Sample *A* is undegraded gelatin; the other four samples are gelatins in different degrees of degradation.
[2] α=Fraction of bonds in the parent molecule which are broken. C_p=weight fraction of unchanged parent molecules. Mn=number average molecular weight. Mw=weight average molecular weight. Mz=z average molecular weight. H_0=intrinsic viscosity.

approach that which would be obtained from infinitely long parent molecules.

Other determinations made with the untracentrifuge showed that the number of amino-acid residues per undegraded molecule was 1,170. The diameter was calculated from viscosity measurements and was equal to 17 A. The length was 800 A.

IV. SWELLING OF COLLAGEN AND GELATIN

Probably no other phase of the physicochemical behavior of collagen is so important in its bearing on technological processes as its swelling properties. Collagen and gelatin both show maximum points of swelling at pH values of approximately 2.5 and 12.5. Closely associated with the swelling phenomena are the amphoteric properties of the protein, which permit its combination with either acids or bases, depending upon the pH value of the environment.

Most of the earlier work on swelling was done with gelatin. Among the earlier workers who should be mentioned are Loeb [6], who applied theory of the Donnan equilibrium in explaining the mechanism of swelling. Previous to this was the work of Proctor [7] and of Proctor and Wilson [8].

E. C. Porter [9] studied the swelling of hide powder by measuring the volumes occupied by the settled powder when equilibrium had been established in buffer solutions of different pH values. He found a minimum at pH 4.8 and two maximums at pH 2.4 and 12.5.

McLaughlin [10] and McLaughlin and Porter [11] did the earliest work on swelling, in which pieces of hide itself were utilized. Similar studies have been made by Wilson and Gallun [12], R. E. Porter [13] and Page and Gilman [14]. All of these investigators found maximum points of swelling at pH values of approximately 2.4 and 12.5, with minimum values in the pH range of 5.0 to 10.0. Wilson and Gallun found two points of minimum swelling at pH 5.1 and 7.6.

Highberger [15] has recently studied the swelling of collagen in unbuffed solutions of hydrochloric acid and sodium hydroxide at 25° C. A curve showing the volume increase of collagen with change in pH value is given in figure 2.

In the swelling of collagen, a definite solution pressure exists and part of the protein goes into solution. This was shown by Highberger [15], and his results are given in figure 3. In general the solution pressure is greatest where swelling is also greatest.

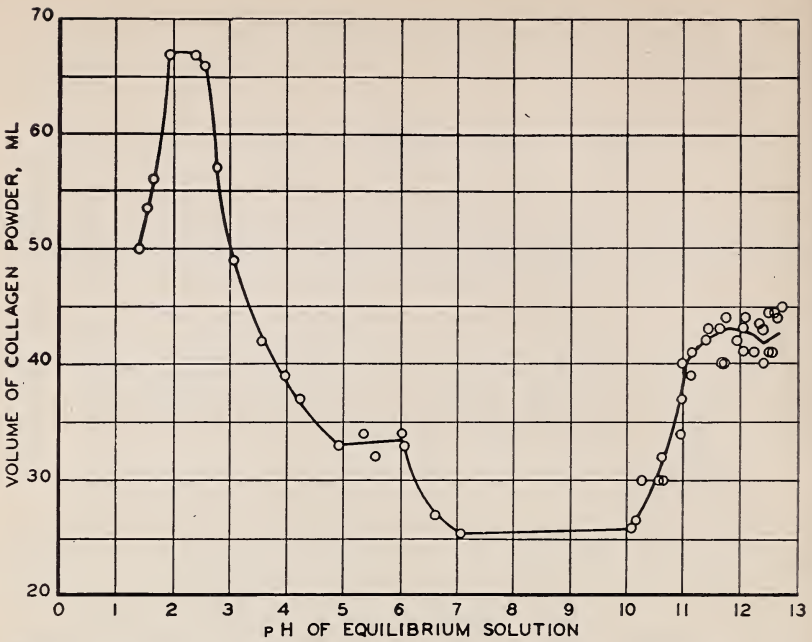


FIGURE 2.—*Effect of* pH *on the volume increase of collagen in unbuffered solutions.*

Volume swelling curve of collagen. 1 g collagen per 100 ml solution, 48 hr at 25° C. From the Journal of the American Leather Chemists Assoc.

The swelling behavior of collagen and gelatin differs mainly in respect to the effect of temperature. Gelatin goes into solution at a lower temperature than collagen, which might be expected from the fact that the former is already in a partially degraded or hydrolyzed condition.

Dorothy Jordan-Lloyd [16] studied the effect of temperature on the swelling of gelatin in hydrochloric acid and sodium hydroxide solutions. The results of some of her work are given in figure 4. At 25° C the gelatins dissolve at those pH values at which the greatest amount of swelling occurs.

The effect of neutral salts on the swelling of gelatin in acidic and basic solutions was studied by Dorothy Jordan-Lloyd and Winifred B. Pleass [17]. They used sodium chloride in the presence of hydrochloric acid and sodium hydroxide, and sodium nitrate in the presence

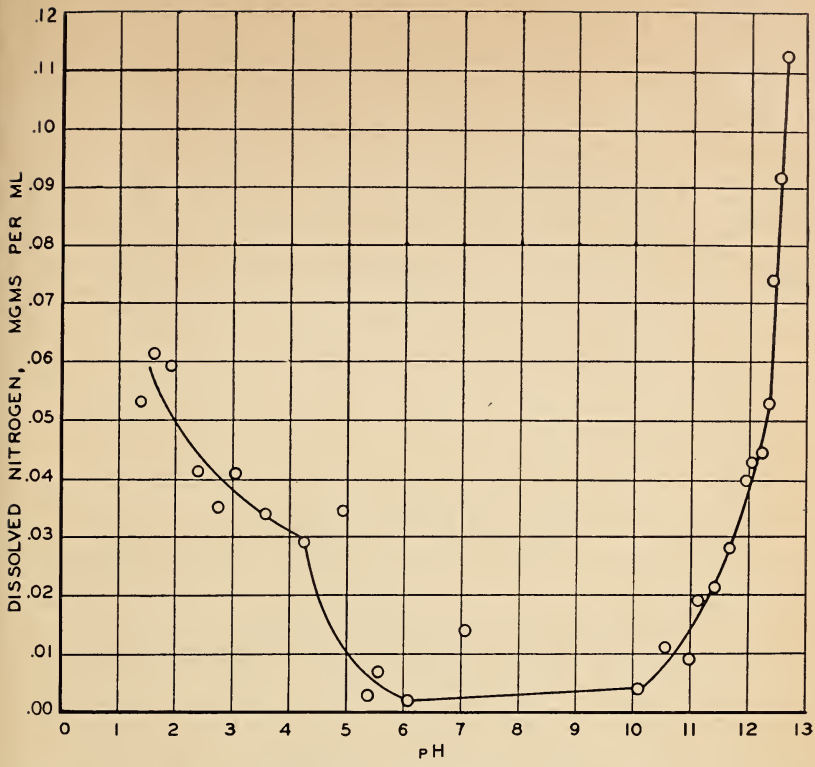


FIGURE 3.—*Solubility of collagen as determined from the total nitrogen in solutions as a function of* pH.

Dissolved nitrogen curve. 1 g collagen per 100 ml solution, 48 hr at 25° C. From the Journal of the American Leather Chemists Assoc.

of nitric acid and sodium hydroxide. The results are shown in figures 5 and 6.

The neutral salts greatly decrease swelling at the points where maximum swelling occurs. The highest concentrations of salts decrease swelling at these points to even less than that which occurs with distilled water. In the pH regions of about 4 to 7, or what is generally referred to as the isoelectric region, increasing concentrations of salt actually increase swelling. This is believed to be due to the solvent action of the salt, and is probably related to the Hofmeister theory of the swelling effect of salts having different valences.

736525—47——2

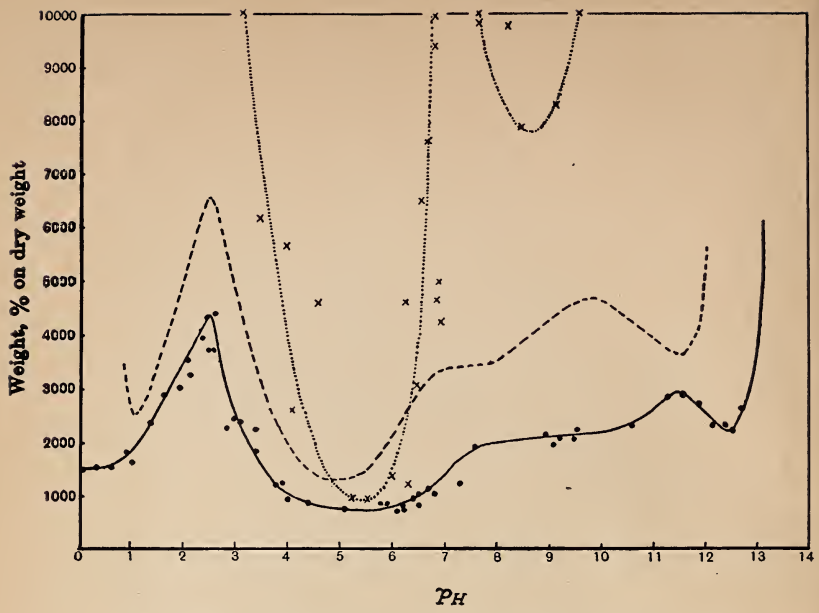


FIGURE 4.—*Swelling of pure gelatin at various temperatures.*

---- 0°C, pH of neutrality 7.35 ---- 18° C, pH of neutrality 7.06 x----x 25° C, pH of neutrality, 6.95. From Cambridge Univ. Press.

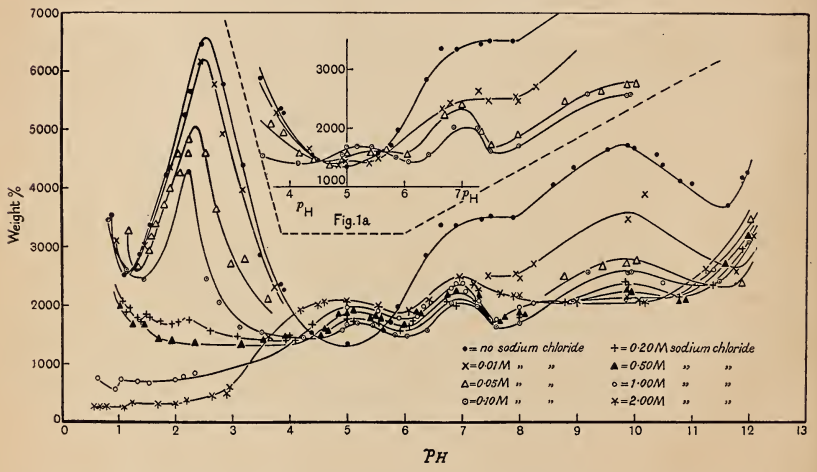


FIGURE 5.—*Effect of hydrochloric acid and sodium hydroxide and these reagents with various concentrations of sodium chloride on the swelling of gelatin.*

pH varying, t constant at 18°C, M constant over a series of concentrations from 0 to 2.0 M. From Cambridge Univ. Press.

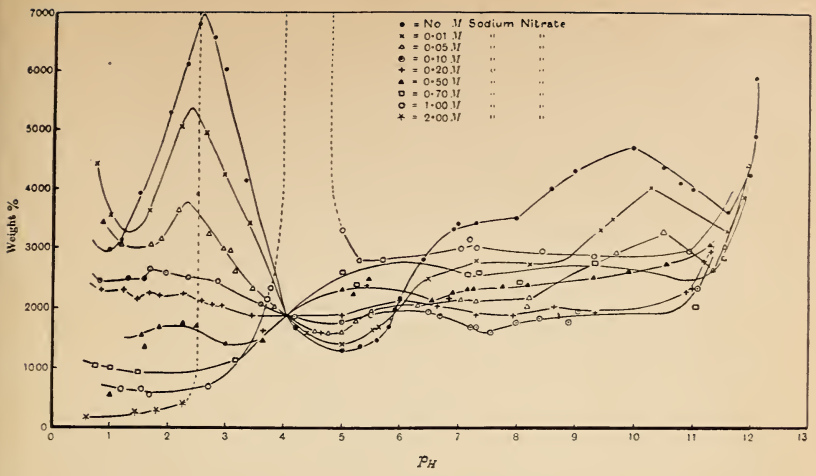


FIGURE 6.—*Effect of nitric acid and sodium hydroxide, and these reagents with various concentrations of sodium nitrate on the swelling of gelatin.*

pH varying, *t* constant at 18° C, *M* constant over a series of concentrations from 0 to 2.0 *M*. From Cambridge Univ. Press.

V. STRUCTURE OF COLLAGEN

1. PREPARATION OF PURE COLLAGEN

The problem of preparing pure collagen is one of considerable complexity. The difficulty is accentuated by the fact that there is no way of determining when the collagen is pure, and also that there may be several different types of this protein. At the present stage of development of this problem, we can speak precisely only in terms of "purified" collagen.

Collagen is obtained from the skin fibers of animals, and is associated closely with small amounts of other proteinaceous as well as fatty materials. Among these proteinaceous materials are reticulin, a fine, filamentous network which surrounds the cellular tissues; elastin, which is the protein of the yellow elastic tissue; glycoprotein, a complex material that may be composed in part of chrondroitin sulfuric acid; and traces of albumins and globulins. The total concentration of these extraneous materials is probably not over 3 percent. The fatty materials consist of triglycerides, phospholipides, and cholesterol. The total amounts of these materials in the skin varies considerably, but on the average they probably do not exceed 1 percent.

The preparation of purified collagen then consists in the removal of most of these materials by chemical and physical methods without appreciably affecting the collagen. Such a procedure has been developed by Highberger [18].

The procedure developed by Highberger for the preparation of a purified collagen from animal skin, consists in the mechanical separation of the entire epidermal system and the flesh layer remaining on

the under side of the skin. There remains then only that part of the skin known as the corium which is composed principally of collagen. Of the protein constituents of animal skin corium, it is possible to effect reasonable separation of all but two from the collagen by the use of water alone or dilute aqueous salt solutions. The two substances referred to are elastin, which exists as fibers of yellow elastic tissue, and a mucoprotein, referred to usually as "mucoid", the nature of which is not well understood.

In the first step of the purification process, the corium, cut into pieces approximately 2 in. by 2 in., is treated with trypsin. This enzyme is supposed to be able to remove elastin without appreciably affecting the collagen. In view of the uncertainty of the action of trypsin, it is questionable as to whether such treatment is justified. However, Highberger reasoned that its inclusion in the process was permissible for several reasons, namely, (1) some thorough washing and soaking procedure is necessary in any event to remove interfibrillary proteins, (2) in case no other means are adopted, such a process must use a salt solution in order to insure the removal of globulins, (3) under properly controlled conditions tryptic digestion can be carried out without appreciable attack on the collagen, (4) by combining the tryptic digestion with the washing process, globulins and other proteins are solubilized and removed without the use of salt, and the benefit of any elastin removal is obtained.

After the digestion with trypsin, the corium material is dehydrated in alcohol, air-dried, and ground in a Wiley mill. It is then extracted with acetone, which is followed by extraction with ethyl alcohol at temperatures no greater than 50° C for the removal of the lipid materials.

For the removal of the mucoid material, half-saturated calcium hydroxide is used. This mucoid may be precipitated by acidifying the calcium hydroxide extract with acetic acid. It appears as a thick, cream-colored paste on centrifuging and washing free of acetic acid.

To remove the calcium accumulated in the collagen in the previous treatment, very dilute acetic acid is used. The treatment with the acid is made at a low concentration in order to prevent gelatinization.

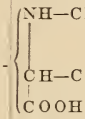
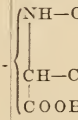
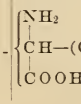
After the final treatment with acid, the collagen is dehydrated in cold 95-percent ethyl alcohol. The alcohol is advantageously washed out of the collagen with ethyl ether before the final air drying, as this promotes a better separation of the fibers.

Collagen prepared by this method is a white, fibrous material. In distilled water it gives a pH of 4.5 to 5.5. The ash content varies with individual preparations, but generally is not over 0.1 percent.

2. AMINO-ACID COMPOSITION OF GELATIN AND COLLAGEN

The structural units of collagen, which are obviously the same as those for gelatin, consist largely of amino acids. The names, formulas, and percentages present of the principal amino acids in collagen or in gelatin are given in table 2. These results were taken from the work of Bergman [19], Dakin [20], and Grossmann [21].

TABLE 2.—*Amino-acid composition of collagen*

Amino acid	Formula	Dry protein weight
		Percent
Glycine		26.5 [19]
Alanine		8.7 [20]
Leucine		7.1 [20]
Proline		17.5 [19]
Hydroxyproline		14.4 [19]
Arginine		8.4 [21]
Lysine		5.9 [21]
Histidine		0.6 [21]
Aspartic acid		3.5 [21]
Glutamic acid		5.7 [21]

Of the 10 different amino acids there may be some doubt as to the existence of histidine. The methods used in the determinations of the amino acids are so cumbersome and tedious that a result of less than 1 percent for any particular amino acid may not be significant. Further, it is difficult to fit an amino acid at such a small percentage into any stoichimetric relationship with the other amino acids of the collagen molecule.

3. OPTICAL STUDIES ON COLLAGEN

Studies with the electron microscope and X-ray diffraction patterns have given some very valuable information about the molecular structure of collagen. For these investigations we are indebted to W. T. Astbury [22] of the University of Leeds, and R. S. Bear [23] and F. O. Schmitt [24] of Massachusetts Institute of Technology. Most of this work has been done with tendons because their high degree of fiber orientation facilitates certain of the analyses.

The maximum resolution obtainable with the electron microscope on protein fibers like collagen is about 50 A and the electron micrographs permit direct visualization of the average density of packing of the protein material within the resolvable regions. The correlation of the results of X-ray diffraction and electron microscopy is highly valuable. The former offers high resolution and is applicable to moist or dry fibers, whereas the latter reveals the larger structures in the submicroscopic fibrils.

In figure 7 is given a characteristic diffraction pattern of collagen fibers. The chief features of the pattern are the equatorial diffractions (E) at about 11 A, the meridional diffractions (M) at 2.86 A, and the long spacing diffractions (L). The equatorial spacing reflects the distance laterally between the polypeptide chains. This distance is increased to 15 A and longer by treatment with water, indicating that water enters between the chains of the submicroscopic fibrils. The meridional spacing is usually identified with the average length of the amino-acid residues along the chain. Astbury [22] suggests that the

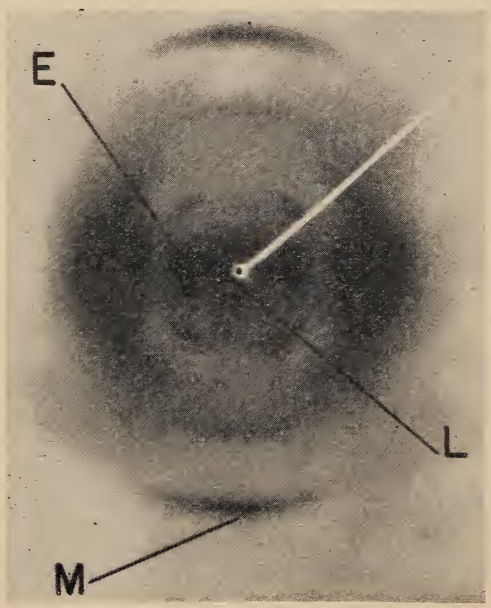


FIGURE 7.—*Short-spacing* X-ray *diffraction pattern of beef tendon.*

Fiber axis is vertical. Characteristic diffractions are: E, equatorial "side chain" spacing; M, meridional diffraction at about 2.86 A; L, meridional long spacings. From MIT.

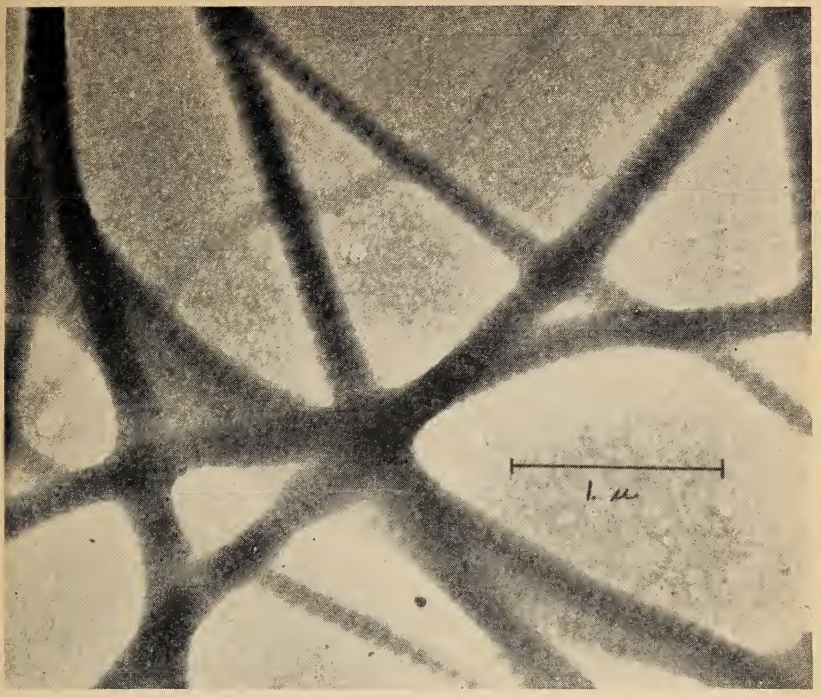


FIGURE 8.—*Electron micrograph of collagen fibers.*

Magnification 32,500. From MIT.

polypeptide chains run in essentially straight lines parallel to the fiber axis, the length of the average residue being less than that in other fully extended protein chains such as silk (3.5 A) because the high concentration of proline and hydroxy proline residues constricts the chains. The long-spacing diffractions are on the fiber axis. Bear [23] has shown that these diffractions are all orders of a fundamental fiber-axis period of about 640 A.

Figure 8 shows an electron micrograph of collagen fibers in which the fibrils show a banded appearance with light and dark bands spaced uniformly and appearing alternately. The average distance from one dark band to another is found to be about 640 A, which coincides with that found for the X-ray long-spacing on the fiber axis. The greater density of the dark bands is believed to be due to a more compact arrangement of the protein. When the fibril is stretched, the light bands increase in length at the expense of the dark bands, consistent with the view that the protein may be more densely packed in the dark than in the light bands. Whole tendon or other collagenous tissues do not show great extensibility. However, for submicroscopic fibrils teased from the tissue, the extensibility may be very great.

Bear [23] has studied the pattern of collagen fibers treated in a number of ways. Some of these results are given in table 3. It may be seen that the long-spacing in the fiber axis varies considerably, from 675 to 550 A, depending on the treatment of the collagen. The merid-

ional spacing shows practically no change, whereas the equatorial spacing varies from 11.1 to 15.5, increasing with treatment of the tendon with water.

TABLE 3.—*X-ray diffraction spacings of collagen fibers*

Collagen source	Pretreatment before drying	X-ray spacings		
		Long spacing	Short spacing	
			Meridional	Equatorial
		A	A	A
Kangaroo tendon	Water	675 [1]	[1] 2.87	[1] 15.5
Do	do	645	2.90	11.5
Beef tendon	do	640	2.85	11.1
Rat tendon	do	640	2.88	11.2
Do	5% tannic acid	625	2.85	11.5
Do	5% chromic chloride	625	2.85	11.9
Do	10% formalin	615		
Rat tendon	10% formalin, maximum heat shortened, reextended on cooling. Final shortening 13%.	550	2.85	11.5
Do	10% formalin, maximum heat shortened. Final shortening 64%.	Lacking	2.87	11.5
Gelatin	Oriented by tension	do	2.85	11.1

[1] These figures are for moist tendon; all other data are for air-dried material.

The fiber-axis long-spacing increases on treatment with water and decreases with treatment with tanning materials. On treatment with heat and reextending, the long-spacing decreases to 550 A. With further heat treatment and shortening, the long-spacing disappears entirely. This evidence, together with the fact that gelatin shows no characteristic long-spacing, indicates that the major portion of the crystal structure in collagen disappears early in the degradation process.

4. PERIODICITY THEORY

As the results of X-ray and electron microscope studies show evidence of crystallinity and a repeated unit structure in the fibrils, there must also be a repeated chemical structure. Such a sequence or orderly arrangement would be expected from our knowledge of other natural organic compounds and the apparent uniformity of the chemical reactions of collagen. Further, the isolation of a tripeptide lysly-prolyl-glycine from gelatin by Grassmann and Riederle [25] is definite evidence of such an arrangement.

Bergmann and Niemann [26] have postulated that the total number of residues in a protein molecule or repeated unit is expressible in the form $2^n \times 3^m$ Bergmann estimates the average residue weight in gelatin to be about 94, which means that 100 g of the protein contains 100/94, or 1.06-g, residues of the average residue. Thus [22] glycine, which has a molecular weight of 75 and is contained in gelatin at a concentration of 26.5 percent, is present in 100-g of protein to the extent of 0.34 g moles. Therefore, it may be inferred that one-third of the residues in gelatin are glycine residues. Proceeding in this manner with the other amino acids, the results given in table 4 are

obtained. All the data in this table, except those for glutamic and aspartic acid, were taken from [22].

It can be seen from the table that at least 72 residues would be required in the collagen molecule to satisfy this hypothesis. If histidine is present, a much higher multiple of 72 would be required. Astbury [22] has suggested 576. If the histidine content were 0.86 percent, its residue would constitute 1/192 of the whole and the least common multiple of this frequency would be $2^6 \times 3^2$, or 576.

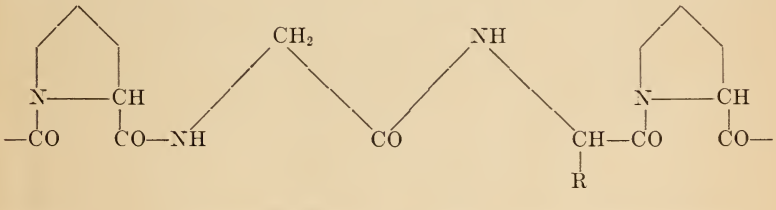
TABLE 4.—*Amino-acid frequencies in gelatin*

Amino acid	Weight	Molecular weight	Gram mole	Frequency
	Percent			
Glycine	26.5	75	0.35	$3(2^0.3^1)$
Proline	17.5	115	.16	$6(2^1.3^1)$
Hydroxyproline	14.4	131	.11	$9(2^0.3^2)$
Alanine	8.7	89	.098	$9(2^0.3^2)$
Arginine	8.4	174	.048	$18(2^1.3^2)$
Leucine	7.1	131	.054	$18(2^1.3^2)$
Lysine	5.9	146	.040	$24(2^3.3^1)$
Glutamic acid	5.8	147	.040	$24(2^3.3^1)$
Aspartic acid	3.4	133	.025	$36(2^2.3^2)$

Astbury further concluded that every third residue, with the exception of 1 in 18, was a proline or hydroxyproline residue, while glycine also occurs as every third residue. The other residues were assumed to be the other amino acids. Thus the general structure was represented as

$$-P-G-R-P-G-R,$$

where P is proline or hydroxyproline, G is glycine, and R is one of the other amino acids. The structural formula may be written as follows:



Astbury also assumed that the large number of proline and hydroxyproline residues exerted an influence on the configuration, so that the length of a residue was 2.85 A rather than 3.5 A, as in the case of silk. He also stated that these residues were on only one side of the chain, and gave this as a reason for the unextensibility of the collagen chains. This hypothesis has, however, been disputed by Huggins [27], who proposed a more probable configuration in which he assumed that the chains were in the form of a spiral, with proline and hydroxyproline existing on all sides of the chain.

5. DETERMINATION OF EQUIVALENT WEIGHT OF COLLAGEN BY HYDROGEN CHLORIDE ADSORPTION

Probably the best fundamental study on the determination of an equivalent weight or a definite chemical unit of structure of collagen was made by Beek [28]. He studied the combination of hydrogen chloride gas with collagen, gelatin, and leather under various conditions, and found that the amounts in every case could be expressed by a definite factor or a multiple of this factor. The results of his study are summarized in table 5.

TABLE 5.—*Adsorption of* HCl *gas by collagen, gelatin, and leather*

Material and condition	f[1]	$f/0.106$[2]	a[3]	f/a[4]
Collagen; HCl for 151 days	0.395	3.75	15	0.02647
Collagen; HCl for 45 days	.398	3.75	15	.02647
Collagen; HCl for 45 days, then brought to equilibrium in presence of sodium for 49 days	.106	1.00	4	.02650
Collagen; HCl removed by sodium in previous experiment	.317	3.00	12	.02642
Collagen; HCl at various pressures:				
12 mm	.184	1.74	7	.02629
30 mm	.236	2.23	9	.02622
Leather; HCl at 15 mm	.343	3.24	13	.02638
Gelatin; HCl for 7 days	.633	5.97	24	.02637
Gelatin; heated at 180° C for 2 days, then heated with HCl for 7 days	.106	1.00	4	.02650

[1] Fraction of total nitrogen combined.
[2] Lowest combining factor.
[3] Integer expressing relative combining fraction.
[4] Average value of f/a=0.02638.

The repeating unit in the collagen structure appears to contain at least 38 nitrogen atoms. This is derived from the reciprocal of the factor 0.02638, which is the average value of f/a from Beek's results, 1/0.02638=37.91, or 38.

The work of Thomas and Kelly [29] on the combination of iron and chromium with collagen may also be considered in connection with the results obtained by Beek. The equivalent weights of collagen calculated from the results given in their paper are given in table 6, together with the corresponding values of f and a. The

TABLE 6.—*Combining weights of collagen with chromium and iron from data of Thomas and Kelly*

Combining weight	f	$a/38$	a
94	0.836	0.842	32
187	.421	.421	16
380	.207	.211	8
750	.105	.105	4
299	.263	.263	10

values of $a/38$ are given for comparison with the observed values of f. The results in the table indicate that chromium and iron combine with proportions of collagen of similar magnitude and variation as HCl.

Beek [30] also studied the adsorption of hydrochloric acid by collagen from saturated sodium chloride solutions. The results of his studies

on collagen and those of similar studies on gelatin by Proctor [31] are given in figure 9. From these results Beek found that 3 nitrogen atoms in each unit of 38 are in groups much more strongly basic than the remainder. The apparent equivalent weight of collagen corresponding to this value when calculated on the basis of 17.86 percent of nitrogen is approximately 1,000.

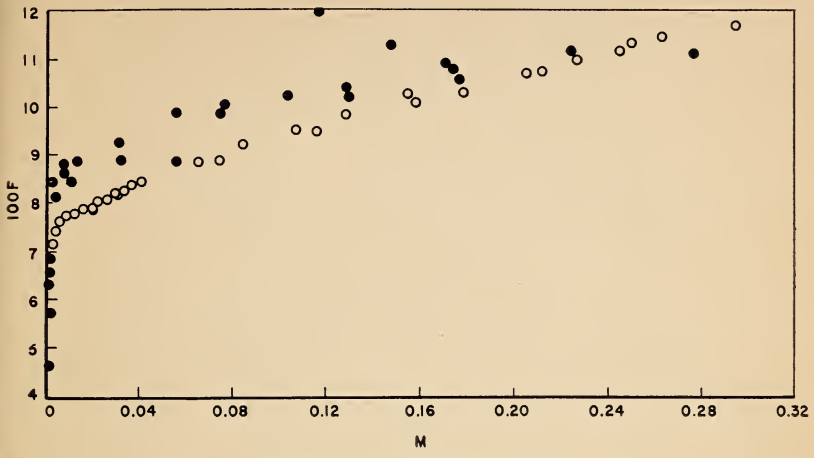


FIGURE 9.—*A comparison of the adsorption of hydrochloric acid from solutions saturated with sodium chloride, by gelatin and collagen.*

●, Proctor, gelatin; ○, collagen. The data for gelatin are taken from the work of Proctor [7], and the data for collagen from that of Beek [30]. f is the fraction of the total nitrogen which reacts with the acid. M is molality of hydrochloric acid.

6. AMINO NITROGEN IN COLLAGEN

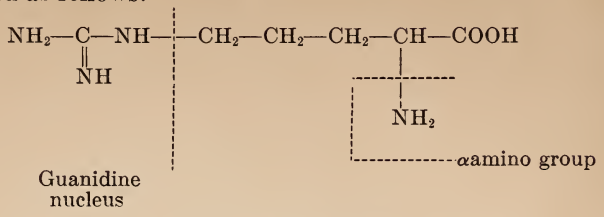
The basic amino acids, arginine, lysine, and histidine, which are of primary importance in proteins, appear to be held in the protein chains by peptide linkages involving their carboxyl and amino groups. Presumably, the remaining basic groups are free in the native protein and as a result, they have been related to the combining of various proteins with acids, dyes, and tannins. A study of the amino-nitrogen content of collagen by the reaction with nitrous acid (Van Slyke method [33]) was made by Kanagy and Harris [32]. The reaction of nitrous acid with primary amino-nitrogen groups proceeds as follows:

$$HNO_2 + RNH_2 \longrightarrow ROH + H_2O + N_2.$$

Thus by determining volumetrically the amount of nitrogen evolved, it is possible to determine the number of such groups in collagen.

The rate of evolution of nitrogen from collagen and from wool with time of reaction is shown in figure 10. The curves indicate that two types of reactions are occurring. One reaction, which undoubtedly is that of a primary amino group, is completed within a very short time, while the other continues slowly at a uniform rate. The only amino acid present in collagen that is known to give off increasing

amounts of nitrogen with time is arginine [34]. Its structural formula is represented as follows:



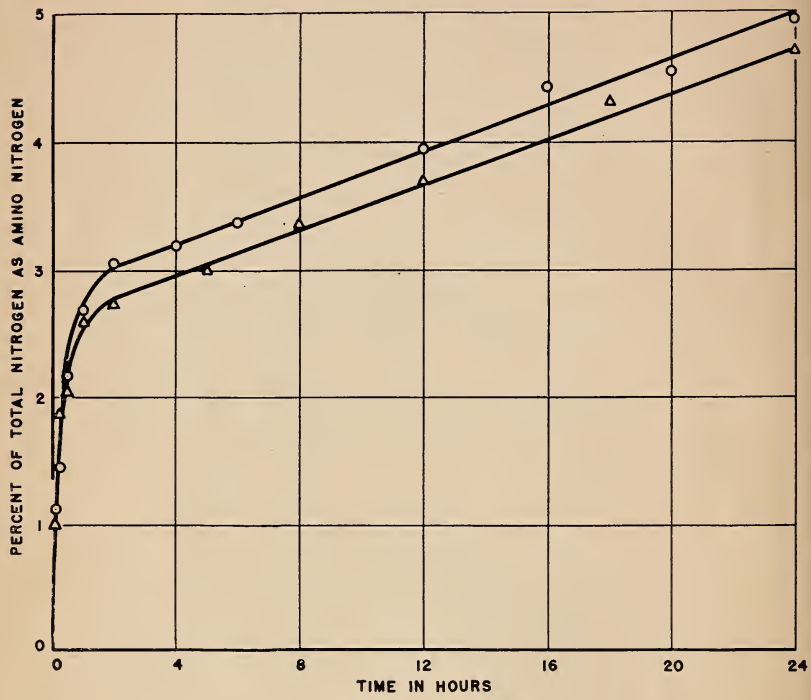


FIGURE 10.—*Rate of evolution of nitrogen from collagen and wool treated with nitrous acid.*

○, collagen; △, wool.

It seems logical, therefore, to study the rate of evolution of nitrogen from this amino acid under conditions identical with those used for collagen. The results are shown in figure 11. As arginine contains four nitrogen atoms, it is obvious that an amount of nitrogen equivalent to one atom is evolved during the first 5 minutes of reaction. This nitrogen comes from the amino group. The nitrogen subsequently evolved is formed by the action of nitrous acid on the guanidine nucleus.

The α-amino group of arginine is, of course, combined in the polypeptide chain of the protein, whereas the guanidine nucleus is free. The primary amino nitrogen in the protein evolved in the rapid initial reaction apparently comes from the ϵ-amino group of lysine, since

this is the only amino group known to be free. This reaction might therefore be used for estimating the amounts of these two amino acids in the protein.

The similarity of the reaction of nitrous acid with arginine to its reaction with collagen indicates that the continued evolution of nitrogen (as represented by the straight line portions of the curves in fig. 10) is due to the action of nitrous acid on the guanidine nucleus of these proteins. If it is assumed that all of this protein of the total nitrogen evolved comes from the guanidine nucleus, then the fraction of the total nitrogen of the protein as arginine nitrogen is equal to the ratio

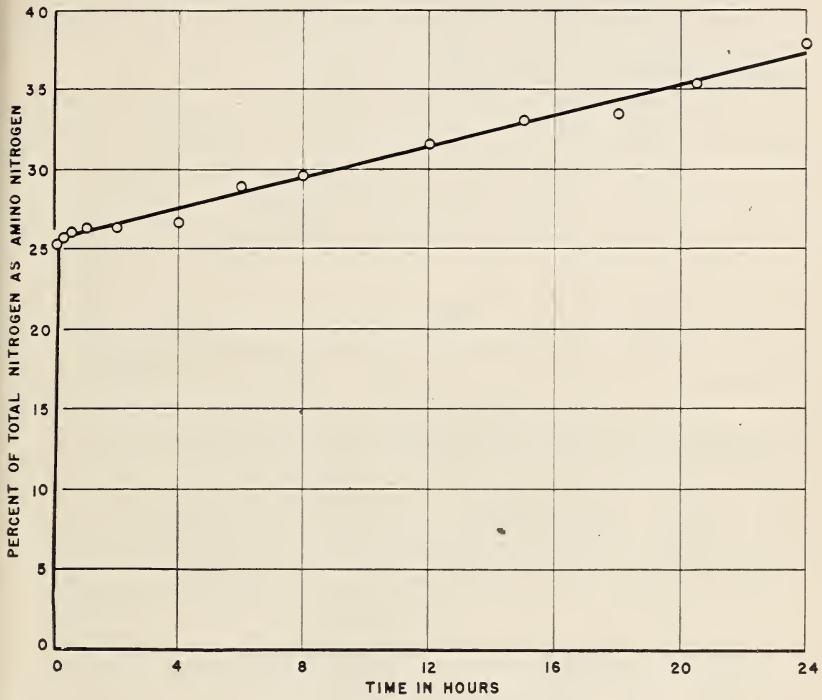


FIGURE 11.—*Rate of evolution of nitrogen from arginine treated with nitrous acid.*

of the slope of the curves for collagen (fig. 10) to the slope for arginine (fig. 11). It is thus possible to calculate the arginine contents of these materials from the following equations:

$$A = \frac{(F_A)(N_p)}{(N_A)} \times 100,$$

where

A = the arginine content of collagen
F_A = the fraction of the total nitrogen of the protein as arginine nitrogen
N_p = the nitrogen content of the protein
N_A = the nitrogen content of arginine.

$$F_A = S_p / S_A,$$

where

S_p=the slope of the straight line portion of the curve for collagen

S_A=the slope of the straight line portion of the curve for arginine.

Substituting in the previous equation,

$$A=\frac{(S_P)(N_P)}{(S_A)(N_A)}\times 100.$$

The slope of the curve for collagen is 0.0875 and for arginine 0.488. The corresponding nitrogen contents are 17.4 and 32.18. Substituting these figures in the above equation the arginine content of collagen was estimated to be 9.7 percent.

Although this method gave good results for the percentages of arginine in collagen and wool, later work [46] with a different type of apparatus indicated that the values obtained may have been the result of a chance, and the method is not believed to be applicable to proteins in general for the determination of arginine. The result for the primary amino nitrogen is, however, definitely of value in estimating the amount of lysine, and this method appears to be applicable also to other proteins.

By correcting for the nitrogen evolved from the guanidine nucleus, the value for the primary amino nitrogen is obtained. This was found to be 4.82 mg, or 2.77 percent, of the total nitrogen, which approximates 1 nitrogen group in 38, corresponding to the unit of Beek. As the amino nitrogen appears to be derived entirely from the ϵ-amino group of lysine, an estimation of the amount of this amino acid may readily be made and is found to be 5.03 percent. The values estimated for arginine and lysine approximate those obtained by separation and determination by quantitative analytical methods. The separation of amino acids by isolation from protein hydrolysates is a long and tedious process and high accuracy can not be claimed.

Assuming that the acid bound by collagen is held by the free basic groups of arginine and lysine, and considering the known analyses for the content of these amino acids in collagen, a ratio of approximately two molecules of arginine to one of lysine is indicated. This value is in fair agreement with known analyses. The three strongly basic nitrogen groups which Beek refers to in each unit of 38 must therefore be the ϵ-amino group of lysine and the two guanidine groups from the two molecules of arginine.

7. BASE-COMBINING GROUPS OF COLLAGEN

In addition to the free basic groups in collagen there are also free acidic groups. These groups are responsible for the amphoteric properties of the protein. In acid solutions, the acid groups fail to ionize and the protein acts as a base, whereas in basic solutions the basic groups do not ionize and the protein behaves as an acid.

The source of the free acidic group is probably the dicarboxylic acids, glutamic and aspartic. In this case, as with lysine and arginine, the α-amino group and the closely associated carboxyl group are combined in the polypeptide chain, whereas the end groups remain

free in the native protein. The end groups referred to are in the case of glutamic and aspartic acids, carboxyl groups.

Beek [35] determined the base-combining capacity of collagen. The results of his work with collagen from hide, tendon, and bone are given in table 7 and figure 12. The determination of base-combining capacity is somewhat more difficult than the determination of the

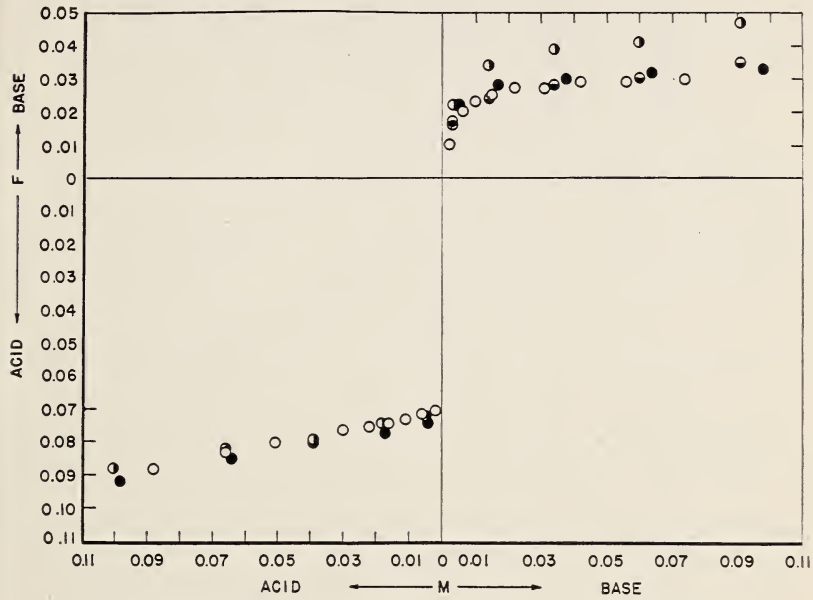


FIGURE 12.—*The combination of collagen with acid and base.*

○, hide; ●, tendon; ◑, bone ◒, unhydrolyzed (calculated). The abscissas (*m*) represent the molal concentration of hydrochloric acid or sodium hydroxide. The ordinates (*f*) represent the amounts of acid or base combined; expressed as the ratio of the number of equivalents combined to the number of equivalents of nitrogen in the sample.

acid-combining capacity, since the protein hydrolyzes more rapidly in basic solutions. However, the results of the base-combining capacity of hide in the ratio of equivalents combined to the equivalents of nitrogen in the sample lies between 0.025 and 0.03 and approximates 1 in 38 equivalents of nitrogen. The result for tendon is slightly higher and that for bone is much higher. The high results for bone are explained by Beek on the basis of the hydrolysis of amide linkages with the evolution of ammonia in the drastic treatment necessary for the purification of the bone collagen. He then determined the amounts of ammonia nitrogen in acid hydrolysates of the hide and bone collagens, which showed that the fraction of the total nitrogen as amide nitrogen was less by 0.010 in the bone collagen. By applying this correction factor to the bone collagen, the curves indicating the amounts of base combined for hide and bone collagen shown in figure 12 become nearly identical.

It is assumed that the source of the acid amide groups in collagen is the carboxyl end groups of aspartic or glutamic acid. This hypothesis is supported by the ease with which amide groups are changed

TABLE 7.—*Combination of collagen with sodium hydroxide*

Hide		Tendon		Bone	
m [1]	f [2]	m [1]	f [2]	m [1]	f [2]
0.002	0.910	0.005	0.022	0.003	0.025
.003	.016	.017	.028	.014	.034
.006	.020	.038	.030	.034	.039
.010	.023	.064	.032	.060	.041
.015	.025	.098	.033	.091	.047
.022	.027	---	---	---	---
.031	.027	---	---	---	---
.042	.029	---	---	---	---
.056	.029	---	---	---	---
.074	.030	---	---	---	---

[1] M is the molal concentration of base in the equilibrium solution.
[2] f is the ratio of the number of equivalents of base combined to the number of equivalents of nitrogen in the sample.

into carboxyl groups, by the residual ammonia in the hydrolysates of proteins and by the fact that the amount of base combined is not of sufficiently high magnitude to neutralize all of the theoretically possible carboxyl groups from aspartic and glutamic acid. Further, it is generally believed that aspartic acid is in the vegetable proteins as asparagine [36], and it is therefore possible that it is also present, as such, in other proteins like collagen.

Assuming that glutamic acid is present in collagen with its end carboxyl group free and able to combine with base, we may estimate the percentage of this amino acid present from the results obtained by Beek. The combination of base equivalent to a fraction of the total nitrogen of 0.03 indicates that 5.50 percent of glutamic acid is in the collagen. Assuming also that aspartic acid is present in the form of the acid amide, and calculating from the value of 2.30 percent of the total nitrogen as determined by Kanagy [37], a value of 3.81 percent is obtained for aspartic acid. Both of these values approximate those obtained for these acids by separation from the hydrolysates of the protein, as given in table 2.

8. ISOELECTRIC POINTS OF GELATIN AND COLLAGEN

The isoelectric point of a protein is that point at which no migration occurs when an electromotive force is applied to it in solution. This is known as the point of minimum properties of the protein where occurs, for example, minimum swelling, minimum viscosity and, of course, minimum ionization. It has also been generally recognized as the point at which the protein ionizes neither as an acid nor as a base.

It has long been known that gelatin in alkaline solution combines with silver to form gelatin-silver compounds, which blacken when exposed to light. Loeb [38] showed that these compounds are formed on the alkaline side of the isoelectric point only, and that the amount of silver bound was a function of the pH of the solution. He added powdered gelatin to a solution of silver nitrate, adjusted a series of gels between pH 6 and 3 with nitric acid, and allowed his experimental tubes to stand for a time. The tubes containing gelatin more alkaline than pH 4.7 blackened slowly in the light; in the tubes containing gelatin more acid than pH 4.7, there was no change.

Loeb showed that other positive ions behave like silver, combining with proteins on the alkaline side of the isoelectric point. Conversely anions combine with proteins on the acid side of the isoelectric point only. If gelatin powder is mixed with sodium bromide and the reaction adjusted. After washing the powder it is found that bromine is retained in all the experimental tubes which had a reaction more acid than pH 4.7, but not in tubes more alkaline.

The method of Loeb just described is cumbersome and at best gives only an estimate of the true isoelectric point, since the pH values at which color extinction is obtained is not as sharp as claimed. The electrophoretic technic developed by Abramson [39] is much superior and in recent work has been used almost exclusively.

The isoelectric points of four collagens and gelatin were determined by the electrophoretic technic by Beek and Sookne [40]. Their results are given in figure 13. A detailed examination of this figure indicates definitely that the isoelectric point of collagen or gelatin depends upon the previous treatment. Unlimed-hide collagen prepared by the Highberger method has an isoelectric point at about

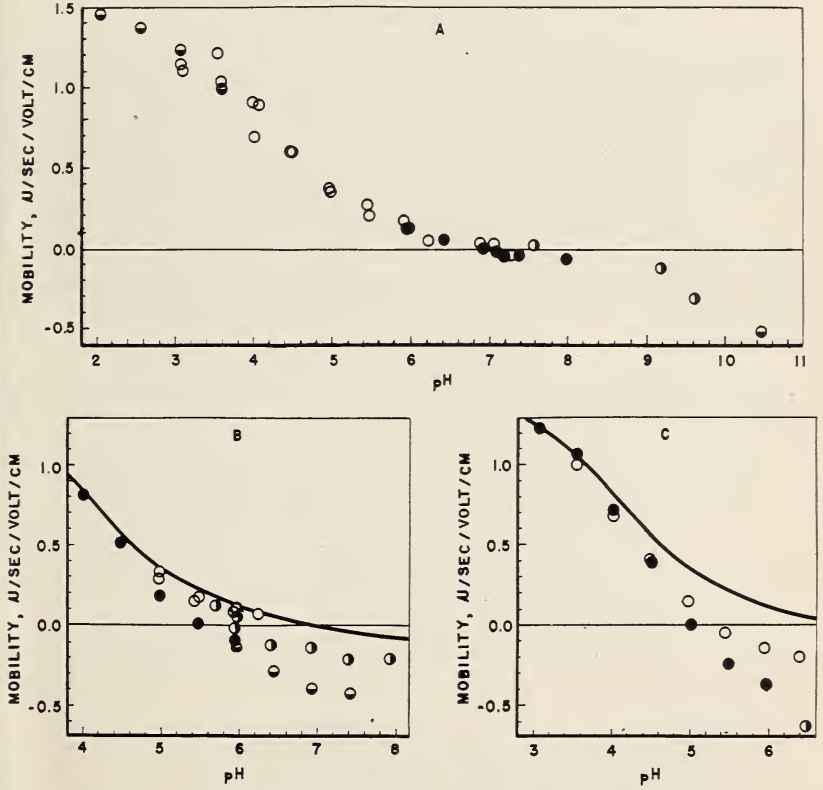


FIGURE 13.—*The mobilities of various collagens as a function of acidity.*

A, Hide collagen; ◒, chloride; ○, acetate; ●, phosphate. *B*, Tendon acetate, ○; tendon phosphate, ◑; bone acetate, ●; bone phosphate, ◒. *C*, Limed hide acetate, ○; gelatin acetate, ●; gelatin phosphate, ◑. Smooth curves as used in *B* and *C* to represent the mobility of the unlimed-hide collagen in order to avoid confusing these graphs with a large number of points.

pH 7, whereas limed hide and gelatin have isoelectric points at about pH 5. The difference is explained by a lower content of amide nitrogen for the gelatin and limed hide than for the unlimed hide. The decrease in amide nitrogen is accompanied by an increase in the number of free carboxyl groups, which increases the amount of acid required to prevent ionization. Consequently, the isoelectric point is found at a lower pH value.

It has generally been assumed that the neutral point and the electrophoretically determined isoelectric point of collagen are identical. This assumption was tested by Beek [35], who added just enough sodium hydroxide to neutralize the mineral acid impurity in collagen and then determined the pH of the solution with the glass electrode. A period of 4 days was allowed for the system to come to equilibrium. The pH of the solution at the end of this time was 7.7. This result is consistent with the fact that collagen contains more basic groups than acidic groups. As electrophoretic measurements give a lower pH value for the isoelectric point, it may be concluded that the assumption that the neutral point and the isoelectric point are identical is not valid for collagen, and therefore is not valid in general.

9. CARBOHYDRATES IN COLLAGEN

Grassmann and Schleich [41] were the first to report that collagen contained a carbohydrate consisting of glucose and galactose. Beek [42] later found that the carbohydrate fraction was not fermented by a galactose active yeast, therefore, neither D-glucose nor D-galactose could be present. He assumed that the carbohydrate consisted of L-glucose and L-galactose. L-Galactose has been found as a component of a polysaccharide of animal origin by Bell and Baldwin [43].

By spectrophotometrical methods, Beek [42] showed that collagen contains about 0.47 percent of a carbohydrate residue consisting of glucose and galactose, which is either bound by a difficultly hydrolyzable linkage or is composed of a mixture of L-glucose and L-galactose.

VI. CONCLUSION

Studies with the electron microscope and X-rays indicate that collagen consists architecturally of a definite unit of structure which is repeated.

It also appears that collagen has a definite chemical unit which is repeated, probably corresponding in some way with the architectural unit. The chemical unit of Beek, consisting of 38 nitrogen atoms, appears to be the most acceptable proposed unit at the present time. It appears to contain two arginine residues to one lysine residue. From the results of the analytical determinations of the amino acids and other evidence given in this circular, a unit consisting of 38 nitrogen atoms or the corresponding 31 amino acids might consist of the following:

Amino acid:	*Number of residues*
Arginine	2
Lysine	1
Glycine	12
Proline and hydroxyproline	8
Aspartic acid	1
Glutamic acid	1
Alanine	4
Leucine	2

This unit does not allow for histidine or carbohydrate. It would be impossible to fit these compounds at their low concentrations into it.

There is a definite need for intensive work on the preparation of highly purified collagen and accurate determinations of the amino acids in this material. Accurate determinations of the elements, carbon, hydrogen, oxygen, and nitrogen, are also necessary. These determinations will be of greater significance if oxygen is determined directly. Studies with the ultracentrifuge on gelatin and application of the modern theories and ideas on the properties of high polymeric systems, in general, should give valuable evidence with regard to the structure and properties of collagen.

The author expresses his appreciation to F. O. Schmitt, of Massachusetts Institute of Technology, for submitting the electron micrograph of collagen fibers, to the Cambridge University Press, and the Journal of the American Leather Chemists Association for permission to use illustrations and data of published articles.

VII. REFERENCES

[1] K. H. Meyer and C. Ferri, Arch. ges. Physiol. (Pflügers) **238,** 78 (1936).
[2] E. Wohlisch and R. du Mesnil de Rochement, Z. Biol. **85** 406 (1927).
[3] W. Grassmann, Kolloid-Z **77,** 205 (1936).
A. Kuntzel, Stiasny-Festschrift Darmstadt, p. 191 (Edward Roether, Verlag, Darmstadt. 1937).
E. Cherbuliez, J. Jeannerat, and K. H. Meyer, Z. physiol. Chem. **255,** 241 (1938).
[4] R. S. Lillie, Am. J. Physiol. **20,** 127 (1907).
J. Eggert and T. Reitstotter, Z. physik. Chem. **123,** 363 (1926).
M. Frankel, Biochem. Z. **249,** 476 (1932).
[5] George Scatchard, J. L. Oncley, J. W. Williams, and A. Brown, J. Am. Chem. Soc. **66,** 1980 (1944).
[6] J. Loeb, Proteins and the theory of colloidal behavior (McGraw Hill Book Co., New York, N. Y., 1924).
[7] H. R. Proctor, J. Chem. Soc. **105,** 313 (1914).
[8] H. R. Proctor and J. A. Wilson, J. Chem. Soc. **109,** 307 (1916).
[9] E. C. Porter, J. Int. Soc. Leather Trades Chem. **5,** 259 (1921); **6,** 83, (1922).
[10] G. D. McLaughlin, J. Am. Leather Chem. Assn. **15,** 228 (1920); **16,** 295 (1921).
[11] G. D. McLaughlin and R. E. Porter, J. Am. Leather Chem. Assn. **15,** 557 (1920).
[12] J. A. Wilson and A. F. Gallum, Jr., Ind. Eng. Chem. **15,** 71 (1923).
[13] R. E. Porter, J. Am. Leather Chem. Assn. **20,** 282 (1925).
[14] R. O. Page and J. A. Gilman, Ind. Eng. Chem. **19,** 251 (1927).
[15] J. H. Highberger, J. Am. Leather Chem. Assn. **31,** 345 (1936).
[16] Dorothy Jordan-Lloyd, Biochem. J. **24,** 1460 (1930).
[17] Dorothy Jordan-Lloyd and W. B. Pleas, Biochem. J. **21,** 1352 (1927).
Dorothy Jordan-Lloyd and W. B. Pleas, Biochem. J. **22,** 1007 (1928).
[18] J. H. Highberger, The chemistry of collagen I, J. Am. Leather Chem. Assn. **31,** 93 (1936).
[19] M. Bergmann, J. Biol. Chem. **110,** 471 (1935).
M. Bergmann and W. H. Stein, J. Biol. Chem. **128,** 217 (1939).
[20] H. D. Dakin, J. Biol. Chem. **44,** 499 (1920).
[21] W. Grassmann, W. F. Schneider, and G. Richter, Collegium **839,** 97, (1940).
[22] W. T. Astbury, J. Int. Soc. Leather Trades Chem. **24,** 69 (1940).
[23] R. S. Bear, J. Am. Chem. Soc. **64,** 727 (1942); **66,** 1297 (1944).
[24] F. O. Schmitt, J. Am. Leather Chem. Assn. **39,** 430 (1944).
[25] W. Grassmann and K. Riederle, Biochem. Z. **284,** 177 (1936).
[26] M. Bergmann and C. Niemann, J. Biol. Chem. **115,** 77 (1936); **118,** 301 (1937).

[27] M. L. Huggins, Chem. Rev. **32,** 195 (1943).
[28] John Beek, Jr., BS J. Research **8,** 549 (1932) RP434.
[29] A. W. Thomas and M. W. Kelly, Ind. Eng. Chem. **14,** 621 (1922); **20,** 632 (1928).
[30] John Beek, Jr., J. Research NBS **14,** 217 (1935) RP765.
[31] H. R. Proctor, J. Chem. Soc. **105,** 313 (1914).
[32] Joseph R. Kanagy and Milton Harris, J. Research NBS **14,** 563 (1935) RP765.
[33] D. D. Van Slyke, J. Biol. Chem. **12,** 277 (1912).
[34] R. H. Plimmer, Biochem. J. **18,** 105 (1924).
[35] John Beek, Jr., J. Am. Leather Chem. Assn. **33,** 621 (1938).
[36] Dorothy Jordan-Lloyd, Chemistry of the proteins, p. 11 (J. A. Churchill, 1926).
[37] Joseph R. Kanagy, J. Research NBS, **17,** 247 (1936) RP909.
[38] Jacques Loeb, Proteins and the theory of colloidal behavior (McGraw Hill Book Co., New York, N. Y., 1924).
[39] H. A. Abramson, L. S. Moyer, M. H. Gorim, Electrophoresis of proteins (Reinhold Publishing Co., New York, N. Y. 1942).
[40] John Beek, Jr., and Arnold M. Sookne, Electrophoresis of collagen, J. Research NBS **23,** 271 (1939) RP1230.
[41] W. Grassmann and H. Schleich, Biochem. Z. **277,** 320 (1935)
[42] John Beek, Jr., The carbohydrate content of collagen, J. Research NBS **27,** 507 (1941) RP1438; J. Am. Leather Chem. Assn. **36,** 696 (1941).
[43] D. J. Bell and E. Baldwin, Nature **146,** 559 (1940).
[44] L. W. J. Holleman, H. G. Bungenberg de Jong, and R. S. Tjaden Modderman, Kolloid-Beihefte **40,** 211 (1934).
[45] A. Kuntzel and K. Doehner, Angew. Chem. **52,** 175 (1939).
[46] Henry A. Rutherford, Milton Harris, and Arthur L. Smith, J. Research NBS **19,** 467 (1937) RP1038.

Washington, September 11, 1946.

○