Determination of crosslink density by end group analysis after partial degradation: 2. Experimental application to keratin

Emory Menefee and Sandra J. Tillin
US Department of Agriculture, Western Regional Research Center, Berkeley, California 94710, *USA*
(*Received* 2 *September* 1980; *revised* 27 *October* 1980)

By a method of partial solubilization and end group analysis, the molar concentration of intermolecularly crosslinked amino acids was determined for a wool sample. The results can be interpreted in two ways. If the wool is assumed to behave homogeneously toward acid hydrolysis, then the crosslink theory indicates that the molar concentration of intermolecular linkages is 7.5%, or that out of a total crosslink concentration of 12.6%, about 40% of the crosslinks are intramolecular. Alternatively, the crosslinking could be entirely intermolecular in a two component model provided the rate of hydrolysis and solubilization of the lightly crosslinked (microfibrillar) component is much slower than that of the densely crosslinked matrix. Partial solubilization was effected with a mixture of hydrochloric and formic acids. The extent of solubilization was found by weighing, and the concentration of amine ends in the soluble fraction determined spectrophotometrically by the absorption, at 422 nm, of the complex formed between amines and trinitrobenzene sulphonic acid (TNBS). The experimental curve of end group concentration *versus* solubility was used to estimate crosslink density by a previously derived theory for both single and two component models.

INTRODUCTION

In Part I, a theoretical method is described for determining the crosslink density of polymers sufficiently crosslinked to be insoluble in solvents that would ordinarily dissolve them. The method involves partial degradation of the polymer, preferably breaking main chain linkages only, followed by leaching out of whatever becomes soluble, with subsequent determination of the number of single end groups appearing in the soluble fraction. Examination of the theoretical relations revealed that the correlation between the end group concentration and the amount dissolved is essentially independent of molecular weight except at very low crosslink densities and/or chain lengths. Since this derived correlation is strongly dependent on the original crosslink density before degradation, the method lends itself to the determination of an 'absolute' intermolecular crosslink density in situations for which the molecular weight of the crosslinked chains is unknown or uncertain. There are a number of applications for more certain knowledge of crosslink density in both synthetic and natural polymers. Among the latter, for example, crosslinks are important in the structure of keratin, collagen, elastin, and other biopolymers. The effect of ageing on these structures is an area of active interest, and considerable circumstantial evidence has implicated crosslinks. The method described in the present paper is simple to carry out, and should be adaptable to routine use without difficulty.

To our knowledge, there is no other method in the literature for the direct determination of intermolecular crosslink density. Estimates are based on secondary correlations of the effect of crosslinking on various measurable physical properties, such as swelling or elastic modulus. In some cases, the total number of disulphide crosslinks can be found with considerable accuracy by chemical means. However, whether these crosslinks exist as intermolecular or intramolecular linkages is uncertain; the effect of such a difference on the interpretation of structure and physical behaviour can be considerable.

As discussed in Part I, two assumptions are needed in deriving theoretical relations sufficiently simple to permit interpretation of partial solubilization data to yield information on crosslink density. Firstly, the degradation is assumed to be random—every linkage between monomers is considered equally liable to attack. Secondly, the crosslinks are assumed to be distributed randomly. The effect on the theory of relaxing either of these assumptions is unknown, although it seems unlikely that deviations from either will be important.

The method of degradation needed to produce a soluble fraction should be chosen carefully. Ideally, it should cause progressive and random breakage of main chain bonds only, with no effect on the crosslinks themselves. Out of numerous methods that could be used to degrade polymers, we can rank some of the more common ones roughly in order of their decreasing randomness of attack: ionizing radiation, heat, hydrolysis or other solvolytic attack, ultrasonic irradiation, and enzymatic attack. Ionizing radiation and heat have the disadvantage of producing new active species that can lead to additional crosslinks. Although this complication can be accommodated in principle by extrapolation of the calculated crosslink density back to zero treatment time, as described in Part I, the great increase in experimental work needed would render these methods less desirable than others. Other methods may not be wholly free from some crosslink scission accompanying main chain clea-

032-3861/81/091219-04$02.00

vage; if a significant amount is suspected, then the extrapolation procedure should be applied for best results.

In this paper, we present our first results using wool as an exemplary crosslinked polymer. Since wool is essentially 100% protein, acid hydrolysis was chosen as the degradation method. The hydrolysis rate constants for various amino acid pairs are not equal; however, the degradation can be assumed to proceed randomly because the various pairs appear pseudorandomly along the protein chain[2]. Furthermore, under relatively mild conditions of acid hydrolysis, it is known that the disulphide crosslinks in keratins are not particularly labile. After solubilization of the degraded protein in formic acid, the concentration of chain ends in the soluble fraction was found by a modified method using trinitrobenzene sulphonic acid to combine with free amine groups, followed by spectrophotometric estimation of the concentration of the adduct. Consistency of the calculated crosslink density over a wide range of soluble fractions gave us confidence that the crosslinks themselves were not appreciably diminished in number during the hydrolysis procedure. Because of the lack of a primary verification of this method for determining crosslink density, it is hoped that future work can be directed toward its use on a crosslinked polymer of known structure, such as a crosslinked insulin. Such a study would be useful in estimating errors that may enter through the degradation procedure or incomplete leaching.

EXPERIMENTAL

Materials

About 30 g of Dubois fine wool top was extracted for 24 h with a 3:2 mixture of benzene and methanol. After air drying, half the wool was ground in a Wiley mill to pass a 20 mesh (1.3 mm) screen, and the other half chopped into pieces approximately 3 mm. To remove any soluble protein initially present or resulting from mechanical degradation, both fractions were soaked for 18 h in a 40% formic acid solution, with occasional stirring. Formic acid is a keratin swelling solvent that has previously been shown[3] to cause only negligible hydrolysis under these conditions. The wool was collected on a filter, rinsed several times with distilled water, and then dialysed against distilled water for 24 h to remove traces of formic acid. After final filtering and rinsing, the wool was dried for 3 h under vacuum at 100°C, and then conditioned for several days at 21°C and 65% relative humidity. The moisture content of the prepared wool was 13.75% by weight.

Formic acid, hydrochloric acid, sodium borate, and sodium hydroxide were standard reagent grades, used as obtained. TNBS (2,4,6-trinitrobenzene sulphonic acid) was used as received from Eastman Chemicals.

Partial hydrolysis

A known weight (~0.5 g) of conditioned wool was mixed with 25 ml of the hydrolysis medium in a stoppered flask. This medium was a solution of equal volumes of 12 M HCl and 88% formic acid, chosen to give reasonably slow rates with enough swelling to facilitate leaching out of solubilized protein fragments. Depending on the extent of partial hydrolysis or solubilization wanted, the flask was shaken at 30°C for 2, 4, 6, 8, 16, 24, or 30 h. Total hydrolysis, to determine the average extinction coefficient and amino acid composition, was done at an elevated temperature with 6 M HCl following the method of Friedman and Noma[4].

To recover the soluble protein released by partial hydrolysis, the contents of the flask were centrifuged for 5 min at maximum speed in a Clay–Adams Safety Head centrifuge. Gel particles resuspended during decanting were filtered out through sintered glass. To estimate how much soluble protein may remain trapped in the undissolved wool, the residues from one series of hydrolyses were additionally treated with 4 ml of 88% formic acid, stirred and ultrasonicated for 1 min. After centrifuging as before, the additional formic acid wash was added to the original supernate, which had been kept cold in an ice bath to prevent further hydrolysis of the solubilized protein. Results indicated no significant amount of protein was released by the additional treatment.

As indicated above, several variations in experimental procedure were used in these runs to reveal any particularly sensitive steps. The variations are:

(1) Chopped fibres, rinsed with formic acid after hydrolysis
(2) Chopped fibres, no rinse
(3) Chopped fibres, ultrasonicated with formic acid
(4) Milled fibres, no rinse
(5) Milled fibres, ultrasonicated with formic acid

To obtain the dry weight of soluble protein per unit volume of supernate, a 10–20 ml aliquot (depending on the estimated amount of protein present) was dried on a rotary evaporator, vacuum dried for 2 h at 100°C, and weighed in a closed container.

End group determination

The number of new amine end groups appearing in the soluble fraction after partial hydrolysis was estimated for an aliquot of the supernate by using the TNBS complexing method of Snyder and Sobocinski[5], modified slightly to use 0.025 M sodium borate, 0.0375 M TNBS, and a reaction time of one hour. Absorption spectra were measured with a Cary 14 recording spectrophotometer, compared with a blank of 0.2 ml of 0.0375 M TNBS in 10 ml of 0.025 M $Na_2B_4O_7$. The absorption maximum of the complex was found at 422 nm.

The orange colour indicating complexing between TNBS and amine groups develops above pH 9. Since the initial pH of the solution of dissolved protein is quite low, owing to the remaining HCl and formic acid, it was first brought to about pH 6 with 6 M NaOH. A quantity of this neutralized solution, ranging from 0.1 ml to 3.5 ml, depending on the estimated amount of protein present, was diluted to 10 ml with 0.025 sodium borate and 0.2 ml of 0.0375 M TNBS added.

The use of 0.025 M sodium borate instead of the 0.1 M recommended by Snyder and Sobocinski was necessitated by precipitation of borate from the neutralized solution at the higher level. No precipitate formed with the more dilute borate buffer. A comparison of spectra for 0.1 ml of neutralized solution showed identical absorbance at 422 nm when diluted with either 0.1 M or 0.025 M sodium borate.

To determine when the reaction between TNBS and the hydrolysates should be terminated, the absorbance of 422 nm was followed for at least 3 h with the 2 h and the 24 h

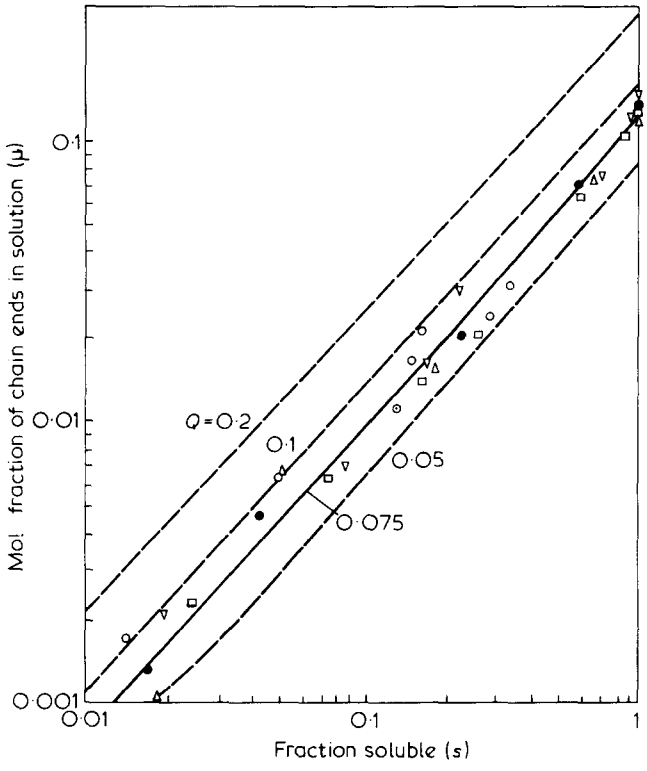

Figure 1 Mole fraction of broken chain ends appearing in the soluble portion, μ, *versus* the total solubility, S. Dashed lines are calculated for a one component model with $U = \infty$, at three levels of crosslink density, $Q = 0.2$, 0.1, and 0.05. Experimental points from the partial hydrolysis of wool include five variations in procedure, identified by number in the text: ○ (1), ● (2), △ (3), ▽ (4), □ (5). The solid line is calculated for the average experimental value of $Q = 0.075$

hydrolysates. Since in all cases the absorbance reached at least 90% of its final value after one hour, this time was chosen for all subsequent runs to expedite the determinations. Errors in end group concentrations caused by using this shorter reaction time should be of second order magnitude.

A known concentration of totally hydrolysed wool was used to determine the average molar extinction coefficient ε at 422 nm. Even though the amino acid composition of partial hydrolysates changes somewhat with the extent of solubilization, calculations of the expected average extinction coefficient from those of the individual amino acids showed that the error caused by using an extinction coefficient from the total hydrolysate was negligible. Because all portions of the hydrolysing chains are not freed with equal probability, the composition of the soluble fraction varies with extent of hydrolysis. Also, lysine and arginine can contribute an extra amine group that would appear as a free end, so the calculation included a correction for the extinction coefficient consisting of taking ε_{lys} and ε_{arg} to be $2\varepsilon_x$, where ε_x is the average molar extinction coefficient of the other amino acids, computed from measured amino acid compositions. Although these compositions showed some change as a function of hydrolysis time, in no instance did the average extinction coefficient need more than 0.9% correction, a variation considered to be entirely negligible.

RESULTS AND DISCUSSION

The observed variation in the number of end groups in the soluble fraction with the amount of wool solubilized can be interpreted in two ways, the choice depending on how the structure of wool is introduced. Together, these approaches illustrate the main ways in which the theory of Part I can be applied. Although wool is a good example of a highly crosslinked protein, it has a complex structure, containing over a hundred different proteins in a few fractions that are fairly distinct chemically and morphologically[2]. The two principal morphological entities are the osmiophobic microfibrils and the osmiophilic matrix. In addition, wool can be fractionated chemically into two portions containing relatively low and high contents of sulphur or cystine. The low sulphur fraction has been associated with the microfibrils, and the high sulphur fraction with the matrix. The kind and nature of interactions between these structural elements is uncertain, though most evidence indicates that at least part of the protein contains low and high sulphur regions on the same molecule, and that there is likely to be considerable intercrosslinking joining the mass into a mechanically interacting whole.

Because of the interrelation among protein chains, the simplest interpretation of the experimental results is a direct application of the equations of Part I to the solubilization data, considering the wool to be homogeneous, with random crosslinking and random chain attack. Results for the acid hydrolysis of the Dubois wool sample are shown in *Figure 1*, where the amine end group concentration in the soluble fraction, μ, is plotted against the amount of solubilization, S. The end group concentration is expressed as the ratio of the number of moles of single ends to the number of moles of amino acid residues in the original sample.

The average crosslink density, Q, can be calculated using the following expressions from Part I:

$$S_0 = \left(\frac{\sqrt{1 + \frac{4}{\mu}\left(\frac{S}{\mu} - 1\right)} - 1}{2\left(\frac{1}{\mu} - \frac{1}{S}\right)} \right)^2 \tag{1}$$

and

$$Q = \frac{S_0 - S}{S_0(1 - S_0)} \tag{2}$$

In these expressions, the sterile coefficient S_0 (essentially a dummy variable) is first calculated from the experimental values of S and μ, and then Q is calculated from equation 2. Even though these equations were derived for infinite chain lengths, they are applicable to wool because its proteins average 100 or more amino acid residues. For all the samples taken, the average Q was 0.075, with a standard deviation of 0.003 for the average.

The lines in *Figure 1* were calculated by methods in Part I. For a given Q, a value of P is assumed, and S_0 calculated iteratively from the relation:

$$S_0 = \left(\frac{P}{1 - a} \right)^2 \tag{3}$$

where $a = (1 - P)(1 - Q + QS_0)$. When S_0 has been found, S is calculated from

$$S = S_0(1 - Q + QS_0) \tag{4}$$

and μ by the relation

$$\mu = S(1-a) \quad (5)$$

The solid line in *Figure 1* is calculated for the average Q of 0.075.

Taking the sulphur content of Dubois wool to be 3.7%[7], and using 109 as the average molecular weight of the amino acid residues in wool, the mole fraction of residues containing a crosslink is 12.6%. By the above calculation, the mole fraction of intermolecularly crosslinked residues is 7.5%. These results suggest that 40% of the disulphide crosslinks may be intramolecular, an estimate that agrees with a previous figure for wool arrived at by indirect arguments[6]. The presence of a substantial proportion of juxtaposed cysteine residues (–CYS–CYS–) in the high sulphur fraction[2] suggests that these pairs may be involved in intramolecular loops. These loops would most likely be small, although longer loops may occur by analogy to other globular proteins. In addition, if the cysteine pair on one chain is intermolecularly linked to a similar pair on another chain, it is possible that the double crosslink so formed would appear in the solubility analysis (as well as in its mechanical contribution) as a single link. Even though the probability of occurrence of such double crosslinks may be extremely low, they should not be ruled out entirely. Also, other non-hydrolysable crosslinks of origins other than cysteine may exist, though there is no evidence for a significant number of such crosslinks in untreated wool.

Because wool has more than one component, the results may be due to differential hydrolysis rates of two or more components, rather than to intramolecular crosslinking. However, because of certain restrictions on such a model, this interpretation is by no means strong. It is apparent that if the sample had two components with nearly the same crosslink densities, these would solubilize much like a single component, even if the two components differed in their hydrolysis rates. If the component crosslink densities differ greatly, but have the same rate of main chain degradation, then the μ *versus* S curves of *Figure 1* would show a jog to the higher crosslink level at some point. As seen in *Figure 1*, this shift does not occur during wool degradation. Hence, if two components of widely differing crosslink density are present, the rate of dissolution of the fraction of lower crosslink density must be considerably lower than that for the highly crosslinked component, so that comparable amounts of each component will appear in the soluble fraction at any given time.

The two component model was examined by methods described in Part I. Fractions X_1 and X_2 or the two components were taken as 0.5. The component crosslink densities, Q_1 and Q_2, are subject to the constraint $X_1Q_1 + X_1Q_2 = 0.126$, the overall mole fraction of crosslinked residues. The fraction of main chain bonds broken for each component was assumed to follow first order kinetics:

$$P_1 = 1 - \exp(k_1 t) \quad (6)$$

$$P_2 = 1 - \exp(k_2 t) \quad (7)$$

For a given ratio k_1/k_2, an assumed value of P_1 determines P_2. The actual values of k_1 and k_2 are not important, since P_1 and P_2 do not enter into the results. When a ratio is chosen, a series of values of Q_1 and Q_2 are assumed (subject to the above constraint), and the solubilities S_1 and S_2 and end group concentrations μ_1 and μ_2 calculated by the same procedures as applied to a single component. The overall S and μ are obtained from the relations:

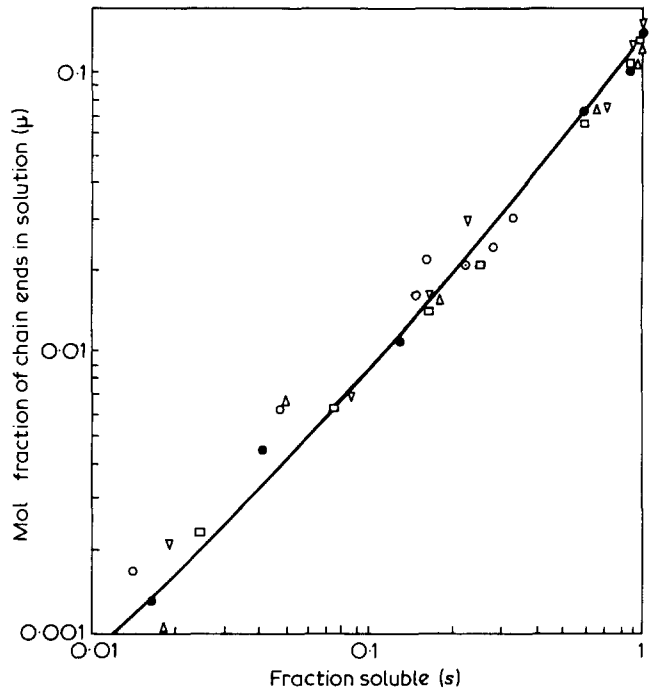

Figure 2 Mole fraction of broken chain ends appearing in the soluble portion, μ, *versus* the total solubility, S. The line is calculated with $U = \infty$, for a model containing equal weight fractions of two components. Parameters were varied to give the best fit, giving $Q_1 = 0.227$, $Q_2 = 0.025$, and a ratio of first order hydrolysis rate constants of $k_1/k_2 = 5$. Experimental points include five variations in procedure, identified by number in the text: ○ (1), ● (2), △ (3), ▽ (4), □ (5)

$$S = S_1X_1 + S_2X_2 \quad (8)$$

and:

$$\mu = \mu_1X_1 + \mu_2X_2 \quad (9)$$

For assumed ratios k_1/k_2 between 2 and 10, the best fit to the experimental S, μ points was found with $k_1/k_2 = 5$, $Q_1 = 0.227$, and $Q_2 = 0.025$. These parameters gave the solid curve shown in *Figure 2*. Crosslink densities of the two fractions are in fair agreement with experimental values for the two major fractions of wool[2]. The relatively slow rate of hydrolysis of the low sulphur fraction may be accounted for by its more organized alpha helical structure, in keeping with the usual behaviour of crystalline phases, though this interpretation needs further examination. It is to be hoped that future hydrolysis experiments with wool will yield less scattered results, which can be subjected to a sufficiently detailed mathematical analysis to help decide some of the questions of structure raised in this paper[8].

REFERENCES

1 Menefee, E. *Polymer* 1981, **22**, 1219
2 Crewther, W. G. *Proc. 5th Int. Wool Textile Res. Conf.* 1976, **1**, 1
3 Menefee, E. and Yee, G. *J. Appl. Polym. Sci.* 1965, **9**, 2835
4 Friedman, M. and Noma, A. T. *Textile Res. J.* 1970, **40**, 1073
5 Snyder, S. L. and Sobocinski, P. Z. *Anal. Biochem.* 1975, **64**, 284
6 Menefee, E. in 'Protein Crosslinking: Biochemical and Molecular Aspects', (Ed. M. Friedman) Plenum Press, New York, 1977, pp 307–327
7 Menefee, E. and Yee, G. *J. Appl. Polym. Sci.* 1965, **9**, 2847
8 Product disclaimer in Part 1 applies for Part 2