JOURNAL OF APPLIED POLYMER SCIENCE VOL. 9, PP. 2819–2827 (1965)

Crosslinking in Keratins. I. Theory for Solubility on Simultaneous Chain Scission and Crosslink Cleavage

E. MENEFEE and J. J. BARTULOVICH, *Western Regional Research Laboratory, Western Utilization Research and Development Division, Agricultural Research Service, U. S. Department of Agriculture, Albany, California*

Synopsis

For a crosslinked polymer with a uniform molecular weight distribution we derive expressions for the amount soluble after any amount of crosslink cleavage or chain scission. Two closed-form results are obtained. One is an approximate form based on the formalism of Charlesby and is usable for most needs. The other, based on a degradation distribution from Montroll and Simha, is exact but rather more cumbersome. We compare numerical results from each approach.

INTRODUCTION

Wool and other keratins contain a number of intermolecular crosslinks along each protein chain. This crosslink density accounts for the extreme insolubility of keratin in ordinary solvents. To obtain usable amounts of soluble keratin proteins for physical–chemical characterization, some primary bonds must first be broken—preferably only those which form the disulfide crosslinks. In practice, however, a certain amount of main-chain scission almost always occurs which can obscure the characterization of the protein extract. By taking into account the details of the solubilization process we not only can avoid many pitfalls in interpreting such experiments, but also obtain considerable new information at the same time.

This paper is the first of a series that explores crosslinking in keratins as it affects the interpretation of various degradation experiments. In this one (designated I), we will derive closed-form expressions for the solubility of a keratin-like solid after specified amounts of crosslink cleavage and chain scission have occurred. The second paper, II,[1] will discuss the sulfur composition to be expected in the soluble fraction and in the residue following various kinds of degradation. In the third paper, III,[2] the results of I and II will be applied to experimental data for acid hydrolysis of keratins in the absence of crosslink cleavage. The final paper of this group, IV,[3] will contain a discussion of experimental data for the thermal cleavage of disulfide crosslinks with a minimum of main chain scission. In prepara-

tion are a discussion of reduction and oxidation experiments, and an application of crosslinking theory to fractional precipitation experiments with proteins.

THEORY

Two general assumptions are made throughout this work which are justified *a posteriori* and which define the "model" to be studied. They are: (*a*) even though keratins may have more than one component, each of these components behaves independently and has, initially, a uniform (monodisperse) distribution of molecular weights; (*b*) main-chain scission and crosslink cleavage are random processes. Although assumption (*b*) is known to be untrue in detail, it is quite reasonable that chain breakage appears to be random, owing to the scattered distribution of various amino acids. Although we may omit assumption (*a*) without undue strain, the mathematical difficulties encountered in relaxing assumption (*b*) are formidable. In III we will discuss the effect of assuming an initially random distribution of molecular weights, and in IV the effect of assuming that the components are intercrosslinked (nonindependent).

Assumption (*b*) permits us to consider the processes of main-chain scission and simultaneous crosslink cleavage serially. In its initial state the keratin component consists of A_0 molecules, each with U monomer units (amino acid residues). A fraction Q of the total number of monomer units, A_1, is engaged in intermolecular (disulfide) crosslinks. For purposes of calculation we consider the following sequence of events. The uncrosslinked molecules of uniform initial length are first broken at random to provide a new distribution. The extent of this degradation is defined by P, the fraction of breakable main-chain bonds that have been cleaved. The molecules are then crosslinked at random until a fraction Q of monomer units is engaged in intermolecular crosslinking. Finally, from knowledge of the molecular weight distribution and Q, the amount of soluble material (sol) and residue (gel) is calculated.

There are at least two methods of derivation which yield closed-form results for this problem. The first involves the straightforward application of a formalism set forth by Charlesby.[4] It yields an approximate result, relying on U being much greater than unity and Q being much less, and is simple to use, since Q and U occur together as the product $QU = \delta$. In the second method we use an expression by Montroll and Simha[5] for the molecular-weight distribution following random degradation of uniform. This result is fitted into Charlesby's formalism for calculation of the soluble fraction via the fraction of sterile crosslinks. The final closed-form expression is exact but rather more cumbersome to use, since Q and U are now separate.

We will give both derivations, marking points in the first at which approximations are introduced, and then show numerical comparisons between the resulting expressions.

Approximate Derivation

For an initially uniform distribution of A_0 molecules, the ith moment is defined as

$$A_i = A_0 U^i \tag{1}$$

If fracture of a molecule is allowed to occur at any residue, then the change in moments in the fractured system is

$$dA'_i/dP = [(i-1)/(i+1)]A'_{i+1} \tag{2}$$

where the primes refer to the fractured system and P is the fracture density (the fraction of possible sites that are broken), defined as

$$P = (A'_0 - A_0)/(A_1 - A_0) \tag{3}$$

Equation (2) is an approximation depending for validity on U being large.

By use of a Maclaurin expansion, A'_i can be written as

$$A'_i = A_i + P\,(dA'_i/dP)_{P=0} + (P^2/2!)\,(d^2A'_i/dP^2)_{P=0} + \ldots. \tag{4}$$

Using eq. (2), we can express the derivatives in the general form

$$d^nA'_i/dP^n = (-1)^n\,[(i-1)\,i/(i+n-1)(i+n)]A'_{i+n} \tag{5}$$

Putting this into eq. (4) and condensing the notation, we have

$$A'_i = \sum_{n=0} (-1)^n\,[i(i-1)/(i+n)\,(i+n-1)]\,(P^n/n!)\,A_{i+n} \tag{6}$$

Using eq. (1) for A_i in eq. (6) gives

$$A'_i = A_0U^i\,(i-1)i \sum_{n=0} (-PU)^n/[(i+n)(i+n-1)n!] \tag{7}$$

Equation (7) represents Charlesby's approximation for the distribution of moments resulting when an initially uniform distribution is randomly fractured. When an arbitrary distribution is crosslinked partially, the approximate gel fraction G is given also by Charlesby as

$$G = (1/A'_1) \sum_{i=2} [(QG)^{i-1}/(i-1)!]\,A'_1(-1)^i \tag{8}$$

Although we will not repeat the derivation here, the approximation in eq. (8) arises mainly from the omission of terms of the form $(QG)^{i-1}A_j$, where $j < i$. For every large U and small Q (light crosslinking) these terms vanish.

Combining eqs. (7) and (8) we obtain

$$G = -\frac{A_0U}{A'_1} \sum_{i=2} \frac{(-QUG)^{i-1}}{(i-1)!}\, i(i-1) \sum_{n=0} \frac{(-PU)^n}{(i+n)(i+n-1)n!} \tag{9}$$

Since there has been no change in the number of residues in the entire system, $A'_1 = A_1$. We may also replace A_1 by A_0U, and define a new

quantity $\delta = QU$ as the number of crosslinks per molecule based on no-fracture conditions. With these simplifications, eq. (9) becomes

$$G = -\sum_{i=2} \frac{(-\delta G)^{i-1}}{(i-1)!}\, i(i-1) \sum_{n=0} \frac{(-PU)^n}{(i+n)(i+n-1)n!} \tag{10}$$

After considerable manipulation, eq. (10) can be put into the following closed form:

$$G = \frac{[1 + (2PU/\delta G)]\,[1 + (PU/\delta G)] - (2PU/\delta^2 G^2)}{[1 + (PU/\delta G)]^3} - \frac{[1 + (PU/\delta G) - (2PU/\delta^2 G^2)] \exp\{-(\delta G + PU)\}}{[1 + (PU/\delta G)]^3} \tag{11}$$

The transcendental nature of the dependent variable G makes immediate use of this expression difficult, so that in the long run we must resort to numerical solutions. These will be presented later when the exact derivation has been described.

By expanding eq. (10) or (11) for small P, we obtain

$$G = -\sum_{i=2} \frac{(-\delta G)^{i-1}}{(i-1)!} + PU \sum_{i=2} \frac{(-\delta G)^{i-1}}{(i-1)!} \frac{i-1}{i+1} - \ldots \tag{12}$$

The two sums in eq. (12) may be evaluated directly:

$$G = 1 - \exp\{-\delta G\} - (PU/\delta^2 G^2) \times [2 - (\delta^2 G^2 + 2\delta G + 2) \exp\{-\delta G\}] - \ldots \tag{13}$$

When there is no crosslink cleavage, δ is constant and the initial gel fraction is given by

$$G_0 = 1 - \exp\{-\delta G_0\} \tag{14}$$

Combining eqs. (13) and (14) we have for this case

$$G = G_0 - (PU/\delta^2 G_0^2)[2 - (\delta^2 G_0^2 + 2\delta G_0 + 2)(1 - G_0)] - \ldots \tag{15}$$

Since the solubility of a keratin in its original condition is usually nil, we may set $G_0 \cong 1$, and obtain from eq. (15)

$$G \cong 1 - (2PU/\delta^2) + \ldots \tag{16}$$

For a given δ it is sometimes useful to know at what point in the degradation the material becomes completely soluble. This is easily found by evaluating eq. (11) for vanishing G. Designating the critical value of PU as X, we find

$$X^2 = 2\delta(e^{-X} + X - 1) \tag{16a}$$

For large δ this is, approximately,

$$X \cong 2\delta[1 - (1/2\delta)] \tag{16b}$$

Exact Derivation

Montroll and Simha[5] have derived the following number distributions which occur after random scission of initially uniform molecules:

$$N(u) = A_0P(1 - P)^{u-1}[2 + (U - 1 - u)P] \tag{17}$$

$$N(U) = A_0(1 - P)^{U-1} \tag{18}$$

Now, following Charlesby if we allow this distribution to be crosslinked until a fraction Q of monomers is engaged in intermolecular crosslinking, then the number of molecules with u monomers, of which C are crosslinked, will be

$$n_C(u) = N(u)(1 - Q)^{u-C}Q^Cu!/(u - C)!C! \tag{19}$$

$$n_C(U) = N(U)(1 - Q)^{U-C}Q^CU!/(U - C)!C! \tag{20}$$

The total number of crosslinks is QA_1. Of these, a fraction $t(C)$ is carried by molecules which have C crosslinks to them:

$$t(C) = (1/QA_1)\left[\sum_{u=1}^{U-1} Cn_C(u) + Cn_C(U)\right] \tag{21}$$

The sterile coefficient S_0 is defined as the probability that a given crosslink is not connected to the gel through other crosslinks. Then

$$S_0 = \sum_{C=1}^{U} S_0^{C-1}t(C) \tag{22}$$

Combining eqs. (19–22) we obtain

$$S_0 = \frac{1}{QA_1}\sum_{C=1}\frac{S_0^{C-1}Q^CC}{C!}\left[\sum_{u=1}^{U-1}\frac{N(u)(1 - Q)^{u-C}u!}{(u - C)!} + \frac{N(U)(1 - Q)^{U-C}U!}{(U - C)!}\right] \tag{23}$$

Incorporating eqs. (17) and (18) into this expression and inverting the summation order, we have

$$S_0 = \frac{1}{U}\sum_{u=1}^{U-1} P(1 - P)^{u-1} \times [2 + (U - 1 - u)P]\sum_{C=1}^{u}\frac{S_0^{C-1}Q^{C-1}(1 - Q)^{u-C}u!}{(u - C)!\ \ (C - 1)!} + \frac{(1 - P)^{U-1}}{U}\sum_{C=1}^{U}\frac{S_0^{C-1}Q^{C-1}(1 - Q)^{U-C}U!}{(U - C)!\ \ (C - 1)!} \tag{24}$$

The sums over C, which may be evaluated immediately from the binomial expansion, are of the form

$$u(1 - Q + QS_0)^{u-1}$$

Hence, eq. (24) now becomes

$$S_0 = (P/U)\sum_{u=1}^{U-1}(1-P)^{u-1}[2+(U-1-u)P](1-Q+QS_0)^{u-1}u + (1-P)^{U-1}(1-Q+QS_0)^{U-1} \quad (25)$$

If we define $a = (1-P)(1-Q+QS_0)$, then eq. (25) becomes

$$S_0 = a^{U-1} + (P/U)\times \left\{[2+(U-1)P]\sum_{u=1}^{U-1} ua^{u-1} - P\sum_{u=1}^{U-1} u^2a^{u-1}\right\} \quad (26)$$

To evaluate the sums over u, we note that

$$\sum_{u=1}^{N} a^u = [a/(1-a)](1-a^N) \quad (27)$$

and obtain the two required sums by successive differentiation with respect to a. Designating the sums as σ_1 and σ_2, we have finally for S_0,

$$S_0 = a^{U-1} + (P/U)\{[2+(U-1)P]\sigma_1 - P\sigma_2\} \quad (28)$$

where

$$\sigma_1 = (1-a)^{-2}[1 - Ua^{U-1} + (U-1)a^U]$$

$$\sigma_2 = (1-a)^{-3}[1 + a - U^2a^{U-1} + (2U^2 - 2U - 1)a^U - (U-1)^2a^{U+1}]$$

and

$$a = (1-P)(1-Q+QS_0)$$

Before proceeding with an examination of this exact closed-form solution, we must connect the sterile coefficient S_0 with the actual soluble fraction $S = 1 - G$. The soluble fraction is the weight fraction of all molecules which have no crosslinks at all or only sterile ones. If we assume the molecular weight per amino acid monomer unit to have an average value m, then

$$S = \sum_{u=1}^{U}\sum_{C=0}^{u} n_C(u)umS_0{}^C/A_1m \quad (29)$$

For S_0, corresponding to eq. (23) we have

$$S_0 = \sum_{u=1}^{U}\sum_{C=1}^{u} n_C(u)CS_0{}^{C-1}/QA_1 \quad (30)$$

Putting in eqs. (19) and (20) for n_C and again evaluating the sums over C by the binomial expansion, we obtain

$$S = \sum_{u=1}^{U} uN(u)(1-Q+QS_0)^u/A_1 \quad (31)$$

and

$$S_0 = \sum_{u=1}^{U} uN(u)(1 - Q + QS_0)^{u-1}/A_1 \tag{32}$$

Therefore

$$S = S_0(1 - Q + QS_0) \tag{33}$$

For small Q this result reduces to the usual approximation that $S = S_0$.

Comparison of Exact and Approximate Solutions

Neither eq. (11) nor eq. (28) permits an easy solution for G or S with a given set of the independent variables P, Q, and U. However, of the two, eq. (11) is the simpler because the variables occur in pairs; that is, as PU or as $QU = \delta$. We prefer therefore to use eq. (11) when conditions of the problem permit it, and now proceed to determine what these conditions are. Probably the clearest way to compare the two equations is to examine a table of values for G computed over a range of P, Q, and U. These results are given in Table I, where G for a specific U is obtained from eq. (28) and that for $U = \infty$ from eq. (11). Sets of numbers to be compared are those within each group of four. We see the disparity in G increasing as P/Q increases and as δ decreases. The approximation [eq. (11)] is certainly safe for $\delta \geqslant 4$ and $P/Q \leqslant 0.1$, and for nearly all cases in which

TABLE I
Comparison of G Calculated for Various U Values From Equation (28) With G Calculated for $U \rightarrow \infty$ from Equation (11)

		G calculated for various P/Q or PU/δ			
δ	U	0.05	0.1	0.4	1.0
2	10	0.748	0.713	0.450	0
2	20	0.759	0.730	0.513	0
2	100	0.766	0.737	0.552	0.145
2	∞	0.768	0.739	0.560	0.184
4	10	0.963	0.937	0.720	0
4	20	0.960	0.937	0.760	0.238
4	100	0.958	0.936	0.782	0.417
4	∞	0.958	0.936	0.788	0.452
10	10	0.988	0.971	0.702	0
10	20	0.988	0.973	0.809	0
10	100	0.989	0.975	0.857	0.487
10	∞	0.989	0.976	0.866	0.559
40	10	—	—	—	0
40	20	—	—	—	0
40	100	0.995	0.987	0.852	0.191
40	∞	0.995	0.988	0.901	0.604

TABLE II
G Calculated from Equation (11)

	G										
PU	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 5$	$\delta = 7$	$\delta = 10$	$\delta = 20$	$\delta = 30$	$\delta = 50$	$\delta = 70$	$\delta = 100$
0.1	0	0.7682	0.9240	0.9856	0.9950	0.9995	0.9979	0.9998	0.9999	1.0000	1.0000
0.2	0	0.7393	0.9073	0.9780	0.9908	0.9958	0.9989	0.9995	0.9998	0.9999	1.0000
0.3	0	0.7101	0.8904	0.9703	0.9865	0.9935	0.9983	0.9993	0.9997	0.9999	0.9999
0.5	0	0.6512	0.8561	0.9544	0.9775	0.9888	0.9971	0.9987	0.9995	0.9997	0.9999
0.7	0	0.5914	0.8210	0.9379	0.9681	0.9838	0.9957	0.9980	0.9993	0.9996	0.9998
1	0	0.5000	0.7672	0.9122	0.9531	0.9757	0.9933	0.9969	0.9989	0.9994	0.9997
2	0	0.1835	0.5778	0.8192	0.8973	0.9442	0.9838	0.9923	0.9971	0.9985	0.9992
3	0	0.0001	0.3759	0.7167	0.8340	0.9073	0.9718	0.9864	0.9947	0.9972	0.9986
5	0	0	0.0002	0.4907	0.6907	0.8212	0.9421	0.9712	0.9885	0.9938	0.9968
7	0	0	0	0.2444	0.5312	0.7228	0.9064	0.9524	0.9805	0.9894	0.9945
10	0	0	0	0.0001	0.2708	0.5591	0.8443	0.9187	0.9657	0.9810	0.9901
20	0	0	0	0	0	0.0002	0.5895	0.7745	0.8985	0.9417	0.9683
30	0	0	0	0	0	0	0.2909	0.5992	0.8127	0.8895	0.9383
50	0	0	0	0	0	0	0	0.1961	0.6069	0.7606	0.8618
70	0	0	0	0	0	0	0	0.0001	0.3722	0.6101	0.7698
100	0	0	0	0	0	0	0	0	0.0009	0.3582	0.6125

$U \geqslant 100$. We will consider the approximate form to be adequate for this series of papers. However, there is evidence (see IV) that wool keratin contains a large fraction in which U is of the order of 25. Hence, for more exact studies later on, we may have to use the rigorous solution [eq. (28)].

Table II gives a series of G values computed from eq. (11) for a range of δ and PU. Although interpolation in this table is not difficult, a desirable alternative is an approximate explicit solution for G in terms of PU and δ; this we have not yet been able to devise in a simple enough form.

References

1. Menefee, E., *J. Appl. Polymer Sci.*, **9**, 2829 (1965).
2. Menefee, E., and G. Yee, *J. Appl. Polymer Sci.*, **9**, 2835 (1965).
3. Menefee, E., and G. Yee, *J. Appl. Polymer Sci.*, **9**, 2847 (1965).
4. Charlesby, A., *Atomic Radiation and Polymers*, Pergamon Press, New York, 1960.
5. Montroll, E. W., and R. Simha, *J. Chem. Phys.*, **8**, 721 (1940).

Résumé

Nous avons déduit des expressions pour déterminer la quantité soluble après chaque rupture de réticulation ou chaque scission de chaine dans le cas d'un polymère réticulé possédant une distribution uniforme du poids moléculaire. On a obtenu deux résultats de forme voisine. L'un est une forme approchée basée sur le formalisme de Charlesby et est utilisable pour la plupart des besoins. L'autre, basé sur la distribution de la dégradation suivant Montroll et Simha, est exacte mais plutôt incommode. Nous comparons les résultats numériques à partir de chaque essai.

Zusammenfassung

Für ein vernetztes Polymeres mit einheitlicher Molekulargewichtsverteilung werden Ausdrücke für den löslichen Anteil nach einem beliebigen Betrag an Vernetzungsstellen- oder Kettenspaltung abgeleitet. Es werde zwei Ergebnisse in geschlossener Form erhalten. Das eine ist eine auf dem Formalismus von Charlesby beruhende Näherung, welche für die meisten Zwecke brauchbar ist. Das andere, auf einer Abbauverteilung von Montroll und Simha beruhende ist streng, jedoch ziemlich schwer auszuwerten. Numerische Ergebnisse beider Beziehungen werden verglichen.

Received February 19, 1965