

ration of the nucleoprotein of the hog thyroid. One of us (R. O. C.) wishes to acknowledge financial assistance received from the Wisconsin Alumni Research Foundation.

### Summary

1. Certain striking physical and chemical properties of the calf thymus nucleohistone protein have been presented. The phosphorus and

nitrogen contents were accurately determined. The specific volume, refractive index increment, and viscosity of this protein in solution, as well as a titration curve are given.

2. The effect of certain preparational methods on the centrifugal and diffusion properties of the hog thyroid nucleoprotein has been briefly discussed.

RECEIVED FEBRUARY 17, 1940

[CONTRIBUTION FROM THE NEW YORK STATE AGRICULTURAL EXPERIMENT STATION]

## The Calculation of Weight Average Molecular Weights during the Course of Decomposition of Homogeneous Linear Polymers

BY WALTER H. DURFEE<sup>1</sup> AND Z. I. KERTESZ<sup>2</sup>

For materials that occur naturally and that are composed of large molecules (in whatever sense the word molecule may be understood), some degree of heterogeneity as regards size of the individual molecules must be assumed. This natural heterogeneity is further increased upon progressing degradation until the lower polymers become dominant and, finally, by complete depolymerization, homogeneity is attained.

Such heterogeneity is obvious in any pectin isolated from plant material.<sup>3</sup> Its heterogeneity is increased by the chemical manipulations necessary for its isolation. On the other hand, there is ample evidence to show<sup>4</sup> that by extracting the crude pectic materials with 50% ethanol the lower polymers originally present (or formed during preparation) are removed, together with some other materials, usually accompanying pectins in nature. It is obvious, however, that even material prepared by this latter method contains a certain range of the various polymers and is not entirely homogeneous.

For several years one of us (Z. I. K.) has been engaged in research on the chemical and physical structure of the pectin molecule. One method of approach to this problem is through a study of the changes which take place in the chemical and physical structure of pectin upon enzymic

hydrolysis. It is hoped that by such a study information can be gained concerning the naturally occurring pectic materials.

It has been reported recently<sup>5</sup> that when the decomposition of pectin is followed by various methods, the conclusions drawn concerning the average degree of polymerization are not in agreement. This led to an investigation of the number and weight averages<sup>6</sup> ( $M_n$ ,  $M_w$ ) for the "molecular weight" during the course of decomposition. During this work a formula was developed for the calculation of the  $M_w$  for any stage of progressive decomposition of a homogeneous linear polymer which, the authors feel, may also be useful to other workers dealing with such materials.

Several articles dealing with similar calculations already have appeared in the literature. W. Kuhn<sup>7</sup> made a study of the kinetics of the decomposition of chain-like polymers and calculated the proportions of various fission products occurring during the course of the hydrolysis. Dostal and Mark<sup>8</sup> investigated the heterogeneity of materials composed of macromolecules and proposed a method for the determination of the distribution of the proportion of fractions having various molecular weights in the mixture. More recently Lassatre<sup>9</sup> considered the degree of polymerization of cellulose during hydrolysis. All these authors approached the problem from the kinetical point of view and we feel that there is need to consider the basic relation between

(1) Department of Mathematics, Hobart College.

(2) Department of Chemistry, New York State Agricultural Experiment Station. Approved by the Director for publication as Journal Paper No. 369, January 30, 1940.

(3) Pectin is considered as a polymer of galacturonic anhydride residues in which the carboxyl groups are mostly changed to the methyl ester.

(4) Schneider, *et al.*, *Ber.*, **69**, 309, 2530, also 2537 (1936); **70**, 1611, also 1617 (1937); **71**, 1553 (1938), etc.

(5) Kertesz, *THIS JOURNAL*, **61**, 2544 (1939).

(6) Kraemer and Lansing, *J. Phys. Chem.*, **39**, 153 (1935).

(7) W. Kuhn, *Ber.*, **63**, 1503 (1930).

(8) Dostal and Mark, *Trans. Faraday Soc.*, **33**, 350 (1937).

(9) Lassatre, not yet published.

$M_n$  and  $M_w$  in a more general manner, without considering the rate at which the progressive decomposition proceeds. In the present work it has been assumed that the number of broken linkages can be chemically determined (as is mostly the case) and thus the progress of the decomposition may be followed in that manner. The purpose of the calculations was to determine the  $M_w$  (and  $M_n$ ) corresponding to any definite extent of hydrolysis as measured by the number of linkages broken irrespective of the velocity of the reaction.

#### Methods Used and Results Obtained.—

Some assumptions have been made in order to increase the general validity of the results as well as for the purpose of defining their limitations. The question whether these assumptions hold in the case of pectins is not the subject of this paper and will be dealt with elsewhere. The molecules have been assumed to be composed of  $n$  identical units coupled by  $n-1$  linkages to form a chain-like compound. The original material (composed of  $n$  units each of the molecular weight  $m$ , so that  $M = mn$ ) has been assumed to be homogeneous in its molecular weight. It has been assumed that fission occurs as easily in a linkage of two units as in a long chain and furthermore that each step in the fission occurs simultaneously in all pieces of all chains, thus ignoring the impossibility of fission occurring in single units.

The calculation of the  $M_n$  at any degree of hydrolysis is obvious and simple. It will always equal  $M$  (original)/ $p = mn/p$ , where  $p$  is the number of groups or pieces into which the original chain of  $n$  units has split. The number of linkages opened is always  $p-1$ .

In the calculation of the  $M_w$  the size of each of the  $p$  groups, as well as all of the various possible ways in which the original chain of  $n$  units may be split into  $p$  groups, must be considered. In order to take all possibilities into consideration, it is necessary to assume the presence of a large number of molecules. While it is true that if a chain is split progressively in such a way that all the resulting groups are composed of the same number of units, the values of  $M_n$  and  $M_w$  will be equal, these values will differ considerably from each other when an average of all possible splittings is calculated, as can be done by formula (4).

We denote by  $K(n,p)$  the number of ways that the  $n$  units of a chain may be split into  $p$  groups,

each containing one or more units. (Obviously,  $1 \leq p \leq n$ ). There being  $n-1$  linkages and  $p-1$  partitions between groups,  $K(n,p)$  is the number of ways that the  $p-1$  partitions may be distributed among the  $n-1$  linkages

$$K(n,p) = {}_{n-1}C_{p-1} = \frac{(n-1)!}{(n-p)!(p-1)!} = \frac{(n-1)(n-2)\dots(n-p+1)}{1 \cdot 2 \dots (p-1)} \quad (1)$$

Now  $M_w = \sum_i f_i M_i$ , where  $f_i$  is "the fractional weight of the constituent of molecular weight  $M_i$ ," and the summation extends over all "constituents."<sup>6</sup> A "constituent" is one of the  $p$  groups: if it contains  $n_i$  units its  $f_i$  is  $n_i/n$ , and its  $M_i = mn_i$ . Thus  $f_i = M_i/mn$ , and for a single chain we have

$$M_w = \sum_{i=1}^p f_i M_i = \sum_{i=1}^p \frac{M_i^2}{mn} = \frac{1}{mn} \sum_{i=1}^p M_i^2 \quad (2)$$

This must be averaged over all  $K(n,p)$  possible arrangements of  $n$  units in  $p$  groups. To do this, suppose  $K(n,p)$  chains arranged one below another. There being  $p$  groups in each chain, there will be  $p$  columns, and in each column a group of  $r$  units will occur as many times as the number of ways in which the remaining  $n-r$  units can be split into  $p-1$  groups, namely

$$K(n-r, p-1) = {}_{n-r-1}C_{p-2}$$

Hence, considering all  $p$  columns, a group of  $r$  units, each with  $f_i = r/n$ ,  $M_i = mr$ , occurs  ${}_{n-r-1}C_{p-2}$  times, contributing to the total  $f_i M_i$  the amount  $\frac{m^2 p r^2}{n} \cdot {}_{n-r-1}C_{p-2}$ . Summing this for all values of  $r$  from 1 to  $n-p+1$ , and dividing by  $K(n,p)$  for the average, we have

$$M_w = \sum_{r=1}^{n-p+1} \frac{m^2 p r^2}{n} \cdot {}_{n-r-1}C_{p-2} / {}_{n-1}C_{p-1} \quad (3)$$

This expression reduces without difficulty to

$$M_w = \frac{2n-p+1}{p+1} m \quad (4)$$

The ratio of  $M_n$  to  $M_w$  is, therefore

$$\frac{M_n}{M_w} = \frac{n(p+1)}{p(2n-p+1)} \quad (5)$$

which is less than unity for  $1 < n < p$ .

This ratio has a minimum value for  $p = \sqrt{2n+2} - 1$ , namely,  $n/(\sqrt{2n+2} - 1)^2$ .

Consider, as an example, a chain of ten units split into seven groups. We may assume  $m = 1$ . We may analyze this situation as follows

(a) 6 groups of 1 unit each, 1 group of 4 units

Units	1	1	1	1	1	1	4
$f_i$	.1	.1	.1	.1	.1	.1	.4
$M_i$	1	1	1	1	1	1	4
$f_i M_i$	.1	.1	.1	.1	.1	.1	1.6

$\Sigma f_i M_i = 2.2$

There are seven such arrangements, corresponding to the seven places in which the 4-group may fall; hence, we have a total  $\Sigma f_i M_i = 7 \times 2.2 = 15.4$ .

(b) 5 groups of 1 unit each, 1 group of 2 units and 1 of 3 units

Units	1	1	1	1	1	2	3
$f_i$	.1	.1	.1	.1	.1	.2	.3
$M_i$	1	1	1	1	1	2	3
$f_i M_i$	.1	.1	.1	.1	.1	.4	.9

$\Sigma f_i M_i = 1.8$

There are 42 such arrangements obtainable by placing the 2-group and 3-group in all possible positions, giving a total  $\Sigma f_i M_i = 42 \times 1.8 = 75.6$ .

(c) 4 groups of 1 unit each, 3 groups of 2 units each

Units	1	1	1	1	2	2	2
$f_i$	.1	.1	.1	.1	.2	.2	.2
$M_i$	1	1	1	1	2	2	2
$f_i M_i$	.1	.1	.1	.1	.4	.4	.4

$\Sigma f_i M_i = 1.6$

There are 35 such arrangements obtainable by

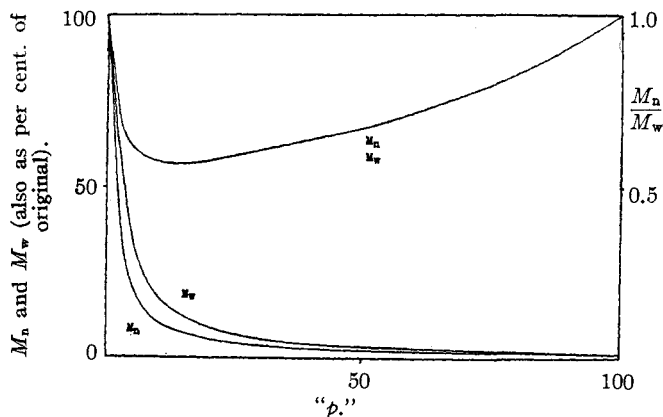


Fig. 1.—The weight average molecular weight ( $M_w$ ), number average molecular weight ( $M_n$ ) and the ratio  $M_n/M_w$  in case of progressive splitting of a 100 unit chain ( $n = 100$ ,  $m = 1$ ).

placing the three 2-groups in all possible positions, giving a total  $\Sigma f_i M_i = 35 \times 1.6 = 56.0$ .

Adding these results we have  $\Sigma f_i M_i = 147.0$ , obtained by using a total of 84 arrangements. The average is  $147.0/84 = 1.75$ , which is the value of  $M_w$ .

This method of analysis quickly becomes impracticable as the value of  $n$  increases. The results may be obtained directly, however, from formulas (1) and (4). In the example considered,  $n = 10$  and  $p = 7$ , hence

$$K(10,7) = {}_9C_6 = 84; \quad M_w = \frac{20 - 7 + 1}{7 + 1} = \frac{14}{8} = 1.75$$

Figure 1 shows the values of  $M_n$  and  $M_w$ , again assuming  $m = 1$ , for a chain originally of 100 units progressively hydrolyzed, that is, for values of  $p$  from 1 to 100. In addition, the values of the ratio  $M_n/M_w$  are plotted. This ratio is a minimum for  $p = 13$ . At this point, we have  $M_n = 7.7$ ,  $M_w = 13.4$ , and  $M_n/M_w = 0.57$ . To the right of this point, the curves slowly converge again, meeting at the point ( $p = 100$ ), indicating complete hydrolysis.

It is realized that in order to have a sufficient number of the various early split-products present the number of molecules required from the starting material is very high; but assuming a *random* distribution in the size of those chains available for the various steps in the fission, the values obtained for the  $M_w$  will not be influenced.

### Summary

A formula has been developed for the calculation of the weight average molecular weight of a linear polymer undergoing progressive splitting. An example is given to show the validity of the formula obtained. By the use of the formula the trend of average number and weight molecular weights for a polymer containing 100 units was calculated and is shown in Fig. 1. At the point representing 13% of complete hydrolysis the ratio between the two averages reaches a minimum, 0.57.

GENEVA, NEW YORK

RECEIVED FEBRUARY 7, 1940