THE MOLECULAR WEIGHTS OF PROTEINS IN PHENOL

BY EDWIN J. COHN AND JAMES B. CONANT

DEPARTMENT OF PHYSICAL CHEMISTRY, HARVARD MEDICAL SCHOOL, BOSTON, AND CHEMICAL LABORATORY OF HARVARD UNIVERSITY, CAMBRIDGE

Communicated June 11, 1926

I. A large body of evidence has been accumulated in recent years which indicates that proteins have very large molecular weights. The molecular weight of egg albumin, the smallest of the proteins that have been adequately studied, is approximately 34,000. This estimate was made by Sørensen[14] on the basis of osmotic pressure measurements in which he corrected for the unavoidable influence of inorganic salts, and is entirely confirmed by analyses of the sulfur and sulfide sulfur and of certain of the amino acids of which the molecule is composed.[3] Since egg albumin contains 1.23% of tryptophane[5] the smallest weight of a gram molecule that could contain one gram molecule of this amino acid would be approximately 16,600. The sulfide sulfur and sulfur contents[10] suggest a molecular weight of twice this minimal molecular weight as the calculations in table 1 demonstrate. The analytical and physical-chemical measurements thus lead to the same result and leave little doubt regarding the size of this protein.

TABLE 1

MINIMAL MOLECULAR WEIGHT OF EGG ALBUMIN

PROTEIN CONSTITUENT	AMOUNT OF CONSTITUENT IN PROTEIN PER CENT	WEIGHT CONTAINING 1 GM. MOL. CONSTITUENT	ASSUMED NO. OF ATOMS OR MOLECULES	MINIMAL MOLECULAR WEIGHT
Cystine content	0.81	29,654	1	29,654
Tryptophane content	1.23	16,593	2	33,186
Sulfide sulfur content	0.491	6,532	5	32,660
Sulfur content	1.616	1,984	17	33,728

The other proteins that have been studied thus far have still greater molecular weights. The molecular weights of those considered in this paper have been calculated from analytical evidence in the following table.

The evidence yields minimal molecular weights for gliadin, zein, gelatin and casein of 72,000, 96,000, 150,000 and 192,000, respectively.

II. These results are in marked contrast to the surprisingly low molecular weights recently reported as calculated from the freezing-point lowering in phenol of certain of these proteins.[8,15] Estimates made by the different investigators who have used this method have ranged from 200 to 600. It has been suggested by them that these values represent the true molecular weight of the units of a large aggregate which is stable in water but which

dissociates in phenol. These results have been summarized in a recent review[7] in which the conflicting interpretations have been considered.

TABLE 2

MINIMAL MOLECULAR WEIGHTS OF GLIADIN, ZEIN, GELATIN AND CASEIN

PROTEIN CONSTITUENT	AMOUNT OF CONSTITUENT IN PROTEIN* PER CENT	WEIGHT CONTAINING 1 GM. MOL. CONSTITUENT	ASSUMED NO. OF ATOMS OR MOLECULES	MINIMAL MOLECULAR WEIGHT
		GLIADIN		
Tryptophane content	1.12	18,223	4	72,892
Cystine content	2.32	10,353	7	72,471
Sulfide sulfur content	0.619	5,181	14	72,534
Sulfur content	1.027	3,123	23	71,892
		ZEIN		
Cystine content	0.75	32,027	3	96,081
Histidine content	0.82	18,915	5	94,575
Sulfide sulfur content	0.212	15,127	6	90,762
Sulfur content	0.60	5,345	18	96,210
		GELATIN		
Cystine content	0.16	150,125	1	150,125
		CASEIN		
Cystine content	0.25	96,080	2	192,160
Sulfide sulfur content	0.101	31,752	6	190,512
Tryptophane content	1.6	12,756	15	191,340

* The most probable values have been taken from tables that have been published elsewhere incorporating the results of Osborne, Dakin, Foreman, Jones, Gersdorff and Mueller and Folin and Looney, and from new unpublished measurements of J. M. Looney.

A real difficulty exists in reconciling such an hypothesis with the analyses described above. One would have to assume, for instance, that these proteins dissociated into unlike units containing different amino acids. For the amino acid composition of certain of the proteins which have been investigated by the cryoscopic method are very well known as a result of recent investigations.[4,5,6,10,11,12,16] The percentage composition of zein, of gelatin and of casein are recorded in table 3. From them, and the molecular weights deduced in table 2, the molecular compositions of these proteins have been calculated. The molecular weights that have been estimated by the cryoscopic method and which have been assumed to represent the true units into which these proteins dissociate in phenol could only represent dipeptids, tripeptids, or in a few instances, molecules containing a somewhat larger number of amino acids. The large number of different amino acids of which these proteins are composed render it certain, however, that these units would be of different composition. It is difficult to imagine a mechanism by which so many dissimilar units might recombine to form the original molecule.

In an effort to understand the discrepancy between these two views regarding the protein molecule, we have been lead to reinvestigate the molecular weights of four proteins, gelatin, zein, casein and gliadin by the cryoscopic method in phenol. These proteins were chosen since their compositions are better known than are those of any other proteins, and since they are not readily denatured.[9] All of them were carefully purified both from other proteins and from electrolytes.

TABLE 3

THE AMINO ACID COMPOSITION OF CERTAIN PROTEINS

AMINO ACID	PERCENTAGE COMPOSITION OF PROTEIN*			MOLECULAR COMPOSITION OF PROTEIN		
	ZEIN %	GELATIN %	CASEIN %	ZEIN MOLS.	GELATIN MOLS.	CASEIN MOLS.
Cystine	0.75	0.16	0.25	3	1	2
Tryptophane	...	...	1.6	..	..	15
Tyrosine	5.6	?	5.36	30	?	57
Histidine	0.82	2.94	2.5	5	28	31
Arginine	1.82	8.22	3.81	10	71	42
Lysine	...	5.92	8.38	..	61	110
Glycine	...	25.5	(0.45)	..	510	(1)
Alanine	3.8	8.7	1.85	41	146	40
Phenylalanine	7.6	1.4	3.88	44	13	45
Serine	1.0	0.4	0.5	9	6	1
Valine	1.9	...	7.93	15	..	130
Leucine	25.0	7.1	9.7	183	81	142
Proline	9.0	9.5	8.7	75	124	145
Hydroxyproline	...	14.1	0.23	..	161	3
Aspartic Acid	1.8	3.4	4.1	13	38	59
β-Hydroxyglutamic Acid	2.5	...	10.5	15	..	124
Glutamic Acid	31.3	5.8	21.77	204	59	284
Ammonia	3.6	0.49	1.61	203**	43**	182**
	96.49	93.63	93.12	647	1,299	1,231

* The convention has been followed of referring to the ratio of the weight of the amino acid recovered, to the weight of the protein analyzed as the percentage composition of the protein. The water absorbed by the amino acids during hydrolysis should, of course, be added to the denominator in order to yield the true percentage composition.

** The number of ammonia molecules have not been added to the total, which therefore represents the number of amino acid molecules thus far isolated from the protein.

The determination of the molecular weight of a protein in aqueous solution by the lowering of its freezing-point is rendered difficult by the presence of even very small amounts of salts. This limitation also applies to non-aqueous solvents and in addition the presence of small amounts of moisture greatly affects the results. Zein was chosen as the first protein to be studied since the small number of reactive groups which it contains[2] suggested that it might be less hygroscopic than many other proteins. A sample dried at

100°C. in vacuo for 3 hours, however, gave freezing-point lowerings in phenol corresponding to a molecular weight of 1800–2200. Further drying by grinding the protein under absolute ether and subsequent evacuation at a pressure of 0.001 mm. Hg for 4 hours yielded a sample whose apparent molecular weight had been increased to 4250.

III. These preliminary experiments revealed the difficulties of removing the last trace of water from proteins. A method of determining molecular weights in phenol was therefore devised which would be unaffected by small quantities of water. This was accomplished by the addition of a second phase consisting of freshly powdered anhydrous calcium chloride. This saturating body was introduced in the belief that as long as the system contained solid anhydrous calcium chloride in equilibrium with its hydration products the freezing-point would remain unchanged, would be independent of added water, and would correspond to the vapor pressure of the solid phases.

Experiments demonstrated that such a system had a definite and sharp freezing-point, far sharper than that of even carefully purified phenol. To test the effectiveness of this system small quantities of water were added. The freezing-point was found unchanged. As much as a tenth of a cubic centimeter of water could be added to a mixture of 20 gm. of phenol and 4 gm. of calcium chloride without affecting the freezing-point. A similar procedure has recently been effectively employed in determining molecular weights in freezing nitro benzene, sodium sulfate being the solid phase.[1,13]

IV. When proteins were added to the system—anhydrous calcium chloride + hydrated calcium chloride + phenol—no measureable change in freezing-point occurred. This was the case with carefully dried zein and gelatin. When samples of proteins, which had not been so carefully dried were studied, a lowering of the freezing-point was observed when insufficient calcium chloride had been employed. The addition of more calcium chloride always raised the freezing-point to the original value. The experiment with zein illustrates this phenomenon (Table 4(*c*)). In connection with the use of this method in similar problems, it might be pointed out that sufficient calcium chloride must be used to combine with all the water present, but also that some water must be present to form the hydrated calcium chloride phase.

In order to demonstrate that the proteins had not been precipitated with the calcium chloride, in several experiments the solid phase was separated by centrifugation and the supernatant liquid poured into ether. In this way copious precipitates of gelatin and zein were obtained.

The determination of the molecular weight of naphthalene in the phenol-calcium chloride system demonstrates the validity of the method (Table 4 (*d*)). In the attempt to apply this method to the determination of the molecular weights of amino acids a complication was encountered. Al-

though the expected initial lowering of the freezing-point occurred, after a few minutes the freezing-point gradually rose to the initial value. This phenomenon was encountered with para-amino benzoic acid, benzamide and arginine. It was shown in the case of para-amino benzoic acid that the material had passed into the solid phase, since the phenol layer contained but a trace as judged by diazotization and coupling with β-naphthol. Judged by the same criterion, the solid phase contained large amounts of para-amino benzoic acid.

TABLE 4

SOME MEASUREMENTS OF THE FREEZING-POINT IN THE SYSTEM, PHENOL + ANHYDROUS $CaCl_2$ + HYDRATED $CaCl_2$

		BECKMAN CALIBRATION:	40.35°C. = 2.000°
(*a*)	15.0 gm. phenol; 2 gm. $CaCl_2$	freezing-point	2.008°
	1 drop H_2O added	first ice	2.010°
		freezing-point	2.000°
(*b*)	17.6 gm. phenol; 3 gm. $CaCl_2$	first ice	1.995°
		freezing-point	1.990°
	0.30 gm. gelatin added	first ice	2.030°
		freezing-point	2.005°
	1.08 gm. more gelatin added	first ice	2.102°
		freezing-point	2.005°
(*c*)	28.3 gm. phenol;* 4 gm. $CaCl_2$	first ice	2.120°
		freezing-point	2.070°
	1.00 gm. zein added	first ice	1.940°
		freezing-point	1.920°
	4 gm. additional $CaCl_2$ added	first ice	2.080°
		freezing-point	2.070°
(*d*)	28.6 gm. phenol;* 6 gm. $CaCl_2$	first ice	2.078°
		freezing-point	2.062°
	0.313 gm. naphthalene added	first ice	1.450°
	(Calculated molecular weight	freezing-point	1.350°
	of naphthalene from these data	freezing-point lowering	0.712°
	= 108; K being taken as 7200:		
	theoretical = 128)		

* A very pure phenol was employed (m.p.t. = 40.85°C.) it was necessary to add 1 drop of water to form the hydrated calcium chloride phase and obtain a constant freezing-point.

It is noteworthy that but few of the naturally occurring amino acids are soluble in phenol, although many proteins are readily dissolved by this solvent.

The behavior of these amino acids made it important to establish quantitatively that the added protein remained in solution during the freezing-point determination. At the conclusion of the measurements upon zein reported in table 4 (*c*), the protein mixture was centrifuged and an

aliquot of the clear supernatant phenol poured into ether. The zein separated and was filtered, washed with ether and the precipitate analyzed for nitrogen by the Kjeldahl method. Although a certain amount of protein might not have been recovered in the aliquot analyzed, 11.27 gm. yielded 22.9 mg. of N, as ammonia, corresponding to 0.14 gm. zein. It is thus evident that 0.14 gm. of zein dissolved in 11.27 gm. of phenol caused a depression in the freezing-point of less than the experimental error. Putting this error even at 0.01°C. yields an apparent molecular weight for zein in phenol of over 10,000. Similar considerations apply to the experiments with other proteins.

We are, therefore, forced to conclude that the study of the freezing-point lowering of proteins in phenol yields no evidence for dissociation of the protein molecule into units of low molecular weight, and that the true molecular weights of the proteins in phenol, as in water, are those revealed by the analytical and physico-chemical methods considered above.*

* Although this method is clearly not suited to the accurate determination of the true molecular weights of the proteins, it offers a simple and accurate procedure for the determination of the impurities in a protein preparation. If the freezing-point lowering of a solution of a protein or even of a tissue were determined both in pure phenol and in the system phenol + calcium chloride, the difference should yield the moisture in the preparation.

[1] Brown, F. S., and Bury, C. R., *J. Phys. Chem.*, 1926, **30**, 694.

[2] Cohn, E. J., Berggren, R. E. L., and Hendry, J. L., *J. Gen. Physiol.*, 1924–25, **7**, 81.

[3] Cohn, E. J., Hendry, J. L., and Prentiss, A. M., *J. Biol. Chem.*, 1925, **63**, 721.

[4] Dakin, H. D., *J. Biol. Chem.*, 1920, **44**, 499; *Z. physiol. Chem.*, 1923, **130**, 159.

[5] Folin, O., and Looney, J. M., *J. Biol. Chem.*, 1922, **51**, 421.

[6] Foreman, F. W., *Biochem. J.*, 1919, **13**, 376.

[7] Haurowitz, F., *Biochemie der Menschen und der Tiere seit*, 1914, 94.

[8] Herzog, R. O., and Kobel, N., *Z. physiol. Chem.*, 1924, **134**, 296.

[9] Mathewson, W. E., *J. Amer. Chem. Soc.*, 1906, **28**, 1482.

[10] Osborne, T. B., *J. Amer. Chem. Soc.*, 1902, **24**, 140; *The Vegetable Proteins*, London and New York, 2nd edition, 1924.

[11] Osborne, T. B., and Guest, H. H., *J. Biol. Chem.*, 1911, **9**, 333.

[12] Osborne, T. B., and Liddle, L. M., *Amer. J. Physiol.*, 1910, **26**, 295.

[13] Roberts, H. M., and Bury, C. R., *J. Chem. Soc.*, 1923, **123**, 2037.

[14] Sørensen, S. P. L., *Compt. rend. trav. Lab. Carlsberg*, 1917, **12**, 262.

[15] Troensegaard, N., and Schmidt, J., *J. physiol. Chem.*, 1924, **133**, 116.

[16] Van Slyke, D. D., and Abderhalden, E., *Handbuch der biologischen Arbeitsmethoden*, Berlin and Vienna, **1**, pt. 7, 1923, 53.