SOME PHYSICAL-CHEMICAL CHARACTERISTICS OF PROTEIN MOLECULES[1, 2]

EDWIN J. COHN

Department of Physical Chemistry, Harvard Medical School, Boston, Massachusetts

Received January 30, 1939

I. INTRODUCTION

The importance of developing an adequate chemistry of the proteins has, perhaps, never been doubted. These vast molecules could not, however, be adequately described by the methods available to earlier generations. The analytical chemistry of the early part of the 19th century accurately revealed, it is true, their elementary composition. A century ago Liebig and Mulder knew that proteins were rich in nitrogen, and contained, besides carbon, hydrogen, and oxygen, small amounts of sulfur and of phosphorus. A few characteristic analyses of this period are worth republishing (see table 1).

The elementary composition of most proteins was thus early estimated to be so nearly alike as not to suggest the vast chemical and biological differences that can be ascribed to the different molecules of this group. Only the analyses for sulfur and phosphorus showed wide variations from protein to protein and led to the controversy as to whether these elements were indeed part of the protein molecule. No sulfur-containing amino acid had been isolated from a protein hydrolysate[3] in 1838, but Mulder

[1] Presented at the Symposium on the Physical Chemistry of the Proteins, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry of the American Chemical Society.

[2] The points of view developed in this communication have in part been derived from the discussions held in our seminar. I am therefore indebted to George Scatchard, John G. Kirkwood, and John D. Ferry for considering with us intermolecular forces; to Hans Mueller, Jeffries Wyman, Jr., J. L. Oncley, John D. Ferry, and J. Shack for considering dielectric constant measurements and their interpretation; to John T. Edsall and John W. Mehl for considering diffusion, viscosity, and double refraction of flow; to Ronald M. Ferry and A. A. Green for considering electromotive force and electrophoretic mobility; to Norval F. Burk for considering osmotic pressure; to J. L. Oncley and Jacinto Steinhardt for considering sedimentation methods of determining molecular weight; and to T. L. McMeekin and Jesse P. Greenstein for considering the amino acid composition of the proteins.

[3] Wollaston (158) discovered cystine in 1810, but it was not isolated from a protein until eighty years later (81, 97).

calculated that on the basis of the analyses for sulfur and phosphorus in table 1, and the atomic weights in use at the time, that the molecular weights of these proteins would be over 50,000.

The isolation of the amino acids from the hydrolysates of proteins occupied the attention of the great organic chemists of the 19th and the early 20th centuries. Something over twenty amino acids are now recog-

TABLE 1

Elementary composition of proteins

COMPOSITION							PROTEIN	FORMULA							MOLECULAR WEIGHT
C	H	N	O	S	Fe	P		C	H	N	O	S	Fe	P	
According to Mulder in 1838* (99)															
54.56	6.90	15.72	22.13	0.36		0.33	Fibrin	400	620	100	120	1		1	55,692
54.48	7.01	15.70	22.00	0.38		0.43	Egg albumin	400	620	100	120	1		1	55,692
54.84	7.09	15.82	21.23	0.68		0.33	Serum albumin	400	620	100	120	2		1	55,893
According to Osborne in 1902† (111)															
52.75	7.10	15.51	23.02	1.616			Egg albumin	696	1125	175	220	8			15,703
55.23	7.26	16.13	20.78	0.600			Zein	736	1161	184	208	3			15,993
52.72	6.86	17.66	21.73	1.027			Gliadin	685	1068	196	211	5			15,568
52.99	7.01	15.93	22.14	1.930			Serum albumin‡	662	1051	171	207	9			14,989
54.64	7.09	17.38	20.16	0.39	0.335		Hemoglobin‡	758	1181	207	210	2	1		16,655
54.57	7.11	16.38	21.03	0.568	0.336		Hemoglobin§	758	1185	195	219	3	1		16,667
52.68	6.83	16.91	22.48	1.10			Fibrin	645	1004	178	207	5			14,708
51.50	7.02	18.69	21.91	0.88			Edestin	624	1021	193	199	4			14,523
53.13	7.06	15.78	22.37	0.80		0.86	Casein	708	1130	180	224	4		4	15,982

* For a critical study see Laskowsky (Ann. **58**, 129 (1846)).

† Osborne quotes the earlier literature in this paper. See, also, Osborne: The Vegetable Proteins, 2nd edition, Longmans, Green and Co. (1924).

‡ Of the horse.

§ Of the dog.

nized to be constituents of protein molecules, as a result of investigations of Kossel (76), Ehrlich (40), Emil Fischer (51), Dakin (36), and most recently of Mueller (98), Rose (120), and Van Slyke (147). Not all proteins contain all known amino acids, and the proportion in which these substances occur varies widely in different proteins. Certain of the amino acids are indispensable for nutrition; certain of them play an important rôle in immunity. Special significance attaches to the sulfur-

containing acids, methionine and cysteine; to the hydroxyl-containing amino acids, serine, threonine, hydroxyproline, and hydroxyglutamic acid, which can form ester linkages with acids; to the phenolic hydroxyl groups of tyrosine; to the hexone bases, histidine, arginine, and lysine, which as free amino acids combine two acid equivalents; and to the dicarboxylic amino acids, aspartic, glutamic, and hydroxyglutamic, which can combine with two equivalents of base per mole when not present as amides.

The characteristic groups of the amino acids must be considered largely responsible for the behavior of the proteins, but the arrangement of these groups, and therefore the structural chemistry of the proteins, cannot yet be considered completely apprehended. The work of Hofmeister and of Emil Fischer leaves no doubt that the amino acids are combined with each other through the peptide linkage. Not all amino acids when bound in peptide linkage are therefore possessed of free amino groups, and not all are possessed of free carboxyl groups, but only those which are trivalent, that is, which have more than one basic, or more than one carboxylic, group. Only the α-amino and carboxyl groups of amino acids are bound in peptide linkage; the side-chain groups—the imidazole, amide, hydroxyl, sulfhydryl, and phenolic groups—are free and give rise to the reactive properties of proteins. The backbone of the protein is presumably that of the peptide chain, with side chains constituting either (*1*) non-polar hydrocarbon chains, as in alanine, valine, or leucine, (*2*) imidazole groups, as in histidine, (*3*) indole groups, as in tryptophane, (*4*) benzene rings, as in phenylalanine and tyrosine, or (*5*) strongly polar guanidine, amino, hydroxyl, or carboxyl groups. Regardless of the nature of the side chains, the distance separating them in the stretched peptide is always the same, since it depends only upon the distance of separation between carbon and carbon or carbon and nitrogen atoms. The 3.5 Å. distance separating side chains is the salient repeating line in x-ray analyses of stretched proteins as of peptides. In the case of peptides in solution, however, it has been demonstrated that there is free rotation around carbon–carbon and carbon–nitrogen bonds. The distances separating the reactive groups may therefore be smaller than 3.5 Å. and not much greater than those calculable from a statistical treatment of probable positions on the basis of free rotation (31, 44, 79, 163).

The elementary composition and the densities of proteins are essentially those that could be calculated from the amino acid residues in the molecule. The numbers of free groups and paraffin side chains of proteins could also be calculated, given the amino acid compositions. The amino acid compositions of many proteins may now be considered known, at least as a first approximation. The spatial relations of the free groups to each other could be revealed by organic chemistry, provided more were known regard-

ing the configuration of the peptide chain in the intact protein. Although various theories have been advanced (159), certain knowledge is still lacking regarding the internal structure of proteins. Meanwhile many physical-chemical methods have become available, in terms of which much can be ascertained regarding the size and shape of the molecule and the number and nature of their reactive groups as well as something regarding their spatial distribution.

II. SIZE AND SHAPE OF THE PROTEIN MOLECULE

The large size of the protein molecule was suspected long before accurate methods were available for the determination of molecular weights. Many membranes, both natural and artificial, were known to be impermeable to given proteins, and their sizes could be approximated from the dimensions of the pores of the membranes. Proteins diffuse slowly, and their molecular weights have been estimated from the rate of diffusion. In fields of force many times that of gravity, they can be centrifuged and their molecular weights calculated from the rate of sedimentation. Although the protein molecules are too large—and therefore present in too small concentration—to have significant effects upon vapor pressure or freezing point, their molecular weights can be estimated from measurements of osmotic pressure, under conditions such that membranes are impermeable to the protein but not to the small concentration of impurities from the last traces of which they are freed with difficulty. Oriented by virtue of asymmetry of shape, or of electrical charge, their times of relaxation to random distribution can be determined; and the relaxation time, like the rate of diffusion and of sedimentation, is a function of the size and shape of the molecule.

Analytical methods and the minimal molecular weight

Any accurate determination of a constituent of a protein—provided the protein preparation is pure and of uniform composition—may be employed in calculating the smallest weight that can be ascribed to this protein if it contains 1 gram-atom or gram-molecule of this component. Thus the early determinations of the iron content of hemoglobin led to the calculation that the molecular weight of this respiratory protein could not be smaller than 16,670 if it contained 1 gram-atom of iron. The iron is present in the prosthetic porphyrin group, heme, and it followed therefore that if hemoglobin contained 2 gram-atoms of iron, its molecular weight would be 33,340; if 4 gram-atoms of iron, 66,680. The minimal molecular weight does not, of course, reveal whether all the molecules in solution are of the same size, or what the average molecular weight is if the protein is not monodisperse, even though the

different molecular species have identical composition. If, however, the protein is of uniform composition, the true molecular weight must be a multiple of the minimal molecular weight calculated from any component.

Only a few proteins contain iron, but most proteins contain sulfur. Moreover the sulfur present is of two kinds, as was clearly understood by the analytical chemists of the 19th century such as Schultz (122), Krüger (77), and T. B. Osborne (111). Krüger was the first to present accurate analytical data for the total and the alkali-labile sulfur. Moreover, he pointed out that for egg albumin and for fibrin the ratio of total sulfur to labile sulfur was, respectively, 4:1 and 3:1[4] (77). The sulfide or labile sulfur therefore represents an integral fraction of the total sulfur. Osborne (111) in 1902 estimated, on the basis of the results in table 1, that the minimal molecular weights of many proteins were in the neighborhood of 15,000 to 30,000,—estimates borne out by recent physicochemical studies.

At the beginning of this century there were no satisfactory analytical methods for determining cystine, and the other sulfur-containing amino acid thus far isolated from a protein hydrolysate, methionine, had not been discovered (98). That part of the total sulfur which is not alkali-labile would appear to represent methionine on the basis of the best analytical measurements at present available (see table 2).

The analyses of the earlier investigators for both total sulfur and alkali-labile sulfur have recently been confirmed for many proteins (10, 17) (table 2). The determination of alkali-labile sulfur is in some cases equal to, in others slightly in excess of, the amount of cystine, or cysteine, estimated to be present. The discrepancy, in the case of the analysis of certain proteins, between alkali-labile sulfur and the sulfur present as cystine or cysteine may represent another thioamino acid, but it may also represent destruction of cystine during hydrolysis (11, 73, 72). Certainly with few exceptions alkali-labile sulfur is now accounted for as cystine (table 2).

Analyses for those amino acids generally present in proteins in small amounts have also been used in the estimation of the minimal molecular weights of the proteins (14, 22, 33). Tryptophane, tyrosine, methionine, and cystine, and occasionally also histidine and arginine, have been particularly useful in this connection, since the analytical methods for their determination are most readily susceptible of being placed on a firm quantitative basis. Improvement in the specificity and the accuracy of analytical methods and in the purity of protein preparations may ultimately be expected to lead to minimal molecular weights and therefore to

[4] That these are characteristic ratios for the occurrence of amino acids in proteins has since been demonstrated repeatedly (14, 15).

TABLE 2

Composition and molecular weights of proteins

COMPOSITION IN PER CENT						PROTEIN	NUMBER OF CONSTITUENTS PER MOLE						ESTIMATED MOLECULAR WEIGHT				
Total sulfur	Sulfide sulfur	Cystine	Methionine	Tryptophane	Tyrosine		Sulfide sulfur	Cysteine	Non-sulfide sulfur	Methionine	Tryptophane	Tyrosine	From composition	From osmotic pressure in concentrated urea solution	From osmotic pressure in aqueous solution	From sedimentation ultracentrifugation	From equilibrium ultracentrifugation
1.62 1.60	0.49	1.22 1.33	5.24 5.07	1.18 1.28	4.00 4.21	Egg albumin	6	4	12	12	2	8	36,000	34,000	34,000	43,800	40,500
3.34		12.5			12.2	Insulin*	36	36				24	35,000			40,900	35,100
†	0.21 0.22	0.91	2.58 2.25		5.66 5.90	Zein	3	3		6		12	39,000	37,000	39,000§	35,000	
1.03 0.99	0.62	2.15 2.40	1.60 2.03	0.83 1.82	3.04 3.50	Gliadin	8	8	6	4 6	2 4	7 8	42,000	44,000	41,000§	26,000	
0.39	0.19 0.18	0.41		1.28	3.15	Hemoglobin (horse)	2	1	2		2	6	33,350‡	34,300	67,000	69,000	68,000
†	†	6.02 5.80		0.52 0.26	4.66 4.79	Serum albumin		36			2 1	18	73,000	73,800	73,000	70,200	66,900
0.76 0.80	0.10 0.09	0.21 0.26	3.25 3.36	1.54 1.60	†	Casein	3	2	21	21	8	†	96,000‡	33,600			

1.10		0.62 0.77	3.4	0.82	3.38	Myosin (rabbit)		3		24	4	18	100,000	100,000	1,000,000		
0.88 0.99	0.35	1.22 1.35	2.30 2.40	1.45 1.56	4.53 4.58	Edestin	6	6	8 9	8 9		13 14	54,000	49,000		309,000	

* Crystalline insulin contains 0.52 per cent zinc (123). Assuming three zinc atoms per insulin molecule leads to a molecular weight of 37,700.

† The values reported are so discrepant that they are not considered here.

‡ Based on the iron analysis.

§ In ethanol-water mixtures.

Note added in proof: Kuhn, Birkofer, and Quackenbush (Ber. **72**, 407 (1939)) have published a new study of sulfur distribution in proteins since this table was compiled. For egg albumin their value for the total sulfur, 1.59 per cent, is in good agreement with Osborne's, and this and their somewhat lower value for methionine, 4.85 per cent, leads to a molecular weight of 36,000 to 37,000 on the basis of twelve residues per mole. The yield of cystine, 1.78 per cent, is moreover far greater than that previously reported, yielding six cysteine residues and removing the discrepancy noted above between sulfide sulfur and cysteine.

Their values for insulin suggest one methionine residue per mole and are therefore in fair agreement with the results of Kassel and Brand (72). It should be pointed out however that Kuhn's study was on amorphous and not on crystalline insulin.

They also studied horse hemoglobin and globin and report values of methionine and cystine suggesting two residues of the former and three of cysteine on the basis of a molecular weight of approximately 33,000. Five atoms of total sulfur are thus demanded by their analysis, as compared with the four previously assumed, and three atoms of sulfide sulfur as compared with the two previously assumed.

The only other protein in the above list studied by Kuhn, Birkofer, and Quackenbush is casein. As in the case of egg albumin and hemoglobin their cystine value, 0.42 per cent, is higher than those tabulated above, removing the discrepancy between cysteine and sulfide sulfur, and their methionine is smaller, namely, 3.03 per cent.

Sulfur and sulfide sulfur.—The earlier literature has been considered by Osborne (111) and in our first study of minimal molecular weights (33). The more recent studies on labile sulfur, the results of which are in many cases in excellent agreement with Osborne's values, are by Clarke and his associates (17, 164), Baernstein (10), and Bailey (11). The relation between total sulfur and cystine sulfur in *insulin* has been considered by Miller and du Vigneaud (95). Assuming thirty-six atoms of sulfur per mole of insulin leads to a molecular weight of 34,488. The cystine analysis suggests that most if not all of the sulfur of insulin is sulfide sulfur.

The new analyses of Blumenthal and Clarke, giving 0.22 per cent of alkali-labile sulfur in *zein*, are in excellent agreement with Osborne's data. Baernstein's value of 0.99 per cent for the total sulfur in *gliadin* is in good agreement with Osborne's early value, but Bailey (11) reports a higher result, 1.19 per cent. Blumenthal and Clarke's result of 0.09 per cent for the alkali-labile sulfur of *casein* is far lower than Osborne's value of 0.76, which is in better agreement with Baernstein's value of 0.80 per cent (10) and Kassel and Brand's value of 0.78 per cent (73). The analyses for sulfur of rabbit *myosin* are due to Bailey (11, 12). Osborne's early value for the sulfur in *edestin*, 0.88 per cent, is slightly lower than the values 0.93 reported by Bailey (11), 0.97 reported by Kassel and Brand (73), and 0.99 reported by Baernstein (10).

Cystine.—The earlier literature, including the work of Folin and Sullivan and their associates, is summarized in table 19 of reference 28. Since then Sullivan has modified his procedure, and Vickery and White (150) have considered the cystine yielded by proteins, as have Baernstein (8) and Bailey (11). The higher value given for *egg albumin* is from Calvery (22). The cystine content in *insulin* is from Miller and du Vigneaud (95). The value 0.91 per cent for *zein* is from Vickery and White (150) and is intermediate between the earlier values quoted (28). Vickery and White's average value for *gliadin* is lower than most values in the literature, but 2.5 per cent is in good agreement with the previous values, although the value is high and is higher than the value 2.40 per cent, which has also been reported (11, 28, 150). The cystine determination on horse *hemoglobin* is by Vickery and White (150). Hewitt reports a cystine content of 5.80 per cent for *serum albumin* (64), a value intermediate between Sullivan's early value of 5.71 per cent (134) and Folin and Marenzi's value of 6.02 per cent (53). The early literature on the cystine content of *casein* is consistent with a large number of values falling between 0.22 and 0.30 per cent (28). Vickery and White's new determinations fall within the same limits, 0.20 and 0.24 per cent. There thus appears to be a discrepancy between the sulfide sulfur and the cystine analyses. This determination renders difficult the estimation of the molecular weight in the neighborhood of 33,000 as demanded by osmotic pressure in urea. Vickery and White (150) have made a large number of analyses on the cystine in *edestin* and their results, which are not far from those of Sullivan and of Folin and Marenzi (28), give an average of 1.25 per cent. Baernstein's values range from 1.14 to 1.36 per cent.

Methionine.—The methods for determining the per cent of methionine in proteins and the values obtained have been considered by Baernstein (9, 10,), Blumenthal and Clarke (17), Bailey (11), and Kassel and Brand (73), and have been reviewed by Toennis (144). In the case of *zein* Baernstein (9) reports values varying from 2.25 to 2.58 per cent, depending upon the source of the preparation and the method of analysis. His value for *gliadin* is far higher than Bailey's. Baernstein (9) reports a value of 3.53 per cent for a *casein* preparation the origin of which he does not give. One supplied by R. A. Gortner had a methionine content of 3.36 per cent and one by

D. B. Jones a methionine content of 3.25 per cent, whereas a Harris preparation, studied by the volatile iodide method by Baernstein (9), had a value of 3.21 per cent and by Kassel and Brand (73) a value of 3.2 per cent. The methionine values for *edestin*, as reported by Baernstein, Bailey, and Kassel and Brand, all run between 2.3 and 2.4 per cent.

Tryptophane and tyrosine.—The earlier literature, including the work of Folin and his collaborators and of Linderstrøm-Lang on casein, has previously been summarized (table 19 of reference 28). Not included in this summary are studies on *egg albumin* by Calvery (22), whose estimate for tryptophane is 1.28 per cent and for tyrosine 4.21 per cent, and on *insulin* by du Vigneaud (151).

On the basis of the tryptophane estimates of a number of investigators the molecular weight of *zein* would have to be many times that observed both by osmotic pressure and ultracentrifugal measurements. No more than 0.2 per cent tryptophane has ever been reported for zein, and a large number of investigators report that zein contains no tryptophane.

The fractions of *gliadin* separated by Haugaard and Johnson (62), although constant in their histidine content, which was close to that reported by Van Slyke (145), vary appreciably in the percentage of certain other of the residues studied. Their least soluble fraction, IV, had a tryptophane content of 1.08 per cent, in excellent agreement with a number of previous determinations (28, page 861). The tryptophane contents of their other fractions were greater than of their fraction IV, and greater the greater the solubility and the greater the acid-insoluble humin nitrogen, reaching a value of 1.82 per cent, or four residues in a 41,000 molecular weight fraction. Conversely, the tyrosine content of their least soluble fraction was highest, 3.02 per cent, and close to that given by Looney (87), 3.04 per cent.

The tryptophane content of *casein* has repeatedly been studied; most results fall between 1.4 and 1.6 per cent, the higher value being due to Linderstrøm-Lang (84). The values for tyrosine reported have been too discrepant to consider. The tryptophane results reported for *edestin* generally range from 1.45 to 1.56 per cent and for tyrosine from 4.53 to 4.58 per cent (28).

The tryptophane content reported by early workers for *serum albumin* was always high, perhaps because of the difficulties in purification, since serum albumin can be separated into a number of crystalline fractions of different solubility (131), different dipole moment (49), different carbohydrate content (64, 131), and different tryptophane content. Fürth and Lieben (56) reported 1.3 per cent, and Hunter and Borsook (69) 1.79 per cent of tryptophane recovered from serum albumin. Folin and Marenzi (53) studied a preparation of serum albumin that had been recrystallized from five to seven times in this laboratory and found 0.52 and 0.53 per cent of tryptophane in the fractions studied. More recently Hewitt (64) has studied a serum albumin recrystallized twelve times and obtained a value just half of that reported by Folin and Marenzi (53), namely, 0.26 per cent.

Although tryptophane appears to be a variable in the various fractions of serum albumin, the tyrosine content would appear to be invariant. Hunter and Borsook (69) reported 4.63 per cent, Folin and Marenzi (53) 4.66 and 4.67 per cent, and Hewitt's crystalbumin preparation yielded 4.79 per cent. Assuming a molecular weight of 73,000 and eighteen tyrosine residues leads to a value of 4.47 per cent tyrosine per mole, which is rather lower than that observed. On the basis of eighteen tyrosine residues, these results thus suggest a molecular weight of between 68,000 and 70,000 rather than 73,000.

Osmotic pressure.—Sörensen estimated the molecular weight of *egg albumin* as 34,000 (130); this value was confirmed by Burk and Greenberg (20) in aqueous solution and by them and Huang and Wu (66) in isoelectric urea solutions. Marrack and Hewitt (89) estimated the molecular weight to be 43,000 at 37°C., but Nichols (105) attributes this higher value to the presence of aggregating material.

Burk (19) has estimated the molecular weight of *zein* as 39,000 in alcohol–water mixtures and 37,000 in urea. I am indebted to Dr. Burk for recalculating the former value, correcting for deviations from the ideal solution law and for the Donnan effect; the value so obtained was 39,000. Burk has also estimated the molecular weight of *gliadin* as being 41,000 in alcohol–water mixtures and 44,200 in urea (19). Adair (1, 2) estimates the molecular weight of *hemoglobin* to be 67,000 in aqueous solution, but Burk and Greenberg (20) and Wu and Yang (160) (for hemoglobin of the horse but not of the cow) found a value approximately half of this, namely, 34,300, in urea solution. Adair and Robinson's (4) estimate for the molecular weight of serum albumin is 72,000, that of Burk (18) 74,600, that of Roche and Marquet (119) 69,000; that of Sörensen, as corrected by Burk (18) and by Adair and Robinson (4) for measurements in 5 per cent ammonium sulfate solution, is 76,300. Burk and Greenberg (20) give the molecular weight of *casein* in urea solution as 33,600 and of *edestin* in this solvent as 49,000 (20). The molecular weight of *myosin* is of the order of one million (153). According to Weber, who has also studied the molecular weight of myosin in urea solution (154) with the method of osmotic pressure, myosin is not so large in this solvent and has a molecular weight of close to 100,000.

Ultracentrifuge.—The new values for *egg albumin* from Upsala are 43,800 with the sedimentation and 40,500 with the equilibrium ultracentrifuge. We are indebted to Professor Svedberg for informing us of the unpublished results of Pedersen (116; also (82). Williams and Watson (157) have studied the molecular weight of egg albumin in urea solution and report a molecular weight in this solvent of approximately 21,000. The sedimentation value for *insulin* is 40,900 and the equilibrium centrifuge value 35,100 (127). The sedimentation value for *zein* is 35,000 (152), for *gliadin* 26,000 (6), and for the *carboxyhemoglobin* of the horse 69,000. The estimate by means of the equilibrium ultracentrifuge for the latter protein is 68,000. The sedimentation value for horse *serum albumin* is 70,200 and the value from the equilibrium centrifuge is 66,900 (102, 116). Svedberg, Carpenter, and Carpenter (136) find that within the limits of experimental error the acid-alcohol soluble portion of *casein* was homogeneous with regard to molecular weight and that therefore it probably was a pure chemical individual. The molecular weight was found to be 375,000 ± 11,000 (137), though Hammarsten casein as a whole was found to consist of protein molecules of different molecular weight. Svedberg, Carpenter, and Carpenter later examined casein prepared by the method of Van Slyke and Baker (146) and found that the bulk of the crude casein prepared by this method had a molecular weight between 75,000 and 100,000 (137, page 710). The new value from sedimentation and diffusion for the molecular weight of *edestin* is 309,000, or approximately six times the weight reported from the osmotic pressure in urea.

true molecular weights of very great accuracy.[5] This is especially true of the beautiful new methods being developed by Bergmann (14).

[5] In the case of smaller molecules physical-chemical methods, such as boiling points and freezing points, yield the number of times the empirical formula weight must be multiplied to give the true molecular weight.

Analytical methods which enable a distinction to be made between —SS— and —SH groups in the intact protein have also been developed in recent years (59, 65, 80, 96), and it would appear that reagents such as urea and guanidine, which bring about changes in molecular weights of certain proteins, in many cases also influence these configurations. Here, too, the simultaneous use of analytical and physicochemical methods may well lead to significant advances in our understanding of the structure of the protein molecule.

Osmotic pressure methods and the average molecular weight

Measurements of osmotic pressure yield the average molecular weight, provided adequate correction is made (*1*) for the unequal distribution of ions across the membrane due to the Donnan equilibrium and (*2*) for the osmotic coefficients of the protein in the solution. In practice the Donnan effect is largely eliminated by carrying out the osmotic pressure measurements not far from the isoelectric point in salt solutions, generally in buffer solutions. But osmotic coefficients deviate appreciably from unity at finite protein concentration, and molecular weights estimated from osmotic pressure measurements therefore involve extrapolation from a series of measurements in which protein concentration is varied. If pH is also varied, so that the protein has a net charge, the osmotic pressure increases with the valence of the protein ion, and estimation of the molecular weight under these circumstances is still more complicated.

Although Hüfner and Gansser (67) attempted to determine the molecular weight of hemoglobin by means of osmotic pressure measurements, it remained for Sörensen (130) adequately to consider the correction necessary for the interpretations of such measurements if they were to yield molecular weights. His study of egg albumin, published just twenty-one years ago, marked a turning point in protein chemistry and suggested that the molecular weight of this protein was in the neighborhood of 34,000. Sörensen has since studied other proteins, and at least two other investigators, Adair (1, 2) and Burk (18, 19, 20), have made large numbers of adequate studies of molecular weights of proteins under a variety of conditions by the osmotic pressure method. The disadvantage of the osmotic pressure method is that at best it can yield only an average molecular weight; its advantage is that it is a thermodynamic measurement, dependent only on the number of particles in solution and not on their shape.[6]

[6] As we shall subsequently see, the activity coefficient, and therefore the osmotic coefficient due to the interaction of protein and electrolyte, are presumably functions of the valence of the latter and of the size and shape and electric moments of the former. The form of the extrapolation for the determination of osmotic pressure may thus also reflect the shape of the protein molecule.

In the course of osmotic pressure measurements Burk and Greenberg (20) noted in 1929 that the molecular weight of horse hemoglobin in concentrated urea solution, instead of being in the neighborhood of 66,680, as in aqueous solution, was half this value, or in the neighborhood of 33,340. This observation, confirmed by Wu (160) for the hemoglobin of the horse but not of the cow, and by Steinhardt (133) in studies at Upsala by means of the ultracentrifuge, has since been extended to a variety of other proteins, many of which are broken down into molecules of smaller size in the presence of this solvent (18, 19, 66, 154, 157). A series of other solvents, including other amides, but also guanidine, certain amino acids, and proteins (91, 113, 115), also possess this property and open a new chapter in protein chemistry. Whether or not these changes in molecular weight are to be considered as "dissociations" or as destruction of "aggregation," and whether or not they are reversible and therefore reveal the state of the native protein, they suggest that certain of the forces binding vast protein molecules together are different from others, and thus they add greatly to our knowledge regarding the structure of the proteins.

Sedimentation methods and the molecular weight

If the molecular weights of proteins were as large as analytical and osmotic pressure measurements suggested, it followed that they would be sedimented from solution by sufficiently great centrifugal forces. It remained for Svedberg (140) in 1926 to develop in the ultracentrifuge a tool adequate to the task. This most important tool, the various improvements it has undergone, and many of the results that have been obtained with it, as well as the equations employed in the calculation of molecular weights from ultracentrifugal measurements, have been reported in two previous contributions to this Journal by Svedberg (135).

The great strides made possible by this development depend in no small part upon the optical systems devised, which render it possible to distinguish more than one boundary and therefore more than one sedimentation rate, and thus to determine the molecular weight of more than one component in solutions that are not monodisperse, that is, that consist of proteins of more than one size. Thus far the method has been used largely to prove that the protein preparations under investigation consisted of molecules uniform with respect to size and to determine the molecular weights of the various proteins in certain natural fluids, such as milk (114) and the blood of certain invertebrates (42, 139). The fact that not only solvents, such as urea and arginine, but also certain proteins may influence each other's sedimentation rate (91) opens a new vista for the determination of the inter- and intra-molecular forces involved and also the possibility of apprehending more regarding the state of proteins in nature.

One cannot overemphasize the great advances that have thus far been achieved and that can be expected in the future from ultracentrifugal studies. The interpretation of the experimentally observed sedimentations involves a theory to which contributions are still being made (82, 116). The equation relating the molecular weight to the sedimentation constant includes a term for the frictional resistance experienced by the molecule. This term, which also appears in the Sutherland–Einstein derivation of the translational diffusion coefficient, is a function of both the size and the shape of the molecule. The combination of the sedimentation constant and diffusion coefficient permits the elimination of the shape factor from molecular weight calculations.[7] Sedimentation velocity measurements alone would give different molecular weights for molecules of the same composition and weight but different shapes. Conversely, molecules of totally different composition, but of the same size, density, and shape, would appear to be monodisperse in the ultracentrifuge if their isoelectric points were not too different. The latter statement has come to have more than theoretical significance, because so many proteins have molecular weights ranging from 34,000 to 42,000 and isoelectric points ranging from 4.7 to 5.7. Moreover, molecules of different sizes and shapes might conceivably have the same sedimentation constant. To deduce identity in the chemical nature of a protein from identity in its sedimentation velocity is a danger clearly envisaged by all those familiar with the methods.

Diffusion methods and the size and shape of the molecule

The slow diffusion of the proteins was one of the early observations which led to their classification as colloids. Determined with less elaborate apparatus than that demanded for ultracentrifugal measurements, the diffusion coefficient has also been used in estimating molecular weights (82, 106).

In one simple form of diffusion apparatus the concentration gradient is within the interstices of a sintered-glass filter, the pores of which must be large in comparison with the diameter of the protein molecule, but small enough to prevent mixing by convection (106). The diffusion coefficient

[7] Polson (118), using a combination of the viscosity equations of Kuhn and Arrhenius, has attempted, by inserting the asymmetries deduced from viscosity measurements in the equation of Herzog, Illig, and Kudar (63) for the diffusion ratio between a spherical and an elongated ellipsoidal particle, to determine the molecular weights of proteins. The results are of the right order of magnitude, and differ by a constant from the best results obtained by the sedimentation method. It would seem, however, that these were known with far greater certainty than the proper form of the viscosity equation. The influence of shape is also eliminated in measurements with the equilibrium centrifuge.

may be calculated from the changes with time of the concentrations on the two sides of the disc, provided the latter has been calibrated with a substance of known diffusion coefficient. If the analysis is specific for a given protein, this method may be employed in the presence of other proteins; if non-specific, measurements upon mixtures of proteins are difficult of interpretation, since they yield diffusion coefficients which are not averages but are weighted in favor of the smaller and more rapidly diffusing protein components.

The optical methods perfected at Upsala for the study of concentration gradients have been applied to diffusion measurements (82) and have greatly improved the classical method of free diffusion. The disturbances occasioned by sampling are avoided, and a more complete analysis of the diffusion is possible. As in the ultracentrifuge, the data reveal whether or not more than one component is present. The study of mixtures presents many difficulties, however, unless the optical method be made specific for a particular molecular species. This method, in contrast to that involving sintered-glass filters, yields not relative but absolute values of the diffusion coefficient.

The diffusion coefficient is a function of the size of the molecule. For the case of spherical molecules the diffusion coefficient is, according to theory, inversely proportional to the radius. A small error in the diffusion coefficient is thus greatly magnified in the estimation of the molecular weight.

The diffusion coefficient is also a function of the shape of the molecule. An ellipsoid encounters more frictional resistance than a sphere of the same volume, and its diffusion coefficient is accordingly smaller. All proteins have in fact smaller diffusion coefficients, D, than those calculated, D_0, from their molecular volumes on the basis of spherical shape. The ratio D_0/D, equal to the Svedberg dissymmetry constant (135), may be interpreted in terms of either departure from spherical shape, or hydration, or both. If the hydration be assumed zero, the equations of Herzog, Illig, and Kudar (63) or of Perrin (117) may be used to calculate from D_0/D the ratio of major to minor axes (s/d) of an ellipsoidal model for the molecule. Such calculations have been reported by Polson (118) and by Neurath (104).

Although the Svedberg dissymmetry constant is often only a little greater than unity, when the values of D and D_0 are substituted in the Perrin equation the asymmetries suggested are very large. Thus, according to this treatment, the dissymmetry constant of 1.18 for egg albumin yields a ratio of major to minor axis of 4:1; thyroglobulin with a dissymmetry constant of 1.41 has, on the same basis, a ratio of major to minor axis of 8:1.

Dielectric constant dispersion methods and the size and shape of the molecule

Proteins, like other molecules, may be expected to assume a random distribution in the body of a dilute solution[8] as a result of thermal motion. All proteins that have thus far been investigated are, however, oriented in an electric field. By virtue of their structure as dipolar ions they follow an alternating potential, provided its frequency is not too great. At low frequencies proteins thus increase the dielectric constants of solutions. At high frequencies, where they are unable to follow the alternations, the

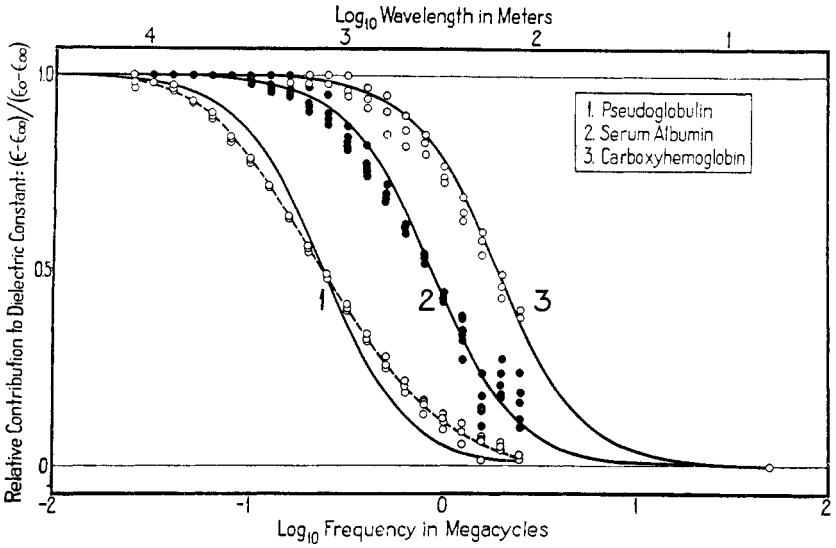

FIG. 1. Dispersion curves of blood proteins (horse)

dielectric constant of a solution is decreased, for the vast protein molecules displace solvent dipoles which could have oriented in the electric field (107, 109).

Plotting the dielectric constant of a solution containing protein against the frequency of the alternating potential yields a dielectric constant dispersion curve, which is characteristic of the protein. The theory of such curves has been developed by Debye (39). The methods that have been employed and the dielectric constants that have been observed have recently been reviewed elsewhere (31, 41, 109, 157). The first dispersion

[8] At the surface of a solution the forces must be considered different, and orientation and layering of the protein result, generally with compression and distortion of the native molecule. The problems associated with surface phenomena are discussed in the previous paper in this Symposium and are therefore not considered here (34). In sufficiently concentrated solutions, or in solutions of low dielectric constant, orienting forces may also be expected to interfere with a random distribution.

curve for a protein was that of Wyman on zein (161). The relation of such curves to the molecular weight has been considered by Wyman, Williams, Arrhenius, Oncley, and Ferry (6, 7, 49, 107, 156, 157, 161, 162).

The extent to which the dielectric constant of a solution is raised by a protein solute depends upon the electrical asymmetry of the molecules; it is a function of their dipole moment, which is determined by the distribution of the electrically charged groups on their surfaces, and has been considered in connection with the number and distribution of the electrically charged groups of proteins (31). The shape of the dielectric constant dispersion curve and the frequencies at which it occurs are, however, functions of the size and shape of the molecule and are therefore considered here.

Studies upon three proteins are graphically represented in figure 1, the ordinate of which is so defined that when the protein molecules are unable to follow the alternating field and contribute nothing to the dielectric constant, its value is zero; when they rotate in phase with the field and make a maximum contribution to the dielectric constant, its value is unity. The curve for hemoglobin of the horse (107) closely follows the form given by Debye's theory. From the critical frequency, ν_c, the relaxation time ($= 1/2\pi\,\nu_c$) is found to be 8.4×10^{-8} sec. For comparison, we may calculate the time of relaxation expected for a spherical molecule of molecular weight 66,670 and specific volume $v = 0.75$ rotating in a medium with the viscosity (η) of water as $\tau = 3\eta Mv/RT = 5.4 \times 10^{-8}$ sec. The discrepancy may again be interpreted in terms of either hydration or spatial asymmetry. If we assume no asymmetry and postulate an amount of hydration equal to that estimated by Adair and Adair (3) for hemoglobin crystals, and assume that the water of hydration rotates with the protein molecule, the calculated relaxation time is 7.6×10^{-8} sec., in fair agreement with that observed. The same amount of hydration would closely account for the observed diffusion, as an alternative interpretation to an elongated ellipsoidal shape with an axial ratio of 4. The latter shape is not consistent with the single relaxation time observed, unless the dipole moment be perpendicular to the long axis. On the other hand, calculations by Mehl and Ferry show that both D_0/D and the observed relaxation time would be given by a postulate of no hydration and the shape of a *flattened* ellipsoid of revolution with an axial ratio of 4. For such a body the relaxation times about the different axes are nearly equal, and even if the dipole were inclined to all axes, the dispersion would be indistinguishable from a simple Debye curve.[9]

[9] The hemoglobin of the pig has the same molecular weight as that of the horse, but the dielectric constant measurements of Arrhenius (7) yield a relaxation time about twice as great, and therefore exceeding much more (by a factor of three) the value calculated from the molecular weight.

The molecular weight of the crystalline serum albumin of the horse is very close to that of horse hemoglobin, being estimated by Svedberg as 67,000 on the basis of ultracentrifugal measurements (135), and by Adair (4) as 72,000 and by Burk (18) as 74,600 on the basis of osmotic pressure measurements. The difference between the relaxation times of these two proteins would therefore be scarcely perceptible on the basis of their molecular weights, and no sufficiently great differences in hydration to account for the observed dispersion curves can be reasonably postulated. The curves of both proteins appear to follow Debye's theory, but the relaxation time obtained for serum albumin, 19×10^{-8} sec., is over twice as great as that for hemoglobin or more than three times that calculated from its molecular weight. Moreover, the viscosities of serum albumin solutions are far greater than those of hemoglobin solutions at the same volume fraction (37, 45, 82). This demonstrates how relaxation times reflect not only the size but also the shape of proteins. Dielectric constant dispersion curves may therefore become of even more use than sedimentation constants in characterizing many proteins, since so many proteins appear to have the same molecular weights. The relaxation time of a protein may readily be employed in calculating its rotary diffusion constant.

Horse serum pseudoglobulin has a far greater effect upon the dielectric constant, reflecting greater electrical asymmetry than either the serum albumin or the hemoglobin of the horse. As judged by viscosity measurements (25, 26, 45), it is also a less spherical molecule. Its molecular weight in aqueous solution has been estimated to be 150,000 or 162,000 (91, 103). The dielectric constant dispersion curve of this protein differs in two respects from the others in figure 1. In the first place, it occurs at much smaller frequencies; in the second place, its shape does not conform to Debye's theory. Analyzable in terms of two or more relaxation times, this curve suggests either that more than one molecular species was present or that the asymmetry of the molecule is sufficient to reveal relaxation times characteristic of rotation about its different axes. Regardless of the interpretation of such a curve, however, it further illustrates the usefulness of dielectric constant studies for the characterization of proteins and for demonstrating differences and similarities between similar proteins from different sources.[10]

The relaxation times of those proteins which appear in figure 1 are summarized in table 3, together with some others. There are as yet no satisfactory experimental estimates of relaxation times of small molecules, such as water, amino acids, and peptides; the order of magnitude to be expected is, however, given by values calculated from the molecular volumes.

In the case of zein, beautifully studied by Williams and his coworkers

[10] For instance, in the comparison of blood and tissue proteins or blood and urinary proteins (55).

TABLE 3

Relaxation times calculated from molecular volumes, from dispersion of the dielectric constant, and from double refraction of flow

SUBSTANCE	MOLECULAR WEIGHT M	MOLECULAR VOLUME V	RELAXATION TIME	
			$\frac{3\eta V^*}{RT}$	$\frac{1}{2\pi\nu_c}$
Water	18	18	2×10^{-11}	
Glycine	75	57	6×10^{-11}	
Diglycine	132	93	10×10^{-11}	
Zein (in 72 per cent ethanol)§ (41)	39,000	29,000	6.9×10^{-8}	32×10^{-8}‡
			58×10^{-8}†	53×10^{-8}
			9.2×10^{-8}†	8.3×10^{-8}
Gliadin (in 60 per cent ethanol) (6)	42,000	31,000	10×10^{-8}	9×10^{-8}
Lactoglobulin (50, 124)	40,000	30,000	3.3×10^{-8}	7×10^{-8}
Egg albumin (108, 124)	34,000	25,500	2.8×10^{-8}	12×10^{-8}‡
			17×10^{-8}†	20×10^{-8}
			3.7×10^{-8}†	4.4×10^{-8}
Hemoglobin (horse) (43, 107)	66,700	50,000	5.4×10^{-8}	8.4×10^{-8}
Serum albumin (horse) (49)	73,000	54,000	5.8×10^{-8}	19×10^{-8}
Serum pseudoglobulin (horse) (49)	162,000	120,000	13×10^{-8}	66×10^{-8}‡
			17×10^{-8}†	33×10^{-8}
			79×10^{-8}†	132×10^{-8}
Tobacco mosaic virus (94)	50,000,000	38,000,000	4.1×10^{-5}	(1×10^{-2})¶
Myosin (cow) (94, 100)				(5×10^{-2})¶

* On the assumption of spherical shape, without correcting for hydration.

† On the assumption of elongated ellipsoidal shape, without correcting for hydration, using the axial ratio given by the ratio of the observed relaxation times.

‡ From the single average critical frequency which best describes the dispersion.

§ The earlier measurements of Wyman (161) in 70 per cent propanol give an average relaxation time of 6.2×10^{-8} sec., as against a calculated value of 10×10^{-8}.

¶ Estimated from double refraction of flow.

(41, 157), and of serum pseudoglobulin and egg albumin, studied by Oncley, Ferry, and Shack (49, 50, 108, 124) in this laboratory, the dispersion, although it can be roughly described by a single "average" critical frequency, is much better represented by two different frequencies which yield different relaxation times. If these be interpreted in terms of rotation of an ellipsoid about different axes, then the asymmetry may be evaluated from the ratio of the relaxation times, *without any arbitrary assumptions concerning hydration.*

The asymmetries, which are themselves independent of hydration, may then be used, with the assumption of zero hydration, to calculate the absolute magnitudes of the different relaxation times from the molecular volumes of the respective ellipsoids. These figures are included in table 3, and for zein and egg albumin are in good agreement with the experimental relaxation times, suggesting that for these proteins hydration has little influence on the rotary diffusion.

Double refraction of flow and the size and shape of the molecule

The phenomenon of double refraction of flow of protein solutions has been discussed so recently (94) that only its use in the study of the size and shape of the molecule need be mentioned here. When ordinary light is passed through a Nicol prism, it is plane polarized and can be extinguished by a second Nicol prism the plane of polarization of which is at right angles to that of the first. If a velocity gradient is produced in a solution of asymmetrical molecules placed between the prisms, light will again pass, as the result of double refraction resulting from the orientation of the molecules. This effect will be greater the more asymmetrical the molecules and the greater the velocity gradient to which they are subjected, that is to say, the larger the proportion of molecules oriented in the stream lines of the solution. From measurements of the double refraction the extent and the principal direction of orientation of the molecules may be readily determined.

The orienting forces in the solution may be considered as opposing disorienting rotatory Brownian motion. Calculations of resistance to rotation of ellipsoidal particles in a viscous medium have been made by Gans (57) and Perrin (117), and more roughly by Werner Kuhn (78). It follows from their calculations that the relaxation time of a very long ellipsoidal particle, rotating about its shorter axis, is approximately proportional to the cube of its length. On the basis of their relations and the studies of Edsall and von Muralt upon myosin (100) it has been estimated that this muscle globulin has a length in the neighborhood of 8500 Å. (94). This value may be compared with estimates of length calculated on the basis of the molecular volume and estimates of the asymmetry, s/d, for less asymmetric proteins. Tentative calculations of this kind yield lengths for egg

albumin of 120 Å., for horse serum albumin of 160 Å., for horse serum globulin of 290 Å., and for tobacco mosaic virus of 5200 Å.

Collagen (94), fibrinogen, ovoglobulin (21), thyroglobulin, tobacco mosaic, and certain other viruses, and the hemocyanins of *Helix* and *Limulus* (83) have also been shown to exhibit double refraction of flow. At higher velocity gradients it is to be expected that less asymmetrical and smaller molecules will also exhibit this phenomenon, and we are at present engaged in reconstructing our apparatus at Harvard University in order to estimate the size and shape of other proteins by this method.

The relaxation times of the asymmetrical molecules thus far studied, calculated as rotation about an axis perpendicular to the longest axis of the molecule from measurements of double refraction of flow by means of Boeder's treatment (94), are given in table 3. For myosin and tobacco mosaic virus they are respectively 1×10^{-1} to 2×10^{-2}. Assuming a molecular weight for this virus of 50,000,000 and calculating its relaxation time on the assumption of spherical symmetry would lead to a relaxation time of 3×10^{-5} or approximately one thousand times smaller than that observed (94).

In table 3 comparison is also made between the relaxation times calculated from dielectric constant dispersion curves. Those observed thus far range from 10^{-8} to 10^{-6} sec., and in every case in which the solvent was water were greater than relaxation times calculated from molecular weights on the basis of spherical symmetry. Extension of the dielectric constant method to far longer wave lengths and of the double refraction of flow method to far greater velocity gradients should render it possible ultimately to study the same protein by these two quite different methods—one electrical, one optical—and thus acquire more satisfactory knowledge regarding the forces involved in the orientation of protein molecules.

Viscosity and the shape of the molecule

According to Einstein's theory, the viscosity of a suspension or solution of uncharged, incompressible spheres is independent of their size, provided they occupy the same volume fraction of the solution and provided their size is large compared with that of solvent molecules. Experimentally it may be shown that when egg albumin and glycine occupy the same volume fraction, the viscosities of their solutions are somewhat similar (37). The radius of the protein calculated as a sphere is approximately 22 Å. and that of the amino acid is 2.8 A., but this difference in size has little influence on viscosity. In the amino acid series, however, we have shown that dilute solutions of the more asymmetrical molecules, such as β-alanine and ϵ-aminocarproic acid, are far more viscous, and the more so the greater the number of carbon atoms between the positively charged ammonium and the negatively charged carboxyl group (37).

Proteins differ widely from each other with respect to the viscosity of their solutions. Harriet Chick (25, 26) reported a large number of viscosity studies,[11] as did Loeb (85) and many subsequent investigators. In Miss Chick's studies serum albumin solutions were always more viscous than egg albumin, and pseudoglobulin and euglobulin still more viscous. The general magnitude of her results has been confirmed (27, 37, 45), and we have assumed therefore, in the light of our studies upon amino acids and peptides of known structure and of Staudinger's studies on other molecules (132), that this was the order in which the shapes of these molecules deviated from spherical symmetry. Deviation between calculated and observed relaxation times is also greater for serum albumin than for hemoglobin, and for pseudoglobulin than for either of these. This is the order of their viscosities.

Although there would appear to be no doubt from these various experimental studies that viscosity is a most sensitive index of asymmetry, a completely satisfactory theoretical equation relating these two properties remains to be developed. The problem has been repeatedly considered (60, 61, 68, 78, 132), and various extensions of the Einstein equation have been suggested and tentatively employed. Estimates of the relation of the two principal axes of the molecule must be adopted with caution, until they prove to be identical with those derived from diffusion measurements, dielectric constant measurements, and measurements of double refraction of flow.

All of these measurements are technically difficult, and the assumptions involved in their theoretical interpretation include such approximations as that molecules are rigid and incompressible, that there is no interaction between solute molecules, and that the solvent may be treated as a structureless continuum. That the degree of asymmetry for the series of proteins thus far studied by such different methods is so nearly of the same order nevertheless suggests that our knowledge begins to be satisfactory regarding not only the size but also the shape of the protein molecule.

III. NUMBER AND DISTRIBUTION OF THE ELECTRICALLY CHARGED GROUPS OF THE PROTEIN MOLECULE

Analytical methods and the number of reactive groups

Proteins as a class are extremely diverse, not only in their biological functions but also in their chemical behavior. Neither the diversity nor the general properties of the proteins can be adequately explained by differ-

[11] Viscosity, like osmotic pressure and a large number of other properties of proteins, passes through a minimum in the neighborhood of the isoelectric point. The present discussion is limited, for lack of space, to the consideration of proteins as dipolar ions. Studies at reactions far from the isoelectric point, where proteins are present as ions, are therefore not discussed here.

ences in the size or shape of their molecules. Thus hydrocarbon molecules of the same size and shape as proteins would presumably be completely soluble in non-polar solvents and relatively unreactive to electrolytes. It is precisely because proteins are possessed of varying numbers of polar groups, distributed on their vast surfaces, that they are polar molecules, insoluble in non-polar solvents but for the most part extremely soluble in water or electrolyte solutions.

There have been great advances in recent years in estimating the amino acid compositions of proteins. The relation between cystine, methionine, and the various forms of sulfur in the protein molecule have been considered in tables 1 and 2 in connection with estimates of the minimum molecular weights of proteins. Tryptophane and tyrosine have also proved invaluable in estimating minimal molecular weights, and these amino acids can be estimated with some accuracy by colorimetric procedures (52, 53). The latter amino acid also appears to be of significance for immunity and for the combination of base by proteins as alkaline reactions.

The analytical methods for the other amino acids of importance in combining acids and bases are also being constantly improved; those for the basic amino acids among others by Vickery (149), Block (16), and Bergmann (14, 15), and those for the dicarboxylic amino acids by Foreman (54), Dakin (36), Breese Jones (70), Calvery (22), and Chibnall (24). It is not within the province of the present communication to consider these methods, but it should be pointed out that every improvement in the estimation of the amino acid composition of proteins has rendered it more certain (32) that it is by virtue of the dissociated carboxyl, sulfhydryl, phenolic hydroxyl, and the positively charged imidazole, guanidine, and ammonium groups that proteins can exist as cations, as anions, or at their isoelectric points as dipolar ions. In large part the diversity in gross behavior between molecules of the same size and shape may be ascribed to differences in the number and in the distribution of these groups.

Electrophoretic mobility, electromotive force measurements, and the number of charged groups

Excepting at their isoelectric points proteins have a net charge and form salts with either acids or bases. According to the theory of dipolar ions the net charge arises at reactions alkaline to the isoelectric point from the loss of a proton from the basic amino acids, leaving the molecules negatively charged, or from the feeble acid dissociation of phenolic hydroxyl or sulfhydryl groups. The regions in which the various free groups of proteins and peptides dissociate and the present state of knowledge regarding amphoteric properties have recently been discussed elsewhere (23, 31). Except for proteins with alkaline isoelectric points, the negative charges of

electrically neutral proteins arise completely, as far as we know, from the dissociation of carboxyl groups.

At reactions acid to the isoelectric point carboxyl groups become undissociated, lose their negative charge, and leave the protein with a net positive charge upon amino, imidazole, and guanidine groups. The number of these is exactly equal, as far as present measurements reveal, to the acid-combining capacity of the proteins. Thus analytical and physicochemical methods combine to give us accurate information regarding the maximum number of positive charges which a protein can bear as a cation or a dipolar ion.

Electrometric titration curves of proteins reveal differences ascribable to the differences in the number and dissociation of the free groups. Despite the general similarity of such curves, those for most proteins (27, 31), even for closely related proteins such as the serum globulins (58), reveal definite differences in the number or strength of dissociating groups.

Such differences in amphoteric properties must lead to differences in electrophoretic mobility, as Sir William Hardy clearly envisaged in his classical studies at the turn of the century. The calculation of net charge from measurements of electrophoretic mobility is far more complicated than from electromotive force measurements because, whereas the latter can be carried out in almost salt-free solution, the boundary conditions in mobility experiments demand the presence of adequate concentrations of buffer. Although theoretical equations for mobility of simpler ions have been developed by Onsager (110), MacInnes (88), and Shedlovsky (125), their application to proteins is not yet complete and extrapolation to salt-free conditions not yet satisfactory (see 38, 129).

Despite this temporary theoretical limitation, electrophoretic mobility measurements are extremely sensitive indications of differences in the charged condition of different proteins. Thus Tiselius, in his recent studies with an apparatus much improved over those previously available, has shown differences in the mobility of the various serum globulin fractions. Moreover, the optical systems that he has developed render possible the following of more than one boundary and can thus be employed not only to prove whether proteins of one or more valence type are present but also, in the latter case, to effect their separation (86, 141).

Whereas the basic amino acids may be present to the same extent in the various serum globulin fractions, both electromotive force measurements (58) and measurements of electrophoretic mobility (142) reveal differences in dissociation, leading to differences in isoelectric point. For these proteins we thus have a method of distinguishing between molecules presumably of the same size (91, 101), although perhaps of somewhat different shape (45). Other examples might be cited. In the case of the

crystalline serum albumins, however, no differences in molecular weight for different fractions have been shown, nor, as judged by viscosity and dielectric constant dispersion curves, in the shape of the molecules. Further, no significant difference in titration curves has been observed since the earliest studies of this kind by d'Agostino and Quagliariello (5, 27, 31). Moreover, the most recent electrophoretic mobility studies upon serum albumins (142) also reveal no difference between different fractions, although it has been known since the classical studies of Sörensen that the solubilities in salt solutions of the various crystalline serum albumins are quite different. In these proteins present methods show no differences in size, shape, or net charge, although gross differences in solubility may none the less be observed.

Dielectric constant increment, dipole moment, and the distribution of electrically charged groups

Not only the number but also the arrangement in space of the charged groups determines behavior. Thus for a tetrapole, like cystine, which might be considered as made up of two dipoles of moments equal to glycine, that is, of 15×10^{-18} electrostatic units, the moment of the whole molecule might be equal to 30×10^{-18} electrostatic units or to zero, depending on whether the dipoles were parallel or antiparallel, or had an intermediate position (34, 75, 93). In the case of an isoelectric molecule like egg albumin with twenty-eight positively and twenty-eight negatively charged groups (31), or of hemoglobin with at least seventy-five positively and negatively charged groups (32), an extremely large number of arrangements of these groups is conceivable, and each arrangement might well lead to a molecule of quite different properties.

The estimation of the dipole moments of amino acids, peptides, and proteins from measurements of dielectric constant have been considered elsewhere (30, 31, 109, 162). Studies upon egg albumin have yielded a moment of approximately 180×10^{-18} electrostatic units. The dielectric constant increments of the various egg albumin preparations that have thus far been studied by Oncley are in good agreement with each other and with the previously published studies of Shutt (126). Similarly the measurements of Oncley (107) upon horse hemoglobin are all consistent with each other and with those of Errera (43). The studies made upon horse serum albumin by Ferry and Oncley (49) reveal differences in moment corresponding to differences in solubility of the various crystalline fractions thus far investigated, ranging from 270 to 510×10^{-18} electrostatic units for fractions of low and high solubilities, respectively, in ammonium sulfate solutions. For these proteins, therefore, differences in behavior which could be ascribed to neither size, shape, nor charge can thus be ascribed to the distribution of the charge on the surface of the vast protein molecule.

Solubility and other measurements of activity coefficients and the electric moments of proteins

The earliest methods of separating proteins from each other depended upon differences in their solubilities in different solvents. Upon these differences the purification of proteins often depends, as well as the crystallization of many of them, and a classification has developed and been adopted which defines proteins as albumins, globulin, prolamines, and glutelins, depending on whether they are soluble in water, salt solutions, alcohol–water mixtures, or only in acid and alkaline solutions.

Solubility depends not only upon the forces between solvent and solute but also upon those between solute molecules in the solid state, that is, upon crystal lattice energies. The crystal lattice energies of proteins and even of amino acids and peptides are difficult to measure, since these substances cannot be sublimed or even fused. Change of solubility with change of solvent should, however, be independent of crystal lattice energy and yield activity coefficients, provided the protein is present in the same solid state in equilibrium with the various solvents investigated.

Where it has been possible to measure activity coefficients of amino acids, peptides, or proteins by means of freezing point (121), vapor pressure (128), or electromotive force measurements (71), the results, as a first approximation, have been satisfactorily in agreement with those derived from solubility measurements, whether the nitrogenous component was present as saturating body and the interaction investigated with solvent molecules (35, 47, 48), or the nitrogenous components were added to insoluble salts (46). Such studies are now too numerous to be considered here,[13] but a few of the pertinent generalizations that have been deduced may be cited. (*1*) In regions of sufficiently low dielectric constant, interaction between ions and dipolar ions depends largely upon the electric moment and also to a small extent upon the size of the molecule. (*2*) Under these circumstances the solubility of elongated amino acids and peptides has been investigated, and the ratio of the logarithm of activity coefficient to ionic concentration has been shown to increase almost proportionately with the electric moment. (*3*) The volume of the molecule diminishes the effect due to its electric moment the more the larger the volume of the solute and the greater the dielectric constant of the solution. Thus the greater the molecular weight in comparison with the moment of the molecule the less will mutual solvent action occur, and the more will mutual precipitating action occur. The "salting-out" of proteins is the

[13] Influence of dipolar ions on the solubility of each other (34) is considered in more detail in the Symposium on Intermolecular Action, held by the Division of Physical and Inorganic Chemistry of the American Chemical Society at Providence, Rhode Island, December, 1938.

best known example of this phenomenon. (4) These interactions would appear both from experience and from theoretical calculations (74, 75) to depend not only upon the sizes but also upon the shapes of the molecules and to depend not only on the dipole moment of the molecule but also upon the number of the charged groups, even where these are symmetrically arranged. Thus the dipole moment does not account completely for molecular interactions between dipolar ions (48), but higher electric moments must presumably also be considered.

IV. DISCUSSION

The adequate development of a theory of protein chemistry demands a formulation of the parameters in terms of which behavior may be explained. In the foregoing we have attempted to discuss certain of these parameters. Certainly the size of the protein molecule must be considered. Knowledge of the size of a protein is not, however, sufficient for its characterization. Egg albumin, pepsin, lactoglobulin, and zein all have nearly the same molecular weight. Yet the first is classified as a water-soluble albumin, the second and third as globulins (insoluble in water but more soluble in salt solutions), and zein as a prolamine insoluble in water but soluble in alcohol–water mixtures.

The shape of a protein is not sufficient for its characterization. Fibrinogen, among the blood proteins, is rod-shaped and so is tobacco mosaic virus. Not all blood proteins nor all viruses are elongated molecules, however. Among the former, hemoglobin is nearly spherical, and there is at least one virus, the bushy stunt virus of tomato (13, 92), which is reported to be nearly spherical.

The isoelectric point of a protein is not sufficient for its characterization. Egg and serum albumins, casein, and gelatin have nearly the same isoelectric point. Although the first two have approximately equal numbers of free acid and basic groups, casein has many more acid than basic, and gelatin many more basic than acid groups.

The dielectric constant increment of a protein is not sufficient for its characterization. Carboxyhemoglobin and the most soluble crystalline fraction of horse serum albumin increase the dielectric constant of water by nearly the same amount, but have quite different relaxation times.

The relaxation time of a protein is not sufficient for its characterization. The different fractions of crystalline horse serum albumin have the same relaxation time, though the dielectric increment of the more soluble is roughly three times that of the less soluble. Moreover, the relaxation time—and also the viscosity—of carboxyhemoglobin is smaller than that of the serum albumins, although all have nearly the same molecular weight.

Whereas no one of these properties will suffice for characterization, these

and other measurements, taken together, yield quantitative information in terms of which we may develop still further both the theory and the practice of protein chemistry.

REFERENCES

(1) ADAIR: Proc. Cambridge Phil. Soc. **1**, 75 (1924); Skand. Arch. Physiol. **49**, 76 (1926).
(2) ADAIR: Proc. Roy. Soc. (London) **A120**, 573 (1928); J. Am. Chem. Soc. **51**, 696 (1929).
(3) ADAIR AND ADAIR: Proc. Roy. Soc. (London) **B120**, 422 (1936).
(4) ADAIR AND ROBINSON: Biochem. J. **24**, 1864 (1930).
(5) D'AGOSTINO AND QUAGLIARIELLO: Nernst Festschrift, Halle, p. 27 (1912).
(6) ARRHENIUS: J. Chem. Phys. **5**, 63 (1937).
(7) ARRHENIUS: Physik. Z. **39**, 559 (1938).
(8) BAERNSTEIN: J. Biol. Chem. **89**, 125 (1930).
(9) BAERNSTEIN: J. Biol. Chem. **97**, 663 (1932); **106**, 451 (1934); **115**, 25 (1936).
(10) BAERNSTEIN: J. Biol. Chem. **97**, 669 (1932); **115**, 33 (1936).
(11) BAILEY: Biochem. J. **31**, 1396 (1937).
(12) BAILEY: Biochem. J. **31**, 1406 (1937).
(13) BAWDEN AND PIRIE: Nature **141**, 513 (1938); British J. Exptl. Path. **19**, 251 (1938).
(14) BERGMANN: Chem. Rev. **22**, 423 (1938).
(15) BERGMANN AND NIEMAN: J. Biol. Chem. **122**, 577 (1938).
(16) BLOCK: The Determination of Amino Acids, Minneapolis (1938).
(17) BLUMENTHAL AND CLARKE: J. Biol. Chem. **110**, 343 (1935).
(18) BURK: J. Biol. Chem. **98**, 353 (1932).
(19) BURK: J. Biol. Chem. **120**, 63 (1937); **121**, 373 (1937); **124**, 49 (1938).
(20) BURK AND GREENBERG: J. Biol. Chem. **87**, 197 (1930).
(21) BOEHM AND SIGNER: Helv. Chim. Acta **14**, 1370 (1931); Klin. Wochschr. **11**, 599 (1932).
(22) CALVERY: J. Biol. Chem. **94**, 613 (1931-32).
(23) CANNAN: Cold Spring Harbor Symposia Quant. Biol. **6**, 1 (1938).
(24) CHIBNALL: Personal communication.
(25) CHICK: Biochem. J. **8**, 261 (1914).
(26) CHICK AND LUBRZYNSKA: Biochem. J. **8**, 59 (1914).
(27) COHN: Physiol. Rev. **5**, 349 (1925).
(28) COHN: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **33**, 781 (1933).
(29) COHN: Naturwissenschaften **20**, 663 (1932).
(30) COHN: Ann. Rev. Biochem. **4**, 93 (1935).
(31) COHN: Cold Spring Harbor Symposia Quant. Biol. **6**, 8 (1938).
(32) COHN, GREEN, AND BLANCHARD: J. Am. Chem. Soc. **59**, 509 (1937).
(33) COHN, HENDRY, AND PRENTISS: J. Biol. Chem. **63**, 721 (1925).
(34) COHN, MCMEEKIN, FERRY, AND BLANCHARD: J. Phys. Chem. **43**, 169 (1939).
(35) COHN, MCMEEKIN, GREENSTEIN, AND WEARE: J. Am. Chem. Soc. **58**, 2365 (1936).
(36) DAKIN: Biochem. J. **12**, 290 (1918).
(37) DANIEL AND COHN: J. Am. Chem. Soc. **58**, 415 (1936).
(38) DAVIS AND COHN: Unpublished work.
(39) DEBYE: Polar Molecules. The Chemical Catalog Company, Inc., New York (1929).

(40) EHRLICH: Z. Ver. deutsch. Zucker-Ind. **53**, 809 (1903).
(41) ELLIOTT AND WILLIAMS: J. Am. Chem. Soc. **61**, 718 (1939).
(42) ERIKSSON-QUENSEL AND SVEDBERG: Biol. Bull. **71**, 498 (1936).
(43) ERRERA: J. chim. phys. **29**, 577 (1932).
(44) EYRING: Phys. Rev. [2] **39**, 746 (1932).
(45) FAHEY AND GREEN: J. Am. Chem. Soc. **60**, 3039 (1938).
(46) FAILEY: J. Am. Chem. Soc. **54**, 576, 2367 (1932): **55**, 4374 (1933).
(47) FERRY, COHN, AND NEWMAN: J. Am. Chem. Soc. **58**, 2370 (1936).
(48) FERRY, COHN, AND NEWMAN: J. Am. Chem. Soc. **60**, 1480 (1938).
(49) FERRY AND ONCLEY: J. Am. Chem. Soc. **60**, 1123 (1938).
(50) FERRY AND ONCLEY: Unpublished work.
(51) FISCHER: Z. physiol. Chem. **33**, 151 (1901); Ber. **35**, 2660 (1902).
(52) FOLIN AND CIOCALTEU: J. Biol. Chem. **73**, 627 (1927).
(53) FOLIN AND MARENZI: J. Biol. Chem. **83**, 89 (1929).
(54) FOREMAN: Biochem. J. **13**, 378 (1919).
(55) FRIEND, FERRY, AND ONCLEY: J. Biol. Chem. **125**, Proc. 39 (1938).
(56) FÜRTH AND LIEBEN: Biochem. Z. **109**, 124 (1920).
(57) GANS: Ann. Physik **86**, 628 (1928).
(58) GREEN: J. Am. Chem. Soc. **60**, 1108 (1938).
(59) GREENSTEIN: J. Biol. Chem. **125**, 501 (1938).
(60) GUTH: Kolloid-Z. **74**, 147 (1936).
(61) GUTH AND MARK: Ergeb. exakt. Naturw. **12**, 115 (1933).
(62) HAUGAARD AND JOHNSON: Compt. rend. trav. lab. Carlsberg **18²**, 96 (1930).
(63) HERZOG, ILLIG, AND KUDAR: Z. physik. Chem. **A167**, 329 (1934).
(64) HEWITT: Biochem. J. **30**, 2229 (1936); **31**, 360 (1937).
(65) HOPKINS: Biochem. J. **15**, 286 (1921); Nature **126**, 328, 383 (1930).
(66) HUANG AND WU: Chinese J. Physiol. **4**, 221 (1930).
(67) HÜFNER AND GANSSER: Arch. Physiol. p. 209 (1907).
(68) HUGGINS: J. Phys. Chem. **42**, 911 (1938).
(69) HUNTER AND BORSOOK: J. Biol. Chem. **57**, 507 (1923).
(70) JONES AND WILSON: Cereal Chem. **5**, 473 (1928).
(71) JOSEPH: J. Biol. Chem. **111**, 479, 489 (1935).
(72) KASSEL AND BRAND: Proc. Soc. Exptl. Biol. Med. **35**, 444 (1936).
(73) KASSEL AND BRAND: J. Biol. Chem. **125**, 435 (1938).
(74) KIRKWOOD: J. Chem. Phys. **2**, 351 (1934).
(75) KIRKWOOD: Chem. Rev. **24**, 233 (1939).
(76) KOSSEL: Z. physiol. Chem. **22**, 176 (1896–7).
(77) KRÜGER: Arch. ges. Physiol. (Pflügers) **43**, 244 (1888).
(78) KUHN: Z. physik. Chem. **A161**, 1 (1932).
(79) KUHN: Kolloid.-Z. **68**, 2 (1934); Z. physik. Chem. **A175**, 1 (1935).
(80) KUHN AND DESNOUELLE: Z. physiol. Chem. **251**, 14 (1938).
(81) KÜLZ: Z. Biol. **27**, 415 (1890).
(82) LAMM AND POLSON: Biochem. J. **30**, 528 (1936).
(83) LAUFFER AND STANLEY: J. Biol. Chem. **123**, 507 (1938).
(84) LINDERSTRØM-LANG: Compt. rend. trav. lab. Carlsberg **17**, No. 9 (1929).
(85) LOEB: Proteins and the Theory of Colloidal Behavior, 2nd edition. McGraw-Hill Book Company, New York (1924).
(86) LONGSWORTH AND MACINNES: Chem. Rev. **24**, 271 (1939).
(87) LOONEY: J. Biol. Chem. **69**, 519 (1926).
(88) MACINNES, SHEDLOVSKY, AND LONGSWORTH: Chem. Rev. **13**, 29 (1933).

(89) MARRACK AND HEWITT: Biochem. J. **23**, 1079 (1929).
(90) McCOY, MEYER, AND ROSE: J. Biol. Chem. **112**, 283 (1935-36).
(91) McFARLANE: Biochem. J. **29**, 407, 660, 1209 (1935).
(92) McFARLANE AND KEKWICK: Biochem. J. **32**, 1607 (1938).
(93) McMEEKIN, COHN, AND BLANCHARD: J. Am. Chem. Soc. **59**, 2717 (1937).
(94) MEHL: Cold Spring Harbor Symposia Quant. Biol. **6**, 218 (1938).
(95) MILLER AND DU VIGNEAUD: J. Biol. Chem. **118**, 101 (1937).
(96) MIRSKY AND ANSON: J. Gen. Physiol. **18**, 307 (1935).
(97) MÖRNER: Z. physiol. Chem. **28**, 595 (1899).
(98) MUELLER: J. Biol. Chem. **56**, 157 (1923).
(99) MULDER: Ann. Chemie **28**, 73 (1838).
(100) VON MURALT AND EDSALL: J. Biol. Chem. **89**, 315, 351 (1930).
(101) VON MUTZENBECHER: Biochem. Z. **266**, 226 (1933).
(102) VON MUTZENBECHER: Biochem. Z. **266**, 250 (1933).
(103) VON MUTZENBECHER AND SVEDBERG: Naturwissenschaften **21**, 331 (1933).
(104) NEURATH: Cold Spring Harbor Symposia Quant. Biol. **6**, 196 (1938).
(105) NICHOLS: J. Am. Chem. Soc. **52**, 5176 (1930).
(106) NORTHROP AND ANSON: J. Gen. Physiol. **12**, 543 (1929); **20**, 575 (1936-37).
(107) ONCLEY: J. Am. Chem. Soc. **60**, 1115 (1938).
(108) ONCLEY: Unpublished work.
(109) ONCLEY, FERRY, AND SHACK: Cold Spring Harbor Symposia Quant. Biol. **6**, 21 (1938).
(110) ONSAGER: Physik. Z. **27**, 388 (1926); **28**, 277 (1927).
(111) OSBORNE: J. Am. Chem. Soc. **24**, 140 (1902); The Vegetable Proteins, Longmans, Green and Co., London and New York (1909).
(112) PALMER: J. Biol. Chem. **104**, 359 (1934).
(113) PEDERSEN: Nature **138**, 363 (1936).
(114) PEDERSEN: Biochem. J. **30**, 948 (1936).
(115) PEDERSEN: Compt. rend. trav. lab. Carlsberg **22**, 427 (1938).
(116) PEDERSEN: Unpublished work.
(117) PERRIN: J. phys. radium **7**, 1 (1936).
(118) POLSON: Nature **137**, 740 (1936).
(119) ROCHE AND MARQUET: Compt. rend. trav. lab. Carlsberg **118**, 898 (1935).
(120) ROSE: J. Biol. Chem. **112**, 283 (1935-36); Physiol. Rev. **18**, 109 (1938).
(121) SCATCHARD AND PRENTISS: J. Am. Chem. Soc. **56**, 2314 (1934).
(122) SCHULTZ: Z. physiol. Chem. **25**, 16 (1898).
(123) SCOTT AND FISHER: Biochem. J. **29**, 1048 (1935).
(124) SHACK: Unpublished work.
(125) SHEDLOVSKY: J. Am. Chem. Soc. **54**, 1405 (1932); J. Franklin Inst. **225**, 739 (1938).
(126) SHUTT: Trans. Faraday Soc. **30**, 893 (1934).
(127) SJOGREN AND SVEDBERG: J. Am. Chem. Soc. **53**, 2657 (1931).
(128) SMITH AND SMITH: J. Biol. Chem. **117**, 209 (1937); **121**, 607 (1937).
(129) SMITH: J. Biol. Chem. **108**, 187 (1935); **113**, 473 (1936).
(130) SÖRENSEN: Compt. rend. trav. lab. Carlsberg **12**, 262 (1917).
(131) SÖRENSEN: Compt. rend. trav. lab. Carlsberg **18**, No. 5, 1 (1930).
(132) STAUDINGER: Die hochpolymeren organischen Verbindungen. Julius Springer, Berlin (1932).
(133) STEINHARDT: J. Biol. Chem. **123**, 543 (1938).
(134) SULLIVAN AND HESS: Pub. Health Repts., Suppl. No. 86 (1930).

(135) SVEDBERG: Chem. Rev. **14,** 1 (1934); **20,** 81 (1937); Kolloid-Z. **85,** 119 (1938).
(136) SVEDBERG, CARPENTER, AND CARPENTER: J. Am. Chem. Soc. **52,** 241 (1930).
(137) SVEDBERG, CARPENTER, AND CARPENTER: J. Am. Chem. Soc. **52,** 701 (1930).
(138) SVEDBERG AND FAHRAEUS: J. Am. Chem. Soc. **48,** 430 (1926).
(139) SVEDBERG AND HEDENIUS: Biol. Bull. **64,** 191 (1934).
(140) SVEDBERG AND NICHOLS: J. Am. Chem. Soc. **48,** 308 (1926).
(141) TISELIUS: Dissertation, Upsala, 1930.
(142) TISELIUS: Biochem. J. **31,** 1464 (1937).
(143) TISELIUS AND GROSS: Kolloid-Z. **66,** 11 (1934).
(144) TOENNIS: Growth **1,** 337 (1937).
(145) VAN SLYKE: Handb. d. biol. Arbeitsmeth., Abt. I, Teil **7,** 51 (1923).
(146) VAN SLYKE AND BAKER: J. Biol. Chem. **35,** 127 (1918).
(147) VAN SLYKE, HILLER, DILLON, AND MACFADYEN: Proc. Soc .Exptl. Biol. Med. **38,** 548 (1938).
(148) VICKERY: Compt. rend. trav. lab. Carlsberg **22,** 519 (1938).
(149) VICKERY AND LEAVENWORTH: J. Biol. Chem. **79,** 377 (1928).
(150) VICKERY AND WHITE: J. Biol. Chem. **99,** 701 (1933).
(151) DU VIGNEAUD, JENSEN, AND WINTERSTEINER: J. Pharmacol. **32,** 367 (1928).
(152) WATSON, ARRHENIUS, AND WILLIAMS: Nature **137,** 322 (1936).
(153) WEBER: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **36,** 109 (1934).
(154) WEBER AND STOVER: Biochem. Z. **259,** 269 (1933).
(155) WILLIAMS: Trans. Faraday Soc. **30,** 723 (1934).
(156) WILLIAMS AND ONCLEY: Physics **3,** 314 (1932).
(157) WILLIAMS AND WATSON: Nature **139,** 506 (1937); Cold Spring Harbor Symposia Quant. Biol. **6,** 208 (1938).
(158) WOLLASTON: Phil. Trans. Roy. Soc., p. 223 (1810).
(159) WRINCH: Proc. Roy. Soc. (London) **A160,** 59; **A161,** 505 (1937); Trans. Faraday Soc. **33,** 1368 (1937).
(160) WU AND YANG: Chinese J. Physiol. **6,** 51 (1932).
(161) WYMAN: J. Biol. Chem. **90,** 443 (1931).
(162) WYMAN: Chem. Rev. **19,** 213 (1936); J. Am. Chem. Soc. **60,** 328 (1938).
(163) WYMAN: J. Phys. Chem. **43,** 143 (1939).
(164) ZAHND AND CLARKE: J. Biol. Chem. **102,** 171 (1933).