ELSEVIER

Biophysical Chemistry 100 (2003) 71–79

Biophysical Chemistry

www.elsevier.com/locate/bpc

Reflections on a century of protein chemistry☆

K.E. van Holde*

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331-7305, USA

Received 6 February 2002; accepted 16 May 2002

Abstract

The development of protein structural chemistry during the twentieth century is briefly reviewed. Emphasis is placed on certain major problems that have defined the field, and how they have been resolved, often as a consequence of technological advances. The ways in which incorrect hypotheses have affected the development of the field are also discussed.

Keywords: Protein; Protein conformation; Protein sequence; History of biochemistry

1. Introduction

In the year 1902 the scientific analysis of protein structure was born with Hofmeister's proposal of the peptide hypothesis [1]. Remarkably, 1902 was also the year of John Edsall's birth. Thus, Dr Edsall's life has been contemporary with the development of this science to which he has contributed so much. It seemed, therefore, appropriate to incorporate a brief history of that development in this commemorative volume.

In attempting to summarize this fascinating and complex history, I am humbled by those who have so ably preceded me, including Dr Edsall himself [2,3], whose papers and books have been of great help. Nevertheless, it seems to me that the moment is appropriate for another overlook. It is just 100 years since Hofmeister [1] and E. Fischer [4] began to make proteins respectable to chemists, and in that century we have come very, very far. I shall not try for an exhaustive history—that has been done, for much of the period, in Fruton's excellent book [5]. Neither do I wish to concentrate on the accomplishments of Dr Edsall, many as they are. Rather, I would like to trace the formulation, analysis, and resolution of certain major questions that have dominated the field, and in so doing, perhaps shed some light on how this unique science of biochemistry advances. I am not concerned here with the immense field of protein function, nor with genomics or the evolution of proteins. Rather, the emphasis will be on that area to which John Edsall had dedicated his scientific life—the structure of globular proteins.

2. Proteins as molecules?

The pioneering researches of Fischer and Hofmeister which centered on the *chemistry* of pro-

☆ I wish to dedicate this essay to the memory of Professor John Edsall, who was an inspiration to me in my earliest professional career, and remains such to this day.

*Tel.: +1-541-737-4155; fax: +1-541-737-0481.

E-mail address: vanholdk@ucs.orst.edu (K.E. van Holde).

0301-4622/03/$ - see front matter
PII: S0301-4622(02)00263-6

teins (see below), seemed hardly relevant to the compelling question at the beginning of the century—were proteins molecules or colloidal aggregates? If they were molecules, they had to be far larger than any molecule recognized at the time, for they exhibited very low osmotic pressures and freezing point depression and could not cross semipermeable membranes that allowed salts or sugars to pass. Such results and the successes of colloid chemistry led many to believe that proteins were colloidal aggregates of smaller molecules. Indeed Fischer himself expressed doubt concerning 'molecular weight' values of 12 000–15 000 that had been claimed for some proteins [6].

The question remained unresolved for almost 30 years. A salient event in this bleak period was Sørensen's measurement of the osmotic pressure of ovalbumin in 1917 [7]. Although the value obtained (34 000) was approximately 25% low, the care with which the work was undertaken and the surprising magnitude of the number must have made a strong impression. That impression was not shared by all, as evident in the fact that T. Svedberg's 1924 book *Colloid Chemistry* [8] does not even mention the result, although it describes Sørensen's osmometer! In 1924, Svedberg did not believe proteins were molecules. A second very important result was reported in 1925, when Adair [9] determined the molecular weight of hemoglobin by osmometry. This result (66 800), was almost exactly four times the minimal weight that had become accepted from careful iron analysis. The calculation of such minimal weights, from measured amounts of either a specific cofactor or adduct (i.e. Fe, O_2, CO or relatively infrequent amino acids) was exploited extensively by Cohn et al., who published in 1925 an extensive list of such calculations [10]. Aside from hemoglobin (16 700) it is notable that the result from ovalbumin (27 000–36 000) essentially agreed with Sørensen's (erroneous) result. As we shall see, this chain of fortuitous errors continues. At the time, consistent errors probably did more good for the field than contradictory results would have!

Although the confluence of results from two methods made more credible the contention that proteins were indeed large *entities*, there remained a gap in the argument. Neither of the techniques used above rule out the possibility that the protein consists of a micellar structure made by the strong association of small peptides. Such micelles might be expected to be heterogeneous in size, but the methods used to this point would not detect this, but yield only a 'number average' molecular weight for such a heterogeneous mixture.

The question of heterogeneity was unequivocally laid to rest by Svedberg's ultracentrifuge. The first model was described in 1924, as 'a new instrument for the determination of the size and distribution of size of particles in amicroscopic colloids' [11], reflecting Svedberg's prejudices at the time. However, even the first application to proteins, in which hemoglobin was studied by the sedimentation equilibrium technique [12] gave a startling result: not only was Adair's value for the molecular weight confirmed, but the hemoglobin particles were all the same size! This homogeneity was confirmed in the following year when Svedberg and Nichols used the new oil turbine ultracentrifuge to study hemoglobin by the sedimentation/diffusion method [13]. Strangely, several sedimentation studies of ovalbumin in this period continued to yield what we now know to be low values (see [14] and [15] (pp. 382–383)).

Over the next decade, a large number of proteins were investigated by ultracentrifugation (see [15] for complete references). Almost always the result was the same—most proteins were very large molecules, homogeneous in size. The overwhelming importance of these studies lay in the finding of homogeneity, for colloidal aggregates are rarely homogeneous in size. Within a few years, the idea of proteins as macromolecules had become firmly established.

Some of the most surprising and important results from this period were observed with the hemocyanins. First, some of these were truly immense molecules, ranging up to several million in molecular weight [16]. Equally startling was the observation that many could be *reversibly* dissociated into homogeneous subunits by changes in pH (for a first example, see [16]; for a detailed and comprehensive study, see [17]). It was soon clear that this phenomenon was not confined to hemocyanins, for by 1930 dissociation of hemoglobin by urea was reported [18]. In other words,

some proteins could be built up in a precise hierarchical fashion, using defined protein building blocks in a fixed stoichiometry. Quaternary structure had been discovered!

These results, together with an apparent clustering of many protein molecular weights into a limited number of size ranges, led Svedberg into a serious error. In 1929 he proposed [19], and for many years defended, the hypothesis that most globular proteins were built up from uniform building blocks of approximately 34 000 Da. Unfortunately for protein chemistry, this idea nicely complemented some (equally erroneous) models for protein amino acid sequence and structure (see below). As we shall see, the resulting confusion required roughly a decade for resolution.

3. The chemical structure of proteins

Although it had long been known that amino acids were released from proteins on hydrolysis, and therefore likely represented, in some way, component parts of these substances, the nature of the linkage was obscure. Indeed, there was so very little evidence upon which to propose a model that Hofmeister's [1] proposal of amide bonds between amino acids seems almost clairvoyant. It is clear, however, that Fischer had already been thinking in the same direction, for as early as 1901, at the beginning of his 'protein' studies, he produced the dipeptide glycylglycine (and named it such, thereby establishing the nomenclature) [4]. Fischer was also the first to refer to such compounds as peptides and polypeptides. In the following years, Fischer synthesized a series of oligopeptides, culminating in 1907 with an octadecapeptide [20]. He recognized these as appropriate models for proteins, but also noted that other bonding structures (e.g. diketopiperazine rings) might be found.

Perhaps more important to protein chemistry than Fischer's syntheses of oligopeptides was the demonstration in 1902 that the dipeptide glycylalanine was one product of hydrolysis of silk fibroin [1]. This along with Fischer's work stands as one of the very few early evidences that the peptide bond actually existed in proteins [20].

After Fischer, there was a hiatus of nearly two decades before sustained research on the chemical structure of proteins was resumed. Unfortunately, the next stage was initially counterproductive. In 1924, Abderhalden [21] proposed that the basic units of protein structure were diketopiperazine rings, in non-covalent association. This postulate drew on a long history of observations of diketopiperazines from protein degradation, as well as the implied support of Fischer himself (see above). It was also welcomed by the colloid chemists of the day, for it provided a suggestion as to how large 'molecular weights' might be accounted for without invoking macromolecules. Although the diketopiperazine theories were soundly criticized in a splendid review by Vickery and Osborne [22] and essentially demolished by the ultracentrifuge studies of Svedberg (see above), acceptance of the peptide theory still came hard. As late as 1938, a no less distinguished scientist than Linderstrøm-Lang [23] suggested that peptide linkages might exist only in denatured proteins!

However, by most biochemists of the early 1930s, the peptide hypothesis was regarded as the most tenable model for protein structure. But if proteins the size of even the smallest investigated by Svedberg consisted of single polypeptide chains, another vexing question arose. How were the amino acid residues *arranged* in the sequence? Homogeneity in the ultracentrifuge implied that the chains for a particular protein (or subunit) were uniform in length. It seemed therefore unreasonable that they would be random in sequence. Yet even Fischer had noted that his octadecamer allowed an enormous variety of sequences. For a chain of 100 residues or more, the number becomes hyper-cosmic. How could just *one* specific sequence be synthesized reproducibly in the cell to make a particular protein? A seemingly attractive idea developed by Bergmann and Niemann suggested [24–26] that the sequences were repetitive; if this were so, syntheses could operate by certain simple rules using a limited number of enzymes, much in the way that complex carbohydrates are now known to be synthesized. Bergmann arrived at the idea of repetitive sequences by a rather circuitous route. Noting Svedberg's observation that many proteins seemed to have molecular weights that were approximate multiples of 34 000 Da, he pointed out that this corresponded

to approximately 288 amino acids or the number $2^5 \times 3^2$, and that virtually any protein molecule weight that had been observed could be approximated by $N_t = 2^n \times 3^n$. They were able to express some of the amino acid composition data available at the time by $N_i = 2^{n^i} \times 3^{n^i}$ where n^i and m^i are zero or integer. (This is neither surprising or convincing, since N_i can have the values 1, 2, 3, 4, 6, 8, 12, 16, 18, etc. allowing *many* possibilities.) The frequency F_i of each amino acid type then becomes

$$F_i = N_t / N_i = 2^{n-n^i} \times 3^{m-m^i}$$

which is of course integral. Different integral frequencies for each amino acid can be most easily explained by a periodic structure.

This erroneous scheme, which came to be called the Bergmann–Niemann hypothesis, gained wide credence in the late 1930s, probably for several reasons:

1. The most widely studied series of peptides came from proteins like silk fibroin or collagen, which *do* have quasi-repetitive structures.
2. The enormous influence of Svedberg's work and the widespread acceptance of his multiplicity theory which fit nicely with the Bergmann–Niemann hypothesis.
3. It also dovetailed quite nicely with the 'cyclol' theory of protein structure (see below), which was enjoying popularity at this time.
4. The Bergmann–Niemann hypothesis at least suggested understandable schemes for protein synthesis in vivo.

The demise of the Bergmann–Niemann hypothesis occurred in only a few years, as better and better compositional data and characterization of more oligopeptides proved to be inconsistent with it [27–29].

Thus, even 40 years after Fischer, our knowledge of the polypeptide structure of proteins had shown very little progress. In most part, this was due to the lack of effective methods for the separation and identification of peptides. Thus there was virtually no *data* upon which to build models properly, and yet models were constructed. The task of demolishing these models had consumed much of the effort of the best protein chemists over two decades.

The development, in the 1940s of a variety of simple and rapid chromatographic techniques [30] was essential for further progress. This was forthcoming in Sanger's great *tour de force*, sequencing insulin. The work began in 1949 [31] and had reached the stage of sequencing of the two chains by 1952 [32]. The completed project is succinctly summarized in a 1956 review [33]. In one stroke, the insulin sequence answered questions that had plagued the field since Fischer's day: Yes, the peptide chains were each of *unique* sequence; no, the arrangement of amino acids was dictated neither by periodicity nor arcane numerology. This result, of course, threw open once more the nagging question—how could such precisely ordered chains be synthesized? Sanger notes this, and with characteristic reserve adds: 'Clearly the present results can provide little information about the mechanisms of protein synthesis, but they do provide certain limitations on those who are speculating on the subject.' [33] Speculators there were in plenty. A number of illustrious scientists had expressed ideas of varying degrees of fantasy on the subject (see, for examples [34–36]). The only thoughtful and sensible example I find from this period is in a remarkable paper by F. Crick [37], who has it *almost* right! Meanwhile, the information for a correct solution to the synthesis problem was rapidly falling in place, and only 5 years after Sanger's insulin structure the beautiful paper by Nirenberg and Matthaei [38] provided the final touch—an RNA *template* was required to direct the synthesis of each protein.

Sanger's work, which has been followed by the sequencing of thousands of proteins, represents nearly the last major advance with respect to the sequence problem. The only truly significant addition was the discovery of post-translational processing, as exampled by the removal of 'signal' sequences [39] or conversion of pro-insulin to insulin [40].

4. Three-dimensional structure of proteins

Until the existence of proteins as macromolecular entities rather than colloidal aggregates was

established in the 1920s, serious consideration of their internal structure was impossible. However, the development of hydrodynamic techniques in the late 1920s quickly established that some proteins could be termed 'globular'. In particular, frictional ratios calculated from sedimentation/molecular weight data showed that a number of proteins were nearly spherical [41]. At about the same time, Muralt and Edsall, in pioneering flow birefringence studies, established that myosin though soluble, was anisotropic [42]. It was clear at this point how the 'globular' proteins (even anisotropic ones) differed from the truly fibrous proteins like silk or keratin. The latter had already been studied by X-ray diffraction with excellent results [43,44] and valid interpretation.

It seems that prior to 1931, little thought had been given to the possible *internal* structures of globular proteins. However, in that year a remarkable paper appeared in the *Chinese Journal of Physiology* [45]. It represented the culmination of a series of experimental papers by Dr H. Wu on protein denaturation, a topic that had been discussed for many years with negligible insight. Wu's paper presents an incisive, prescient analysis of globular protein structure and denaturation. It is succinctly stated in the Summary, which could hardly be improved upon today:

> 'Evidence is adducted in support of the hypothesis that the molecule of natural, soluble protein is not a flexible open chain of polypeptide but has a compact structure. The force of attraction between the polar groups in a single molecule of protein holds them together in an orderly ways, just as the force of attraction between different molecules holds many molecules together in a crystal.In denaturation or coagulation the compact and orderly structure is disorganized. If denaturation occurs in acid or alkali or in urea solution, the individual molecules are disrupted but they remain separate. In coagulation they interpenetrate and are entangled. The facts known about denaturation and coagulation in diverse ways are explained and correlated by the theory.'

Probably because of its obscure venue, Wu's paper seems to have been unknown to most in the West. For example, the 1936 paper by Mirsky and Pauling [46] which covers essentially the same ground, does not cite Wu. Mirsky notes, in Discussion during the 1938 Cold Spring Harbor Symposium, that he had only learned of the Wu paper a few months earlier. Mirsky and Pauling did go a bit further than Wu, in specifically suggesting hydrogen bonds as playing a major role in stabilizing the folded structure. The years following Wu's paper also saw the first solid experimental studies of protein denaturation, which included demonstrations that some such reactions were reversible ([47,48], for examples). Yet the full implications of such results were slow to be realized.

Just as the Bergmann–Niemann hypothesis complicated attempts to understand protein sequences, there arose in 1936 a model for protein structure that was to cause confusion and misdirected effort. This is the 'cyclol' hypotheses, put forward by Dr D. Wrinch [49]. In brief, Wrinch proposed that polypeptides did not exist in globular proteins as linear chains, but instead folded into a sheet-like structure of six-membered rings, with the NH group of one amino acid forming an N–C bond with another. Certain sizes of such cyclols were favored, with that involving 288 residues being one. *Mirabile dictu*, this corresponds to the Bergmann–Niemann/Svedberg magic number for a protein of approximately 34 000 Da! Thus, a synergism between three (erroneous) ideas sprang up, which for a short time drew the allegiance of some very prominent scientists, including I. Langmuir [50]. However, the cyclol theory was soon demolished—first by Haurowitz who could find no evidence for the multitudinous hydroxyl groups predicted by the cyclol theory either by chemistry or spectroscopy [51], and second by a closely-reasoned theoretical analysis by Pauling and Niemann [52].

Although X-ray diffraction had been discovered many years earlier, and considerable success had been had with diffraction studies of protein fibers like hair and silk (see above), attempts with globular protein crystals had consistently failed. In 1934, Bernal and Crowfoot [53] discovered that pepsin crystals, kept in their mother liquor, gave excellent diffraction patterns. Although neither the analytical techniques nor computational power were then available to interpret the data in detail, Bernal and Crowfoot noted the high degree of

Table 1
A condensed chronology of protein structure, 1900–1961

Year	Major contributions	Other papers of note
1901	Fischer and Fourneau: glycylglycine synthesis [4]	
1902	Hofmeister: statement of peptide hypothesis [1]	
1906	Fischer: overview of peptide synthesis [6]	(20)
1917	Sørensen: M of ovalbumin by osmotic pressure (O.P.) [7]	
1924	Abderhalden: diketopiperazine hypothesis [21]	(11)
1925	Adair: hemoglobin M by O.P. [9]; Cohn et al., minimal molecular weights [10]	
1926	Svedberg and Fahraeus: SE of hemoglobin [12]; Summer, XL of urease [62]	(14)
1927	Svedberg and Nichols: M of hemoglobin by sedimentation-diffusion [13]	
1928	Vickery and Osborne review [22]; Meyer and Mark: fiber X-ray studies [43]	(22)
1929	Svedberg: multiplicity hypothesis [19]	
1930	Burke and Greenberg: hemoglobin dissociation [18]; Muralt and Edsall: myosin [42]	(41)
1931	Wu: theory of denaturation [45]	
1932	Astbury and Street: α-β-keratin [44]; Northrop: reversibility of denaturation [47]	
1933		
1934	Bernal and Crowfoot: X-ray of pepsin [53]; Anson and Mirsky: denaturation thermodynamics [48]	
1935		
1936	Mirsky and Pauling: denaturation [46]; Wrinch: cyclols [49]; Bergmann and Niemann: $2''\times3'''$ [24]	(17)
1937	Bergmann and Niemann: fibroin peptides [25]	
1938		
1939	Pauling and Niemann [52], Haurowitz and Astrup: refutation of cyclols [51]	(50)
1940	Svedberg and Pedersen book [15]	
1941	Gordon et al.: protein hydrolysis does not support Bergmann [27]	(34)
1942	Chibnall questions ‘protein numerology’ [28]	
1943	Synge: sequence data questions Bergmann [29]; Cohn and Edsall book [61]	
1949	Sanger: first insulin chain; definite yet irregular [31]	
1950		(35)
1951	Pauling, Corey and Branson: two hydrogen-bonded structures [60]	
1952	Sanger: two insulin chains sequenced [32]	
1953	Green, Ingram and Perutz: isomorphous replacement in hemoglobin [54]	
1954		(36)
1956	Sanger: the complete structure of Insulin [33]	
1957		
1958	Kendrew et al.: low resolution structure of myoglobin [55]; Crick: on protein synthesis [37]	
1959		
1960	Kendrew et al.: 2 Å structure of myoglobin [55]	
1961	Nirenberg and Matthaei: RNA templates for protein [38]; White, [57], Anfinsen and Haber: RNAse folding [58]	(54)

order, and evidence that the structure was compact and closely packed.

During the next 10 years techniques for crystal growth and data collection steadily improved, but analysis was blocked by the lack of methods to determine the phases of reflections. Many methods were attempted, but the major breakthrough came in 1953 when Perutz et al. found that isomorphous replacement using mercury atoms allowed phasing of hemoglobin reflections [54]. By 1958 the first low resolution myoglobin structure was published [55] and refined 2 years later [56].

As with sequencing, subsequent years have seen the X-ray analysis of thousands of proteins, with immeasurable aid to the understanding of many biochemical processes. Remarkably, however, no major new *general* insights about proteins seem to have emerged. A series of papers on the spontaneous folding of proteins caused considerable interest in the early 1960s (see, for example [57–59]). However, the recent interest in 'chaperones' and 'the protein folding problem' makes it clear that we still do not fully understand the determinants of globular protein structure.

5. Overview

What can we deduce concerning the ways in which science progresses from this one slice of history? It is helpful, I believe, to prepare a rough chronology as shown in Table 1. Here I have listed what I consider to be the most important contributions for a period of 60 years. More recent papers are not included, for it is difficult to get a proper perspective on the more recent work, and, in my opinion, by 1961 the fundamental questions about protein structure had been resolved. The citations are in two categories: Those I judge to have been of the very greatest importance are numbered in square brackets, to left, and briefly identified. The remaining numbers (in parentheses, to right) I consider to be of considerable, but lesser importance. I apologize to any who feel that vital papers have been left out; the choice is certainly somewhat subjective and in any event reflects my very limited knowledge in some areas.

In glancing at this Table, one is struck by the fact that there have been, in essence, two 'golden decades' of protein chemistry in the twentieth century (see vertical bars). One extends from approximately 1925 to 1935, the other from approximately 1951 to 1961. The first corresponds to the period when it was clearly established what proteins really are, and what questions could be formulated concerning their fine structure. The second golden decade begins when the techniques finally became available to answer those questions. This decade also opens with the groundbreaking proposal by Pauling et al. [60] of secondary structures in proteins.

At the risk of offending many of my contemporaries, I will argue that very little *fundamentally* new about proteins has been learned since 1961. Certain large problems still exist—for example how proteins fold and how they function dynamically at the single molecule level. These problems are under active pursuit, and may (hopefully) yield some surprises.

Finally, I am struck by two other observations. The first is how important a role *techniques* have played in this development. Ultracentrifugation broke the protein-identity problem. The sequencing of proteins could not proceed until a number of chromatographic and chemical techniques were available. These, in turn, could not have been developed without the full understanding of the ionic behavior of proteins and peptides described in the seminal monograph by Cohn and Edsall [61]. Edsall's pioneering work in this field represents one of his major contributions to protein chemistry. Similarly, we could not have unraveled protein structure without a solution to the X-ray phase problem and the development of computers. In every case, it was technological advance that allowed a critical breakthrough. Second, it is remarkable how much was contributed by so *few*. A surprisingly small number of scientists seem to have provided the essential keys to solve the few really critical problems. Among these we must certainly include John Edsall.

References

[1] F. Hofmeister, Über Bau und Gruppierung der Eiweisskörper, Ergeb. Physiol. 1 (1902) 759–802.
[2] J.T. Edsall, Proteins as macromolecules: an essay on the

development of the macromolecular concept and some of its vicissitudes, Arch. Biochem. Biophys. (Suppl. I) (1962) 12–20.
[3] J.T. Edsall, The development of the physical chemistry of proteins, 1898–1940, Ann. NY Acad. Sci. 325 (1979) 53–73.
[4] E. Fischer, E. Fourneau, Über einige Derivate des Glykocolls, Ber. Deutsch. Chem. Ges. 34 (1901) 2868–2871.
[5] J.S. Fruton, Molecules and Life, Wiley-Interscience, NY, 1972.
[6] E. Fischer, Untersuchungen über Aminosäuren, Polypeptide und Proteine, Ber. Deutsch. Chem. Ges. 39 (1906) 520–610.
[7] S.P.L. Sørensen, Studies on proteins. V. On the osmotic pressure of egg-albumin solutions, C.R. Trav. Lab. Carlsberg 12 (1917) 202–372.
[8] T. Svedberg, Colloid Chemistry, first ed., Chemical Catalog Co, NY, 1924.
[9] G.S. Adair, The osmotic pressure of haemoglobin in the absence of salts, Proc. R. Soc. Lond. A 109 (1925) 292–300.
[10] E.J. Cohn, J.I. Hendry, A.M. Prentiss, Studies in the physical chemistry of the proteins. V. The molecular weights of the proteins. Part I. The minimal molecular weight of certain proteins, J. Biol. Chem. 63 (1925) 721–766.
[11] T. Svedberg, H. Rinde, The ultracentrifuge, a new instrument for determination of the size and distribution of size of particles in amicroscopic colloids, J. Am. Chem. Soc. 46 (1924) 2677–2693.
[12] T. Svedberg, R. Fahraeus, A new method for the determination of the molecular weight of the proteins, J. Am. Chem. Soc. 48 (1926) 430–438.
[13] T. Svedberg, J.B. Nichols, The application of the oil turbine type of ultracentrifuge to the study of the stability region of CO–hemoglobin, J. Am. Chem. Soc. 49 (1927) 2920–2934.
[14] T. Svedberg, J.B. Nichols, Molecular weight of egg albumin in electrolyte-free condition, J. Am. Chem. Soc. 48 (1926) 3081–3092.
[15] T. Svedberg, K.O. Pedersen, The Ultracentrifuge, Clarendon Press, Oxford, UK, 1940.
[16] T. Svedberg, E. Chirnoaga, The molecular weight of hemocyanin, J. Am. Chem. Soc. 50 (1928) 1399–1411.
[17] I.B. Eriksson-Quensel, T. Svedberg, The molecular weights and pH-stability regions of the hemocyanins, Biol. Bull. 7 (1936) 498–547.
[18] N.F. Burk, D.M. Greenberg, The physical chemistry of the proteins in non-aqueous and mixed solvent. I. The state of aggregation of certain proteins in urea–water solutions, J. Biol. Chem. 87 (1930) 197–238.
[19] T. Svedberg, Mass and size of protein molecules, Nature 123 (1929) 871.
[20] E. Fischer, Synthetical chemistry in its relation to biology, J. Chem. Soc. 91 (1907) 1749–1765.
[21] E. Abderhalden, Das Eiweiss als eine Zusammungfassung assoziierter, Anhydride enthaltender Elementarkomplexe, Naturwiss. 12 (1924) 716–720.
[22] H.B. Vickery, T.B. Osborne, A review of hypotheses of the structure of proteins, Physiol. Rev. 8 (1928) 393–446.
[23] K. Linderstrøm-Lang, Peptide bonds in globular proteins, Nature 142 (1938) 990.
[24] M. Bergmann, C. Niemann, On blood fibrin. A contribution to the problem of protein structure, J. Biol. Chem. 115 (1930) 77–85.
[25] M. Bergmann, C. Niemann, On the structure of silk fibroin, J. Biol. Chem. 122 (1937) 577–596.
[26] M. Bergmann, The structure of proteins in relation to biological problems, Chem. Rev. 22 (1938) 423–435.
[27] A.H. Gordon, A.J.P. Martin, R.L.M. Synge, A study of the partial acid hydrolysis of some proteins with special reference to the mode of linkage of the basic amino acids, Biochem. J. 35 (1941) 1369–1387.
[28] A.C. Chibnall, Amino acid analysis and the structure of proteins, Proc. R. Soc. Lond. B131 (1942) 136–160.
[29] R.L.M. Synge, Partial hydrolysis products derived from proteins and their significance to protein structure, Chem. Rev. 32 (1943) 135–172.
[30] A.H. Gordon, Electrophoresis and chromatography of amino acids and proteins, Ann. NY Acad. Sci. 325 (1979) 95–105.
[31] F. Sanger, Some chemical investigations of the structure of insulin, Cold Spring Harb. Symp. Quant. Biol. 14 (1949) 153–160.
[32] F. Sanger, The arrangement of amino acids in proteins, Adv. Prot. Chem. 7 (1952) 1–69.
[33] F. Sanger, in: D.E. Green (Ed.), The Structure of Insulin, in Currents in Biological Research, Interscience, NY, 1956, pp. 434–459.
[34] M. Delbrück, A theory of autocatalytic synthesis of polypeptides, Cold Spring Harb. Symp. Quant. Biol. 9 (1941) 122–126.
[35] J.S. Fruton, The role of proteolytic enzymes in the biosynthesis of peptide bonds, Yale J. Biol. Med. 22 (1950) 263–271.
[36] G. Gamow, Possible relations between deoxyribonucleic acid and protein structures, Nature 173 (1954) 318.
[37] F.H.C. Crick, On protein synthesis, Symp. Soc. Exp. Biol. 12 (1958) 138–163.
[38] M. Nirenberg, J.H. Matthaei, The dependence of cell free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, Proc. Natl. Acad. Sci. USA 47 (1961) 1588–1602.
[39] C. Milstein, G.G. Brownlee, T.M. Harrison, M.B. Mathews, A possible precursor of immunoglobulin light chains, Nature New Biol. 289 (1972) 117–120.
[40] D.F. Steiner, D.E. Oyer, The biosynthesis of insulin and a possible precursor of insulin in a human islet cell carcinoma, Proc. Natl. Acad. Sci. USA 57 (1967) 473–480.

[41] T. Svedberg, Ultracentrifugale Dispersitätsbestimmungen von Eiweisskörpern, Kolloid Zeitschr. 51 (1930) 10–24.
[42] A.L. Muralt, J. Edsall, Studies of the physical chemistry of muscle globulin. IV. The anisotropy of myosin and double refraction of flow, J. Biol. Chem. 89 (1930) 351–386.
[43] K. Meyer, H. Mark, Über den Aufbau des Seiden-Fibroins, Ber. Deutsch. Chem. Gesellschaft. 61 (1928) 1932–1936.
[44] W.T. Astbury, A. Street, X-ray structure of hair, wool, and related fibers. I-General, Phil. Trans. R. Soc. Lond. A230 (1932) 75–101.
[45] H. Wu, Studies on denaturation of proteins. XIII. A theory of denaturation, Chin. J. Physiol. 5 (1931) 321–344.
[46] A.E. Mirsky, L. Pauling, On the structure of native, denatured, and coagulated proteins, Proc. Natl. Acad. Sci. USA 22 (1936) 429–447.
[47] J.M. Northrop, Crystalline trypsin. IV. Reversibility of the inactivation and denaturation of trypsin by heat, J. Gen. Physiol. 16 (1932) 323–337.
[48] M.L. Anson, A.E. Mirsky, The equilibrium between native and denatured hemoglobin in salicylate solutions and the theoretical consequences of the equilibrium between native and denatured protein, J. Gen. Physiol. 17 (1934) 393–398.
[49] D. Wrinch, The pattern of proteins, Nature 137 (1936) 411–412.
[50] I. Langmuir, The structure of proteins, Proc. Phys. Soc. 51 (1939) 592–612.
[51] F. Haurowitz, T. Astrup, Ultraviolet absorption of genuine and hydrolysed proteins, Nature 143 (1939) 118–119.
[52] L. Pauling, C. Niemann, The structure of proteins, J. Am. Chem. Soc. 61 (1939) 1860–1867.
[53] J.D. Bernal, D. Crowfoot, X-ray photographs of crystalline pepsin, Nature 133 (1934) 794.
[54] D.W. Green, V.M. Ingram, M.F. Perutz, The structure of hemoglobin. IV. Sign determination by the isomorphous replacement method, Proc. R. Soc. Lond. A 225 (1953) 287–307.
[55] J.C. Kendrew, G. Bode, H.M. Dintzis, R.C. Parrish, H. Wykoff, A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, Nature 181 (1958) 660–662.
[56] J.C. Kendrew, R.E. Dickerson, R.E. Strandberg, R.G. Hart, D.R. Davies, D.C. Phillips, V.C. Shore, Structure of myoglobin. A three-dimensional Fourier synthesis at 2Å resolution, Nature 185 (1960) 422–427.
[57] F.H. White, Regeneration of native secondary and tertiary structure by air oxidation of reduced ribonuclease, J. Biol. Chem. 236 (1961) 1353–1358.
[58] C.B. Anfinsen, E. Hata, Studies on the reduction and re-formation of protein disulfide bonds, J. Biol. Chem. 236 (1961) 1361–1363.
[59] C.B. Anfinsen, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, Proc. Natl. Acad. Sci. USA 47 (1961) 1308–1314.
[60] L. Pauling, R.B. Corey, H.R. Branson, The structure of proteins: Two hydrogen-bonded helical conformations of the polypeptide chain, Proc. Natl. Acad. Sci. USA, 37 (1951) 205–211.
[61] E.J. Cohn, J.T. Edsall, Proteins, Amino acids and Peptides as ions and Dipolar ions, Reinhold, NY, 1943.
[62] J.B. Sumner, The isolation and crystallization of urease, J. Biol. Chem. 69 (1926) 435–441.