On the Laboratory That Produced the Book *Proteins, Amino Acids and Peptides*

John T. Edsall
Dept. of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

A celebration of a scientific book, first published 50 years ago, is a rare event in our time. Given the speed with which science is progressing today, a backward look toward half a century ago might seem like an excursion into archeology. So, I was startled and appreciative of an unusual honor, when I learned that the American Institute of Chemical Engineers had chosen to organize a symposium in recognition of a book published in 1943, *Proteins, Amino Acids and Peptides* by Edwin J. Cohn and myself. What gives such special status to this book that it is still remembered after half a century?

Certainly our knowledge of protein structure and function, 50 years ago, was a tiny fraction of what it is now. Yet, several biochemists, considerably younger than myself, have told me that they still turn to the book on occasion for information that is hard to find elsewhere; at least two librarians have told me that they kept the copy in their own library under lock and key and released it to students only after telling them to be careful to bring it back.

In this article I shall be concerned with personal recollections. I will look briefly at some aspects of the history of protein chemistry and will add a few remarks at the end about the state of protein chemistry today. This is in essence a talk, not a scientific article. I make no attempt to provide references for what I say. A properly referenced talk would call for hundreds of references, and I cannot undertake that.

For me, of course, the book has a deep personal meaning, as it certainty had for my former chief, Edwin J. Cohn (1892–1953) who headed the Department of Physical Chemistry at Harvard Medical School for many years. Cohn had decided, after receiving his PhD in 1917 to devote his scientific career to the study of proteins and their constituents, the amino acids, and peptides. In the 1920s, the work of the lab centered on the physical chemistry of proteins, particularly the effects of ionic strength and pH on protein solubility, and on the nature of proteins as acids and bases. The experimental results were very interesting, but their interpretation was, in many respects, baffling. This led Cohn, shortly before 1930, to recognize that we needed to know more about the physical chemistry of the amino acids and peptides, from which the proteins are built up to understand proteins better. This became the central focus of the work of the laboratory for the decade of the 30s, although studies on proteins never ceased.

Status of Protein Chemistry about 1930

By 1930 it was becoming clear that purified proteins—at least the "globular" proteins that were soluble in water or saline solutions—were well defined macromolecules, with molecular weights of many thousands. The Svedberg in Sweden, who had developed the ultracentrifuge, found that purified proteins sedimented as molecules of uniform size. He soon became a strong advocate of the view that proteins are true macromolecules.

Their sedimentation rates and the low viscosity of their solutions indicated that native proteins were compact molecules, somewhat like small spheres or ellipsoids, not like long, open, twisting chains, such as those of synthetic polymers in a good solvent. Denatured proteins, however, did behave rather like the random coils of synthetic polymer chains.

In the late 1920s, Mona Spiegel-Adolf in Vienna found she could reconvert acid-denatured serum albumin to native albumin, primarily by restoring it to a pH near 7, favorable to the native structure, and waiting for a short time. Soon after this, M. I. Anson and A. E. Mirsky at the Rockefeller Institute achieved success in reversing the denaturation of hemoglobin. Similar work on other proteins followed. In 1931, Hsien Wu at Peking Union Medical College, who had done extensive experimental work on the subject, published the first good theory of denaturation, or protein unfolding, as we now generally call it. He pictured the process as a reversible transition between the semicrystalline internal order of the native protein and the disorganized state that we would now call a random coil. He marshaled good evidence to support his view. This article was published in the *Chinese Journal of Physiology*, but many chemists in America and Europe remained unaware of it, although it was written in good English.

"Core Group" of Investigators in 1930–40

About 1930, Cohn's laboratory shifted its major research

emphasis from proteins to amino acids, peptides and related substances, as I have mentioned above. Work on proteins never stopped, but for a decade the emphasis was on smaller compounds. A number of young investigators came and went, staying perhaps a year or two, but a "core group" was developed, which gradually came together over the decade. All of us in that group became deeply involved in problems concerning the nature of proteins and their constituents. I would say that the group consisted of nine members, six in our laboratory at the medical school were Cohn, Thomas L. ("Mac") McMeekin (1900–1979), Jesse P. Greenstein (1902–1959), myself (1902–), John Lawrence ("Larry") Oncley (1911–) and John D. Ferry (1912–). I would also recall the names of two technicians of exceptional quality—Muriel H. Blanchard and John H. Weare—who showed such ability and skill that we listed them as coauthors in a number of articles from the laboratory.

McMeekin was an organic chemist who did important physicochemical experiments on the solubility of amino acids and peptides in various media, studying their interactions with ions in solution, and also with organic solvents, throwing much light on what we now call hydrophobic interactions. Greenstein was both an organic and a physical chemist. He was a master of peptide synthesis and did much on the physical chemistry of peptides; he was a man of exceptional energy and enthusiasm, and also a devoted teacher. Later, he became the first head of the chemical division of the National Cancer Institute in Bethesda and contributed much to the chemistry of cancer.

Oncley developed a powerful method for studying the dielectric constants and relaxation times of proteins, which provided important information about their electrical properties and shapes. John Ferry, the youngest member of the group, also worked on dielectric properties of proteins. He became later one of the greatest authorities on the viscoelastic properties of polymer systems, including proteins. Both Oncley and Ferry contributed chapters to the book, and both played key roles in the development of human plasma proteins for clinical use during World War II.

The three "outside" members of the group were as deeply involved as the others. Two were at MIT. One was George Scatchard (1892–1973), a great authority on the physical chemistry of solutions and an intimate friend of Cohn. He was deeply interested in proteins, at a time when few of the "pure" physical chemists took an interest in such substances. He was a most helpful adviser and critic for younger members of the group like myself. He also wrote one of the most important chapters in our book and became coauthor with me in another chapter, in which he greatly improved an earlier draft of mine.

Another very important member was John G. Kirkwood (1907–1959), then a young faculty member at MIT. He was highly gifted mathematically and developed theoretical interpretations of the data that the experimenters in the group provided, and his calculations led to new experiments. He summarized his contributions well in his chapter in the book (Chapter 12). He later had a highly distinguished career, culminating at Yale, where he headed the Chemistry Department for years.

Finally, there was my closest scientific friend, Jeffries Wyman (1901–), in the Harvard Biology Department. His great contribution to the work was in the study of the remarkable dielectric properties of amino acids and proteins. Later, he was to become one of the world's great authorities on the physical chemistry of hemoglobin and on the thermodynamics of multicomponent interacting systems.

There was a constant interplay of discussion of ideas and planning of new experiments among the members of the group. It was for me, and I think for all of us, a memorable time for creative activity in the midst of the great depression. Harvard provided some of the funds that kept us going, but our main source of support was the Rockefeller Foundation, for we fitted very well into Warren Weaver's plans for promoting work in biology that had a strong foundation in chemistry and physics. We also had some support from a couple of industrial firms, but that was a minor factor. With Cohn's fund-raising ability and Weaver's backing, we younger workers did not have to spend time in writing grant applications during those years. When I see the problems of young workers today, I realize that we were, in some respects at least, a privileged group.

Ionic Dipoles and Their Properties

Around 1930 there was still some doubt concerning the acidic and basic properties of amino acids and proteins. To take the simplest case of all was isoelectric glycine at pH 6 to 7, properly described as $H_2N \cdot CH_2 \cdot COOH$ or as $^+H_3N \cdot CH_2 \cdot COO^-$?

The first of these two formulas had been generally accepted as correct in the early 20th century, but in 1916 E. Q. Adams argued that the second one, with both the amino and the carboxyl group in ionized form, was in much better accord with the evidence from the ionization constants of related amines and carboxylic acids. In 1923, Niels Bjerrum in Copenhagen independently came to the same conclusion (he evidently had not seen the article by Adams, since he did not cite it) and gave a more extensive examination of the problem. He pointed out, for instance, that the high melting points of the amino acids, of the order of 300°C (usually with decomposition), could be due to the strong electrostatic forces between the charges in the amino acid acid molecules when packed together in the crystal.

This class of molecules, with the two ionic charges of opposite sign, separated by a considerable distance, had been called *Zwitterionen* (hybrid or hermaphrodite ions) by the Germans. We called them "dipolar ions" in our book, but I would now suggest a change. Linus Pauling once wrote me a letter objecting to the term. He thought that the use of the noun "ion," for what was in fact a dipole of high electric moment, was misleading. I think Pauling was right, and I would now suggest calling them "ionic dipoles." The use of "dipole" as the noun, with "ionic" as the adjective (since the constituents of the dipole are actually ionic) is, I believe, more suggestive of the actual structure than the term we used in the book.

Implications of the Ionic Dipole Structure

The arguments of Adams and Bjerrum for the ionic dipole structure of simple amino acids were compelling. To understand these arguments was practically to accept them, but acceptance carried further implications that needed to be ex-

plored. The most obvious was the fact that ionic dipoles must have high electric moments compared to ordinary polar molecules. The two charge centers, in an α-amino acid such as glycine must, from the geometry of the molecule, be about 3.0 Å apart. This, given the magnitude of the charges, ($\pm 4.8 \times 10^{-10}$ e.s.u.) gives an electric moment of $3 \times 4.8 = 14.4$ D (in debye units). This is much larger than the values for most polar organic molecules, such as alcohols, ketones, and amides, which are generally of the order of magnitude of 1 to 5 debye units.

Wyman's studies of amino acid and peptide solutions in water showed that in all cases the dielectric constant of the solution was higher than that of pure water (78.5 at 25°C) and increased linearly with the concentration of solute. At 1 M concentration the molar increment (δ) was about 23 for α-amino acids, 70 for dipeptides, 159 for tetraglycine, and 234 for hexaglycine. Lysyl glutamic acid, with two positive and two negative charges, gave a δ value of 345. I should note that the Italian chemist G. Devoto was independently studying dielectric properties of these substances, getting results in good accord with those of Wyman.

Wyman's theoretical analysis led to the conclusion that the δ value was roughly proportional to the square of the dipole moment, μ, in debye units. Kirkwood, in his chapter in our book, gave a more sophisticated discussion of the problem (pp. 290–296) and tabulated distances, in Å, between the two charges of the dipoles. For instance, he reported a value of 3.27 Å for glycine and 5.76 Å for glycylglycine. The latter value indicates that there is a significant amount of free rotation around the CH_2 groups in glycylglycine.

Interactions of Charged Groups with Water: Effects on Apparent Molar Volumes and Heat Capacities

Other characteristic differences between the properties of ionic dipoles and isomeric polar molecules emerged from the work in the Cohn lab. In 1894, P. Drude and W. Nernst had pointed out that the ionization of 1 mol of water involves a volume decrease in the system of about 21 cm^3. Similar contractions occur in the ionization of organic acids and bases. They attributed these effects to electrostriction of the solvent; the intense electric field in the immediate neighborhood of the ions attracts water molecules and squeezes them close to the ions, thus decreasing the volume of the system.

We showed that electrostriction also occurs in solutions of ionic dipoles by comparing the molar volume increment of such a dipole in water with that of an isomer of low polarity. The first comparison was between glycine and its isomer glycolamide: $HO \cdot CH_2 \cdot CONH_2$, a polar molecule with a small moment, not an ionic dipole. The increment in volume, per mole of gylcolamide added to the system was 56.2 cm^3; per mole of glycine it was 43.5. The difference of 12.7 cm^3 represents the electrostriction due to the charged groups.

For ionic dipoles with more widely separated charges, the amount of electrostriction is greater, as might be expected on electrostatic grounds. For glycylglycine, for instance, with a chain of four atoms between the ionic groups, the value estimated was 17.0 cm^3/mol, probably close to the limiting value for a chain of infinite length.

There were also striking differences in apparent molar heat capacities between ionic dipoles and their polar but uncharged isomers. This was to be expected. The ionic groups not only squeezed neighboring water molecules closer to them; they also largely immobilized them, thereby diminishing their contribution to the heat capacity of the system. The heat capacities were investigated by Professor Frank T. Gucker and his associates at Northwestern University, using the same preparations we had used. The apparent molar heat capacities, at a given temperature and concentration, were always much lower for ionic dipoles than for their uncharged isomers; the difference was of the order 30 $cal \cdot K^{-1} \cdot mol^{-1}$ at all temperatures, but the apparent molar heat capacities for all these substances varied with temperature far more than the molar volumes did.

Effects of Nonpolar Side Chains on Apparent Molar Heat Capacity

In 1935, I examined published data on the apparent molar heat capacities of some fatty acids and alcohols of varying chain length, as pure organic liquids and in aqueous solutions. For the pure liquids the increment in molar heat capacity near room temperature for each added CH_2 group in the chain was around 6–8 cal; the value for solutions of the same compounds dissolved in water was 20–30 cal. The large difference was remarkable, but no one apparently had commented on it before. The introduction of nonpolar residues into water clearly produced some important interactions at the interface between the two components. Such effects have been intensively studied by a number of authors in more recent years, including Julian Sturtevant, Peter L. Privalov, Ingemar Wadso, and Stanley J. Gill. They are of obvious importance in relation to protein folding and unfolding, since the latter process exposes many hydrophobic side chains to direct contact with water—side chains that were buried inside the compact structure of the native protein.

Discovery in England with Great Implications: Work of Bernal and Crowfoot on Protein Crystals

In 1934, 60 years ago, J. D. Bernal and Dorothy Crowfoot (later Dorothy C. Hodgkin) discovered how to obtain X-ray diffraction spectra of high quality from protein crystals. It was essential to keep them moist, for water is an integral component of such crystals. Bernal and Crowfoot placed the crystals in capillary tubes, in moist air, and sealed the tubes. When exposed to a suitable X-ray beam, the crystals gave sharply defined diffraction patterns with thousands of reflections whose intensities could be measured. Potentially the amount of information in these patterns concerning protein structure was enormous, but the structures could not be solved from the intensities alone; it was necessary to know their phases also. The phase problem had been solved for crystals of many small molecules, but then and for years thereafter there was no good way of solving the phase problem. Dorothy Hodgkin (1910–1994) was to become one of the supremely great X-ray crystallographers of our time; her work on cholesterol, penicillin, the vitamin B12 coenzyme, and insulin showed her great insight into difficult and unusual structures. She was not only a great scientist, but as a person she was universally admired and beloved. Her recent death was a loss to all who knew her.

My Indebtedness to Linus Pauling and E. Bright Wilson

About 1930, I first became aware of the work of Linus Pauling (1901–1994) on the nature of the chemical bond. I was planning to work, as I later did, on the Raman spectra of aqueous solutions of amino acids, and other acids and bases, over a substantial range of pH. The Raman spectra in such solutions would reveal the characteristic vibrational frequencies of these compounds at various pH values. It should be a straightforward matter to observe the characteristic vibrations of the amino and carboxyl groups, in the ionized and the unionized state, in simple carboxylic acids and amines. One could then study the amino acids also as a function of pH and see directly from the vibrational frequencies of the ionizable groups whether the amino acids, near their isoelectric points, behaved as ionic dipoles or as ordinary polar molecules. Many people were obtaining Raman spectra of organic substances as are pure liquids, but I think I was the first person to use aqueous solutions in this way. It meant that I had to detect the spectra of compounds at the order of only 1 molar concentrations in water, much lower than for pure liquids. This required a spectrograph of high light-scattering power, which I got from Adam Hilger in London, with a grant from a fund at Harvard. Once I got the system working, the results gave direct proof of the ionic dipole structure of the amino acids. I also studied the effects of deuterium substitution for ordinary hydrogen on the spectra of the ionizable groups and found many other correlations between Raman frequencies and structure.

To properly understand Raman spectra, one needed to know some quantum mechanics. Fortunately for me, the *Introduction to Quantum Mechanics* by Linus Pauling and E. Bright Wilson, Jr. just appeared. Wilson, who had got his PhD with Pauling at Caltech, had just come to Harvard and offered a seminar course in quantum mechanics, using the book as his text. I attended the course regularly and found it very valuable. Wilson, who became one of the great chemists of the 20th century, was also an excellent teacher. He gave me the insight I needed for the study of Raman spectra and provided a valuable introduction to quantum mechanical concepts in general, although I never attempted to master the subject fully.

Pauling's great book, *The Nature of the Chemical Bond*, was of fundamental importance for me and greatly enlarged my view of chemistry in general. The two concepts in the book that were most directly important for me were the role of hydrogen bonds and the concept of resonance. Pauling was not the first person to discover the hydrogen bond, but he made more powerful use of it than anyone else. I remember his coming to Harvard Medical School about 1933 and lecturing on the hydrogen bond, saying that he was convinced that it must play an important part in the functioning of biological systems. That talk made a great impression on many of his listeners, and his prophecy concerning its biological significance was soon verified. Certainly it is fundamental to the understanding of the nature of ice and liquid water, and countless other phenomena in biochemistry. The concept of resonance was a valuable guide to the understanding of structures that contained bonds intermediate in character between the classical double and single bonds. Resonance stabilized such structures as the carbonate and nitrate ions, the guanidinium group in arginine, and the urea molecule, as flat structures, with all the atoms coplanar. This is because of the large amount of double bond character, imparted by resonance, to bonds that had previously been considered to be single bonds. For protein chemists the most important example of such a system is the peptide bond, which appears in every link between amino acid residues in proteins and peptides. For me at the time this concept was vital for the interpretation of my findings in the Raman spectrum studies. Later, it was crucial for Pauling himself, when he proposed the α-helix and the pleated sheet structures, both of which proved to be of such central importance in the structures of proteins.

In 1940–41, I had my first sabbatical year since I had started teaching at Harvard in 1928. With the aid of a Guggenheim Fellowship I was able to work that year in Pauling's laboratory at Caltech and live in Pasadena with my family. That was a wonderful experience, and it was also my first opportunity to come to know the American West. It was also crucial for the completion of the manuscript of the book, which was my principal activity during that academic year, with frequent letters and queries to Cohn and Scatchard, back in Boston, as the final stages of the writing were reached. Caltech was a wonderfully stimulating environment with many gifted people around me, especially (for me) in the chemistry and biology departments. The Paulings welcomed us, and Linus was full of stimulating ideas. R. B. Corey, E. W. Hughes, and others were doing basic X-ray work on the structures of amino acids and peptides, which was to serve as a major foundation for Pauling's great later contributions, which I have mentioned above. To me, as to many others, Pauling was the greatest chemist of our time, and I owe him a great debt.

That year at Caltech, with the freedom to concentrate on the completion of the book, was indispensable. Without that, the book might never have been finished. When I came back to Harvard in August 1941, Cohn had essentially mobilized the laboratory, on a war footing, to work out methods for fractionation of blood plasma proteins and to supply the armed forces with fractions suitable for clinical use. The members of our laboratory, with aid from others, were to work out the methods. Then, a group of seven major pharmaceutical firms would apply them, in consultation with the lab, for large-scale production. Substantial progress was already under way when the U. S. entered the war, after Pearl Harbor. That was the beginning of nearly four years of work under constant pressure to meet those goals. In the midst of it all, the manuscript of the book went off to the Reinhold Publishing Corp. in New York, as an American Chemical Society monograph. I read and corrected the proofs as they flowed in, and eventually the book appeared, still in the midst of the war, in 1943. Few people had time to read it then, but it made its way among the protein chemists and others. The publishers had underestimated its appeal; the first printing was only 1,500 copies, and a second printing was needed before long. Some 20 years later, another firm (Hafner, New York) reprinted it again.

Great Days of Protein Chemistry

The really great days of protein chemistry were still to come. In 1943, we did not even have a single complete amino

acid analysis for any protein. Fred Sanger's pioneering work on the sequence of insulin was completed shortly after 1950. About the same time, Max Perutz discovered how one could label a protein with a few heavy atoms in defined places to solve the phase problem in X-ray diffraction: John Kendrew's structure of myoglobin and Perutz's of hemoglobin followed in a few years. The true nature and function of nucleic acids, the genetic code, and the process of protein synthesis became understood in what now seems like an amazingly short time. Today we know perhaps a thousand three-dimensional protein structures and many thousands of sequences, and are beginning to understand the principles that govern protein structure and function. Engineering of designed protein structures is now in its beginnings and seems sure to develop rapidly in the next decade. The contrast with the early history I have been discussing is immense, but I have precious memories of the group in which I was involved, 50 years and more ago.