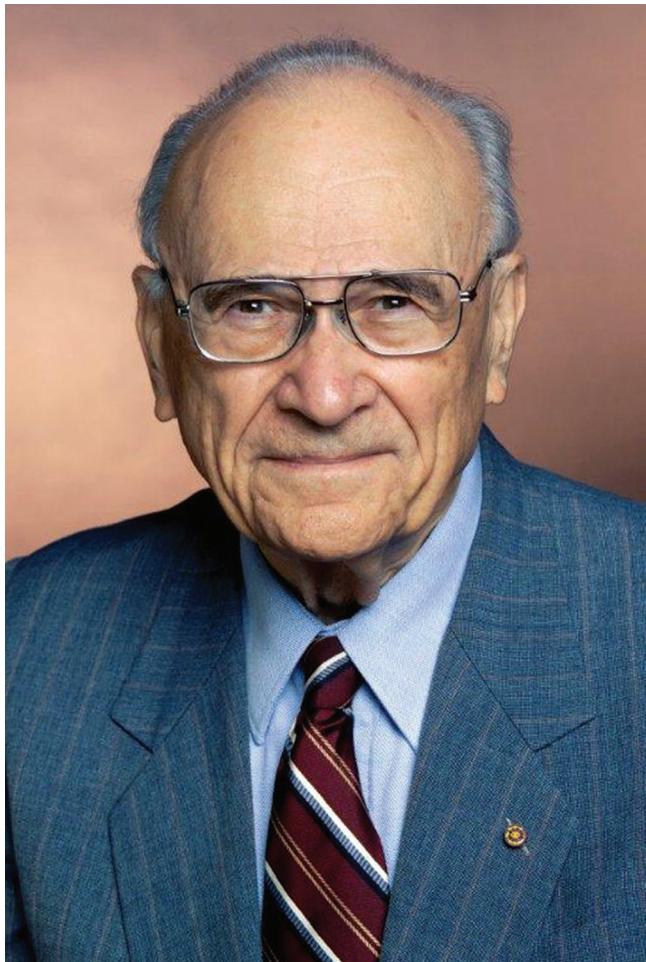


## Career Accomplishments of Harold A. Scheraga



Harold Scheraga has been a pioneer in the application of physical chemical methods to biological systems. Remarkably, his research program has spanned over six decades and is still going strong. His entire academic career has been at Cornell University where he received his initial academic appointment in 1947. He has combined both experimental and theoretical approaches to produce a new understanding of molecular interactions within and between proteins. In addition, he has been an exceptional mentor of young scientists both within his research program and in the classroom. He has also been an exemplary citizen of the university and scientific community, having served as Chair of the Department of Chemistry and on study sections and other review committees. What follows is a brief summary of Harold's many accomplishments. His career has also been highlighted in a recent *Annual Review of Biophysics* (40, 1–39 (2011)).

When Harold joined the chemistry faculty at Cornell University, protein biophysics was in its infancy; much of the knowledge and understanding we take for granted today was unknown. Proteins had not been sequenced, nor was there a detailed understanding of the three-dimensional structures and functions of DNA and proteins. Most especially for proteins,

Harold Scheraga would play a major role in elucidating the physical principles underlying their molecular behavior. He would go on to provide a molecular understanding of the forces responsible for protein structure and stability, how to predict the native structure of a protein, and how proteins adopt that state, i.e., the mechanism of protein folding. These have been the long-term objectives that Harold has pursued throughout his long and illustrious career.

His initial efforts at Cornell concentrated on hydrodynamic experiments to determine the size and shape of proteins. At the time, there were three-competing theories, Kirkwood–Riseman, Debye–Bueche, and Flory–Fox, for explaining the hydrodynamic properties of solutions of synthetic polymers. In collaboration with Nobel Prize winning polymer chemist Paul Flory, based on results from ultracentrifugation and viscosity measurements on polyisobutylene, they concluded that Flory–Fox theory was closest to experiment. With confidence in the Flory–Fox hydrodynamic theory of spherical polymers, Harold and a Flory postdoc, Leo Mandelkern, extended the theory to flexible ellipsoidal models of proteins, and found that many earlier hydrodynamic studies of protein solutions were misinterpreted. For example, fibrinogen from blood plasma was previously alleged to be a rod-like molecule with an axial ratio of 18:1. However, Scheraga–Mandelkern theory showed that it was 5:1. This was in agreement with the value proposed by Cecil Hall on the basis of electron microscopy.

At the same time, Harold, with his first graduate student, Michael Laskowski, Jr., investigated the thrombin-induced conversion of fibrinogen to a fibrin clot, a fundamental process in blood clotting, and deduced a reversible kinetic mechanism for fibrin clot formation. These studies also provided molecular insight into a bleeding disorder arising from a mutation in fibrinogen's amino acid sequence. These studies of fibrinogen were extended in later years to provide more of the molecular details of the blood clotting process.

A dramatic series of events occurred in the early 1950s that transformed the view of a protein from a colloidal particle (for which one could determine its size and shape from hydrodynamic measurements) to a real molecular entity: Fred Sanger's determination of the amino acid sequence of insulin and Linus Pauling's proposal of the  $\alpha$  and  $\beta$  secondary structures, with emphasis on backbone hydrogen bonds. Simultaneously, Crick and Watson proposed the double-helical structure of DNA. This led to a large effort in many laboratories to study  $\alpha$ -helix–coil transitions in polypeptides (and polynucleotides) as a model of protein unfolding and folding. In a tour de force, Harold, together with many students, used random copolymers of amino acids in a host–guest technique to determine the helix-forming tendency of all 20 naturally occurring amino acids. With Douglas Poland, Harold analyzed the statistical mechanical basis for phase transitions in

**Special Issue:** Harold A. Scheraga Festschrift

**Published:** June 14, 2012



polyamino acids and DNA models. This work culminated in the well-known book, *Theory of Helix–Coil Transitions in Biopolymers: Statistical Mechanical Theory of Order–Disorder Transitions in Biological Macromolecules* (1970), by Poland and Scheraga that still serves as the fundamental reference for the field. In the past few years, many physicists have taken up the Poland–Scheraga model of poly- and oligonucleotide DNA melting to consider salt and other effects on melting.

For polyamino acids, the  $\alpha$ -helix–coil transition reflects short-range interactions with a temperature distribution of many helical and coil states; thus, the helix–coil transition is continuous. However, globular proteins generally undergo a two-state transition, where the protein is either fully unfolded or folded with intermediates at best sparsely populated. Such a cooperative first-order phase transition reflects the relative roles that both short- and long-range interactions play in determining the character of the folding transition. Thus, as pointed out by Hao and Scheraga, for protein denaturation and folding, the short-range Ising model is invalid.

Whereas Pauling focused on hydrogen bonding of the backbone to propose the  $\alpha$  and  $\beta$  structures, Laskowski and Scheraga considered the role of hydrogen bonds involving ionizable side chains. They had previously deduced a role for such side chain–side chain hydrogen bonding in the thrombin-induced conversion of fibrinogen to fibrin. They pointed out that modified pK's suggest that ionizable groups may be involved in side-chain hydrogen bonding. This concept was applied to locate side-chain hydrogen bonds in proteins from observations of modified pK's. From a large series of physical and related biochemical experiments on bovine pancreatic ribonuclease A (RNase A), in which three tyrosines and three carboxyl groups had abnormal pK's, it was possible to determine three Tyr…Asp distance constraints. Since there are six tyrosine and eleven carboxyl groups in RNase A, over 19,000 possible ways of pairing the three tyrosines with three carboxyls exist. The three pairs predicted by Scheraga were found in the subsequently determined X-ray crystal structure of RNase A.

While focus was placed on the effect of hydrogen bonding on ionizable side chains, consideration also was given to the interactions between various nonpolar side chains, i.e., the hydrophobic effect, which is one of the dominant forces driving protein folding. George Némethy and Harold Scheraga formulated a statistical mechanical model of the hydrophobic effect, in which water plays a key role. This investigation was carried out in three parts. First, the structure and thermodynamic properties of liquid water and liquid D<sub>2</sub>O were considered. Then, aqueous solutions of aliphatic and aromatic hydrocarbons were treated to characterize the hydrophobic hydration of these nonpolar solutes and the hydrophobic interactions between them in aqueous solution. Finally, the thermodynamics of pairs and multiples of the hydrophobic interactions between the side chains of all naturally occurring nonpolar amino acids in proteins were computed. A subsequent series of experiments verified the magnitudes of the computed thermodynamic quantities. Years later, with the introduction of molecular mechanics, their model was confirmed by molecular dynamics simulations of aqueous solutions of hydrocarbons.

For many years, an unanswered question was how side chain–side chain hydrogen bonding could help stabilize proteins because hydrogen bonds between side chains and water have to be broken on forming side chain–side chain hydrogen bonds, resulting in essentially no free energy change.

With the insight that the nonpolar portions of so-called polar side chains can be involved in hydrophobic interactions with nearby nonpolar side chains, increased strength could be provided to the hydrogen bond, thereby affecting protein structure and stability.

The acquisition of the three Tyr…Asp distance constraints in RNase A motivated Harold to develop a molecular mechanics approach to compute protein structure. This was viewed as an impossible goal by many prominent protein chemists. Together with George Némethy, he initiated pioneering computational research with this approach, using a hard-sphere potential to eliminate unrealistic backbone conformations. This soon evolved into a more realistic all-atom force field, ECEPP (empirical conformational energy program for peptides) which, together with global optimization of the potential energy (later free energy) to obtain the thermodynamically most stable state, was applied to compute stable structures of cellulose chains, linear and cyclic peptides, fibrous proteins such as collagen and silk, enzyme–substrate complexes, and the entropy change accompanying hydrophobically induced protein–protein interactions. The largest globular protein structure that Harold, together with Jorge Vila and Daniel Ripoll, was able to compute on a large, parallel computer with the all-atom ECEPP force field, plus one of his global optimization algorithms, electrostatically driven Monte Carlo, was the 46-residue three-helix protein A, with good agreement with the experimental structure.

Despite the success with protein A, he realized that large proteins could not be treated with the then available, or foreseeable, computer power. Thus, together with Adam Liwo, he developed a coarse-grained, rather than an all-atom, model and force field, namely, UNRES (a united residue potential function). An early achievement with UNRES, together with a conformational space annealing global optimization algorithm (developed with Jooyoung Lee) was the computation of the structure of the protein HDEA which had an rmsd from the experimental structure of 4.2 Å for residues 25–85 and of 2.9 Å for residues 16–42.

Because UNRES averages out unnecessary degrees of freedom, the success with a constant-temperature UNRES model encouraged the development (together with Mey Khalili and Adam Liwo) of an algorithm to simulate protein folding pathways by molecular dynamics with the UNRES model. With a new molecular dynamics procedure that employs Lagrange equations of motion with Langevin dynamics in the canonical ensemble, the structures of single-chain proteins up to 75 residues were obtained with rmsd's of about 4 Å. This molecular dynamics procedure was extended (with Ana Rojas) to protein dimers, and ultimately to A $\beta$ , a multichain complex. A further development was the introduction of temperature into UNRES to optimize the free energy rather than potential energy, and thereby compute not only structures along folding pathways but also thermodynamic properties. Further enhancement of UNRES was the introduction by Liwo of a high degree of parallelization that extended its applicability to proteins containing several hundred residues. This facilitated the treatment of large protein–protein complexes. The force field and theoretical treatments have been refined throughout the years to permit the treatment of large proteins and protein–protein complexes. The methods developed have proven to be among the very best for predicting protein structures and folding pathways.

In parallel with these theoretical protein folding simulations, Harold carried out an experimental program to determine the structural pathways and kinetics of the oxidative folding of the four-disulfide protein RNase A, with David Rothwarf, and a ribonuclease homologue, the four-disulfide protein onconase, with Robert Gahl. Despite similarities in sequence and three-dimensional structure, these two proteins fold along very different (and multiple) pathways.

Harold's most recent new activity, together with Jorge Vila, is the use of density functional theory to compute the  $^{13}\text{C}^\alpha$  chemical shifts of proteins as a means of validating ensembles of NMR-determined protein structures and for use of NMR to determine protein structure. To facilitate the use of density functional theory-computed  $^{13}\text{C}^\alpha$  chemical shifts by NMR spectroscopists, a server, CheShift, has been made available. Any NMR-determined conformation can be entered and the density function theory-calculated  $^{13}\text{C}^\alpha$  chemical shifts of all its residues obtained, thereby facilitating the evaluation of conformational ensembles.

Not only has Harold excelled in research, he has also been an outstanding teacher. Harold has trained an extraordinary number of research scientists. Over 385 students and postdoctorals have passed through his laboratories. Many of these people have gone on to outstanding careers in biophysical chemistry. He always takes the time to interact closely with the people in his laboratory and to know the details of all of the individual research projects of his students and postdoctorals. He was also a brilliant class room teacher. His *Physical Chemistry of Proteins* course was popular throughout Cornell and was taken by over 1000 students. His teaching was not confined to advanced students. He loved to teach introductory physical chemistry and received accolades from the students. He also volunteered to teach freshman chemistry, and, as would be expected, excelled at this also.

Harold served as Chair of the Department of Chemistry at Cornell for seven years. As with research and teaching, his accomplishments were remarkable. He recruited a large number of faculty, both senior and junior, oversaw the construction of a major research wing, and laid the groundwork for a complete renovation of the old chemistry building. He expanded the outlook of the department to include chemical treatments of both biology and material sciences. The reputation of the Cornell Chemistry Department was greatly enhanced under his guidance, and the department was consistently ranked among the top ten in the country.

Finally, Harold served with distinction on NSF and NIH study sections, was a member of the editorial boards of multiple journals, and coauthored a NRC report on the importance of training in macromolecular science in a chemistry curriculum.

The career accomplishments of Harold Scheraga in research, teaching, and service are truly exceptional. He has more than 1250 publications in his bibliography. Moreover, he publishes very few short notes and never puts his name on a publication in which he did not play a major role. He has been very well recognized for his achievements, as detailed in his curriculum vitae. These awards include election to the National Academy of Sciences and the American Academy of Arts and Sciences and many honorary degrees and professorships. The specific awards for his research include the ACS Eli Lilly Award in Biological Chemistry, the Townsend Harris Medal, the Nichols Medal, the Pauling Medal, the ACS Kendall Award in Colloid Chemistry, the Linderström-Lang Medal, the ACS Award in Polymer Chemistry, the ACS Repligen Award for Chemistry of

Biological Processes, the Stein and Moore Award (Protein Society), the ACS IBM Award for Computers in Chemistry, the ACS Ralph Hirschman Award in Peptide Chemistry, and the ACS Murray Goodman Award.

At the tender age of 90, Harold maintains a very vigorous and dynamic research program in his quest to solve protein folding and structure problems. He continually develops new methods and ideas that reflect his ongoing passion for science. The protein community is most fortunate to have one of its founders who is not only active but is still at the forefront of cutting edge science.

Gordon G. Hammes

Jeffrey Skolnick