

The 1985 Nobel Prize in Chemistry

WAYNE A. HENDRICKSON

THE ROYAL SWEDISH ACADEMY OF SCIENCES HAS CHOSEN to honor Herbert A. Hauptman and Jerome Karle with this year's Nobel Prize in Chemistry. Karle and Hauptman are recognized for their pioneering development of direct methods for the determination of crystal structures. "Direct methods," in the parlance of crystallographers, concern the use of mathematical relationships among the diffraction data from a crystal in order to solve the "phase problem" and thereby produce an image of the atomic structure. Computer programs based on direct methods are now used routinely to solve several thousand crystal structures each year.

The announcement from Stockholm specifically cites work done by the prizewinners between 1950 and 1956. Today's successful programs are the fruit of intellectual seed sown in those years. However, these early ideas took some while in gestation. If the letter of Nobel's will were to have been followed, the present award should probably have made in 1954 (although that year's choice of Linus Pauling can hardly be questioned). It was in 1953 that Hauptman and Karle published a monograph entitled "Solution of the Phase Problem I. The Centrosymmetric Crystal." This recondite mathematical treatise seems to bear scant resemblance to chemistry, and its boldly pronounced solution of a previously intractable problem was greeted with skepticism and some resistance. However, the impact in the long run has been tremendous. The monograph first plowed the new and fertile ground of using probability theory to attack the crystallographic phase problem. This probabilistic approach was generalized by Karle and Hauptman in 1956 to include non-centrosymmetric crystals and it has prevailed in practical implementation. The routine solution announced then is indeed a reality today for structures of considerable complexity.

Fortunate it is, then, that the will of Alfred Nobel to recognize "the most important chemical discovery or improvement" contributed "during the year immediately preceding" is not taken too literally: the Code of Statutes established in 1900 to govern the Nobel Foundation interprets Nobel's intention regarding current eligibility so as to include "works or inventions of older standing" if "their importance have not previously been demonstrated." A goodly measure of the celebrated stature of the Nobel Prizes in science can safely be ascribed to the wisdom of tempering the wish to single out specific accomplishments of immediate impact with a test of experience as a proof of significance.

Crystal structures in chemistry. Three-dimensional structure is a critically important ingredient of chemistry. It is essential to know how the atoms are disposed within a molecule in order to understand its chemical bonding and to make sense of its reactions and interactions with other molecules. There are several ways by which such structures can be deduced, but the analysis of x-ray diffraction data from crystals is far and away the most significant of these

techniques. Results from x-ray crystallography are both comprehensive and definitive—one obtains a detailed picture of the entire structure, and there is near certainty that a properly done structure will be correct.

The importance of crystal structures in chemistry is certainly not new. Indeed, much had already been learned from crystal structures before Karle and Hauptman published their first paper together in 1950. Most of these analyses were based on Patterson methods, often with heavy atoms, and many involved exceptionally clever work. However, the implementation of direct methods in effective programs has truly transformed the field. Whereas in the early 1960's a single crystal structure could constitute a Ph.D. thesis project, now that same structure might be solved and refined within a couple of hours after the data collection (which itself is timed in hours) is completed. The advance is not only in speed. Direct methods are effective with equal atom structures, and this obviates concerns about chemical perturbations from heavy-atom labels introduced for phasing. Moreover, quite large structures can now be tackled: organic molecules with up to 50 nonhydrogen atoms are almost always routine, those of 50 to 100 atoms generally cause little trouble, and structures have been done with more than 200 atoms in the asymmetric unit of the crystal.

As complicated structures have come to be worked out quickly, reliance on crystallography has increased. It remains indispensable for details of stereochemistry, but now a crystal structure is often also the approach of choice for simple chemical characterization of natural products or intermediates in a chemical synthesis. For some time, the main chemical journals have carried an expectation that reports of novel compounds should be substantiated by crystal structure results. One measure of the impact of crystal structures is the nearly universal presence of crystallography laboratories (often a single scientist) in major chemical and pharmaceutical companies. These units each produce 50 to 100 structures per year (many never see the light beyond company doors), and in most places 90 to 95 percent are done by direct methods.

Atomic coordinates of published structures are maintained in various data bases. The largest of these is the Cambridge Structural Database of organic and metallo-organic crystal structures. This file now has 45,000 entries and 4000 to 5000 more are added each year. It is estimated that about three-quarters of these structures were determined by direct methods—in any case, only 4000 of the entries date from before 1970 which is about when direct methods programs came into widespread use. Another 25,000 structures are recorded in an inorganic data base, and a great many of these were also worked out with direct methods. Computers and diffractometers contribute greatly to present-day productivity, but it is the phase-determining methods that are most responsible for making crystallography a reliable "nonsporting" science.

The crystallographic phase problem. Before one can appreciate the achievement that is honored with this Prize, it is first necessary to have some understanding of the so-called phase problem. This problem arises because x-ray waves propagate through matter with

W. A. Hendrickson is a professor in the Department of Biochemistry and Molecular Biophysics, Columbia University College of Physicians and Surgeons, New York, NY 10032.

barely perceptible deviation from straight lines. Consequently, there are no x-ray lenses with which to form direct images of the atomic structures in crystals. The scattered radiation can be collected experimentally, but the measured intensities depend only on the magnitudes of the scattered x-ray waves while the phases of these waves are lost in the observational interaction. Once this phase information is recovered, an image of the structure as represented by the electron density distribution can be computed by a straightforward Fourier synthesis.

The coherent scattering from a crystal is restricted to discrete directions, known as reflections, that are governed by the crystal lattice. Reflection intensities depend on the atomic structure as well as on various experimental factors. It is structural dependence that interests us here, and this is given by diffraction theory as

$$F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\phi_{\mathbf{h}}) = \sum f_j \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j)$$

whereby the structure factor, $F_{\mathbf{h}}$, for an allowed direction, \mathbf{h} , depends on the sum of atomic scattering, f_j , from the atoms located at positions \mathbf{r}_j in the crystal. For simplicity, atomic motions are ignored here. Each structure factor depends on all atoms, and for a typical small molecule one measures about 70 independent $|F|^2$ values per unique atom.

Mathematically, then, the crystal structure problem can be viewed as a highly overdetermined system of equations with atomic coordinates as the unknown parameters. The transcendental character of these equations foils an analytical solution, but alternatively the atomic positions will be specified as the maxima in a Fourier series with the structure factors as coefficients. This approach requires that the phases be known. What Karle and Hauptman recognized from the start was that the redundancy in this system of equations implies that relationships must exist among the structure factors and that this interdependence might afford a solution to the phase problem.

Development of direct methods—the years together. Hauptman and Karle, Karle and Hauptman—these by-lines recur some 30 times during the 1950's and early 60's, and nearly all appear in the pages of *Acta Crystallographica*. The first in the series (also the first crystallographic paper by either author) is a landmark contribution from 1950. In it Karle and Hauptman examine the diffraction consequences of the physical constraint that the electron density within a crystal is positive everywhere. The elegant result is a certain set of determinantal inequality relationships among the structure factors. Harker and Kasper had already found some of these inequalities in 1948 in what can be considered the first direct methods paper. However, the determinants are comprehensive, and an important new relationship, later found to have profound implications, was discovered in the third rank inequality. By this, the phase of a reflection with indices $\mathbf{h}_1 + \mathbf{h}_2$ is bounded in relation to the phases of the two reflections \mathbf{h}_1 and \mathbf{h}_2 . How restrictive the bound might be depends on the magnitudes of the three structure factors.

The inequality formalism gave only subtle guidance for judging the likelihood that a particular phase indication be correct. Thus the next step by Hauptman and Karle was to develop the appropriate probability theory. First, in 1952 and 1953 following similar work by Wilson, they found the a priori probability distribution for a particular structure factor, assuming that, apart from symmetry considerations, all atomic positions are equally likely. This set the stage for the 1953 monograph discussed earlier. Symmetry dictates two distinctive situations: the phases for centrosymmetric structures are restricted to two possible values, 0 or π , whereas they can assume any value between 0 and 2π in the non-centrosymmetric case. The monograph considers the simpler centrosymmetric case, and a 1956 paper by Karle and Hauptman makes the generalization to non-centrosymmetric crystals. In both treatments, conditional



Herbert A. Hauptman



Jerome Karle

joint probability distributions are used to find the a posteriori probability for a structure factor once the intensities and certain phases are known. A theory of structure invariants and seminvariants was introduced to select reflections of interest. The bottom line of the monograph is a step-by-step phasing procedure based on formulas that derive from the joint probability distributions. The formulas of the monograph, called sigma-1, sigma-2, sigma-3, and so on, correspond to certain of the inequalities and each has an associated probability measure.

Others have also played important roles in the development of direct methods. A particularly influential contribution was made by David Sayre in 1952 when he discovered that, for a structure composed of equal and resolved atoms, a simple Fourier convolution relates a given structure factor to all of the others. The terms of Sayre's equation are similar to those of the Hauptman-Karle sigma-2 formula and to the third-order Karle-Hauptman determinant. Despite their diverse origins—atomicity, conditional probability, positivity—all point to the same fundamental triplet phase relation:

$$\phi_{\mathbf{h}} = \phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}}$$

The great advantage of the probability approach is that it allows ready assessment of the likelihood that single-term phase indications are correct. This is critical in the early stages of a phase determination. Woolfson and Cochran made improved estimates of these probabilities. Another important practical consideration concerns the dependence of the strength of a relation on the number of atoms and the order of the relation. Here, Klug contributed in 1958.

After their 1956 paper on the non-centrosymmetric problem, which contains the famous tangent formula for phase refinement, Karle and Hauptman embarked on an alternative probability approach that leads directly to values for phase invariants (particular sums of phases such as $\phi_{\mathbf{h}} + \phi_{-\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}}$) in terms of the full set of all magnitudes. However, the most significant contributions after 1956 concerned the practical implementation of the direct methods, especially for application to non-centrosymmetric crystals. Isabella Karle played a very critical role in this regard. The procedures of the monograph had already been used, starting with colemanite in 1954, but it was Isabella's striking successes with cyclohexaglycyl (1963) and L-arginine (1964) that really convinced the skeptics. A recipe for structure solution by this new symbolic addition procedure was published by Karle and Karle in 1966. A next step was the dissemination of direct methods by way of computer programs. Here Michael Woolfson played an essential role. The multisolution techniques that he and his co-workers (particularly Main and Germain) introduced led to MULTAN, the most widely used of direct methods programs. Continued improvement in the theory

and implementation of direct methods goes on today in many centers around the world.

Karle and the Naval Research Laboratory. The prizewinning work that Karle and Hauptman did together was carried out at the U.S. Naval Research Laboratory (NRL) in Washington, D.C. Jerome and Isabella Karle came to NRL in 1946 and have been there ever since. In the beginning they followed up their doctoral research training with Lawrence Brockway and continued in gas electron diffraction. Indeed, Karle credits his use of physical constraints in the analysis of electron diffraction data for the concept of non-negativity as applied to the crystallographic phase problem. Eventually, the crystal structure problem consumed him, and Isabella as well, but he continued an active electron diffraction group until very recently.

In 1968 NRL established a Chair of Science for Karle as Chief Scientist of the Laboratory for the Structure of Matter. In that context he expanded to a strong research group of about a dozen scientists. He occasionally collaborates with these associates in diverse projects ranging from organic crystal structures to amorphous scattering and on to protein crystallography. However, for the most part he has fostered the independent careers of this highly regarded group. Karle's personal research effort continues to focus primarily on the theoretical aspects of direct methods. After Hauptman left to take up another position at NRL, this work has been done alone or, if tied to application, with Isabella. He has paid special attention to enhancing the procedures that are already in operation. However, he has also been heavily involved with new methods such as those for evaluating invariants. Most recently he has made several very promising contributions to the analysis of anomalous dispersion data.

Jerome Karle was born in New York City on 18 June 1918 where he attended Abraham Lincoln High School in Brooklyn and was graduated from the City College of New York in 1937 with a B.S. in chemistry and biology. He received an M.A. in biology from Harvard in 1938 and then, after a year and a half with the New York State Health Department in Albany, he entered the University of Michigan where he received his Ph.D. in physical chemistry in 1943. He and Isabella met and married at Michigan, and during the interim before moving to Washington he worked there on a research project for the Navy. During his early years at NRL until 1961, Karle also served as a professorial lecturer at the University of Maryland. Karle has served in many offices and been accorded many honors. Notable among these are his 1981–84 term as president of the International Union of Crystallography, his election to the National Academy of Sciences in 1976, and his receipt, with Hauptman, of the 1984 Patterson Award of the American Crystallographic Association.

Hauptman and the Medical Foundation of Buffalo. Hauptman joined the NRL in 1947 and not long afterward began his fruitful collaboration with Karle. Then, in 1965 he became head of the Mathematical Physics Branch of NRL. During this period he continued with direct methods developments as well as the new administrative responsibilities. He also initiated a collaboration with Dorita Norton of the Medical Foundation of Buffalo in applications of the new methods of steroids. Then in 1970 after 30 years with the federal government, Hauptman moved to the Foundation in Buffalo, New York. The Foundation is a small independent research institute that specializes in endocrine research. Since 1972 Hauptman has been its executive vice president and research director.

After moving to Buffalo, Hauptman participated vigorously in the structure determination work on steroids and other biological compounds. He also established a molecular structure group that is internationally renowned. However, he remains a mathematician at

heart, and his main activity continues in probability theory as applied to the phase problem. During the 1970's he made a number of significant advances in structure invariant evaluations, alone and together with associates. Notable among these are his investigations into quartet and quintet invariants and introduction of the neighborhood concept. More recently, Hauptman has attacked the problem of integrating the techniques of direct methods with anomalous dispersion and isomorphous replacement.

Herbert Hauptman was born on 14 February 1917 in New York City and attended Townsend Harris High School before going to City College where, like Karle, he graduated in the class of 1937. Hauptman's B.S. degree was in mathematics and he went on to obtain an M.A. in mathematics from Columbia University in 1939. Hauptman began his career with the government in Washington in 1940 and held various posts before joining NRL in 1947. In 1955 he was awarded a Ph.D. in mathematics from the University of Maryland based on research suggested by his work in direct methods. He maintained an association with Maryland as part-time professor while in Washington and he has been a research professor at the State University of New York at Buffalo since then. Hauptman was president of the Association of Independent Research Institutes in 1979 and 1980, and he shared the Patterson Award of 1984 with Jerome Karle.

Reactions and prospects. It would be very wrong to carry away an impression from this Prize that all crystal problems are routine or that crystallographers as such are made redundant by the successes of direct methods. While the programs do succeed with impressive regularity, they can and do fail. It is clever crystallographers who turn these failures into atomic coordinates. Direct methods are limited both by the inherent statistical character of the process and by the inverse dependence of phasing power on the number of atoms. The past decade has seen steady progress in the ability to cope with "pathological" cases and a gradual extension of the range of applicability. However, it appears that an asymptote for the routine of about 100 atoms is being approached and that radically new developments are needed to bring macromolecular structures into reach. Several exciting possibilities are active: one integrates anomalous scattering directly into the formulation of direct methods, another involves the methods of maximum entropy, and a third couples direct method into an algebraic analysis of multi-wavelength anomalous dispersion data. There is also the elusive prospect of exploiting prior knowledge about components of the molecular structure.

For crystallographers these cautions and hopes are important but secondary reactions. Within this community, the announcement of the 1985 Prize in Chemistry elicited a wave of pride and delight. Although neither Karle nor Hauptman has been an educator per se, both have been teachers to many of us through meetings and workshops as well as by example. Herb and Jerry are such fixtures at meetings of the American Crystallographic Association that possibly none has been held without one or both in attendance. The Nobel Prize is not new to this field as the impact of x-ray diffraction results in chemistry has been tracked in prizes to Pauling in 1954, Perutz and Kendrew in 1962, Crick, Watson, and Wilkins in 1962, Hodgkin in 1964, Hassel and Barton in 1969, Lipscomb in 1976, and Klug in 1982. However, not since the early prizes to von Laue in 1914 and to the Braggs in 1915 had contributions to diffraction analysis been directly recognized. This overlooks monumental achievements by Ewald and Patterson among others. It is a special pleasure now to see this fitting tribute to Karle and Hauptman. Their work is so basic that it affects many fields, but most profoundly chemistry—not least by transforming chemical crystallographers into crystallographic chemists.