CROATICA CHEMICA ACTA CCACAA **74** (1) 1–35 (2001)

ISSN-0011-1643
CCA-2712

Feature Article

(Bio)Crystallography at the Turn of the Millennium

Biserka Kojić-Prodić[a,*] *and Jan Kroon*[b]

[a] *Ruđer Bošković Institute, P. O. Box 180, HR-10002 Zagreb, Croatia*

[b] *Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

Received May 17, 2000; revised September 4, 2000; accepted September 11, 2000

This review is focused on the present status of X-ray crystallography and its impact on chemistry and the life sciences. The discovery of the DNA double-helix structure from a fibre X-ray diagram in 1953 marked the last century, whereas the human genome project (to be completed in the year 2003, the 50-th anniversary of Crick's and Watson's discovery) and the proteomics will open the new millennium. The crystal structure determinations of very complex biological assemblies such as viruses and ribosomes illustrate the power of contemporary X-ray structure analysis. The historical background is given, which touches on some of the important steps from the early days of the discovery of X-rays. The advanced and new methods and technologies of the modern era, such as third-generation synchrotrons, sensitive area detectors, up-to date computer technologies with computer graphics, cryo-techniques, microcrystallization methods, and genetic engineering, all contributing to the development of X-ray crystallography, are evaluated in separate paragraphs. Some examples that illustrate the power of the methods are selected from the contemporary research in the field. Systematics of known crystal structures of small molecules and macromolecules and their assemblies collected over the years has developed our perception of the nature of the chemical bond, and the interactions between atoms and molecules, which determine the chemical, physical, and biological properties. This fundamental knowledge enables structure and property predictions, useful in many branches of science and technology. A method, starting with the structure of table salt, has developed into a very fine tool for

* Author to whom correspondence should be addressed. (E-mail: kojic@rudjer.irb.hr)

looking into complex living systems. Time-resolved crystallography and cryo-electron microscopy are capable of recording biological events and they inform us about the dynamics of enzymes and living cell activities. In combination with bioinformatics, very fast events, recorded experimentally or proposed theoretically, can be interpreted by the methods of molecular dynamics simulations. The vivid interplay of ideas coming from different scientific fields and technologies has led to capital discoveries, which have opened roads to new disciplines, such as molecular genetics and molecular medicine. These two branches, in particular, enriched by findings on DNA and protein structures, can provide efficient therapies for many diseases, health prevention, and reduce ageing problems. Our views on the nature of the chemical bond have been revised and our horizons will be extended and clearer in the years to come. Applying an ethical approach in science, humanity will learn how to improve the quality of life all over the world.

Key words: review-X-ray crystallography, biocrystallography, protein crystallography, synchrotron, cryo-technology, macromolecular structures, molecular assemblies, structural chemistry, molecular modelling.

The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

W. L. Bragg

INTRODUCTION

Use of X-rays has revolutionized many branches of science, such as chemistry, biology, medicine, pharmacology, physics, and particularly material sciences and a variety of technologies. This review is focused on its impact on chemistry and the life sciences. Scattering, diffraction and absorption of X-rays can be used for examination of small molecules, but also for proteins and complex assemblies, such as viruses, and even living cells. X-ray structure analysis is the predominant technique for investigating molecular structure at atomic resolution in the solid state. The molecular topology in the crystal provides insight into the interatomic and intermolecular interactions, which determine the physical, chemical, and biological properties. Let us give a few striking examples that illustrate how X-ray structure analysis has revealed the nature of the chemical bond. The X-ray structure analysis of diamond, published in 1913,[1] revealed tetrahedrally disposed carbon atoms and thus confirmed the assumption of Le Bel[2] and van't Hoff [3] in the

last decade of the 19th century. In 1891, E. Fischer and A. von Bayer[4,5] employed a tetrahedral concept to understand the chemistry of sugars, which led them to the problem of optical isomerism and conformational differences. They were able to assign the relative configurations for enantiomers of various sugars and derivatives and their work established a new scientific discipline – stereochemistry. In 1951 came the evidence that, by mere chance, Fischer's assignment of glyceraldehyde was correct. J. M. Bijvoet[6] discovered how X-ray anomalous scattering can be used to determine the absolute configuration of chiral molecules exemplified by the structure of sodium rubidium tartrate. The Bijvoet method and Barton's principles of conformational analyses accelerated the development of stereochemistry. These findings were followed by Noble prizes awarded to R. B. Woodward, D. H. R. Barton, O. Hassel, V. Prelog, and J. W. Cornforth for stereochemistry related research. It became clear that molecules of life, such as amino acids, sugars and many natural compounds, are chiral. The chiral origin of life still remains an enigma. There are more examples that illustrate how our perceptions are subject to changes. Table salt was treated as a molecule but after its X-ray analysis it turned to be a three-dimensional array of sodium and chlorine ions held together by Coulomb interactions. In the early days of the 20th century, it was difficult to reconcile the new view of ionic compounds with the classical chemical ideas of that time. The crystal structure of penicillin, solved by D. Hodgkin in 1949, revealed for the first time the presence of a four-membered β-lactam ring,[7] unknown to organic chemists of that time. Her determination of the vitamin B-12 structure showed the existence of a metal-carbon bond and a corrinoid ligand system, not previously encountered in natural products.[8] A completely new field of organometallic chemistry emerged from this finding. Many examples, including myoglobin and hemoglobin structures, can be added to the list, but the elucidation of the DNA structure from an X-ray-fibre diagram in 1953[9] was undoubtedly of paramount importance. It opened up a new era of molecular biology and genetics, which led to the understanding of normal and pathological processes of living cells at the molecular level. Linus Pauling, using a rational approach to understanding the cause of sickle-cell anaemia at the molecular level,[10] envisaged molecular medicine in 1949. On the basis of X-ray structures of a wild-hemoglobin and its mutant-proteins, molecular pathology of hemoglobin was understood. It was discovered that in sickle-cell hemoglobin (Hbs) hydrophobic Glu6 A3(6)β is replaced by hydrophobic Val residue.[11,12] As a consequence the nature of interactions between two valine residues of two subunits, a hydrophobic patch is generated and molecules aggregate into long filaments. Such aggregates influence the red blood cells to distort into sickle shape, which leads to anaemia. This is among the first examples of understanding a disease at the molecular level. These days, molecular

medicine turns out to be a modern discipline aimed at understanding diseases, genetic anomalities, ageing problems at even the atomic level, and accordingly designing novel classes of molecules to combat illness.[13] The discovery of DNA structure and its genetic function and the human genome project strongly relate molecular biology and genetics with X-ray crystallography, enormously influencing life sciences. The human genome project[14] brings many scientific excitements and benefits, but it also calls for high ethical standards.

X-ray structure analysis has resulted in more than 250 000 known structures of small molecules and 11 466 macromolecular structures (stored in various databases such as the Cambridge Structural Database, Protein Data Bank, Nucleic Acid Database, Inorganic Crystal Structure Database, Metals Data File), which are accumulating every day. They constitute rich sources of information, which can be used in many ways.[15] In addition to simple bibliographic searches, data mining can also be employed in sophisticated ways, like for structure predictions, to verify theoretical models, to design molecules of predicted properties (*e.g.* design of drugs and substrate analogues for enzyme catalysts), for structural classification of proteins and correlation with their functions.

HISTORICAL BACKGROUND

Inspired by Maxwell's theory of electromagnetic waves, W. C. Röntgen discovered at the University of Würzburg, on November 8, 1895, a radiation of unknown origin and properties, which he called X-rays.[16] In order to make his discovery more convincing to himself and to many others, he asked his wife Bertha to put her hand over a film into a beam of »unknown rays«. The shadowgraph of a hand skeleton with a golden ring taken on December 22, 1895 circulated quickly all over the world, as it did a hundred years later at its centennial anniversary. The discovery gained enormous publicity because everybody realized the power of X-rays to reveal »the architecture of the invisible«. The long lasting wish to see the interior of the human body without using a scalpel became reality. The world realized the importance of Röntgen's discovery and the first Nobel prize for physics was awarded to him in 1901. X-rays were named röntgen radiation after him. Though scientists tend to shun superlatives, H. H. Seliger[17] nevertheless classified the discovery of X-rays as the most important application of a physical phenomenon in the history of science. The application of X-rays started only three months after the discovery. The ability to produce a shadowgraph of the internal structure of a wide variety of animate and inanimate objects has been extensively exploited in medicine and various technologies. Simple ra-

diography, used by M. Curie in movable units on the battlefields during the first world war in order to help wounded soldiers, has now developed into sophisticated computerized tomography with digital imaging. Powerful betatrons, gammatrons and mevatrons used in cancer therapy are also based on Röntgen’s discovery. Small high-energy synchrotrons adopted by hospitals are used as angiography diagnostic devices.[18–20] They are among the most advanced technologies in medicine. The early application of fluoroscopes for passenger control, started in France in 1897, has become common practice at airports. Fluoroscopy became an efficient method for catching smugglers of Egyptian mummies and archeological objects at the beginning of the 20th century, as it still is today in detecting armed terrorists. X-ray fluorescence analysis has developed into a powerful analytical method for the determination of trace elements, which has found its application in mineralogy, geology, archeology, ecology, pollution control of soil and water, and many other fields. Even astrophysics benefits from Röntgen’s discovery. The world’s most powerful X-ray telescope carried by Ariane 5’s rocket was launched in December 1999 within the European Space Agency programme.[21]

Seventeen years after Röntgen's discovery, Max von Laue suggested that X-rays can be diffracted or scattered by a crystal.[22] The experiments performed immediately after that by W. Friedrich and P. Knipping[22] were aimed at resolving the controversy over the nature of X-rays. The experiments revealed its wave nature and confirmed von Laue's idea,[23] for which he won the Noble prize in 1914. A new branch of science based on this discovery saw the light in November 1912 when the 22-year-old Lawrence Bragg interpreted the complicated pattern of Laue's photograph. He applied the very simple idea coming from the optics of visible light that X-rays are reflected from the atomic planes in the crystal (Bragg's law).[24] Together with his father, W. H. Bragg, he built the first spectrometer,[25] which enabled the structure determination of some minerals, such as diamond and table salt. The most striking idea in that area, put forward in 1915 by W. H. Bragg,[26] involved the representation of the periodic repeat of atomic patterns in the crystal by Fourier series. This mathematical formulation was a milestone in X-ray structure analysis. In 1915, the Nobel Committee recognized the great scientific discovery of the Braggs and awarded them a prize. According to W. L. Bragg, »the development of X-ray crystallography since 1912 has more than fulfilled our early expectations. It not only has revealed the way atoms are arranged in many diverse forms of matter, but also has cast a flood of light on the nature of the forces between atoms and the large-scale properties of matter. In many cases this new knowledge has led to a fundamental revision of ideas in other branches of science«.[27] W. L. Bragg's predictions of X-ray crystallography given in 1968 have been amply confirmed by the determination of many complex protein structures such, as vi-

ruses, *e.g.* the blue tongue virus[28] and Norwalk virus,[29] and the ribosome structure.[30–42] During its 88 years of developement, X-ray crystallography has opened up new ways of studying matter, providing scientists with a very special kind of powerful »spectacles« to see the invisible architecture of the molecules in crystals and it has so far harvested 15 Nobel prizes. On the occasion of the centennial anniversary of Röntgen's discovery, the University of Würzburg established a foundation for the Centennial Röntgen Award for Life Sciences, Medicine, and Physics; the first was given to Prof. Axel T. Brünger of the Howard Hughes Medical Institute, Yale University, for the application of molecular dynamics simulations in the refinement of protein structures.

FUNDAMENTALS OF X-RAY DIFFRACTION

The principles of X-ray diffraction are well established. X-ray scattering originates from electrons (Thomson scattering). An incident wave induces other scattering centres to produce this effect and these individual waves interfere. The resulting scattered wave coming from the object (array of atoms in the crystal) depends on its internal structure, *i.e.* geometrical pattern and the scattering power of each atom. First, consider a crystal as consisting of periodically arranged building blocks (called unit cells). Because of this periodicity, no scattering will be observed when scattered waves from these unit cells are not exactly in phase; this is because the interference is completely destructive. Therefore we need to consider the scattering from the unit cell only. When a volume element at position $\boldsymbol{r}$ (defined with respect to the origin of the unit cell) is irradiated by a beam that propagates in the direction $\boldsymbol{s}_0$ ($s_0 = 1/\lambda$), the amplitude of the scattered beam is proportional to $V\rho(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$ [$\rho(\boldsymbol{r})$ being the electron density at the position of vector $\boldsymbol{r}$; vectors are printed in bold face italic type and scalars in italic type]. In order to take into account the interference with beams scattered by electron densities at other positions, we should consider its phase (conveniently expressed with respect to the origin of the unit cell). When scattering occurs in direction $\boldsymbol{s}$, this phase is calculated to be $2\pi\,\boldsymbol{S}\cdot\boldsymbol{r}$ (with $\boldsymbol{S} \equiv \boldsymbol{s} - \boldsymbol{s}_0$).

The total scattering over all elements of the volume V of the unit cell, normalised to the scattering of a free electron, takes the functional form of a Fourier transform:

$$F(\boldsymbol{S}) = \iiint \rho(\boldsymbol{r}) \exp\,(2\pi i\ \boldsymbol{S}\cdot\boldsymbol{r})\ \mathrm{d}V$$

This represents the structure factor; it is a fundamental expression in X-ray crystallography.

In X-ray structure analysis, the Fourier inversion is used.

Because $F(\boldsymbol{S}) = \mathscr{F}[\rho(\boldsymbol{r})]$

$$\rho(\boldsymbol{r}) = \mathscr{F}[F(\mathbf{S})]$$

The crucial problem is that $F(\boldsymbol{S})$ is a complex quantity; $F(\boldsymbol{S}) = |F(\boldsymbol{S})|e^{i\varphi_{\boldsymbol{s}}}$; it has a magnitude and a phase. We are measuring only the scattered intensity, and this is proportional to $|F(\boldsymbol{S})|^2$, so the phase is not accessible by the diffraction experiment. This central problem of X-ray analysis is known as the phase problem. To obtain the desired electron density $\rho\boldsymbol{r}$**,** whose maxima reveal the positions of the atoms in the unit cell, the phase problem has to be solved. Diffraction from the crystal, by its periodic nature, is restricted to certain directions. They are subject to the so-called Laue conditions: $\boldsymbol{S} \cdot \boldsymbol{a} = nh$, $\boldsymbol{S} \cdot \boldsymbol{b} = nk$, $\boldsymbol{S} \cdot \boldsymbol{c} = nl$; h, k, l are integers, relatively prime to each other, n is an integer factor.

Bragg derived a more manageable equation, considering diffraction as »reflection« from a series of equidistant and equivalent planes in a lattice:

$$2\, d_{hkl} \sin \Theta = n\lambda$$

d_{hkl} is the distance between neighbouring lattice plans; h, k, l (identical to the integers appearing in the Laue equations) specify the lattice planes and are termed Miller indices, n is called the order of reflection. In diffraction, it is common practise to absorb n in h, k, l so that the latter symbols represent the Laue indices. Θ is the angle of incidence and reflection with respect to the lattice plane.

Structure factors obtained for a large number of reflections with Laue indices $\boldsymbol{h} = (h,k,l)$ give the desired electron density [summed over all atoms (j) in the unit cell at the positions x_j, y_j, z_j]

$$\rho(\boldsymbol{r}) = 1/V \sum_{\boldsymbol{h}} F(\boldsymbol{h}) \exp(-2\pi i\ \boldsymbol{h} \cdot \boldsymbol{r}) =$$

$$= 1/V \sum_{\boldsymbol{h}} F(\boldsymbol{h}) \exp[-2\pi i\,(x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c}) \cdot (h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*)]$$

where $\boldsymbol{r} = x\boldsymbol{a} + y\boldsymbol{b} + z\boldsymbol{c}$, and $\boldsymbol{a}^*$, $\boldsymbol{b}^*$ and $\boldsymbol{c}^*$ are the vectors of the reciprocal lattice (mathematically defined by the relations $\boldsymbol{a}^* \cdot \boldsymbol{a} = 1$, $\boldsymbol{b}^* \cdot \boldsymbol{b} = 1$, ..., $\boldsymbol{a}^* \cdot \boldsymbol{b} = 0$, *etc*).

The phase problem can be tackled in two different ways: a) using vector methods requiring a heavy atom or a sufficiently large molecular fragment (strong scattering element) to be present in the structure, and b) application of direct methods (based on algebraic relations combined with information from internal crystal symmetry, and probability theory). In macromolecular crystallography, various combinations of heavy-atom methods are used: the single isomorphous replacement method (SIR), the multiple isomorphous

replacement method (MIR), and the multiple wavelength anomalous dispersion method (MAD). We shall briefly discuss the physical base of the anomalous effect in order to show its implications in X-ray diffraction. Electrons can be treated as oscillators; if the frequency of the primary beam approaches some of their eigen frequencies, resonance will occur with a concomitant additional phase shift in the scattered beam. The scattering under these conditions is considered anomalous and its contribution is added to the normal atom scattering factor ($f = f_0 + \Delta f' + i f''$, the last two terms being called the real and the imaginary dispersion corrections, the values of which depend on the wavelength and the anomalously scattering atom in the structure). One consequence of this effect is that Friedel's law is disobeyed ($F_{hkl} \neq F_{-h-k-l}$), as discovered by Bijvoet.[6] The Bijvoet differences are wavelength dependent and maximal at the absorption edge. Radiation of different wavelengths can be easily produced by synchrotrons. Thus, the departure from this law can be used for phasing (MAD)[43,44] and for the determination of the absolute configuration of molecules.

Molecular replacement is an additional possibility for phasing, using known structure fragments (homology search *via* Protein Data Bank) combined with translation and rotation functions to match fragments of a known structure with those in the unsolved structure. Non-crystallographic symmetry for phase averaging of noncrystallograhic units can be applied even between different crystal forms for improvement of poorly determined phases rather than for *ab initio* phase determination. Rossmann and Blow 1963[45] initiated the method while determinating virus structures. These days, Rossman phasing is a common practise in protein crystallography, particularly in solving virus structures.

The development of mathematical methods for obtaining the model of a crystal structure from experimental data and the improvement of its accuracy have been directed towards evaluation of measured data (data reduction), solution of the phase problem, density-map interpretation, and refinement of the model. Mathematical formulations and adequate algorithms are the basic elements for software development, which has to keep pace with the quickly advancing hardware. To illustrate the impact of computer technology, let us mention an example: in the late sixties the determination of the structure of vitamin B-12 ($C_{63}H_{88}N_{14}PCo$, totally 181 atoms in a molecule)[8] required more than eight years. Intensity evaluation was done by naked eye, data reduction and two-dimensional Fourier syntheses were calculated manually (using Beevers-Lipson strips invented in 1936, and a somewhat advanced punch-card method). Nowadays, such a determination would hardly take a day.

Among the mathematical formulations that have advanced crystallographic software are the fast Fourier algorithms, least-squares procedures

(such as simulated annealing), various combinations of Patterson searches (for the global shape of a molecule and its orientation, applying rotation and translation functions and a number of combinations and modifications thereof), and maximum entropy enhancement. There is quite a number of software packages available for X-ray structure analysis of small and large molecules. Great progress in the solution of complex protein structures has been achieved with the program packages CCP4,[46] AmoRe[47] and CNS (a new version of X-PLOR).[48]

SYMMETRY

The natural beauty of crystals, observed by the naked eye or light microscopy, derives from their symmetrical regularity, colours and brightness (Figure 1). However, the inner symmetry of the three-dimensional pattern within a crystal, revealed by X-ray diffraction, implies much more; it affects the properties of all crystalline material. The inherent symmetry of a molecule can also be depicted from the X-ray analysis (Figures 2a[49] and 2b[50]). In addition to recognising the 230 ways of periodical packing of a pattern (space groups), X-ray structure analysis can show internal non-crystallographic symmetry originating either from pattern irregularities or formation of molecular assemblies. In the regular crystal symmetry, there are only 1, 2, 3, 4 and 6-fold rotation axes and no five-, seven- or eight-fold rotation operations. However, five-fold symmetry[51] common in the living world (flowers, a sea-star, spider web) has been detected in virus structures, some protein structures,

Figure 1. A natural sample of the sulphate mineral celestite ($SrSO_4$) from Poland, private collection.

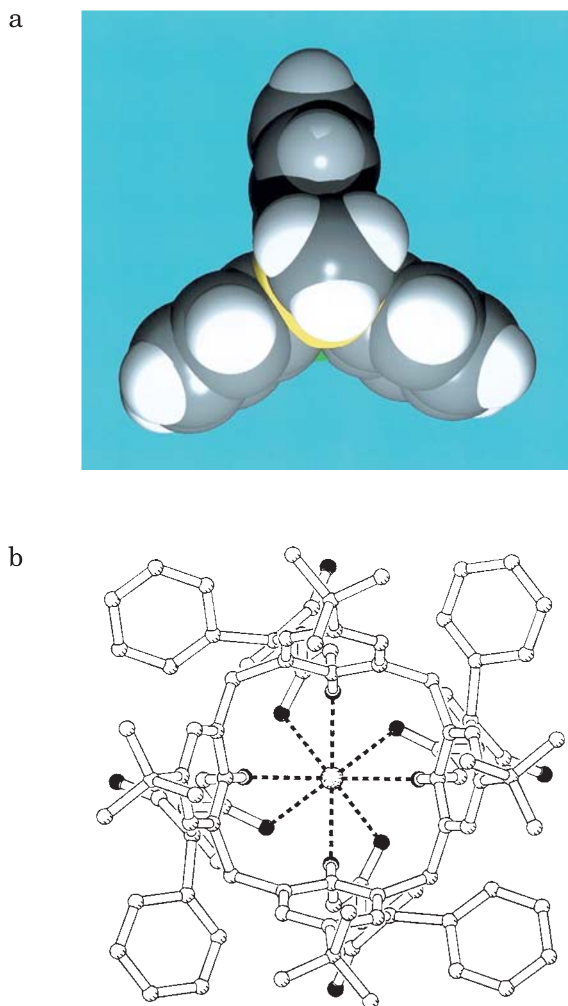

Figure 2. X-ray structure analysis reveals the molecular symmetry: a) molecule of As, Ge-triptycene with a three-fold symmetry axis,[49] and b) 5,11,17,23-tetra-*tert*-butyl-26,28,25,27-tetrakis(*O*-methyl-*R*-α-phenylglycylcarbonylmethoxy)calix[4] arene sodium perchlorate with a four-fold symmetry axis.[50]

and also fullerene molecules. The spherical virus morphology of icosahedral symmetry includes pentamers[52] represented here by a detail of the crystal structure of the human hepatitis B virus capsid[53] solved at 3.3 Å resolution (Figure 3). Illustrative examples of the biological function and assembly formation using seven- and eight-fold symmetry are also given. The α-haemolysin toxin produced by the pathogenic bacterium *Staphylococcus aureus* forms a

heptameric pore structure (solved at 1.9 Å resolution)[54] (Figures 4a and b). The heptameric membrane associated complex inserts through the lipid bilayers, forming transmembrane pores for rapid calcium influx, which can induce cell lysis.[55]

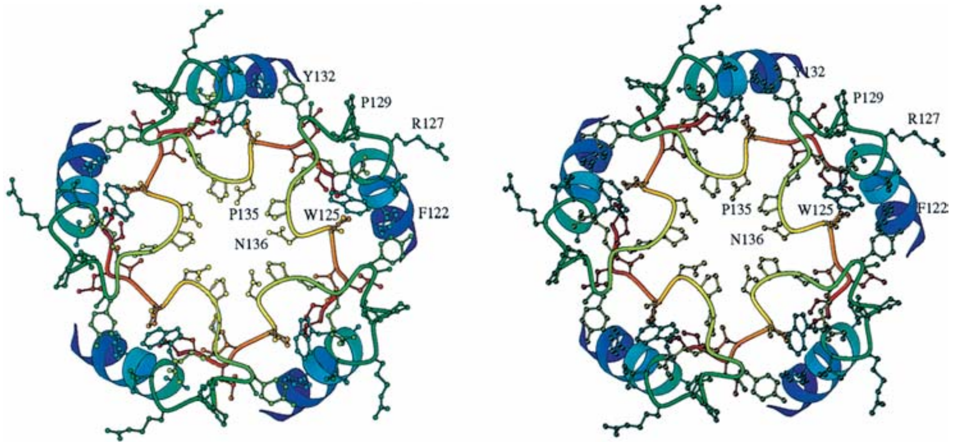

Figure 3. A stereo view from outside the capside of the human hepatitis B virus showing the interactions between adjecent subunits with five-fold symmetry.[53] Reprinted by permission from Cell[53] copyright (2000) Cell Press.

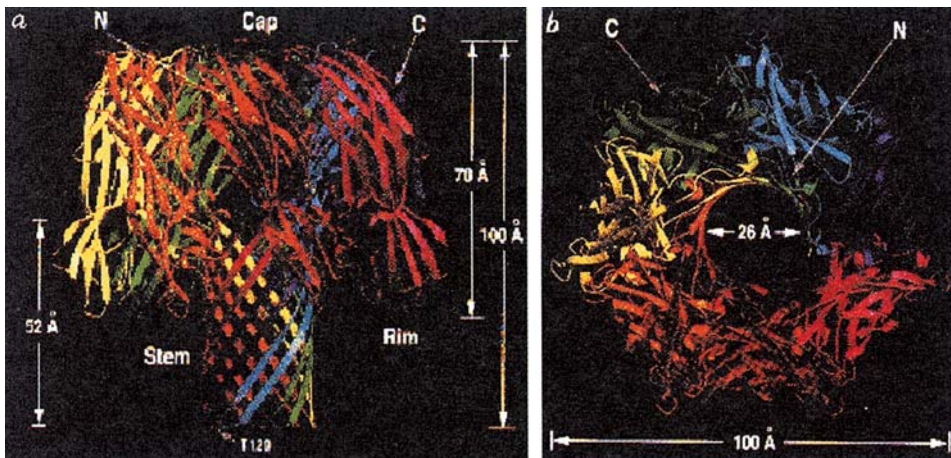

Figure 4. The α-haemolysin heptamer at 1.9 Å resolution (each promer has been coloured differently): a) view perpendicular to the seven-folded axis, b) top-view along the seven-fold axis.[54] Reprinted by permission from Science[54] copyright (2000), American Association for the Advancement of Science.

The porphobilinogen synthase family of metalloenzymes catalyzes the first common step in the biosynthesis of tetrapyrroles such as corrins (containing Co), chlorophyll (containing Mg), and porphyrin (with Fe). In the crystal, it assembles into an octamer (Figure 5).[56]

For enzyme stereospecificity and the receptor-substrate complementarity nature has denied life from mirror symmetry. As a consequence, crystals containing proteins and viruses belong to space groups that do not include inversion or mirror symmetry, which would otherwise lead to inversion of amino acid or carbohydrate chirality.

Figure 5. Yeast porphobilinogen synthase octamer.[56] Reprinted by permission from the International Union of Crystallography, Acta Crystallographica[56] copyright (2000).

ADVANCED AND NEW METHODS AND TECHNOLOGIES SUPPORTING THE DEVELOPMENT OF X-RAY CRYSTALLOGRAPHY

The tremendous progress of X-ray crystallography is based on the use of powerful computers, employed in on-line experiments such as robotic machines for crystallization experiments, controlled data collection,[57,58] data processing,[59] interpretation of Fourier maps,[60] refinement of an experimentally derived model of the structure,[48] visualization of the three-dimensional structure and modelling using computer graphics, handling databases (data mining) for structure systematics and predictions, and study of protein folding and structure/activity correlations.

For the determination of macromolecular structures, more intense sources of high brilliance and a high degree of collimation are required for data

collection. Some improvement in this respect was made by the use of rotating anode tubes, and a significant step forward was achieved by the first exploitation of synchrotron X-ray radiation in the late seventies. The constant progress of technology led to the third generation of synchrotrons[61–66] operating at 6 GeV and above, providing crystallographers with a more intense (many orders of magnitude in intensity than classical X-ray tubes) monochromatic beam of high brilliance and collimation, suitable for very small, weakly diffracting samples (about 50 μm cross-sectional diameter) as well as very large unit cells (such as the blue tongue virus core of $795 \times 822 \times 753$ Å^3 (Ref. 28) and the 50S ribosomal subunit of $212 \times 301 \times 576$ Å^3 (Refs. 31, 37)). An additional experimental advantage of such sources is the wavelength tuneability, which makes it possible to use the method of multi wavelength anomalous dispersion (MAD) over a wide range of atomic absorption edges for phase determination of structure factors. By virtue of such a powerful X-ray source it is also possible to perform time-resolved experiments, *e.g.* using polychromatic radiation on stationary crystals producing Laue patterns. This idea might seem paradoxical to the basic concept of X-ray crystallography. In general, X-ray crystallography reveals only the space average structure over all the molecules in the crystal for the time of exposure to the X-ray beam. However, for time-resolved crystallography using a polychromatic synchrotron beam and a stationary crystal, and an exposure time about three orders of magnitude shorter than that of a monochromatic experiments, ultra-fast experiments are possible.[67] In reality, a series of events can be recorded. With a powerful source and fast detector, but also using smart software for the interpretation of the rather complex Laue patterns, it is possible to study enzymatic reactions *in situ* at nanosecond time resolution.[68,69] This facility induces studies of various photo-activated processes. It is possible to use real-time *in situ* observations of many chemical reactions, like oxidation processes, dimerization, hydrogen-deuterium exchange (using neutron diffraction) and many others.[70] An example of wide interest is the reaction dynamics of myoglobin. A nanosecond time-resolved data set on the carbon monoxide complexes of myoglobin revealed a rupture of the covalent bond between the iron atom of the heme and carbon monoxide by a nanosecond light flash.[71] The carbon monoxide migrates away from the heme, and the heme and the protein relax towards the stable deoxyhemoglobin structure. At room temperature, the entire reversible cycle is completed in the crystal in some hundreds of microseconds. This has been studied by various spectroscopic methods as well. Still, a few details relating to the dynamics of the whole process remain unclear.

There is another aspect of the application of intensive X-ray beams, which has also been considered. In addition to nonspecific radiation damage of the crystal that might appear at cryogenic temperatures (generally mani-

fested as decay of intensities), there also appears to be specific radiation damage. This can be related to particular chemical reactions initiated by radiation. Very recent data recorded on the crystal of acetycholinesterase from *Torpedo californica* (*Tc*AChE) at 100 K using an unattenuated beam with a wavelength of 0.93 Å (experimental details are given in Ref. 72), revealed time-dependent changes in the electron density in the region of the disulfide bridge between Cys 254 and Cys 265 (Figure 6). Sequential Fourier maps clearly revealed continuously diminishing electron density with time, which ends with complete cleavage of the Cys 254 – Cys 265 disulfide bond. Two other disulfide bonds, Cys 402 – Cys 521 and Cys 67 – Cys 94, appeared to

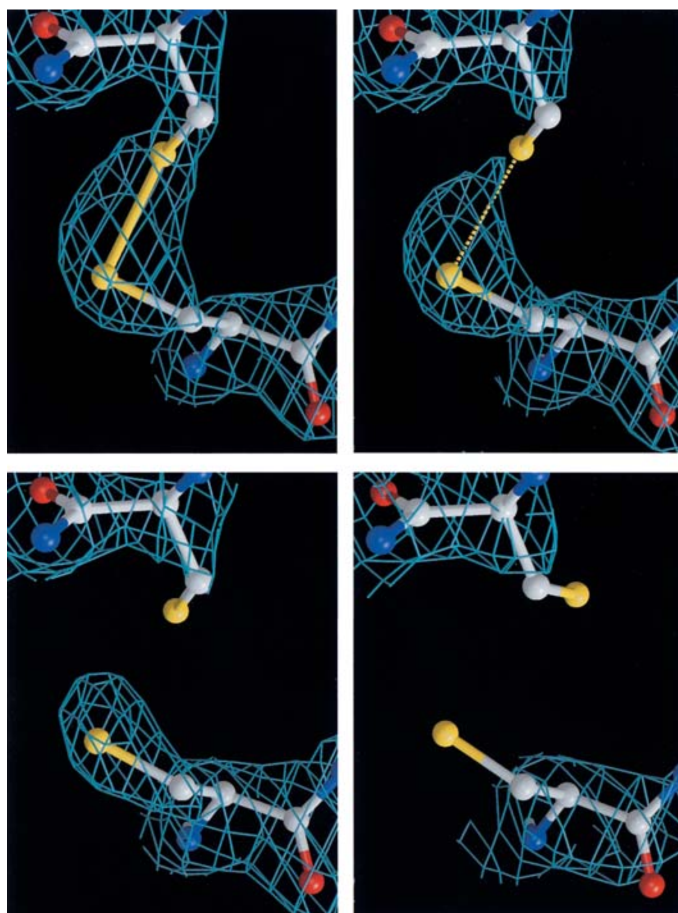

Figure 6. Difference Fourier maps of acetylcholinesterase from *Torpedo californica* revealed time-dependent changes in electron density in the region of disulfide bridge between Cys 254 – Cys 265, which end with a rupture of disulfide bond.[72]

be more stable over the same period of time. It was also noticed that the catalytic triad residue His 440 is more sensitive to radiation damage than the other histidine residues. Such changes should be recognized in order to be clearly distinguished from the changes related to intrinsic enzymatic activity.

The crucial part of data collection is associated with recording diffracted intensities. They should be measured as fast as possible (short exposure time) and be equally sensitive to weak and strong intensities (dynamic range) over a wide space angle, having no dead time (to record promptly the intensity, *i.e.* with near photon-counting statistics). Two alternative types of detectors have been used: point detectors measuring the energy of a single reflection (classical detectors) and area detectors, which collect numerous reflections on the same two dimensional device (such as films, wire detectors, imaging plates, charge-coupled devices (CCDs)).[73,74] Image plate (IP) and CCD-based systems have been developed primarily for protein crystallography and they reasonably meet the above requirements. Further improvements for particular applications are expected. The properties of the crystal (its size and unit cell dimensions) and diffraction pattern (peak to background intensity ratio and the angular width of the diffraction features) are intrinsic to the crystal.[75] Thus, experimental conditions, such as the properties of the X-ray beam and the type of detector, should match the sample.

Among various techniques applied for improving X-ray diffraction experiments, cryo-techniques[76,77] deserve to be discussed. Cooling a crystal with liquid nitrogen or helium (down to 100 K and 3–5 K, respectively) during intensity measurements significantly improves the quality of the experimental data, but also makes it possible to obtain data of thermally unstable or air-sensitive samples. Freezing of the sample considerably reduces radiation damage and also prevents crystal slippage from the fixed position on the goniometer-head during intensity measurements. Flash-cooling to near 100 K extends the crystal life-time, even during measurements in the synchrotron beam, and contributes to better resolution of the data. By cooling the sample, thermal vibrations of atoms in the crystal are reduced (electron density becomes more concentrated) and thus the ratio between peak intensity and background increases (better signal to noise ratio). As a consequence, it is much easier to locate the atoms in a Fourier map. In crystals with translational defects, it is possible to discriminate between static and dynamic disorders. Dynamic disorder is reduced by cooling, whereas static disorder is not significantly affected, unless there is a phase transition. Cryo-techniques are also inevitable for studying phase transitions for various compounds of interest in material sciences. For such purposes, cooling with liquid helium is essential if phase transitions can be expected below 100 K.

During the last eight years, the use of cryo-techniques in macromolecular crystallography has significantly improved the quality of intensity data and makes it possible to work with very sensitive proteins. Specially manufactured tools for selecting and mounting a micro-dimensional sample in a drop of cryoprotectant liquid supported by a fibre loop sitting on the top of a pin holder at the goniometer-head are for sale, but the »green-fingers« needed to mount a crystal properly are not commercially available.

To produce a good diffracting crystal requires even more knowledge and skills. The first protein crystals produced 150 years ago were of hemoglobin and the first X-ray diagram of a protein (pepsin) was recorded by D. Hodgkin and J. D. Bernal in 1934.[78] Twenty years ago, crystallization was treated as art rather than science, but in the last decade it has been a quickly developing branch of science, based on a large number of experiments, some of them being performed in space shuttles to eliminate the influence of gravitation on crystal growth. Indeed, microgravity experiments often produce crystals diffracting at higher resolution. Diffraction experiments performed on samples of insulin crystals grown under earth and microgravity conditions revealed a significant difference in quality. The reflection profiles of earth-grown crystals showed multiple gaussian shapes and were 0.060–0.100 degrees wide at the FWHM. Microgravity crystals were perfect, with single gaussian profiles and FWHM width of only 0.002–0.006 degrees.[79] The preparation of crystals involves nucleation and growth. When experimentators realized that these processes were governed by the same physicochemical rules for small and large molecules, considerable improvements were made. More than twenty microtechniques were invented. Among the most common are vapour diffusion, liquid diffusion, and dialysis.[80,81] All of these methods explore the protein solution, searching for conditions that increase the degree of saturation to the stage where protein-protein interactions become adequate for initial nucleation. It is important to find the most suitable technique and optimize all experimental conditions, which are protein and method dependent. Thus, special screening procedures according to experimentally established protocols (stored in the biomacromolecular crystallization database)[82,83] can be applied. In order to screen the enormous numbers of experimental parameters, robotic machines have also been designed. Generally, these experiments take care of the purity of the preparation, concentration of the protein, concentration and type of the precipitant (which reduces the solubility of the protein), pH, temperature, selection of buffer and additives (salts, organic components, and detergents) in low concentrations. A typical experiment can be run in a solution volume of 2–4 μL, although there are experiments with droplets of 1–2 μL. For a »macroscopic« crystal suitable for X-ray analysis, about 7–15 μg of protein is needed. Even

commercially available screening protocols such as Wizard-I[84] cannot solve all problems. In order to succeed in growing good quality crystals, one has to know as much as possible about the protein characteristics. A reasonable amount of substance should be provided to complete the biochemical characterization and to crystallize the protein of interest. Usually, proteins from natural sources are not available in amounts larger than a few mg.

Methods of molecular genetics support tremendously the development of X-ray structure analysis of macromolecules, with recombinant DNA techniques being used for cloning and expressing proteins of interest in tens of milligrams. However, there are more aspects of genetic engineering that are useful for protein crystallography. Once an initial protein structure is known, understanding its function is the final goal of the research. By site-directed mutagenesis of specified amino acid(s) responsible for interactions, functional hypotheses can be tested. This procedure is common practice for almost each new protein studied.

This technique makes it also possible to substitute sulphur with selenium (as an attractive anomalous scatterer of X-rays) in methionines, which then serve as a phasing label in the experimental phasing of structure factors by multiple anomalous dispersion (MAD method), using variable wavelengths in synchrotron experiments.[85,86] Using this method, preparation of several heavy-atom derivatives and the subsequent data collection can be avoided. This is not only a more sophisticated approach, but it also prevents possible protein destruction by the incorporated metal atoms. According to the evidence[87] about 40% of new protein structures are being determined by the MAD phasing method. For the structures solved by synchrotron radiation, MAD phasing exceeds 70% (very often using SeMet phase label).

HIGHLIGHTS OF (BIO)CRYSTALLOGRAPHY

Revision of various aspects of chemical bonding relies very much on charge density studies applied in chemistry, biology, and solid-state physics.[88–90] X-ray analysis performed for charge-density studies differs somewhat from the conventional structure determination in its theoretical concept. Technical developments (already discussed in the previous paragraph) enable high accuracy of data collection, and advanced software, specifically designed least-squares procedures (based on an aspherical atom model), provide very precise electron density maps. Use of low temperature (< 20 K) experiments reduces the thermal vibrations of the atoms in the sample with their detrimental obscuring effects on electron density maps. The electron maps obtained using such data treatment reveal fine details about the elec-

tronic structure and its modifications, influenced by partners connected by the chemical bond or participating in intermolecular interactions. As the method stands today, such fine details match perfectly the pictures obtained by theoretical *ab initio* calculations. Charge-density studies can reveal overlap between atomic orbitals, lone pairs, the bending of bonds in strained rings, π bonding, polarization effects, deformation of electron density in intermolecular interactions, particularly in hydrogen bonds. Due to the high chemical and biological relevance of the properties of sulphur – sulphur interactions, the charge density study of hexagonal L-cystine[91] is selected to illustrate the achievements of the method. The deformation density distribution and topological analysis of charge density reveals that the S–S bond of disulfide bridge is a weak single covalent bond (Figure 7).[91,92] The valence shell charge concentration suggests sp^3 hybridization of sulphur atom. The analysis also reveals hydrogen bonds *via* COO^- and NH_3^+ (Figure 8)[91,93] which contribute dominantly in the dipolar moment.

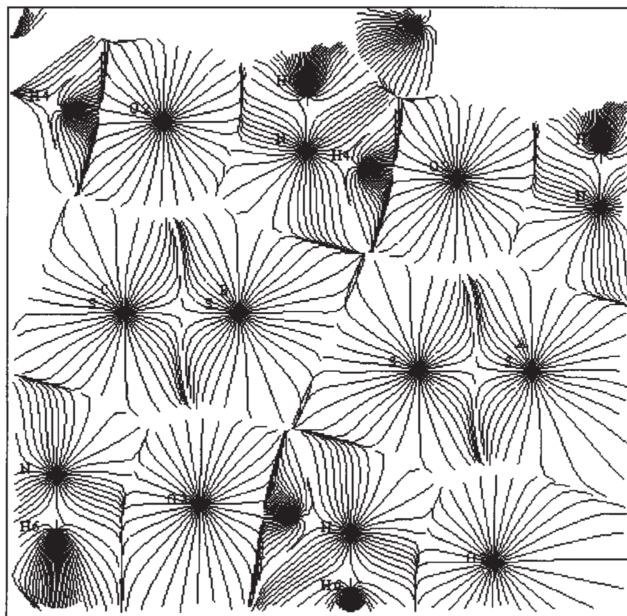

Figure 7. $\vec{\nabla}, \rho(\mathbf{r})$ gradient vector field of the electron density of L-cystine in the plane containing the S–S bonds.[91] S$\cdots$S^B bond is along the [0 1 0] direction. The $\vec{\nabla}, \rho(\mathbf{r})$ was calculated from the NEWPROP Program.[92] A disulfide bridge occurs between two cystine molecules connected by a two-fold axis parallel to [1 0 0] or [1 1 0] (S–S^A = 2.0472(4) Å). The S$\cdots$S^B (S^B: *x*, 1+*x*–*y*, 1/6–*z*) intermolecular contact found in L-cystine (3.4264(4) Å) is much shorter than the sum of Van der Waals radii (3.7 Å). The picture was obtained from Prof. dr. C. Lecomte.

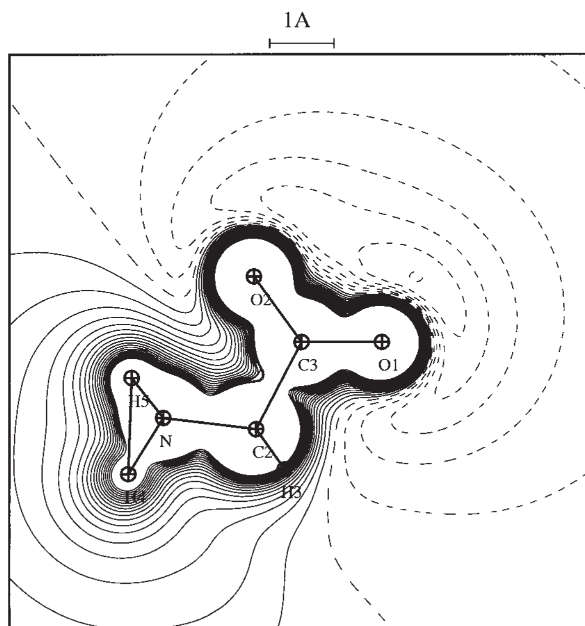

Figure 8. Electrostatic potential in the plane of the COO^- group of the isolated L-cystine molecule. Contours intervals are at 0.05 e $Å^{-1}$; solid lines are positive. Dotted line are negative, and zero contour is a broken line (ELECTROS Program).[93] This map gives the electrostatic potential generated by the L-cystine molecule removed from its crystal lattice in its carboxyl plane. The oxygen atoms are surrounded in the outer region by a wide and deep negative region (–0.30 e $Å^{-1}$; 1 e $Å^{-1}$ = 332 kcal mol^{-1} = 1390 kJ mol^{-1}), which would favor the approach of electrophilic agents as well as hydrogen bonds.[91] The picture was obtained from Prof. dr. C. Lecomte.

Understanding the nature of non-covalent interactions is essential for reaching the secrets of molecular aggregation in the biotic world, but also in synthetic chemistry which has now expanded into the fascinating field of supramolecular chemistry.[94,95] An important step in the process of aggregation is molecular recognition, which is based on chemical and geometrical complementarity between interacting molecules. One can design a supramolecular synthon, a small molecule provided with a particular functionality suitable for intermolecular interactions, *e.g.* containing points of recognition such as proton-donor and proton-acceptor sites, which form infinite connections of one-, two- or three-dimensional arrays of hydrogen bonds. An illustrative example is a discrete spherical supramolecule of icosahedral symmetry, composed of six calix[4] resorcinarenes and eight water molecules connected by sixty hydrogen bonds with a well-defined cavity of 1373 $Å^3$ (Figure 9), which can encapsulate various guest molecules.[96] Such types of su-

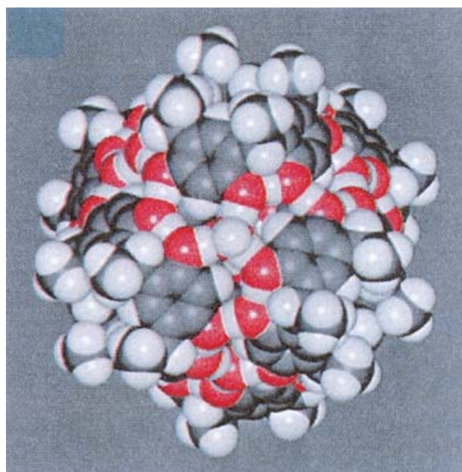

Figure 9. A chiral spherical molecular assembly of six calix[4]resorcinarenes and eight water molecules along a three-fold axis with topology similar to spherical viruses.[95] Reprinted by permission from Nature[95] copyright (2000) Macmillan Magazines Ltd.

pramolecules of adjusted lipophilicity can imitate biological vesicles, and some of them might be used for controlled release of encapsulated drugs into body fluids. In general, supramolecules of spherical and tubular morphologies have potential fundamental and practical implications; through nanotechnology they are related to material science, biomimetic chemistry, and molecular electronics. The topology of such molecules (Figure 10)[97] is very close to that of the spherical polyhedral shells of the capsids of the blue tongue virus,[28] simian virus 40,[98] human hepatitis B virus (Figure 11),[53] reovirus[99] or tubular protein coats of viruses, and fullerene molecules,[100] which also involve covalent interactions. In spherical viruses, capsids with icosahedral symmetry would provide the most compact protein assembly, serving as a protective coat around the nucleic acids. Capsids with strict icosahedral symmetry would allow a maximum of sixty chemically identical subunits. However, it has been known from the structures of numerous spherical viruses that the number of coat proteins significantly exceeds 60. For example, in simian virus 40 (polyoma virus),[98] including a double-stranded DNA genome, there are 360 copies of the major structural protein, arranged as 72 pentameric building blocks on the viral surface. The subunits also contain two more proteins. In order to illustrate the complexity of the virus structures and the advances made, both scientific and technical ones, the biggest virus core structure so far known, of the blue tongue virus is briefly described.[28,101] The BTV particle (Figure 12) (54 MDa) contains 780 copies of a

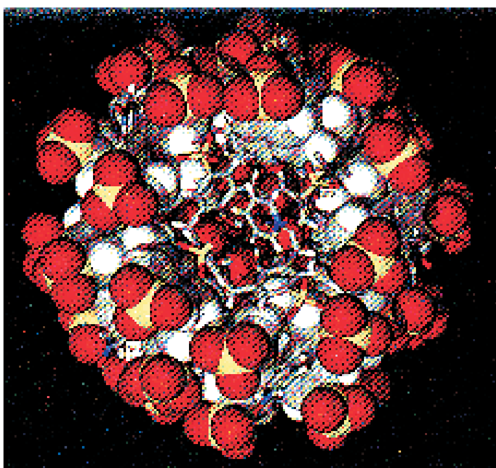

Figure 10. Spherical arrangement with a pseudo five-fold symmetry of amphiphilic *p*-sulphonato calix[4]arene building blocks with the geometry similar to a spherically shaped coat of virus.[97] Reprinted by permission from Science[97] copyright (2000), American Association for the Advancement of Science.

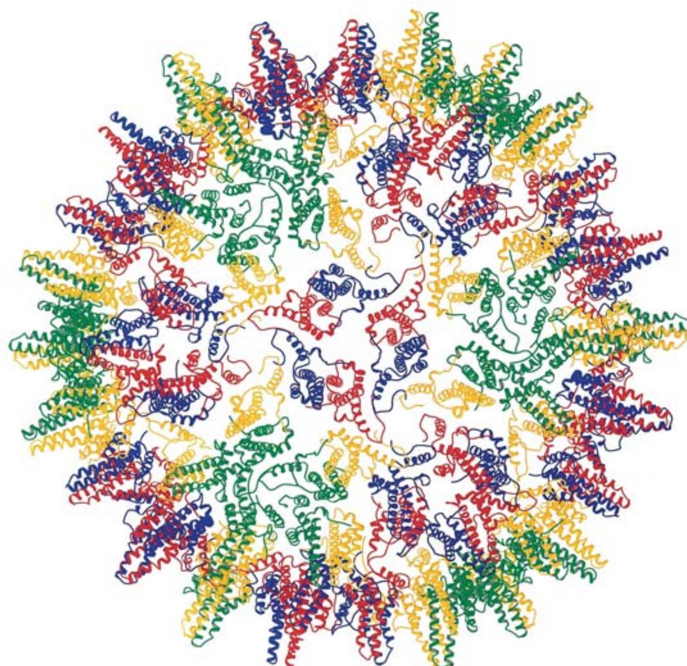

Figure 11. The capsid of the human hepatitis B virus revealing an icosahedral symmetry.[53] Reprinted by permission from Cell[53] copyright (2000) Cell Press.

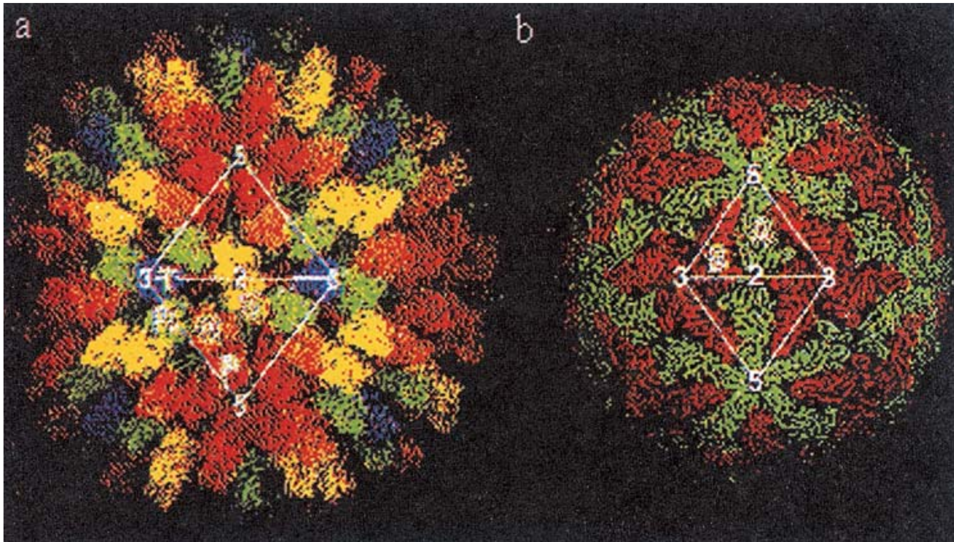

Figure 12. The architecture of the blue tongue virus (BTV) core: a) the icosahedral asymmetric unit contains 13 copies of VP7 (T13) arranged as 5 trimers, P, Q, R, S, and T (coloured red, orange, green, yellow, and blue, respectively), and b) the VP3 (T2) scafold with icosahedrally unique molecules A (green) and B (red), having a structurally different environment. The two-fold, three-fold and five-fold symmetry axes are marked.[28] Reprinted by permission from Nature[28] copyright (2000) Macmillan Magazines Ltd.

VP7 (T13) protein subunit (38 kDa) arranged as basic trimers (assembling into hexamers and pentamers) on a T13 (T corresponding to the number of proteins in the icosahedral asymmetric unit) quasi-equivalent lattice, which form the core surface. This VP7 (T13) subunit of the outer layer covers a subcore shell constructed of 120 copies of the VP3 (T2) protein (100 kDa). The VP3 (T2) molecules bind RNA. The core also includes some enzymatically active proteins and ten double-stranded segments of (ds) RNA (about 19 000 base pairs). Thus, the icosahedral building block is composed of 13 subunits of VP7 (T13) attached to two molecules of VP3 (T2). To collect 21 million intensities, about 1000 crystals had to be used (due to radiation damage occurring also at cryogenic temperatures). The set of unique data of about 3.3 million was used in Fourier syntheses and refinement. 49 000 atoms were located within an icosahedral asymmetric unit while the entire BTV particle contains 2.9 million atoms. The elucidation of this structure depended on many creative steps, among which was electron density averaging (already described in the section *Fundamentals of X-ray Diffraction*) over non-crystallographic 30-fold symmetry. There is also the fundamental

question relating to the biological function of the virus. How does Nature efficiently organize such a »living polyhedron«? This is realized by reducing exact symmetry to an approximate one; the subunits in clusters are *quasi*-equivalent. Such associations are sufficiently flexible to accommodate distortions and a higher number of subunits can be built in.[52] Grimes and coworkers[28] took the challenge to explain the nucleation of such a large assembly. It is remarkable that the self-assembly mechanism of primitive biological entities (viruses), according to Klug, resembles the nucleation and growth of a crystal lattice.

»The crystal is, in a sense, the supramolecule *par excellence*: a lump of a matter, of macroscopic dimensions, millions of molecules long, held together in a periodic arrangement by just the same kind of noncovalent bonding interactions as are responsible for molecular recognition and complexation at all level.«[102] Even more, the crystallization process itself exhibits supramolecular assembly. Thus, crystal packing patterns with their infinite motifs represent aggregations of supramolecular types. It is very important to remember that the properties of a synthon (a single molecule) are different from those of its supramolecular aggregates. This approach is the basic concept of crystal engineering, which is also useful in generating molecules for nanotechnologies.

Molecular nanotechnology draws on cutting-edge advances in physics, chemistry, biology, and computer science to build structures measured in nanometers. Its start and development is based on the very simple statement made by R. Feynman, Nobel laureate in physics (for quantum electrodynamics). »The principles of physics, as far as I can see, do not speak against the possibility of manoeuvring things atom by atom.« Nature leaves us to learn how to arrange atoms the way we want and to develop molecular manufacturing. The properties of products obtained depend on how the atoms are arranged. A very simple example: if we rearrange the atoms in coal, we can make diamond. Thus, we are not far from the dreams of medieval alchemists. Very sophisticated maneuvering is based on DNA self-assembled molecules created by human mind and hands. N. Seeman and coworkers[103–107] created DNA molecules whose helix axes have the connectivities of a cube and of a truncated octahedron (Figure 13)[108] and DNA knots. For developing ways to construct three-dimensional objects, including cubes and more complex polyhedra, from synthetic DNA molecules, N. Seeman won the 1995 Feynman Prize in Nanotechnology. The periodic arrays (crystals) of DNA objects could be used to direct the assembly of biochip computers. But the list of application is wider. By developing DNA assemblies experimentalists aim to construct mini-biorobots for monitoring, repair, construction and control of the human biological system at the molecular level (nanomedicine).

Figure 13. DNA-based truncated octahedron; a view down the four-fold axis. Each edge of the polyhedron contains two double helical turns of DNA. The molecule has 14 cyclic strands of DNA and its molecular weight is about 790 000 Da. Legend employed: each nucleotide is shown with a coloured dot corresponding to a backbone, and a white dot corresponding to a base. Printed by permission of Prof. dr. N. C. Seeman.[108]

Crystallographers have a very direct interest in synthesizing periodic matter in a rational fashion in order to avoid the weak points of current crystallization protocols. The idea is to build DNA assembled molecules into cages that would contain oriented guest molecules and thus to assist regular molecular orientation in a 3-D space and grow crystals.

Molecular recognition regarded by many as a recent fashionable concept, is in fact quite an old mental picture of biological events described by E. Fischer in 1894[109] as »Schlüssel-Schloss-Prinzip« (Key-Lock analogy). E. Fischer's global concept (free of details) on the specificity of enzyme activity survived a century, and has been one of the most fertile metaphors. It was used in 1897 by P. Ehrlich[110] in the area of immunology and its contemporary vision is one of the best examples of the molecular recognition approach. E. Fischer was aware of enzyme specificity and built a picture of enzyme-substrate interaction, which was verified by examples of the X-ray structures of numerous enzymes and their complexes with substrates and/or inhibitors. These X-ray structures have provided essential data for understanding the mechanism of biological activity at the atomic level and have

led to the design of new drugs. Developments of X-ray analysis have offered more reliable data on protein structures and their assemblies, and revised pictures of many biological events. The static lock-key model was expanded to a dynamic one based on the induced fit theory, illustrated by the hand-glove analogy.[111] It means that in the process of a substrate or inhibitor binding into an enzyme active site, some conformational adaptations can take place. Molecular dynamics simulations can evaluate the energy changes of such processes and identify the most stable conformations.

One of the most striking examples is the three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1 (at 3.0 Å resolution,[112] 2.7 Å[113]). This pepsin-like enzyme is the smallest among the retroviral proteases with 99 amino acids. The protease is a dimer with a crystallographic two-fold axis through the active site (Figure 14). Quite a number of complexes with transition state inhibitors and protease mutants were analysed (Ref. 114 and references therein). The knowledge of these structures and details of the active site mechanism enabled the design of specific inhibitors; some of them have become new anti-HIV drugs. This research has contributed significantly to the decrease in the number of fatal cases, by two thirds during the past ten years (Figure 15). Knowledge of an enzyme structure is also a prerequisite for site-specific mutagenesis, which might be used in gene therapy. Structure-based drug design depends on the target

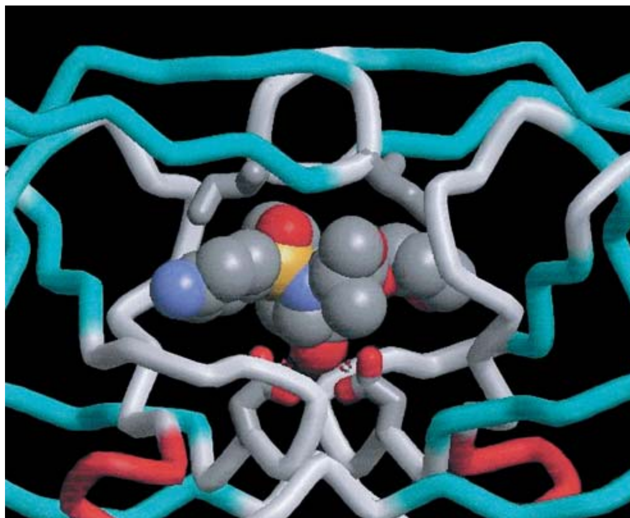

Figure 14. The HIV-1 aspartyl protease appears as the active dimer with a crystallographic two-fold axis.[113,114] The active site is shown with an inhibitor. The picture was obtained from dr. M. A. Navia.

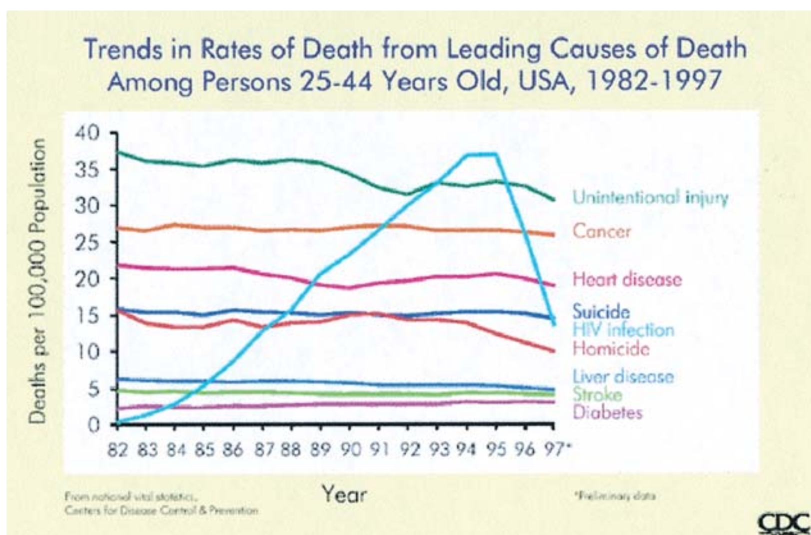

Figure 15. Diagram illustrating the reduction of fatal cases of AIDS-patients after discovery of the structure and function of aspartic HIV-1 protease, which enabled design of its inhibitors serving as useful drugs (copied from http://www.cdc.gov./hiv/graphics/mortalit.htm).

protein, the location of that protein in the host and/or pathogen, the flexibility of the target protein upon ligand binding, and the nature of the site to which these compounds have to bind. The combination of X-ray analyses and molecular modelling,[115,116] supported by molecular genetics including the proteomics approach (protein genomics),[117] in the search for new drugs has become a standard procedure employed in pharmaceutical industry.

Methods of contemporary crystallography are capable of probing deeply into cell biology events, *e.g.* in an attempt to elucidate the structure of the bacterial ribosome (weight of 2.6 million Da, consisting of two thirds of RNA and one third of proteins), the factory in the cell cytoplasm that makes proteins.[30–42] For about 40 years, research has been focused on the structures of the bacterial ribosome components. The resolution of the crystal structure electron density of the 50S ribosome subunit (at 2.4 Å) (Figure 16)[31,37,40,41] reveals a new fact about the ribosome function, which enables to propose the mechanism of the peptide synthesis catalysed by ribosome; actually, the ribosome is a ribozyme.[41,42] These, among the largest asymmetric objects ever determined by X-ray crystallography, are: 30S (900 kDa),[32,36,118,119] 50S (1600 kDa),[31,37,40,41] and 70S (2500 kDa, Refs. 33, 34 and references therein). These structures have been essential for establishing the location

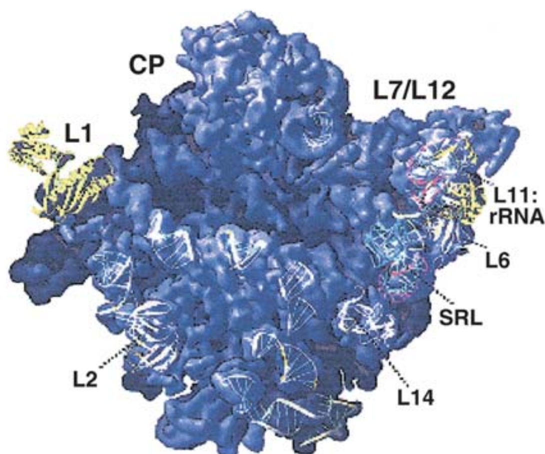

Figure 16. »Crown« view of the 50S subunit of ribosome with ribbon representations of protein and RNA-fragment structures; identified units are labeled.[31,40,41] The picture was obtained from dr. N. Ban.

of the 54 proteins and more than 4500 nucleotides of RNA and their relative positions and functions within the global ribosome unit. Cryo-electron microscopy[120–122] has also contributed significantly to the mapping of the ribosome particle and to the location of the heavy atom sites required for phasing structure factors.

Among the high achievements of cryo-electron microscopy with regard to structure elucidations are the dazzling examples of the lipid-containing (ds) DNA viruses: *Chilo iridescent* and *Paramecium bursaria chlorela* viruses of about 1900 Å in diameters. Both have a size of about 1 GDa with characteristic icosahedral symmetry.[123]

PROSPECTS

»As everyone knows, long-term predictions in science are hopeless and even short-term predictions are usually wrong«.[124] One might quote Daniel Kleppner's words dedicated to the 50th anniversary of *Physics Today* (1998) and skip this last paragraph. However, one can expect that scientists will continue to make new discoveries that will take away our breath and widen our horizons. On the basis of the results achieved, one can try to predict the most promising research lines.

Cryo-electron microscopy,[125,126] time-resolved crystallography,[68,69] protein crystallography with fourth-generation synchrotrons, and computer model-

ling are on the way to reveal dynamic pictures of crucial living-cell processes. Elucidating the mechanism of protein synthesis in the ribosome means getting closer to understanding how cells are capable of controlling the ongoing processes, and to prevent mistakes or to correct them. Cells must avoid misfolded or incorrectly assembled proteins, because only the correct ones should travel further to their site of action. It is obvious that, knowing the mechanisms at the molecular level, one can try to solve the problems in pathological and degenerative processes such as Alzheimer and Kreuzfeld-Jacob's diseases, cancer, and ageing. Molecular medicine will be more and more oriented towards gene therapy.

It is certain that the human genome project[14] and the many recently completed genome sequences of several microorganisms and three plants will keep protein crystallographers busy for quite a long period ahead. The *Drosophila* melangaster (fruit fly) genome sequencing can be added to this list.[127,128] Its cell biology and development have much in common with mammals and »this sequence may be the Rosetta stone for deciphering the human genome«.[129] Envisaging the solving of protein structures related to the human genome requires special tactics and a selection list of core proteins has to be made for X-ray crystallography. The priority list has been a topic of intense discussion among scientists. Whatever we may say about the impact of the human genome project on the life sciences and humanity in general may sound like science fiction.

The accumulation of known protein structures will enable us to establish more precisely the structure-function correlations, which will be used in drug design and biotechnology. It will also be possible to apply mutagenesis even more precisely. Looking into chemically analogous enzyme mechanisms of various species (from bacteria to mammals) at the molecular level, we will learn more about the biological evolution and Darwin's theory.

Definitely, the numerous new fields offer great challenges to the new generation of scientists to discover the secrets of Nature and to use them to the benefit of human beings.

The prediction of the crystal and molecular structures of small molecules and of the secondary structures of proteins (protein folding) can be viewed as the study of a complex organizational problem,[130] but prediction will become more routine and accurate in the future. Being a good »prophet« in structure predictions means finding a solution of the global free energy minimum for the system studied. The potential energy of the molecule or its assembly depends on the topological arrangement of the atoms and has to be optimized. Thus, the complexity of the nontrivial global optimization problem increases for large systems, but parallel computers and efficient methods and their combinations (simulated annealing, hypersurface deformation

approach, molecular dynamics simulations, genetic algorithms and modifications) are very promising. For further improvements of algorithms and potential energy functions, more accurate structural data need to be extracted from structural databases. The enormous potential of protein folding studies does not only concern secondary structure predictions but also dynamics, equally applicable to real and to theoretical systems at femto seconds and even further. These studies will reveal how simple concepts can emerge from complex systems.

How it can be that mathematics, being after all
a product of human thought independent of experience,
is so admirably adapted to the objects of reality.

A. Einstein

Acknowledgment. – The auhors thank the publishers (IUCr-Acta Crystallographica, Cell Press, Nature and Science) for copyright agreements to use the illustrations, previously published, and quoted in this paper. Thanks are due to dr. M. A. Navia, The Althexis Company, Inc., Waltham, MA. for his suggestions and discussion and an illustration of HIV-1 protease, to dr. N. Ban, Department of Molecular Biophysics and Biochemistry and Chemistry, Yale University and the Howard Hughes Medical Institute, New Haven, USA for valuable comments and the photograph of the S50 subunit of the ribosome particle. Prof. dr. C. Lecomte, Univ. Henri Poincaré, Nancy, generously donated a few illustrations from his charge density studies. Prof. dr. N. C. Seeman, department of Chemistry, The New York University, New York for a permisson to use his illustarations of DNA assemblies. The authors appreciate the kind donation of the photograph of a crystal by Prof. V. Bermanec, Faculty of Science, Department of Mineralogy, University of Zagreb, Croatia. One of the authors (B. K.-P.) thanks A. Višnjevac, Mr. Sc., Ruđer Bošković Institute, Zagreb, Croatia for assisting in the preparation of the illustrations.

REFERENCES

Introduction

1. W. H. Bragg, W. L. Bragg, *Proc. Phys. Soc. London, Sect. A* **89** (1913) 277–292.
2. J. A. Le Bel, *Bull. Soc. Chim. Fr.* **22** (1874) 337–347.
3. J. H. van't Hoff, *La chimie dans l'espace*, Bazendijk Rotterdam, 1875.
4. E. Fischer, *Ber. Dtsch. Chem. Ges.* **24** (1891) 1836–1845.
5. E. Fischer, *Ber. Dtsch. Chem. Ges.* **24** (1891) 2683–2687.
6. J. M. Bijvoet, *Proc. K. Ned. Akad. Wet.* **52** (1949) 313–314.
7. D. Crowfoot Hodgkin, C. W. Bunn, B. W. Rogers-Low, and A. Turner-Jons, *The X-ray Crystallographic Investigation of the Structure of Penicillin*, Oxford University Press, Oxford, 1949.
8. D. C. Hodgkin, *Les Prix Nobel* en 1964, Nobel Foundation, 1965, pp. 71–91.
9. J. D. Watson and F. H. C. Crick, *Nature* **171** (1953) 737–738.

10. L. Pauling, *Science* **110** (1949) 543–548.
11. V. M. Ingram, *Nature* **180** (1957) 326–328.
12. B. C. Wishner, K. B. Ward, E. E. Lattman, and W. E. Love, *J. Mol. Biol*. **98** (1978) 179–194.
13. B. J. Strasser, *Science* **286** (1999) 1488–1490.
14. R. L. Strasberg, E. A. Feingold, R. D. Klasner, and F. S. Collins, *Science* **286** (1999) 455–457. (http://www.ornl.gov/hgmis)
15. F. H. Allen, *Acta Crystallogr*., *Sect. A* **54** (1998) 758–751.

Historical Backgrounds

16. W. C. Röntgen, *Sitzungber. Phys. Med. Ges. Würzburg* **52** (1895)132.
17. H. H. Seliger, *Phys. Today,* November (1995) 25–31.
18. H. D. Zeman, *The Processing of Intravenous Coronary Angiograms Produced Using Synchrotron Radiation*, in: B. V. S. Ebashi, E. Rubenstein, and M. Koch (Eds.), *Handbook of Synchrotron Radiation*, Vol. 4, Elsevier, Amsterdam, 1991, pp. 697–718.
19. H. D. Zeman, *Imaging Apparatus and Technique for Synchrotron Radiation Angiography*, in: B. V. S. Ebashi, E. Rubenstein, and M. Koch (Eds.), *Handbook of Synchrotron Radiation*, Vol. 4, Elsevier, Amsterdam, 1991, pp. 679–696.
20. S. Achenbach, W. Moshage, D. Ropers, J. Nossen, and W. G. Daniel, *N. Engl. J. Med*. **339** (1998) 1964–1971.
21. D. Graham-Rowe, *New Sci.* December 18 (1999) 12.
22. W. Friedrich, P. Knipping, and M. Laue, *Sitzungber. K. Bayer. Akad. Wiss.* (1912) 303–322.
23. M. Laue, *Phys. Z.* **14** (1913) 421–423.
24. W. L. Bragg, *Nature* **90** (1912) 410.
25. W. H. Bragg and W. L. Bragg, *Proc. R. Soc. London, Ser. A* **88** (1913) 428–438.
26. W. H. Bragg, *Philos. Trans. R. Soc. London, A* **215** (1915) 254–275.
27. W. L. Bragg, *Sci. Am.* **219** (1968) 58–70.
28. J. M. Grimes, J. N. Burroughs, P. Gouet, J. M. Diprose, R. Malby, S. Ziéntara, P. P. C. Mertens, and D. I. Stuart, *Nature* **395** (1998) 470–478.
29. B. V. Prasad, M. E. Hardy, T. Dokland, J. Bella, M. G. Rossmann, and M. K. Estes, *Science* **286** (1999) 287–290.
30. A. Ramakrishnan and S. W. White, *Trends Biochem. Sci.* **23** (1998) 208–212.
31. N. Ban, P. Nissen, J. Hansen, M. Capel, P. B. Moore, and T. A. Steitz, *Nature* **400** (1999) 841–847.
32. W. M. Clemons Jr., J. L. C. May, B. T. Wimberly, J. P. McCutcheon, and M. S. Capel, *Nature* **400** (1999) 833–840.
33. J. H. Cate, M. M. Yusupov, G. Zh. Yusupova, T. N. Earnest, and H. F. Noller, *Science* **285** (1999) 2095–2104.
34. G. M. Culver, J. H. Cate, G. Zh. Yusupova, M. M. Yusupov, and H. F. Noller, *Science* **285** (1999) 2133–2135.
35. A. Yonath, J. Harms, H. A. S. Hansen, A. Bashan, F. Schlünzen, I. Levin, I. Koelln, A. Tocilj, I. Agmon, M. Peretz, H. Bartels, W. S. Bennett, S. Krumbholz, D. Janell, S. Weinstein, T. Auerbach, H. Avila, M. Piolleti, S. Morlang, and F. Franceschi, *Acta Crystallogr*., *Sect. A* **54** (1998) 945–955.
36. A. Tocilj, F. Schlünzen, D. Janell, M. Glühmann, H. A. S. Hansen, J. Harms, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, and A. Yonath, *Proc. Natl. Acad. Sci. USA* **96** (1999) 14252–14257.

37. N. Ban, P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz, *EMBO Millennium Symposium on Structural Biology,* Book of Abstracts, EMBL-Heidelberg, March 26–29, 2000.
38. A. Liljas, *Science* **285** (1999) 2077–2078.
39. C. Davies and S. W. White, *Structure* **8** (2000) R41–R45.
40. N. Ban, P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz, *Science* **289** (2000) 905–920.
41. P. Nissen, J. Hansen, N. Ban, P. B. Moore, and T. A. Steitz, *Science* **289** (2000) 920–930.
42. T. R. Cech, *Science* **289** (2000) 878–879.

Fundamentals of X-ray Diffraction

43. M. A. Walsh, G. Evans, R. Sanishvili, J. Dementieva, and A. Joachimiak, *Acta Crystallogr., Sect. D* **55** (1999) 1726–1732.
44. C. C. H. Chen, A. Kim, H. Zhang, A. J. Howard, G. M. Sheldrick, D. Dunaway-Mariano, and O. Herzberg, Golden Anniversary Meeting 1950–2000, ACA 2000, July 22–27, 2000, St. Paul, Minn. USA, p. 45, 02.06.03.
45. M. G. Rossmann and D. M. Blow, *Acta Crystallogr.* **16** (1963) 39–45.
46. CCP4, Program for protein crystallography, *Acta Crystallogr., Sect D* **50** (1994) 760–763.
47. J. Navaza, *Acta Crystallogr., Sect. A* **50** (1994) 157–163.
48. A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jieang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren, *Acta Crystallogr., Sect. D* **54** (1998) 905–921.

Symmetry

49. N. Rot, W.-J. A. de Wijs, F. J. J. de Kanter, M. A. Dam, F. Bickelhaupt, M. Lutz, and A. L. Spek, *Main Group Met. Chem.* **22** (1999) 519–526.
50. L. Frkanec, A. Višnjevac, B. Kojić-Prodić, and M. Žinić, *Chem. Eur. J.* **6** (2000) 442–453.
51. I. Hargittai (Ed.), *Quasicrystals, Networks, and Molecules of Five-fold Symmetry*, VCH, New York, 1990.
52. D. L. Caspar and A. Klug, *A Cold Spring Harbor Symp. Quant. Biol.* **27** (1962) 1–24.
53. S. A. Wynne, R. A. Crowther, and A. G. W. Leslie, *Molecular Cell* **3** (1999) 771–780.
54. L. Song, M. R. Hobaugh, Ch. Shustak, S. Cheley, H. Bayley, and J. E. Goauax, *Science* **274** (1996) 1859–1866.
55. B. Walker, O. Braha, S. Cheley, and H. Bayley, *Chem. Biol.* **2** (1994) 99–105.
56. E. K. Jaffe, *Acta Crystallogr., Sect. D* **56** (2000) 115–128.

Advanced Methods and Techniques

57. Z. Dauter, *Acta Crystallogr., Sect. D* **55** (1999) 1703–1717.
58. A. G. W. Leslie, *Acta Crystallogr., Sect. D* **55** (1999) 1696–1702.
59. M. G. Rossmann, and C. G. van Beek, *Acta Crystallogr., Sect. D* **55** (1999) 1631–1640.

60. H. R. Powell, *Acta Crystallogr.*, *Sect. D* **55** (1999) 1690–1695.
61. J. R. Helliwel, *Methods Enzymol.* **276** (1997) 203–217.
62. J. R. Helliwel, *Nature Structural Biology*, Synchrotron Supplement, August (1998) 614–618.
63. P. F. Lindley, *Radiat. Phys. Chem.* **45** (1995) 367–383.
64. P. F. Lindley, *Acta Crystallogr., Sect. D* **55** (1999) 1654–1662.
65. R. F. Service, *Science* **285** (1999) 1344–1346.
66. E. Pennisi, *Science* **285** (1999) 1342–1343.
67. J. R. Helliwel and P. M. Rentzepis (Eds.), *Time-Resolved Diffraction*, Clarendon Press, Oxford, 1997.
68. K. Moffat, *Acta Crystallogr., Sect. A* **54** (1998) 833–841.
69. K. Moffat, *Nature Structural Biology*, Synchrotron Supplement, August (1998) 641–643.
70. Y. Ohashi, *Acta Crystallogr.*, *Sect. A* **54** (1998) 842–849.
71. T.-Y. Teng, V. Srajer, and K. Moffat, *Biochemistry* **36** (1997) 12087–12100.
72. M. Weik, R. B. G. Ravelli, G. Kryger, S. McSweeney, M. L. Raves, M. Harel, P. Gros, I. Silman, J. Kroon, and J. L. Sussman, *Proc. Natl. Acad. Sci. USA* **97** (2000) 623–628.
73. Y. Amemiya, *Methods Enzymol.* **276** (1997) 233–286.
74. S. W. Muchmore, *Acta Crystallogr., Sect D* **55** (1999) 1669–1671.
75. C. Nave, *Acta Crystallogr., Sect D* **55** (1999) 1663–1668.
76. D. W. Rodgers, *Methods Enzymol.* **276** (1997) 183–203.
77. E. Garman, *Acta Crystallogr., Sect. D* **55** (1999) 1641–1653.
78. J. D. Bernal and D. Crowfoot, *Nature* **133** (1934) 794–795.
79. G. Borgstahl, E. Snell, H. Bellamy, J. Lovelance, and W. Pangborn, Golden Anniversary Meeting 1950–2000, ACA 2000, July 22–27, 2000, St. Paul, Minn. USA, p. 73, P021.
80. P. Weber, *Methods Enzymol.* **276** (1997) 13–22.
81. J. R. Luft and G. T. DeTitta, *Methods Enzymol.* **276** (1997) 190–130.
82. G. L. Gilliland, M. Tung, D. M. Blakeslee, and J. E. Ladner, *Acta Crystallogr., Sect. D* **50** (1994) 408–413.
83. G. L. Gilliland, *Methods Enzymol.* **277** (1997) 546–556.
84. T. M. Bergfors (Ed.), *Protein Crystallization: Techniques, Strategies and Tips, Laboratory Manual,* International University Line, La Jolla, CA.
85. N. C. Gassner and B. W. Matthews, *Acta Crystallogr., Sect. D* **55** (1999) 11967–1970.
86. G. M. Sheldrick, Golden Anniversary Meeting 1950–2000, ACA 2000, July 22–27, 2000, St. Paul, Minn. USA, p. 66, 10.01.04.
87. P. Lindley, European Synchrotron Radiation Facility, Grenoble, personal communication.

Highlights

88. P. Coppens, *X-Ray Charge Densities and Chemical Bonding*, Oxford University Press, Oxford, 1997.
89. P. Coppens, *Acta Crystallogr., Sect. A* **54** (1998) 779–788.
90. A. Volkov, Y. Abramov, P. Coppens, and C. Gatti, *Acta Crystallogr., Sect. A* **56** (2000) 332–339.

91. S. Dahaoui, V. Pichon, J. A. K. Howard, and C. Lecomte, *J. Phys. Chem. A* **103** (1999) 6240–6250.
92. M. Souhassou, R. H. Blessing, *J. Appl. Crystallogr.* **32** (1999) 210–217.
93. N. E. Ghermani, N. Bouhmaida, and C. Lecomte, ELECTROS: Computer program to calculate electrostatic properties from high-resolution X-ray diffraction, Internal report URA CNRS 809, Université Henri Poincaré, Nancy 1, France, 1992.
94. J.-M. Lehn, *Supramolecular Chemistry, Concepts and Perspectives,* VCH, Veinheim, 1995.
95. L. R. MacGillivray and J. L. Atwood, *Nature* **389** (1997) 469–472.
96. L. R. MacGillivray and J. L. Atwood, *Angew. Chem., Int. Ed. Engl.* **38** (1999) 1018–1033.
97. G. W. Orr, L. J. Barbour, and J. L. Atwood, *Science* **285** (1999) 1049–1052.
98. R. C. Liddington, Y. Yan, J. Moulai, R. Sahli, and T. L. Benjamin, *Nature* **354** (1991) 278–284.
99. K. M. Reinisch, M. L. Nibert, and S. C. Harrison, *Nature* **404** (2000) 960–966.
100. H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and E. R. Smalley, *Nature* **318** (1985) 162–163.
101. J. E. Johnson and V. S. Reddy, *Nature Struct. Biol.* **5** (1998) 849–852.
102. J. D. Dunitz, *Thoughts on Crystals as Supermolecules,* in: G. R. Desiraju (Ed.), *The Crystals as a Supramolecular Entity, Perspectives in Supramolecular Chemistry,* Vol. 2, Wiley, New York, 1996, pp. 1–30.
103. N. C. Seeman, *Chem. Intelligencer* **1** (1995) 338–347.
104. N. C. Seeman, *Angew. Chem., Int. Ed. Engl.* **37** (1998) 3220–3238.
105. N. C. Seeman, *Trends Biotechnol.* **17** (1999) 437–443.
106. N. C. Seeman, in: J.-P. Sauvage and C. Dietrich-Buchecker (Eds.), *Synthetic DNA Topology, Molecular Catenanes, Rotaxanes, and Knots*, Wiley-VCH, Weinhem, 1999, pp. 323–356.
107. N. C. Seeman, in: A. Carbone, M. Gromov, and P. Pruzinkiewitcz (Eds.), *DNA Nanotechnology: From Toplogical Control to Structural Control, in Pattern Formation in Biology, Vision, and Dynamics*, World Scientific Publishing Company, Singapore 2000, pp. 271–309.
108. N. C. Seeman, http://seemanlab.4.chem.nyu.edu./homepage.html.
109. E. Fischer, *Ber. Dtsch. Chem. Ges.* **27** (1894) 2985–2993.
110. P. Ehrlich, *Proc. R. Soc. London* **66** (1900) 424–448.
111. D. E. Kosland, *Angew. Chem., Int. Ed. Engl.* **33** (1994) 2375–2378.
112. M. A. Navia, P. M. D. Fitzgerald, B. M. McKeever, C.-T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer, *Nature* **337** (1989) 615–620.
113. R. L. Lapatto, T. L. Blundell, A. Hemmings, J. Overington, A. Wilderspin, S. Wood, J. R. Merson, P. J. Whittle, D. E. Danley, K. F. Geoghegan, S. J. Hawrylik, S. E. Lee, K. G. Sheld, and P. J. Hobart, *Nature* **342** (1989) 299–302.
114. A. F. Wilderspin and R. J. Surgue, *J. Mol. Biol.* **239** (1994) 97–103.
115. G. Klebe, *New Tools for Drug Design Based on Protein Ligand Recognition Principles,* in: F. Diedrich and H. Künzer (Eds.), *Recent Trends in Molecular Recognition*, Springer, Berlin, 1998.
116. G. Klebe, *J. Mol. Med.* **78** (2000) 269–281.
117. M. R. Wilkins, K. L. Williams, R. D. Appel, and D. F. Hochstrasser, *Proteome Research: New Frontiers in Functional Genome*, Springer, Berlin, 1997.

118. B. T. Wimberly, D. E. Brodersen, W. M. Clemons, Jr., R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan, *Nature* **407** (2000) 327–339.
119. A. P. Carter, W. M. Clemons, D. E. Brodersen, R. J. Morgan-Warren, and B. T. Wimberly, *Nature* **407** (2000) 340–348.
120. R. K. Agrawal, P. Penczek, R. A. Grassucci, N. Burkhardt, K. H. Nierhaus, and J. Frank, *J. Biol. Chem.* **274** (1999) 8723–8729.
121. H. Stark, E. V. Orlova, J. Rinke-Appel, N. Junke, F. Mueller, M. Rodnina, W. Wintermeyer, R. Brimacombe, and M. van Heel, *Cell* **88** (1997) 19–28.
122. J. Frank, J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Rademacher, R. Grassucci, R. K. Lata, and R. K. Agrawal, *Nature* **376** (1995) 441–444.
123. X. I. Yan, N. H. Olson, J. L. van Etten, M. Bergoin, M. G. Rossmann, and T. S. Baker, *Nature Struct. Biol*. **7** (2000) 101–103.

Prospects

124. D. Kleppner, *Phys. Today*, November (1998) 11–13.
125. H. R. Saibil, *Acta Crystallogr., Sect. D* **56** (2000) 1215–1222.
126. M. H. B. Stowell, A. Miyazawa, and N. Unwin, *Current Opinion in Structural Biology* **8** (1998) 595–600.
127. G. M. Rubin *et al*., *Science* **287** (2000) 2185–2195.
128. G. M. Rubin and E. B. Lewis, *Science* **287** (2000) 2216–2218 and references therein.
129. T. B. Kornberg and M. A. Krasnow, *Science* **287** (2000) 2218–2220 and references therein.
130. D. J. Wales and H. A. Scheraga, *Science* **285** (1999) 1368–1372.

SAŽETAK

(Bio)kristalografija na prijelomu tisućljeća

Biserka Kojić-Prodić i Jan Kroon

Prikaz je usmjeren na suvremenu ulogu rentgenske kristalografije i njezin utjecaj na kemiju i znanosti o životu. Otkriće strukture dvostruke zavojnice DNK iz dijagrama vlakna 1953. označilo je proteklo stoljeće, a projekt ljudskog genoma (koji će biti dovršen 2003. na 50. godišnjicu otkrića Cricka i Watsona) i proteomici, otvorit će novo tisućljeće. Određivanje kristalnih struktura čak i složenih bioloških agregata, kao što su virusi i ribosomi, ilustriraju snagu današnje rentgenske strukturne analize. U prikazu povijesne osnove dodirnute su neke od važnih etapa, počevši od samog otkrića rentgenskih zraka. Usavršene, nove metode i tehnologije modernog vremena – kao što su sinkrotroni treće generacije, osjetljivi detektori, suvremena računalska grafika, kriotehnike, metode mikrokristalizacije i genetičko inženjerstvo – doprinos su razvoju rentgenske kristalografije, što je opisano u posebnom poglavlju. Primjeri koji predočuju snagu metode odabrani su iz aktualnog istraživanja područja. Sistematika poznatih struktura »malenih« i makro-molekula te njihovih agregata, izgrađivana godinama, uvećala je naše znanje o kemijskoj vezi i međudjelovanjima atoma i molekula, koja određuju kemijska, fizikalna i biološka svojstva. Takve osnovne spoznaje

omogućuju predviđanje struktura i svojstava molekula što se pokazalo korisnim u mnogim područjima znanosti i tehnologije. Metoda je krenula sa strukturom kuhinjske soli, a razvila se u vrlo fini pribor za gledanje unutrašnjosti složenih živih sustava. Kristalografija vezana uz praćenje strukturnih promjena u vremenu i krioelektronska mikroskopija mogu bilježiti biološke procese i obavijestiti nas o dinamici enzima i o aktivnostima živih stanica. Zahvaljujući bioinformatici vrlo brzi događaji – zabilježeni eksperimentalno ili pretpostavljeni teorijski – mogu se interpretirati simulacijama molekulske dinamike. Veoma intenzivno prožimanje ideja iz različitih znanstvenih područja i tehnologija dovodi do kapitalnih otkrića koja otvaraju putove novim disciplinama kao što su molekulska genetika i molekulska medicina. Potonje dvije grane, obogaćene otkrićem struktura DNK i proteina, mogu ponuditi efikasne terapije za mnoge bolesti, očuvanje zdravlja i ublažiti probleme starenja. Naši pogledi na prirodu kemijske veze mijenjaju se, i naši će se vidici proširiti i postati jasniji u godinama koje dolaze. Poštujući etičke principe u znanosti, čovječanstvo će naučiti kako poboljšati kvalitetu života svim stanovnicima Zemlje.