Protein Crystallography from the Perspective of Technology Developments

Xiao-Dong Su[1], **Heng Zhang**[1], **Thomas C. Terwilliger**[2], **Anders Liljas**[3], **Junyu Xiao**[1], and **Yuhui Dong**[4]

[1] State Key Laboratory of Protein and Plant Gene Research, and Biodynamic Optical Imaging Center (BIOPIC), School of Life Sciences, Peking University, Beijing 100871, China

[2] Bioscience Division, Los Alamos National Laboratory, Mail Stop M888, Los Alamos, NM 87545, USA.

[3] Department of Biochemistry and Structural Biology, Lund University, Lund, Sweden.

[4] Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, People's Republic of China

Abstract

Early on, crystallography was a domain of mineralogy and mathematics and dealt mostly with symmetry properties and imaginary crystal lattices. This changed when Wilhelm Conrad Röntgen discovered X-rays in 1895, and in 1912 Max von Laue and his associates discovered X-ray irradiated salt crystals would produce diffraction patterns that could reveal the internal atomic periodicity of the crystals. In the same year the father-and-son team, Henry and Lawrence Bragg successfully solved the first crystal structure of sodium chloride and the era of modern crystallography began. Protein crystallography (PX) started some 20 years later with the pioneering work of British crystallographers. In the past 50-60 years, the achievements of modern crystallography and particularly those in protein crystallography have been due to breakthroughs in theoretical and technical advancements such as phasing and direct methods; to more powerful X-ray sources such as synchrotron radiation (SR); to more sensitive and efficient X-ray detectors; to ever faster computers and to improvements in software. The exponential development of protein crystallography has been accelerated by the invention and applications of recombinant DNA technology that can yield nearly any protein of interest in large amounts and with relative ease. Novel methods, informatics platforms, and technologies for automation and high-throughput have allowed the development of large-scale, high efficiency macromolecular crystallography efforts in the field of structural genomics (SG). Very recently, the X-ray free-electron laser (XFEL) sources and its applications in protein crystallography have shown great potential for revolutionizing the whole field again in the near future.

Keywords

2014 is celebrated as the International Year of Crystallography (IYCr-2014). About 100 years ago Max von Laue in Germany and the father-and-son Bragg team (William Henry Bragg and William Lawrence Bragg) in England pioneered the use of X-ray to irradiate crystals for atomic structure determination, and for those pioneering investigations they were awarded Nobel prizes for physics in the years 1914 and 1915 [1, 2, 3, 4, 5] (see Table 1). This was the birth of X-ray crystallography which has profoundly changed the way we perceive the atomic structures of the world built of organic and inorganic materials and molecules. With the X-ray crystal diffraction technique one can really "see" the atomic arrangements in a crystal lattice and accurately measure atomic bond distances, bond and dihedral angels and other parameters. This paper will review some important technical developments relevant to crystallography in general and in particular focus on the application of X-ray crystallography to study the structure and function of proteins and other bio-macromolecule complexes. X-ray crystallography, though a very specialized field, has received very significant attention from the Nobel committees compared with other similar scientific fields. Table 1 lists the milestone discoveries and technique developments recognized by Nobel prizes related to protein crystallography and protein structures (X-ray, electron and neutron crystallography). This list does not contain all important contributions that made crystallography as it is today of course. A recent review on the Nobel-crystallography topics has covered the work of most of the Nobel laureates who made major contributions to crystallography [6]. Our paper will pay attention to other important contributors to protein crystallography as well.

X-ray crystallography has emerged as one of the most important scientific and technical tools in the 20th century. Most solid state matter (rocks and soils) on earth is in the crystalline state, and crystallography is particularly well-suited to study the formation, growth mechanism, atomic structures and functions of all kinds of crystalline matter. Further, crystallography is an interdisciplinary research field interfacing with mathematics, physics, chemistry, mineralogy, metallurgy, biology, earth sciences, material sciences, medical sciences and drug discovery research. The core techniques for crystallography are the methods for determining atomic structures of any diffracting crystalline material with high accuracy and high resolution. Since three-dimensional atomic structural information is useful in almost all branches of science, it is fair to say that crystallography is a foundation of modern science. There are two types of crystallographic applications, one is to investigate the atomic structures of naturally occurring crystals. These applications are mostly related to mineralogy, metallurgy and earth/planet sciences and can be said to study the natural atomic structures in crystalline materials. The other type of crystallography is to determine molecular structures at atomic resolution of mostly "artificial" crystals. This branch includes structure determination of organic compounds, particularly proteins and biological macromolecules, for structure-function relationship investigation and applications. The key step for this latter application of crystallography is to obtain high quality diffracting crystals. This step requires careful preparation of samples and crystals.

Biological macromolecular crystallography (often shortened as protein crystallography, PX; or macromolecular crystallography, MX) is an interdisciplinary research field interfacing crystallography (not limited to X-ray, as electron and neutron waves are also commonly used) and biology. This field has experienced tremendous and nearly exponential growth

over the past 20 years (Figure 1). By now, over 100,000 three-dimensional structures of bio-macromolecules (including proteins, nucleic acids, carbohydrates and their complexes) have been determined, overwhelmingly by crystallographic methods. Based on these structures, our understanding of important life processes has reached an unprecedented level of detail. The structure determination of an important bio-macromolecule often leads to an improved understanding of its structure-function relationship. The DNA double helix by Watson and Crick; the myoglobin structure by John Kendrew; hemoglobin by Max Perutz; the structures of penicillin and vitamin B_{12} by Dorothy Hodgkin, all gave crucial insights into the functions of these molecules. These high resolution and accurate structures further led to the understanding of important life processes, the onset mechanisms of diseases, and the discoveries of medicines. With such developments, protein crystallography has entered into a new phase in which the combination of crystallography and biological research has given rise to an important frontier scientific research field, structural biology. Currently, membrane-spanning proteins and bio-macromolecular complexes and assemblies are the main cutting-edge research areas in the field. In recent years, with the maturation of high throughput, large-scale and low-cost structural determination techniques such as structural genomics (SG), the use of protein crystallography has become more and more widespread in many branches of life sciences, and the number of structures determined has increased exponentially in the recent 20 years. This trend will promote our understanding of life processes into a newer and deeper level based on the three-dimensional structures of many if not all biological macromolecules.

The first documented crystalline protein was hemoglobin, identified more than 170 years ago (lower-left corner of Figure 1). In 1840, Friedrich Hünefeld occasionally discovered protein crystals when worm blood was pressed against glass surfaces [15]. Various "blood crystals" from the blood of invertebrates and vertebrates were described in the following years [16, 17, 18, 19, 20]. Between 1885 and 1894, crystallization of plant seed proteins was discovered by Hartig and extended by Osborne and others [21, 22]. In the 1890s, Hofmeister (who discovered the Hofmeister series that describes the effects of different salt ions on protein solubility) and Hopkins crystallized hen egg albumin and horse serum albumin, respectively [23, 24]. Systematic work on enzyme studies enabled Sumner to crystallize the first enzyme (urease) and to correlate urease enzymatic activity with the presence of this protein [25]. At almost the same time, crystals of the hormone insulin were obtained by Abel [26, 27]. Following urease, many other enzymes were isolated in crystalline form by Sumner and Northrop and they shared the 1946 Nobel Prize in Chemistry with Stanley, who claimed to have obtained the first virus (Tobacco mosaic virus, TMV) crystal in 1935 [28, 29, 30]. Bernal and Crowfoot (later changed to Hodgkin after marriage) obtained the first X-ray diffraction pattern of protein (pepsin) crystals kept hydrated in a capillary in 1934 and demonstrated that it should be possible to solve its atomic structure. This was the event that really started protein X-ray crystallography [31]. Kendrew determined the first three-dimensional protein structure (myoglobin) following Perutz's general strategy of multiple isomorphous replacement (MIR) for obtaining phase angles by X-ray crystallography in 1958 [9]. Subsequently, Perutz solved the crystal structure of hemoglobin by MIR [11, 32]. In the 1950s-1970s, several great steps were made and furthered our mechanistic understanding of biology, including Watson and Crick's description of the structure of DNA

double helix in 1953 [33], giving rise to modern molecular biology. X-ray diffraction of DNA crystalline fibers contributed greatly to the discovery of the DNA double helix. David Philips' laboratory determined the structure of lysozyme [34, 35]; this was the first enzyme structure and it opened up the field of structural enzymology and made a great leap forward in understanding enzyme catalytic mechanism as reviewed by [36]. Antibody structural studies opened the field of structural immunology [37, 38, 39, 40], and the first RNA structure was transfer RNA (tRNA) [41, 42]. In the early years of protein crystallography, proteins and nucleic acids studied were all purified from naturally abundant sources. Rare but important proteins could only be studied after the invention of recombinant DNA techniques by Stanley Cohen in 1970s [12, 13, 14] and the resulting progress in expression of exotic proteins in hosts such as *E. coli*. Other crucial techniques responsible for the exponential growth of PX included synchrotron radiation (SR), computer algorithms for obtaining crystallographic phase information, and structural genomics (SG) as demonstrated in Figure 1. Finally, the X-ray Free-Electron Laser (XFEL) is an emerging technique showing great promise for changing PX and structural biology's future again [43, 44, 45, 46].

As early as the 1840s Friedrich Hünefeld wrote an extensive book on his observations of earth worm haemoglobin crystallization that would indicate that proteins, or at least haemoglobins, must have definitive molecular homogeneity and structures. In fact, crystallization had been a method used by organic chemists to characterize their compounds before 1840, and re-crystallization was used as an important means for purification of small molecule compounds. Nevertheless, the molecular and structural implications of haemoglobin crystallization did not lead scientists until much later to realize that proteins must possess well-ordered structures, and furthermore that these structures would be important for understanding their functions, for example, enzymatic catalysis activity. Richard Willstätter, a Noble laureate based on studies of plant pigments studies and a leader among enzyme chemists in the 1920's, concluded that the protein ('colloidal') component was a non-specific carrier of adsorbed chemical catalytic agents such as metal ions which contain the catalytic powers [47]. When James Sumner crystallized jack bean urease in 1926 [25], the biochemical community refused to believe that the crystals represented the enzyme, and it took more than 10 years and a few more enzymes purified and crystallized by Northrop to convince the community that the protein entity itself possessed the enzymatic activity. It took much longer to understand the catalytic mechanisms of enzymes, mostly from crystal structure studies, as insightfully reviewed by David Blow [36].

As mentioned earlier, the major bottleneck for protein crystal structure determination is to obtain diffraction-quality crystals. In the first half of the 19th century, pioneer crystallographers had realized the importance of pH, temperature, salts, and organic solvents and attempted to control and optimize protein crystallization conditions. Producing high quality crystals, however, was much more of an art than a science during that period. Crystallographers made progress on crystallization methods and theory (the vapor-diffusion method, screening strategies, new precipitants) from the 1960s, accelerating the development of protein crystallography [48, 49, 50]. The situation did not dramatically change and even now we can not really crystallize every protein as we desire. However,

with lots of effort made and lessons summarized, we have an increasingly clear picture of protein crystallization as well-reviewed recently [51].

TECHNOLOGY DEVELOPMENTS FOR PROTEIN CRYSTALLOGRAPHY

Crystallography is one of the major intellectual endeavors that span human history for centuries. After generations of consistent efforts by scientists and mathematicians, it has been transformed from a mathematical hypothesis to physical reality, mainly due to the contributions of the X-ray diffraction technique. In recent decades it has become a mature technical field. For protein crystallography, there are several crucial scientific/technical developments that have made the field possible. These developments can be categorized as soft, hard and wet-lab techniques.

SOFT TECHNOLOGY

Theoretical foundations, computational algorithm, and Software development

Theoretical background—Crystallography started from ancient time when early human beings paid attention to the symmetry and reproducibility of rock (mineral) materials with different sizes and symmetric shapes. The highly symmetric morphology led to the idea of periodic growth of units (atoms or molecules) into three-dimensional space and the concept that any kind of crystals could be composed of some kind of periodic lattice of units. For such lattices, Miller indices were introduced in 1839 by the British mineralogist William Hallowes Miller [52]. Simultaneously in France, the Bravais lattice, studied by Auguste Bravais (1850) was defined as an infinite array of discrete points generated by a set of discrete periodic translation operations [53]. Mathematicians such as Sohncke, Fedorov and Schonflies studied possible rotational and translational symmetries and their combination that could exist in a three-dimensional lattice. In 1879 they listed the 65 space groups (sometimes called Sohncke space groups or chiral space groups) whose elements preserve orientation. The full list of 230 possible space groups finally emerged in 1892 based on a correspondence between Fedorov and Schonflies. British geologist William Barlow later enumerated the groups with a different method. Burckhardt describes the history of the discovery of the space groups in detail [54]. A definitive source regarding three-dimensional space groups is the *International Tables for Crystallography* [55].

Fourier analysis (Fourier transform and Fourier series analysis) is another crucial mathematical foundation of crystallography. About 200 years ago, Joseph Fourier invented Fourier analysis theory for heat transduction [56]. Since then, Fourier analysis has become a very powerful and general mathematical tool for analysis of periodic functions. However, for crystallography, the Fourier space (reciprocal space) is more than just mathematical terms, it is a physical reality due to plane waves scattering (diffraction) through crystals under Fraunhofer conditions and the intensities of Fourier terms can be measured directly.

Although atomic structures in a unit cell must be expressed as non-negative real functions, their Fourier transforms contain complex components. In general, the experimental measurements of X-ray intensities can only provide the magnitude of these "Fourier components", but not the phases. Therefore, the most serious step in solving a crystal

structure *via* a Fourier transform is the so-called "phase problem", i.e. to restore those phases lost in data collection experiment by whatever means. Many theoreticians in crystallography have tried to use a "non-negative real functions" criteria to solve the "phase problem". Among them, David Sayre was able to conclude that, for an equal-atom structure (atomicity hypothesis), the phase of F(h) is related to that of the product $\Sigma F(k)/F(h-k)$. This triple product sign relationship is all there in the equation that he included in his article and became known as 'Sayre's equation,' exact for an equal-atom structure, and an important advance in our understanding of direct methods [57]. Many more have contributed greatly to the development of direct methods, some of them even obtained Nobel prizes (see table 1) for their work. "Direct methods" have already become a routine magic box to solve most small-molecule the structures automatically. However, for protein crystallography the use of direct methods is still limited.

Computational algorithm—Before the invention of computers in 1940s, Fourier transforms were calculated manually, but even after the computer was used for crystal structural calculation after the World War II, the speed and fidelity of the Fourier transform calculation by computers was still not satisfactory. A fast Fourier transform (FFT) is an algorithm to compute the discrete Fourier transform (DFT) and its inverse with high speed and fidelity. There are many different FFT algorithms involving a wide range of mathematics, from simple complex-number arithmetic to group theory and number theory. The Cooley-Tukey FFT algorithm was popularized in 1965 [58], but some FFTs were known as early as 1805. Fast Fourier transforms have been described as "the most important numerical algorithm of our lifetime", it is certainly the single most important numerical algorithm implemented in crystallographic computations. In addition to crystallography, FFTs are widely used for many applications in mathematics, engineer and other branches of science and technology.

CCP4 (Collaborative Computational Project No. 4: Software for Macromolecular X-Ray Crystallography4) is by far one of the most popularly used software packages for protein crystallography. CCP4 was set up in 1979 to support collaborations between researchers working on such software in the UK, and to assemble a comprehensive collection of software to satisfy the computational requirements of the relevant UK groups. CCP4 has since become more and more widespread among European protein crystallographic laboratories and also worldwide since the 1990s [59, 60]. One of the most important programs in CCP4 is the so-called "fft" (later developed into more popular used sfall) adapted by Ten Eyck, L. and others from the above FFT algorithm for PX applications [61, 62, 63]. There are about 200 different stand-alone programs currently in the CCP4 package.

Phasing and computational developments—As stated above, the most difficult problem for *ab initio* determination of protein crystal structures is the so-called "phase problem", i.e. to determine the phase angle of each measured reflection. This problem is easier in small molecule crystallography than PX since the small molecule crystals normally diffract to much higher resolution with smaller unit cells, which makes the number of phase angles to be solved much smaller, whereas many thousands of phase angles need be determined for PX structure determination. Therefore, computational methods have been

indispensable for PX development. Major breakthroughs in PX phasing methods development are reviewed below.

MIR (multiple isomorphous displacement) method: The multiple isomorphous replacement (MIR) approach was the first standard practical method to resolve "the phase angle problem" in PX, and it was originally pursued and established by Max Perutz (Perutz, 1985) at the Cavendish Laboratory, Cambridge. The theoretical basis of heavy-atom phasing [64, 65, 66, 67] was extended to implement more statistically powerful approaches later based on maximum likelihood, and a number of powerful and general programs (including some stand alone versions) are currently in use with CCP4 software suite: such as MLPHARE [68], SHARP [69], SOLVE [70], PHASES [71] and so on.

MAD/SAD methods: During 1950s, the Dutch crystallographer Bijvoet had first solved the absolute configuration of several small molecule crystal structures using intrinsic anomalous signals. This method formed the theoretical basis for the application of anomalous scattering signals in solving the "phase problem" in general. With improved data-collection techniques (particularly after the application of strong intensity and tunable wavelength X-ray sources from synchrotrons) developed during 1970s-80s, it was possible to accurately measure diffraction intensities in protein crystals. This gradually shifted PX towards the use of the anomalous signal as a primary source of phase estimations by the multiple or single-wavelength anomalous diffraction approaches (MAD or SAD). Protein crystals generally lack strong anomalous scatterers, which made the MAD/SAD methods not generally applicable. However, the introduction of selenomethionine (providing selenium atoms as strong anomalous scatterers) into proteins by Wayne Hendrickson's lab has been a great breakthrough in the application of MAD/SAD methods in PX [72], and resulted in the MAD/SAD phasing as the dominating methods in *ab initio* phasing for protein crystal structures [73].

MIRAS/SIRAS (multiple/single isomorphous replacement plus anomalous scattering) [74] is also used with heavy atom derivatives since all heavy atoms are usually good anomalous scatterers at certain wavelength, many programs can be used for this approach, e.g. in SHARP [69] and PHASER [75]. Single-wavelength SAD phasing, pioneered by Hendrickson & Teeter in 1981 [76] and Wang in 1985 [77], is technically simpler than MAD, and gained much more popularity during recent years. In terms of data collection, SAD may need more accuracy than MAD, especially when weak anomalous scatterers such as phosphorus or sulfur are used [78, 79], but SAD does not require precise wavelength tuning and can be performed even in home laboratories with Cu or Cr X-ray sources. Furthermore, protein crystal diffraction data decays with radiation damage over prolonged collection times. In some cases, the diffraction decay can be significant even within a data set, making it impossible to measure multiple anomalous data sets accurately. Faster and more accurate data collection strategies focusing on one single data set can however be applied for SAD data collection.

SSAD method: Sulfur-SAD (S-SAD) is a type of SAD method using very weak anomalous signals from intrinsic sulfur atoms presented in cysteine and methionine residues in proteins. Since it only requires one set of native diffraction data, BC Wang gave it a name "direct

crystallography"; the crystal structure can be calculated directly from one set of such native crystal data by S-SAD method in theory [77]. In fact, Hendrickson & Teeter were the first to solve a small protein crambin structure using S-SAD method using an in-house X-ray source [76]. In practice it is not always easy to solve protein structures by S-SAD due to low signal-to-noise ratio. A few examples of successful S-SAD applications are listed from SR source data collection and from home X-ray source [80, 81]. The introduction of the chromium rotating anode by Rigaku/MSC [82] and designs of new synchrotron beamlines optimized for the use of long wavelength (e.g. at Diamond; http://www.diamond.ac.uk/Activity/Beamlines/) are deliberately intended for the application of SAD phasing with light elements such as sulfur.

MR (molecular replacement) method: The molecular replacement method, originally initiated and developed by Michael Rossmann [83], is now responsible for more than half of all structures deposited in PDB (protein data bank, http://www.rcsb.org/pdb/home/home.do). Since almost all protein structure folds are thought to have been discovered already, thanks in part to structural genomics efforts, and as structural representatives of many protein families are known, it is quite clear that in the near future the vast majority of macromolecular structure determinations will be based on MR methods. This will be aided by the availability of protein sequences from vast numbers with the development of NGS (next generation sequencing) techniques.

In the early stages of MR development, the numerical calculation of the rotation function was slow and inaccurate. Tony Crowther then developed a fast rotation function that transformed the rotation function into an FFT algorithm and completely changed the field [84, 85] by making MR routine and fast. Due to lack of computing power, the early applications of MR usually involved two stages, first performing the three dimensional rotational search followed by the three dimensional translation function, as programmed in AMoRe [86] and MOLREP [87]. With the development of computing power following Moore's law (Figure 3), newer programs that can perform six-dimensional searches [88, 89], and software implementing maximum-likelihood principles [75], MR is now more powerful and successful in difficult cases even with relatively dissimilar search models [90]. Recent developments of MR methods include more accurate model generation and calculations combining Phenix and Rosetta[91, 92, 93].

Direct methods and assisted methods: Direct methods, first proposed by Sayre, Karle and Hauptman in the 1950s, have generally and successfully been applied to small molecule crystallography. This led to a Nobel Prize in chemistry for Karle and Hauptman in 1982 (Table 1). In PX, the location of heavy or anomalous atoms has heavily relied on direct methods, firstly by the early popular program MULTAN [94], later by using the direct method program SHELXD [95]), although complete solution of atomic resolution structures has also been accomplished for some small proteins. The most powerful direct method programs are based on the dual space (direct- and reciprocal-space) recycling principle, as implemented in SnB [96], SHELXD [95], ACORN [97] and CRUNCH [98].

A particular effort of using direct methods not to locate heavy atoms, but to improve phasing was made by the program OASIS developed by Haifu Fan and colleagues at the Chinese

Academy of Sciences [99, 100, 101]. OASIS is especially powerful in breaking phase ambiguities during SIR or SAD phasing.

Solvent flattening/flipping and density modification programs—In the early 1980s, B.C. Wang proposed the so-called solvent flattening algorithm as a powerful method of phase improvement and extension [77]. It is based on that idea that the packing of the protein in the crystal lattice can be divided into two different regions, the protein region with stronger electron density and the solvent region, with weak, noise-like electron-density that could be iteratively filtered out. Other restraints based on the known expected behavior of the electron density are also very powerful, such as the use of non-crystallographic symmetry (NCS) averaging [102] and histogram matching [103]. These density-modification procedures have been incorporated into many programs such as DM [104], and closely-related algorithms are implemented in RESOLVE [105] and SOLOMON [106].

Computer graphics—The presentation of a protein structure was very difficult before the invention of computer graphics (Perutz, 1985). The first widely used graphics program FRODO [107] was a real breakthrough for understanding protein atomic structure and for the rapid PX development in general. The PX computer graphics guru, Alwyn Jones, is a fan of J.R.R. Tolkien's "The Lord of the Rings", and used many names from Tolkien's books and figures to name his computer graphics programs and routines. The most widely used computer graphics systems include O [108], COOT [109], PyMol by the late Warren DeLano, XtalView [110], and QUANTA [111]. Computer graphics programs may often show users of protein structures with false precision in non-bonded distances particularly for inexperienced beginners, a detailed analyses of precision in non-bonded distances can be found in [112].

Computer graphics programs are not only used for display the molecules and electron-density maps. More importantly they are also used in combination with other programs for map interpretation and construction of molecular models, manually or more recently by automatic fitting and model building functions. The construction of the protein chain is very often performed without human intervention by automatic model building programs such as ARP/wARP [113], RESOLVE [105], and Buccaneer[114].

Structure refinement and structure validation programs—Earlier protein crystal structures were solved and published without refinement. This situation continued until late 1980s when structural refinement became a routine for PX, although least-squares refinement with geometry regularization had been already introduced in 1970s [63, 115]. This was followed by least-squares refinement with built-in geometry restraints as implemented in PROLSQ [116, 117], and other refinement programs such as TNT [118] .

A major breakthrough in PX structural refinement became possible when Axel Brunger programmed least-square refinement with geometric restraints combined with molecular dynamics (MD) simulation into X-PLOR [119]. Later X-PLOR was expanded to include NMR structure refinement and was superseded by the program suite Crystallography and NMR System (CNS) [120]. Since then, X-PLOR and CNS have become highly popular protein structural refinement tools and have been used all over the world. The application of

MD simulation in every structural laboratory could be arguably claimed as the most widespread application of MD simulation methods pioneered by Martin Karplus and colleagues and recognized very recently by a Nobel prize (see table 1). The over-fitting problem of protein structural refinement has been ingeniously solved by Brunger through introducing the objective validation tool of R-free calculated using a small fraction of reflections never used in the refinement calculation [121].

Other crystal structural refinement programs often used include SHELXL [122] which evolved from small-molecule crystallography and which has many options not available in other refinement programs. SHELXL is vital for full matrix inversion calculation which yield standard deviations on atomic coordinates, but it requires relatively high resolution, to overcome the resolution challenge, semi-quantitative validation methods have been developed [123, 124]. Other refinement software systems such as REFMAC [125] and BUSTER [126] are based on maximum-likelihood principles. REFMAC is also extensively utilized by the automatic model building program ARP/wARP during structural solution and automatic model building. Recently, the program solution and refinement suite Phenix has gained lots of popularity due to its automatic and user-friendly features [127, 128].

Structural model validation has always been important before a structure is finally finished and deposited into the PDB since the structural data might be misinterpreted and any errors will propagate into the relevant fields . Several programs have been especially developed for checking the correctness of protein models. The most popular and comprehensive are PROCHECK [129], WHAT_CHECK [130], SFCHECK [131] and MOLPROBITY [132].

HARDWARE TECHNOLOGY: Light sources, detectors and computer hardware

X-ray Light Sources

The home sources—In the early days after the discovery of X-ray by Röntgen until the 1970s, sealed tubes were the standard laboratory X-ray sources. In these sealed tubes electromagnetic radiation was produced by bombardment of anodes made of metals such as copper or molybdenum. Electrons were emitted by a thermionic cathode in the form of a lamp filament and were accelerated to energies of tens or hundreds of kilovolts. The sealed tube setup for X-ray crystallography was gradually replaced by rotating-anode machines in which the anode, kept in high-speed rotation, could withstand much more bombardment and deliver more intense X-ray beams (reviewed in [133]). The rotating-anode machines remain standard laboratory equipment for most X-ray crystallography labs worldwide where synchrotron radiation is not available. Uli Arndt, a guru of crystallographic X-ray source technology who developed many types of X-ray home sources, also developed a micro-focusing sealed tube that is almost as intense as rotating anode sources at the focusing spot and that has gained popularity in the recent years [134].

Copper anode targets have been the most popular among PX labs due to its strong intensity at Kα wavelength of 1.5418 Å, matching the diffraction limits of about 2-3 Å for many protein crystals. Molybdenum targets producing X-rays of 0.71 Å wavelength have mostly been used for small molecule crystallography requiring sub- Å resolution. The routine

application of chromium-based (wavelength = 2.23 Å) rotating anodes, useful for the sulfur SAD data collection and phasing on protein crystals, have been a recent commercial product in home source technology [82]. Although home X-ray sources are becoming more and more compact, automated and user-friendly for operation, they are also turning into accessory tools for screening protein crystals and to confirm crystallization and freezing conditions. The major drawbacks of home sources are their relatively weak intensity and fixed wavelength. This can be compared with synchrotron radiation (SR) X-ray sources which are becoming the "must" tool for anyone who carries out macromolecular crystallography.

Synchrotron radiation (SR) and X-ray free electron laser (XFEL)

The emergence and application of SR to biological samples and protein crystallography started with the pioneering work of Rosenbaum ,Holmes and Phillips [135, 136]. The SR produced by the first generation synchrotron (cyclotron) was once regarded as a hideous energy-wasting side effect needed to be avoided by design. This changed after Rosenbaum, Holmes and other crystallographers demonstrated that the SR could be used as a very strong X-ray source to get much higher resolution X-ray diffraction and improved signal-to-noise ratio. Since then, many applications have been found for SR. Dedicated synchrotrons for the generation of stronger SR (the second generation of synchrotrons) and specialized beamlines had been built for X-ray crystallography, particularly for PX applications. In the beginning, fixed-wavelength bending-magnet beamlines were the common solution. However, the majority of beamlines in current use are tunable-wavelength lines, often equipped with insertion devices such as undulators or wigglers that enhance the generation of X-rays.

In the 1990s, the advances in technology led to the construction of so-called the third-generation synchrotron sources, characterized by larger ring sizes (some of their diameters approach 1 km) and much higher beam brightness and stability. Such machines were first built in Grenoble, France (ESRF), Chicago, USA (APS), Harima Science Park City, Japan (SPring-8) and Berkeley, USA (ALS). Other recent examples include synchrotrons in Didcot, UK (DIAMOND), Shanghai, China (SSRF: Shanghai Synchrotron Radiation Facilities), Saint-Aubin, France (SOLEIL), Villigen, Switzerland (SLS), Cerdanyola, Spain (ALBA) and Hamburg, Germany (PETRA III). Figure 2 graphically describes the developments of the major events in SR and XFEL compared with other rapid growing techniques.

Data collection: detectors and methods

Diffraction data collection using sensitive X-ray detectors is a crucial step for experimental crystallography since all the following steps are computational and can always be repeated easily. Photographic films and precession cameras [137] were the first methods that could be used for PX data collection and data reduction and were successfully applied to early protein crystal structure determinations. Later, automatic three- and four-circle diffractometers [138, 139] were designed and used in crystallography. This kind of system was most successful for crystallographic data collection of small molecule crystals due to the inefficiencies of limited spatial resolution and collection of only one diffraction spot at once. Protein crystals normally diffract much more weakly and decay much faster than small molecule crystals,

and protein crystals have much larger unit cells as well. Simultaneous collection of the weak diffraction spots over a large area with high spatial resolution for PX data is absolutely required and this could be achieved by X-ray-sensitive film. The later application of area detectors from high energy detecting techniques (see Table 1) with sufficient sensitivity and spatial resolution has been a crucial step for PX development. During the development of detectors for PX applications, automatic computer-controlled two-dimensional detectors, such as multi-wire proportional counters, TV cameras, phosphorus-based imaging plates, charge-coupled devices (CCDs) and complementary metal oxide semiconductor (CMOS) detector have been applied in many PX labs and in general increased the accuracy and speed of measuring diffraction intensities. However, many of those types of detectors are becoming obsolete; currently widely-used detectors are CCD and CMOS based, and the pixel array detectors such as the Pilutas detector by DECTRIS Ltd. [140, 141] are becoming the new standard detectors in PX applications. Kazuya Hasegawa at Spring-8 has recently developed an X-ray CMOS detector (Hamamatsu Photonics KK C10158DK) with a new shutterless continuous rotation method for protein crystallography [142]. The new detectors mentioned above, together with the shutterless continuous rotation method, will likely to bring PX data collection into a new era.

Data reduction programs—The first thing to do after data collection by current oscillation or shutterless continuous rotation method is indexing and integration of the Bragg reflections, then reducing the raw data into intensity or amplitude files for later structure determination. The most widely used data reduction programs are HKL 2000 (earlier version called Denzo) and later versions [143] with powerful auto-indexing routines; CCP4 suite derived MOSFLM [144]; simple but elegant XDS (which may not be regarded as user-friendly for beginners since it has no good graphics interface to look at the spots) with many versatilities [145]; d*TREK [146] and XGEN [147](A. J. Howard; http://xgen.iit.edu/). The most popular data reduction program for neutrons is Laugen [148].

Cryo-crystallography

Cryo-crystallography, i.e. collecting diffraction data at temperatures about 100 K, is a crucial technique development for PX, particularly for SR data collection since radiation damage is usually much more severe at synchrotron sources. The low temperature can protect the protein crystals during data collection. Cryoprotectants, composed of small molecules such as glycerol, ethanol, MPD, glucose, low MW PEG are small organic molecules, normally with many hydroxyl groups or certain salts at high concentration. Solutions containing cryoprotectants can be solidified in a state of amorphous glass upon freezing and this glass state may preserve crystalline order and the diffraction properties of the protein crystals [149]. Several types of cryo jets that inject a gaseous stream from a liquid nitrogen container to maintain the low temperature during data collection are commercially available. The protein crystals are usually scooped from the mother liquor with a small nylon fiber loop (or other equivalent small loop-like structures made from X-ray transparent materials) and quickly plunged in the liquid nitrogen or placed in the stream of the cold nitrogen at temperatures about 100 K [150].

WET LAB TECHNOLOGY

Protein purification, crystallization and recombinant DNA techniques

From the 1950s through the end of 1960s, obtaining proteins was gradually becoming a limiting factor for protein crystallography, and only a few protein structures were determined worldwide. A major breakthrough came in the early 1970s, when three scientists, Berg, Boyer and Cohen at Stanford University invented recombinant DNA technology [12, 14, 151]. It was soon found that the amounts of target proteins derived from recombinant sources could be sufficient to produce crystals. At almost the same time, the establishment of Protein Data Bank (PDB) was announced at a meeting held at Cold Spring Harbor Laboratory (CSHL). The application of molecular biology has contributed in a major way to the rapid growth of the PDB.

Early on, protein crystallization would mostly be carried out by salting-out, for example using ammonium sulfate. Today there are many ways to crystallize proteins (and nucleic acids) [48, 152]. The most popular crystallization setups are vapor diffusion methods using the hanging-drop [153] or sitting-drop approaches. Membrane proteins have been crystallized ways similar to soluble proteins. One specific crystallization method, 'in cubo', i.e. in the scaffold of the cubic lipid phase for membrane proteins was been developed more recently [154, 155].

Organic solvents and polymer molecules such as PEG have been found very useful for protein crystallization, and various sets of crystallization screening conditions including selected conditions by sparse-matrix sampling [50] have been developed. Currently many ready-made solutions are available commercially, including for example products by: Hampton Research, http://www.hamptonresearch.com/; Emerald Biostructures, http://www.emeraldbiostructures.com/; Jena Bioscience, http://www.jenabioscience.com/; Molecular Dimensions, http://www.moleculardimensions.com/.

Automation for setting up crystallization conditions has been successfully introduced [156, 157], and currently many academic and industrial laboratories are equipped with crystallization robots.

The developments of wet lab technologies: Protein production and purification

Recombinant expression of exotic proteins in *Escherichia coli* (*E. coli*) is the most commonly used method in the structural biology research field. Though straightforward and cost- and time-effective, this prokaryotic system has its limitations. Many eukaryotic proteins such as secretory molecules, membrane proteins, and large protein complexes require post-translational modifications and an advanced chaperone machinery for proper folding and assembly. Eukaryotic expression systems including yeasts, insect cells, and mammalian cells have been established to produce such proteins (reviewed in [158, 159, 160]). The well-studied baker's yeast, *Saccharomyces cerevisiae* and the methylotroph *Pichia pastoris* are frequently used as alternatives to *E. coli*, owing to their ability to grow in simple and inexpensive media [161, 162]. For example, the *P. pastoris* expression system has been successfully applied to the structural studies of the potassium channels [163, 164]. The baculovirus expression system (BVES) and insect cell system (Sf9/Sf21 cells from

Spodoptera frugiperda, and Tn5 cell from *Trichoplusia ni*) were developed more than thirty years ago, and is the most commonly used eukaryotic expression system [165, 166]. Notably, G-protein coupled receptors are often generated this way [167, 168]. To circumvent problems caused by the lytic nature of baculoviruses, S2 cells from *Drosophila melanogaster* have been increasingly used [169]. Mammalian cells, including the CHO (chinese hamster ovary) and the HEK 293 (human embryonal kidney) cell lines are also often used [170]. As a reflection of the maturity of the mammalian expression technology, the mTOR kinase structure was recently solved using protein samples produced from large amounts of HEK293 cells [171]. In the post structural genomics era, we expect that eukaryotic expression systems will be brought into the mainstream in order to tackle difficult proteins and biological questions.

Cell free protein production has been exploited and widely applied as reviewed by Yokoyama and colleagues [172, 173].

RECENT AND FUTURE PERSPECTIVES: Structural genomics and X-ray free-electron laser

Structural Genomics 2000-2015

At the end of the 20th century many macromolecular crystallographers were inspired by the Human Genome project and the rapid developments in methods for structure determination. They envisioned that further automation of structure determination would allow large-scale and systematic structure determination of macromolecules. This genomics-inspired approach became known as Structural Genomics (or Structural Proteomics). Several centers worldwide started to be engaged in structural genomics projects in Europe [174], the USA [175], and Japan [176].

The Human Genome project and the associated massive genomic sequencing of a wide range of organisms became a major impetus for Structural Genomics [177, 178, 179, 180]. The availability of genomic sequences for a rapidly-increasing number of organisms made it possible to think about systematically determining the structures of groups of macromolecules. These sequences showed the extent of homology among genes from related organisms. These same sequences made the effort required to clone genes and express them in suitable host organisms such as *E. coli* vastly simpler than it had been previously. These sequences further opened up the possibility of testing members of a group of homologous genes and identifying which ones were the most suitable for structure determination.

The approaches and philosophy of these genome projects also contributed to the strategy and philosophy of Structural Genomics. The genome projects showed that technology development and a high level of automation could make a process such as DNA sequencing that was slow but gave very solid information into a high-throughput and prodigious information source. The genome projects showed that early release of information could spur rapid use of this information and that a high level of public data sharing was compatible with cooperative academic work. The Wellcome Trust, the Japanese ministry MEXT, and the US National Institutes of Health, major players in the human genome project, led efforts

to coordinate early Structural Genomics planning at the turn of the 21st century. These efforts led to a vision for Structural Genomics and the founding of the International Structural Genomics Organization (http://www.isgo.org) with a goal of fostering international cooperation and sharing in Structural Genomics[181].

A second major impetus for Structural Genomics was the rapid advance in technologies for structure determination. For macromolecular X-ray crystal structure crucial factors included the availability of tunable and powerful synchrotron radiation, methods for reducing the damage to crystals from that same radiation, and methods for obtaining crystallographic phase information from anomalous scattering. The systematic use of selenomethionine as an anomalously scatterer that could be readily incorporated into proteins as a powerful agent for obtaining crystallographic phases has dramatically changed structural biology and greatly contributed to the vision and realization of Structural Genomics [72]. The development of algorithms for automatic interpretation of macromolecular diffraction data and generation of atomic, combined with the extensive automation in gene sequencing and the beginnings of automation in gene cloning and protein expression suggested to many that ever more automated and rapid structure determination could be carried out [182, 183, 184, 185, 186].

Technologies from Structural Genomics—The accomplishments of the Structural Genomics efforts between 2000 and 2015 include the development of a suite of technologies and approaches for large-scale structure determination that have dramatically changed structural biology as a whole, along with the determination of thousands of new protein structures.

During the first decade and a half of the 21st century, many Structural Genomics groups around the world developed 'pipelines' for structure determination. These were systematic and often highly automated procedures for targeting which proteins to focus on, expressing genes, purifying proteins, crystallizing proteins, and determining protein structures by X-ray diffraction and NMR [182, 187, 188, 189]. This systematic approach to structure determination later became the standard approach. Today, many structural biology laboratories have pipelines for structure determination. These pipelines are based on systematic targeting of groups of proteins for structure determination, new technologies for every aspect of protein production, robotics for automated cloning, protein expression, systematic approaches for X-ray data collection and structure determination, systematic evaluation of technologies, and automated procedures for analysis and reporting of the resulting structures.

Some of these pipelines for structure determination were highly integrated and systematic. Fig. 3 illustrates the pipeline from the Joint Center for Structural Genomics showing the linkage of technologies and data through a central database and the systematic targeting and analysis made possible by the integration of approaches.

Target selection in Structural Genomics and many individual laboratories in this period consisted of highly systematic analyses of sequences of many genomes, followed by selection of a wide range of sequences homologous to the protein of greatest interest so that

the chances of obtaining new structural insights were maximized. In many Structural Genomics laboratories, particularly in the US, the target selection also included an attempt to maximize the likelihood of obtaining novel structural information by solving the structures of proteins that were substantially different in sequence from any proteins with known structures.

Highly automated procedures for cloning, often using specialized cloning vectors optimized for systematic cloning and robotics and multi-channel pipetting for repetitive operations, were used in many Structural Genomics laboratories and subsequently used widely in individual laboratories. Testing the expression of proteins and identification of optimal expression vectors, organisms for expression (*E. coli*, yeast, viral expression in eukaryotic cells), and expression conditions became increasingly carried out with robotics or at least in highly systematic procedures. Protein purification robotics capable of highly parallel purifications of proteins were used in Structural Genomics, while individual laboratories used simpler automated systems for protein purification [190].

Crystallization and crystal observation robots were developed by many Structural Genomics groups at the turn of the 21st century and became widespread in the structural biology community over the next decade [191, 192, 193, 194]. Automated methods for analysis of crystallization droplets and identification of crystals were developed during this period. The reliable automatic identification of small crystals proved to be much more difficult than many initially imagined and these automated procedures most commonly have been used to prioritize droplets for human visual examination rather than as a decision-making tool for choosing exactly what crystals to use in X-ray diffraction experiments.

An important contribution of Structural Genomics to structural biology technologies was providing a strong impetus for development and optimization of highly automated beamlines for X-ray diffraction analysis at synchrotrons. This included the critical automation of mounting of cryo-cooled crystals on a beamline and the development of standardized crystal mounting systems so that the same crystal holders could be used on many different beamlines [195, 196]. It also included the development of highly automated systems for X-ray data collection, including many that could be used remotely so that users could send crystals to a beamline and collect their data from their own laboratories [197]. Automated software for processing the images from X-ray diffraction experiments was available before this period but major improvements and new software was developed in large part for Structural Genomics and became widely used in individual laboratories as well.

Some automation of structure determination by multi-wavelength methods and significant automation of model-building had been developed just prior to the beginning of Structural Genomics and provided part of the foundation for the effort [91, 198, 199]. The development of technologies for structure determination from X-ray diffraction data of macromolecules was then accelerated by the needs of Structural Genomics groups for rapid and automated structure determination. Integrated systems for initial structure solution from multi-wavelength and other experimental X-ray diffraction data followed by structure determination and nearly-finished molecular models were developed and became widely used in the Structural Genomics and the broader crystallographic community.

The highly systematic structure determination carried out by many Structural Genomics groups made it feasible to partially automate the validation, deposition, and annotation of structures. Many improvements in automated validation of protein and nucleic acid structures were made during this period, and were applied to the checking and improvement of structures from Structural Genomics groups. The use of comprehensive databases to track data for Structural Genomics projects made deposition of structures much easier for both the depositors and for the Protein Data Bank receiving the structures. Several Structural Genomics groups partially automated procedures for generating annotations of crystal structures and displaying them on public web sites. The US Protein Structure Initiative developed a publicly available "KnowledgeBase" that provided systematic access to much of the data generated by US Structural Genomics projects and individual Structural Genomics projects around the world developed websites with comprehensive information on their targets, status, and results [200].

The high-throughput (HTP) and automation oriented technology and pipelines developed and commercialized by the big consortia and relevant companies can now be easily adopted by any "standard" structural biology laboratory at lower costs, but with similar scale and efficiency, as exemplified by a laboratory in Peking university, China.

The Su lab has started to establish an SG platform several years ago at the level of a university laboratory [201, 202], adopting automation and HTP techniques from other SG centers, and further developing crystallization tools and imaging equipment etc. methods. We have applied the SG approaches to the structural and functional characterization of proteins from a dental caries pathogen *Streptococcus mutans*. Using the platform, we have cloned about 1,400 genes encoding non-membrane proteins from *S. mutans* genome containing 1963 ORFs [183, 203] into a variety of pET vectors using mainly conventional cloning techniques. We then over-expressed and purified soluble proteins by using an *E. coli* expression system. The purified proteins were screened for crystallization. Data collection and structure determination were performed on the target proteins capable of forming diffracting crystals. So far, we have cloned about 4000 genes from different organisms on this platform, crystallized more than 300 different proteins and solved over 100 protein crystal structures.

XFEL Applications

In 2009, SLAC (Stanford Linear Accelerator Center) National Accelerator Laboratory launched a new facility, Linac Coherent Light Source (LCLS), which provided the world's first hard X-ray free-electron laser (XFEL). XFEL is a brand new kind of X-ray source, complementary to SR sources but with approximately 10 orders of magnitude increase of the peak intensity. The development and application of XFEL sources will definitely open new fields and provide novel opportunities for protein crystallography and structure biology [204].

Obtaining high quality crystals has been a major bottleneck for protein crystallography, but with the development of XFEL this bottleneck may be dramatically reduced or even completely overcome in the near future. Due to the very strong and coherent pulses of XFEL, even very small crystals, with sizes of 100 nm (typically around ten layers of

macromolecules), can provide enough diffraction signal for structure determination. In addition, weaker non-Bragg scattering can potentially be collected with XFEL sources to provide additional (over-sampled) information from the tiny protein crystals [205, 206]. Therefore, the current requirement for the sizes and qualities of protein crystals may be markedly reduced. The extremely high peak intensities of XFEL completely remove the possibility of protecting the samples from radiation damage, each crystal can be only used once for data collection. However, the short durations of XFEL pulses, ranging from a few to hundreds of femto-seconds, offer the chance to collect diffraction data before the sample is destroyed. This is called "diffraction-before-destruction". In such experiments where one sample can only contribute one diffraction pattern, automated experimental devices are needed for development for efficient data collection.

The feasibility of "diffraction-before-destruction" had been proven by Henry Chapman in 2006 [207], before the appearance of LCLS. In 2011, Chapman reported diffraction of photosystem I nanocrystals [205] at low resolution. The sizes of the crystals he used ranged from 200 nm to a few μm. The experimental setting in using XFEL is very different from the conventional protein crystallography, as shown in Fig. 5. In order to distinguish with the conventional diffraction methods, this kind of experiment is now called as "nanocrystallography" or "serial femtosecond crystallography".

In 2012, the group of Henry Chapman demonstrated again the high-resolution structure determination of lysozyme crystals [43]. The phases of the reflections were solved by molecular replacement. A novel structure determination of cysteine protease cathepsin B from *Trypanosoma brucei* was also reported in the same year by the Chapman group [208]. These two experiments have clarified beyond any doubt the feasibility and capability of XFEL in structure determination by crystallographic methods.

For *de novo* crystal structure determination, isomorphous replacement or anomalous scattering are necessary for phasing. In 2013, the group of Ilme Schlichting reported the structure of Gd-loaded lysozyme solved by single-wavelength anomalous scattering [209]. However, due to the ambiguity in indexing caused by multiple crystal settings, for the crystals with certain space groups, such as trigonal, hexagonal and cubic crystals of the 23 point group, the method for extracting anomalous scattering signals still needs to be further developed.

The main obstacle of nanocrystallography in XFEL now is the huge number of samples required and efficient data collection methods needed for the experiments. Because of the pulse characteristic of XFEL, the frequency that XFEL pulses hit the randomly injected nanocrystals is pretty low, usually around 1-3% with current experimental setups. The property of "diffraction-before-destruction" required a totally new method, so-called Monte-Carlo integration, for data reduction. This method needs a large number of diffraction patterns from randomly-oriented crystals in order to obtain accurate structure factors for structure determination, especially for anomalous scattering. Usually, the number of diffraction patterns needs to be in the order of 10 thousand or larger; therefore, the number of crystals used in an experiment may in the millions. Significant efforts in methodological research will need to be launched before XFEL can be used routinely in structural biology.

Some efforts in decreasing the number of patterns used in serial femtosecond crystallography have been in development. The group of Dong from Beijing Synchrotron Radiation Facility (BSRF) works on improving the Monte-Carlo integration [210]. By analyzing the sizes of the crystals and the profiles of the diffraction spots in diffraction patterns obtained from the geometry factors of the diffraction can be determined, allowing accurate estimation of the structure factor amplitudes and removing the need for the Monte Carlo integration method.. The new method only requires thousands of diffraction patterns to obtain the structure factors precisely enough for structure determination, which is at least one order of magnitude reduction of the conventional Monte-Carlo integration method.

The ultimate goal for XFEL is not merely solving the structures of tiny crystals, but obtaining structures without crystals. The extreme brilliance and coherence of XFEL are the key elements allowing for this possibility. In 2011, the group of Janos Hajdu demonstrated the single-particle scattering of mimivirus via XFEL [206]. Although the result of this experiment is far from solving the atomic-resolution structure of single particle, it depicts the feasibility of this approach.

In summary, the situation of XFEL is very similar to that of synchrotron radiation in 1980s. The feasibilities of its applications in many fields are authenticated but great amount of efforts in methodological research remain.

The first 100 years of crystallography and 80 years (counting from Dorothy Hodgkin's first pepsin crystal diffraction in 1934) of protein crystallography have created structural chemistry and structural biology, and several other related scientific fields that have made us to really "see" the atomic structures that had been conceptualized by the Greek sages thousands of years ago, but never had the opportunity to actually see them. Although our current crystallography lenses should allow us to precisely see the "static" atomic positions with sub-atomic accuracy, the atoms themselves are never static, they constantly move and vibrate with different degrees of flexibility determined by their bonded and more importantly, non-bonded molecular structures and dynamic features, with important functional consequences for biological processes. With current technologies, not limited to crystallography, we have just started to understand and to "see" the dynamic structures no matter how indirectly. It is reasonable to postulate that the technical and theoretical developments in the next 100 years may allow us to "see" the dynamic structures of biological macromolecules, particularly with emerging new tools such as XFEL.

Acknowledgements

References

1. Friedrich W, Knipping P, Laue M. Interferenz-Erscheinungen bei Röntgenstrahlen. Sitz Bayer Akad Wiss. 1912:303–22.

2. Bragg WH. X-rays and crystals. Nature. 1912; 90:219.

3. Bragg WH, Bragg WL. The reflection of X-rays by crystals. Proc Roy Soc Lond A. 1913; 88:428–38.

4. Bragg WL. The structure of some crystals as indicated by their diffraction of X-rays. Proc R Soc Lond A. 1913; 89:248–77.

5. Thomas JM. William Lawrence Bragg: the pioneer of X-ray crystallography and his pervasive influence. Angewandte Chemie (International ed in English). 2012; 51:12946–58. Epub 2012/12/06. [PubMed: 23212904]

6. Jaskolski M, Dauter Z, Wlodawer A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. The FEBS journal. 2014 Epub 2014/04/05.

7. Helliwell, JR. Macromolecular Crystallography with Synchrotron Radiation. Cambridge University Press; 2005.

8. Helliwell, JR. International Tables for Crystallography. IUCr; F. Chester, UK: 2012. Crystallography of Biological Macromolecules.; p. 189-204.

9. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature. 1958; 181:662–6. Epub 1958/03/08. [PubMed: 13517261]

10. Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DC, Shore VC. Structure of myoglobin: A three-dimensional Fourier synthesis at 2 A. resolution. Nature. 1960; 185:422–7. Epub 1960/02/13. [PubMed: 18990802]

11. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North AC. Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature. 1960; 185:416–22. Epub 1960/02/13. [PubMed: 18990801]

12. Chang AC, Cohen SN. Genome construction between bacterial species in vitro: replication and expression of Staphylococcus plasmid genes in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America. 1974; 71:1030–4. Epub 1974/04/01. [PubMed: 4598290]

13. Cohen SN. DNA cloning: a personal view after 40 years. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:15521–9. Epub 2013/09/18. [PubMed: 24043817]

14. Cohen SN, Chang AC, Boyer HW, Helling RB. Construction of biologically functional bacterial plasmids in vitro. Proceedings of the National Academy of Sciences of the United States of America. 1973; 70:3240–4. Epub 1973/11/01. [PubMed: 4594039]

15. Hunefeld, FL. Der Chemismus in der tierescher Organization. Leipzig; FA Brockhouse: 1840.

16. Kolhker Z. Wlssensch Zoologic. 1849; 1:266.

17. Leydlg Z. Wissensch Zoologic. 1849; 1:116.

18. Relchert KE. Mullers Arch Anat Physlol Wmsensch Medlcm. 1849:197–251.

19. Funke OZ. Über das milzvenenblut. Z Rat Med. 1851; 1:172–218.

20. Lehmann, CC. Lehrbuch der Physlologlschen Cherme. 2nd. Leipzig; p. 1853

21. Hartig T. Ueber das Klebermehl. Botanische Zeitung. 1855; 13:881.

22. Osborne TB. Crystallized vegetable proteids. Amer Chem J. 1892; 14:662–89.

23. Hofmeister T. Uber die Darstellung von krystallisirtem Eiralbumin und die Krystallisirbarkeit colloider Stoffe. Z Physiol Chem. 1890; 14:165.

24. Hopkins FG, Pincus SN. Observations on the crystallisation of animal proteins. J Physiol. 1898; 23:130–6.

25. Sumner JB. The isolation and crystallization of the enzyme urease: preliminary paper. J Biol Chem. 1926; 69:435.

26. Abel JJ. Crystalline Insulin. Proc Natl Acad Sci USA. 1926; 12:132–6. [PubMed: 16587069]

27. Abel JJ, Geiling EMK, Roultier OA, Bell FM, Wintersteiner O. Crystalline insulin. J Pharmacol Exp Ther. 1927; 31:65–85.

28. Northrop JH. CRYSTALLINE PEPSIN. Science (New York, NY). 1929; 69:580. Epub 1929/05/31.

29. Stanley WM. ISOLATION OF A CRYSTALLINE PROTEIN POSSESSING THE PROPERTIES OF TOBACCO-MOSAIC VIRUS. Science (New York, NY). 1935; 81:644–5. Epub 1935/06/28.

30. Sumner JB, Dounce AL. CRYSTALLINE CATALASE. Science (New York, NY). 1937; 85:366–7. Epub 1937/04/09.

31. Bernal JD, Crowfoot DC. X-ray photographs of crystalline pepsin. Nature. 1934; 133:794–5.

32. Perutz M. Early days of protein crystallography. Methods Enzymol. 1985 1985/01/01 ed.

33. Watson JD, Crick FH. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature. 1953; 171:737–8. [PubMed: 13054692]

34. Blake CC, Koenig DF, Mair GA, North AC, Phillips DC, Sarma VR. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature. 1965; 206:757–61. Epub 1965/05/22. [PubMed: 5891407]

35. Johnson LN, Phillips DC. Structure of some crystalline lysozyme-inhibitor complexes determined by X-ray analysis at 6 Angstrom resolution. Nature. 1965; 206:761–3. Epub 1965/05/22. [PubMed: 5840126]

36. Blow D. So do we understand how enzymes work? Structure (London, England : 1993). 2000; 8:R77–81. Epub 2000/05/10.

37. Poljak RJ, Amzel LM, Chen BL, Phizackerley RP, Saul F. The three-dimensional structure of the fab' fragment of a human myeloma immunoglobulin at 2.0-angstrom resolution. Proceedings of the National Academy of Sciences of the United States of America. 1974; 71:3440–4. Epub 1974/09/01. [PubMed: 4215080]

38. Poljak RJ, Amzel LM, Chen BL, Phizackerley RP, Saul F. Structure and specificity of antibody molecules. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 1975; 272:43–51. Epub 1975/11/06.

39. Epp O, Lattman EE, Schiffer M, Huber R, Palm W. The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0-A resolution. Biochemistry. 1975; 14:4943–52. Epub 1975/11/04. [PubMed: 1182131]

40. Sarma R, Zaloga G. Structure studies on styrene-treated immunoglobulin crystals. Journal of molecular biology. 1975; 98:479–84. Epub 1975/11/05. [PubMed: 1195395]

41. Jack A, Ladner JE, Klug A. Crystallographic refinement of yeast phenylalanine transfer RNA at 2-5A resolution. Journal of molecular biology. 1976; 108:619–49. Epub 1976/12/25. [PubMed: 798036]

42. Sussman JL, Kim S. Three-dimensional structure of a transfer rna in two crystal forms. Science (New York, NY). 1976; 192:853–8. Epub 1976/05/28.

43. Boutet S, Lomb L, Williams GJ, Barends TR, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, Kassemeyer S, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, van Bakel N, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang M, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang D, Zatsepin NA, Schafer D, Defever J, Neutze R, Fromme P, Spence JC, Chapman HN, Schlichting I. High-resolution protein structure determination by serial femtosecond crystallography. Science (New York, NY). 2012; 337:362–4. Epub 2012/06/02.

44. Hirata K, Shinzawa-Itoh K, Yano N, Takemura S, Kato K, Hatanaka M, Muramoto K, Kawahara T, Tsukihara T, Yamashita E, Tono K, Ueno G, Hikima T, Murakami H, Inubushi Y, Yabashi M, Ishikawa T, Yamamoto M, Ogura T, Sugimoto H, Shen JR, Yoshikawa S, Ago H. Determination of damage-free crystal structure of an X-ray-sensitive protein using an XFEL. Nature methods. 2014 Epub 2014/05/13.

45. Fisher SJ, Blakeley MP, Cianci M, McSweeney S, Helliwell JR. Protonation-state determination in proteins using high-resolution X-ray crystallography: effects of resolution and completeness. Acta crystallographica Section D, Biological crystallography. 2012; 68:800–9. Epub 2012/07/04.

46. Helliwell JR. Biochemistry. How to solve protein structures with an X-ray laser. Science (New York, NY). 2013; 339:146–7. Epub 2013/01/12.

47. Willstätter, R. Problems and Methods in Enzyme Research. Cornell University Press; Ithaca: 1927.

48. McPherson, A. Preparation and Analysis of Protein Crystals. Wiley and Sons; New York: 1982.

49. Bergfors, TM. Protein Crystallization: Second Edition. International University Line; La Jolla, California: 2009.

50. Jancarik J, Kim S-H. Sparse matrix sampling: a screening method for crystallization of proteins. Journal of Applied Crystallography. 1991; 24:409–11.

51. Giege R. A historical perspective on protein crystallization from 1840 to the present day. FEBS Journal. 2013; 280:6456–97. [PubMed: 24165393]

52. Miller WH. A Treatise on Crystallography. 1839

53. Bravais A. Memoire sur les systemes formes par les points distribues regulierement sur un plan ou dans l'espace. J Ecole Polytech. 1850; 19:1–128.

54. Burckhardt JJ. Zur Geschichte der Entdeckung der 230 Raumgruppen. Archive for History of Exact Sciences. 1967; 4:235–46.

55. Hahn, T. International Tables for Crystallography, Volume A: Space-Group Symmetry. 5th ed.. Springer-Verlag; Berlin: 2002.

56. Fourier J. Theorie analytique de la chaleur, par M. Fourier: Chez Firmin Didot, pere et fils. 1822

57. Sayre D. The squaring method: a new method for phase determination. Acta Cryst. 1952; 5:60–5.

58. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. Math Comput. 1965; 19:297–301.

59. The CCP4 suite: programs for protein crystallography. Acta crystallographica Section D, Biological crystallography. 1994; 50:760–3. Epub 1994/09/01.

60. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A. Overview of the CCP4 suite and current developments. Acta Crystallographica Section D: Biological Crystallography. 2011; 67:235–42.

61. Ten Eyck LF. Crystallographic fast Fourier transforms. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1973; 29:183–91.

62. Ten Eyck LF. Efficient structure-factor calculation for large molecules by the fast Fourier transform. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1977; 33:486–92.

63. Agarwal RC. A new least-squares refinement technique based on the fast Fourier transform algorithm. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1978; 34:791–809.

64. Green DW, Ingram VM, Perutz MF. The structure of hemoglobin. IV. Sign determination by the isomorphous replacement method. Proc R Soc London A. 1954; 225:287–307.

65. Blow DM, Crick FHC. The treatment of errors in the isomorphous replacement method. Acta Cryst. 1959; 12:794–802.

66. Rossmann MG. The position of anomalous scatterers in protein crystals. Acta Cryst. 1961; 14:383–8.

67. Matthews BW. The extension of the isomorphous replacement method to include anomalous scattering measurements. Acta Cryst. 1966; 20:82–6.

68. Otwinowski, Z. Isomorphous replacement and anomalous scattering. Publisher; 1991.

69. de La Fortelle E, Bricogne G. SHARP: A maximum-likelihood heavy-atom parameter refinement program for the MIR and MAD methods. Methods in enzymology. 1997; 276:472–94.

70. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallographica Section D: Biological Crystallography. 1999; 55:849–61.

71. Furey W, Swaminathan S. PHASES-95: a program package for processing and analyzing diffraction data from macromolecules. Methods in enzymology. 1997; 277:590–620. [PubMed: 18488326]

72. Hendrickson WA, Horton JR, LeMaster DM. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. The EMBO journal. 1990; 9:1665–72. Epub 1990/05/01. [PubMed: 2184035]

73. Hendrickson WA. Maturation of MAD phasing for the determination of macromolecular structures. Journal of Synchrotron Radiation. 1999; 6:845–51.

74. Ramakrishnan V, Biou V. Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. Methods in enzymology. 1997; 276:538–57. [PubMed: 9048381]

75. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. Acta Crystallographica Section D: Biological Crystallography. 2004; 60:432–8.

76. Hendrickson WA, Teeter MM. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. Nature. 1981; 290:107–13.

77. Wang BC. Resolution of phase ambiguity in macromolecular crystallography. Methods in enzymology. 1985; 115:90–112. Epub 1985/01/01. [PubMed: 4079800]

78. Dauter Z, Adamiak DA. Anomalous signal of phosphorus used for phasing DNA oligomer: importance of data redundancy. Acta Crystallographica Section D: Biological Crystallography. 2001; 57:990–5.

79. Ramagopal UA, Dauter M, Dauter Z. Phasing on anomalous signal of sulfurs: what is the limit? Acta Crystallographica Section D: Biological Crystallography. 2003; 59:1020–7.

80. Liu ZJ, Vysotski ES, Chen CJ, Rose JP, Lee J, Wang BC. Structure of the Ca2+ - regulated photoprotein obelin at 1.7 A resolution determined directly from its sulfur substructure. Protein Science. 2000; 9:2085–93. [PubMed: 11152120]

81. Ren H, Wang L, Bennett M, Liang Y, Zheng X, Lu F, Li L, Nan J, Luo M, Eriksson S. The crystal structure of human adenylate kinase 6: An adenylate kinase localized to the cell nucleus. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:303–8. [PubMed: 15630091]

82. Yang C, Pflugrath J, Courville D, Stence C, Ferrara JD. Away from the edge: SAD phasing from the sulfur anomalous signal measured in-house with chromium radiation. Acta Crystallographica Section D: Biological Crystallography. 2003; 59:1943–57.

83. Rossmann, MG. The molecular replacement method. Gordon & Breach; New York: 1972.

84. Crowther RA, Amos LA. Three-dimensional image reconstructions of some small spherical viruses. Cold Spring Harbor symposia on quantitative biology. 1972; 36:489–94. Epub 1972/01/01.

85. Crowther R. The fast rotation function. Int Sci Rev Ser. 1972; 13:173–8.

86. Navaza J. AMoRe: an automated package for molecular replacement. Acta Crystallographica Section A: Foundations of Crystallography. 1994; 50:157–63.

87. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. Journal of applied crystallography. 1997; 30:1022–5.

88. Kissinger CR, Gehlhaar DK, Fogel DB. Rapid automated molecular replacement by evolutionary search. Acta Crystallographica Section D: Biological Crystallography. 1999; 55:484–91.

89. Glykos NM, Kokkinidis M. Multidimensional molecular replacement. Acta Crystallographica Section D: Biological Crystallography. 2001; 57:1462–73.

90. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. Journal of applied crystallography. 2007; 40:658–74. [PubMed: 19461840]

91. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta crystallographica Section D, Biological crystallography. 2002; 58:1948–54. Epub 2002/10/24.

92. Das R, Baker D. Macromolecular modeling with rosetta. Annu Rev Biochem. 2008; 77:363–82. [PubMed: 18410248]

93. DiMaio F, Terwilliger TC, Read RJ, Wlodawer A, Oberdorfer G, Wagner U, Valkov E, Alon A, Fass D, Axelrod HL. Improved molecular replacement by density-and energy-guided protein structure optimization. Nature. 2011; 473:540–3. [PubMed: 21532589]

94. Mukherjee A, Helliwell J, Main P. The use of MULTAN to Locate the Positions of Anomalous Scatterers. Acta Crystallographica Section A: Foundations of Crystallography. 1989; 45:715–8.

95. Sheldrick GM, Gould RO. Structure solution by iterative peaklist optimization and tangent expansion in space group P1. Acta Crystallographica Section B: Structural Science. 1995; 51:423–31.

96. Miller R, Gallo SM, Khalak H, Weeks C. SnB: crystal structure determination via shake-and-bake. Journal of Applied Crystallography. 1994; 27:613–21.

97. Foadi J, Woolfson M, Dodson E, Wilson K, Jia-xing Y, Chao-de Z. A flexible and efficient procedure for the solution and phase refinement of protein structures. Acta Crystallographica Section D: Biological Crystallography. 2000; 56:1137–47.

98. de Graaff RA, Hilge M, van der Plas JL, Abrahams JP. Matrix methods for solving protein substructures of chlorine and sulfur from anomalous data. Acta Crystallographica Section D: Biological Crystallography. 2001; 57:1857–62.

99. Wang J, Chen J, Gu Y, Zheng C, Jiang F, Fan H, Terwilliger T, Hao Q. SAD phasing by combination of direct methods with the SOLVE/RESOLVE procedure. Acta Crystallographica Section D: Biological Crystallography. 2004; 60:1244–53.

100. Yao, Dq; Huang, S.; Wang, Jw; Gu, Yx; Zheng, Cd; Fan, Hf; Watanabe, N.; Tanaka, I. SAD phasing by OASIS-2004: case studies of dual-space fragment extension. Acta Crystallographica Section D: Biological Crystallography. 2006; 62:883–90.

101. He Y, Yao D-Q, Gu Y-X, Lin Z-J, Zheng C-D, Fan H-F. OASIS and molecular-replacement model completion. Acta Crystallographica Section D: Biological Crystallography. 2007; 63:793–9.

102. Bricogne G. Methods and programs for direct-space exploitation of geometric redundancies. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1976; 32:832–47.

103. Zhang K, Main P. The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. Acta Crystallographica Section A: Foundations of Crystallography. 1990; 46:377–81.

104. Cowtan K. DM: an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM newsletter on protein crystallography. 1994; 31:34–8.

105. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallographica Section D: Biological Crystallography. 2000; 56:965–72.

106. Abrahams J, Leslie A. Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallographica Section D: Biological Crystallography. 1996; 52:30–42.

107. Jones TA. A graphics model building and refinement system for macromolecules. Journal of Applied Crystallography. 1978; 11:268–72.

108. Jones TA, Zou J-Y, Cowan St, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallographica Section A: Foundations of Crystallography. 1991; 47:110–9.

109. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallographica Section D. 2004; 60:2126–32.

110. McRee DE. XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. Journal of structural biology. 1999; 125:156–65. [PubMed: 10222271]

111. Oldfield T. From maps to molecules in minutes. Acta Crystallographica Section A. 2000; 56:s27.

112. Gurusaran M, Shankar M, Nagarajan R, Helliwell JR, Sekar K. Do we see what we should see? Describing non-covalent interactions in protein structures including precision. IUCrJ. 2014; 1:74–81. Epub 2014/07/31.

113. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nature structural biology. 1999; 6:458–63. Epub 1999/05/20.

114. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallographica Section D: Biological Crystallography. 2006; 62:1002–11.

115. Isaacs N, Agarwal R. Experience with fast Fourier least squares in the refinement of the crystal structure of rhombohedral 2-zinc insulin at 1.5 A resolution. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1978; 34:782–91.

116. Konnert J. A restrained-parameter structure-factor least-squares refinement procedure for large asymmetric units. Acta Crystallographica Section A. 1976; 32:614–7.

117. Konnert JH, Hendrickson WA. A restrained-parameter thermal-factor refinement procedure. Acta Crystallographica Section A. 1980; 36:344–50.

118. Tronrud DE, Ten Eyck LF, Matthews BW. An efficient general-purpose least-squares refinement program for macromolecular structures. Acta Crystallographica Section A. 1987; 43:489–501.

119. Brunger AT, Kuriyan J, Karplus M. Crystallographic R factor refinement by molecular dynamics. Science (New York, NY). 1987; 235:458–60. Epub 1987/01/23.

120. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. Acta Crystallographica Section D. 1998; 54:905–21.

121. Brunger AT. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature. 1992; 355:472–5. Epub 1992/01/30. [PubMed: 18481394]

122. Sheldrick GM, Schneider TR. SHELXL: high-resolution refinement. Methods in enzymology. 1997; 277:319–43. Epub 1997/01/01. [PubMed: 18488315]

123. Cruickshank D. Remarks about protein structure precision. Acta Crystallographica Section D: Biological Crystallography. 1999; 55:583–601.

124. Blow D. Rearrangement of Cruickshank's formulae for the diffraction-component precision index. Acta Crystallographica Section D: Biological Crystallography. 2002; 58:792–7.

125. Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallographica Section D. 1997; 53:240–55.

126. Bricogne G, Irwin J. Maximum-Likelihood Refinement of Incomplete Models with BUSTER+ TNT. Macromolecular refinement Proceedings of the CCP4 Study Weekend at Chester College. 1996:85–92.

127. Zwart, PH.; Afonine, PV.; Grosse-Kunstleve, RW.; Hung, L-W.; Ioerger, TR.; McCoy, AJ.; McKee, E.; Moriarty, NW.; Read, RJ.; Sacchettini, JC. Structural Proteomics. Springer; 2008. Automated structure solution with the PHENIX suite.; p. 419-35.

128. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix. refine. Acta Crystallographica Section D: Biological Crystallography. 2012; 68:352–67.

129. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography. 1993; 26:283–91.

130. Hooft R, Vriend G, Sander C, Abola EE. Errors in protein structures. Nature. 1996; 381:272. [PubMed: 8692262]

131. Vaguine AA, Richelle J, Wodak SJ. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallographica Section D. 1999; 55:191–205.

132. Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic acids research. 2004; 32:W615–9. Epub 2004/06/25. [PubMed: 15215462]

133. Arndt UW. Personal X-ray reflections. Methods in enzymology. 2003; 368:21–42. Epub 2003/12/17. [PubMed: 14674267]

134. Arndt UW, Long JVP, Duncumb P. A Microfocus X-ray Tube Used with Focusing Collimators. Journal of Applied Crystallography. 1998; 31:936–44.

135. Holmes KC, Rosenbaum G. How X-ray Diffraction with Synchrotron Radiation Got Started. Journal of Synchrotron Radiation. 1998; 5:147–53. [PubMed: 15263469]

136. Phillips JC, Wlodawer A, Yevitz MM, Hodgson KO. Applications of synchrotron radiation to protein crystallography: preliminary results. Proceedings of the National Academy of Sciences of the United States of America. 1976; 73:128–32. Epub 1976/01/01. [PubMed: 1061106]

137. Buerger, MJ. The precession method in x-ray crystallography. Wiley; 1964.

138. Arndt UW, Phillips DC. The linear diffractometer. Acta Crystallographica. 1961; 14:807–18.

139. Arndt, UW.; Willis, BTM. Single-Crystal Diffractometry. Cambridge University Press; 1966.

140. Schleputz C, Herger R, Willmott P, Patterson B, Bunk O, Bronnimann C, Henrich B, Hulsen G, Eikenberry E. Improved data acquisition in grazing-incidence X-ray scattering experiments using a pixel detector. Acta Crystallographica Section A: Foundations of Crystallography. 2005; 61:418–25.

141. Trueb P, Sobott B, Schnyder R, Loeliger T, Schneebeli M, Kobas M, Rassool R, Peake D, Broennimann C. Improved count rate corrections for highest data quality with PILATUS detectors. Journal of synchrotron radiation. 2012; 19:347–51. [PubMed: 22514168]

142. Hasegawa K, Hirata K, Shimizu T, Shimizu N, Hikima T, Baba S, Kumasaka T, Yamamoto M. Development of a shutterless continuous rotation method using an X-ray CMOS detector for protein crystallography. Journal of applied crystallography. 2009; 42:1165–75. [PubMed: 22477775]

143. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode.. In: Charles, W.; Carter, editors. Methods in enzymology. Vol. 276. Academic Press; 1997. p. 307-26.

144. Leslie A. Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr. 1992; 26:27–33.

145. Kabsch W. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. Journal of Applied Crystallography. 1988; 21:916–24.

146. Pflugrath J. The finer things in X-ray diffraction data collection. Acta Crystallographica Section D. 1999; 55:1718–25.

147. Howard, A. Data processing in macromolecular crystallography.. In: Bourne, PE.; Watenpaugh, KD., editors. Proceedings from the Macromolecular Crystallographic Computing School; Oxford: Oxford University Press; 1996. 2000. Chapter in: Crystallographic Computing 7

148. Campbell JW. LAUEGEN, AN X-WINDOWS-BASED PROGRAM FOR THE PROCESSING OF LAUE X-RAY-DIFFRACTION DATA. Journal of Applied Crystallography. 1995; 28:228–36.

149. Hope H. Cryocrystallography of biological macromolecules: a generally applicable method. Acta Crystallographica Section B. 1988; 44:22–6.

150. Teng T-Y. Mounting of crystals for macromolecular crystallography in a free-standing thin film. Journal of Applied Crystallography. 1990; 23:387–91.

151. Jackson DA, Symons RH, Berg P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proceedings of the National Academy of Sciences. 1972; 69:2904–9.

152. Ducruix, A.; Giegé, R. Crystallization of Nucleic Acids and Proteins: A Practical Approach. IRL Press at Oxford University Press; 1992.

153. Davies DR, Segal D. Protein crystallization: Micro techniques involving vapor diffusion. Methods in enzymology. 1971; 22:266–9.

154. Landau EM, Rosenbusch JP. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. Proceedings of the National Academy of Sciences. 1996; 93:14532–5.

155. Chiu ML, Nollert P, Loewen MeC, Belrhali H, Pebay-Peyroula E, Rosenbusch JP, Landau EM. Crystallization in cubo: general applicability to membrane proteins. Acta Crystallographica Section D: Biological Crystallography. 2000; 56:781–4.

156. Chayen NE, Shaw Stewart P, Maeder D, Blow D. An automated system for micro-batch protein crystallization and screening. Journal of applied crystallography. 1990; 23:297–302.

157. Weselak M, Patch MG, Selby TL, Knebel G, Stevens RC. Robotics for automated crystal formation and analysis. Methods in enzymology. 2003; 368:45–76. [PubMed: 14674268]

158. Aricescu AR, Assenberg R, Bill RM, Busso D, Chang VT, Davis SJ, Dubrovsky A, Gustafsson L, Hedfalk K, Heinemann U, Jones IM, Ksiazek D, Lang C, Maskos K, Messerschmidt A, Macieira S, Peleg Y, Perrakis A, Poterszman A, Schneider G, Sixma TK, Sussman JL, Sutton G, Tarboureich N, Zeev-Ben-Mordehai T, Jones EY. Eukaryotic expression: developments for structural proteomics. Acta crystallographica Section D, Biological crystallography. 2006; 62:1114–24. Epub 2006/09/27.

159. Lee JK, Stroud RM. Unlocking the eukaryotic membrane protein structural proteome. Curr Opin Struct Biol. 2010; 20:464–70. Epub 2010/08/27. [PubMed: 20739007]

160. Nettleship JE, Assenberg R, Diprose JM, Rahman-Huq N, Owens RJ. Recent advances in the production of proteins in insect and mammalian cells for structural biology. J Struct Biol. 2010; 172:55–65. Epub 2010/02/16. [PubMed: 20153433]

161. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the Pichia pastoris expression system. Yeast. 2005; 22:249–70. Epub 2005/02/11. [PubMed: 15704221]

162. Boer E, Steinborn G, Kunze G, Gellissen G. Yeast expression platforms. Applied microbiology and biotechnology. 2007; 77:513–23. Epub 2007/10/10. [PubMed: 17924105]

163. Long SB, Campbell EB, Mackinnon R. Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. Science (New York, NY). 2005; 309:897–903. Epub 2005/07/09.

164. Long SB, Tao X, Campbell EB, MacKinnon R. Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. Nature. 2007; 450:376–82. Epub 2007/11/16. [PubMed: 18004376]

165. Vaughn JL, Goodwin RH, Tompkins GJ, McCawley P. The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). Vitro. 1977; 13:213–7. Epub 1977/04/01.

166. Smith GE, Summers MD, Fraser MJ. Production of human beta interferon in insect cells infected with a baculovirus expression vector. Mol Cell Biol. 1983; 3:2156–65. Epub 1983/12/01. [PubMed: 6318086]

167. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature. 2007; 450:383–7. Epub 2007/10/24. [PubMed: 17952055]

168. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science (New York, NY). 2007; 318:1258–65. Epub 2007/10/27.

169. Brillet K, Pereira CA, Wagner R. Expression of membrane proteins in Drosophila Melanogaster S2 cells: Production and analysis of a EGFP-fused G protein-coupled receptor as a model. Methods Mol Biol. 2010; 601:119–33. Epub 2010/01/26. [PubMed: 20099143]

170. Aricescu AR, Owens RJ. Expression of recombinant glycoproteins in mammalian cells: towards an integrative approach to structural biology. Current opinion in structural biology. 2013; 23:345–56. Epub 2013/04/30. [PubMed: 23623336]

171. Yang H, Rudge DG, Koos JD, Vaidialingam B, Yang HJ, Pavletich NP. mTOR kinase structure, mechanism and regulation. Nature. 2013; 497:217–23. Epub 2013/05/03. [PubMed: 23636326]

172. Yokoyama S. Protein expression systems for structural genomics and proteomics. Current opinion in chemical biology. 2003; 7:39–43. [PubMed: 12547425]

173. Kimura-Soyema, T.; Shirouzu, M.; Yokoyama, S. Cell-Free Protein Synthesis. Springer; 2014. Cell-free membrane protein expression.; p. 267-73.

174. Heinemann U. Structural genomics in Europe: slow start, strong finish? Nature Structural & Molecular Biology. 2000; 7:940–2.

175. Terwilliger TC. Structural genomics in North America. Nature Structural & Molecular Biology. 2000; 7:935–9.

176. Yokoyama S, Hirota H, Kigawa T, Yabuki T, Shirouzu M, Terada T, Ito Y, Matsuo Y, Kuroda Y, Nishimura Y. Structural genomics projects in Japan. Nature Structural & Molecular Biology. 2000; 7:943–5.

177. Burley SK, Almo SC, Bonanno JB, Capel M, Chance MR, Gaasterland T, Lin D, Sali A, Studier FW, Swaminathan S. Structural genomics: beyond the human genome project. Nature genetics. 1999; 23:151–7. [PubMed: 10508510]

178. Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. Science (New York, NY). 2003; 300:286–90.

179. Consortium GP. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–73. [PubMed: 20981092]

180. Siva N. 1000 Genomes project. Nature biotechnology. 2008; 26:256.

181. Stevens RC, Yokoyama S, Wilson IA. Global efforts in structural genomics. Science (New York, NY). 2001; 294:89–92.

182. Lesley SA, Kuhn P, Godzik A, Deacon AM, Mathews I, Kreusch A, Spraggon G, Klock HE, McMullan D, Shin T. Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline. Proceedings of the National Academy of Sciences. 2002; 99:11664–9.

183. Li L, Nan J, Li D, Brostromer E, Wang Z, Liu C, Hou Q, Fan X, Ye Z, Su XD. Structural genomics studies of human caries pathogen Streptococcus mutans. J Struct Funct Genomics. 2014 Epub 2014/01/30.

184. Chance MR, Fiser A, Sali A, Pieper U, Eswar N, Xu G, Fajardo JE, Radhakannan T, Marinkovic N. High-throughput computational and experimental techniques in structural genomics. Genome research. 2004; 14:2145–54. [PubMed: 15489337]

185. Joachimiak A. High-throughput crystallography for structural genomics. Current opinion in structural biology. 2009; 19:573–84. [PubMed: 19765976]

186. Jeon WB, Aceti DJ, Bingman CA, Vojtik FC, Olson AC, Ellefson JM, McCombs JE, Sreenath HK, Blommel PG, Seder KD. High-throughput purification and quality assurance of Arabidopsis thaliana proteins for eukaryotic structural genomics. Journal of structural and functional genomics. 2005; 6:143–7. [PubMed: 16211511]

187. Montelione GT, Zheng D, Huang YJ, Gunsalus KC, Szyperski T. Protein NMR spectroscopy in structural genomics. Nature Structural & Molecular Biology. 2000; 7:982–5.

188. Adams MW, Dailey HA, DeLucas LJ, Luo M, Prestegard JH, Rose JP, Wang B-C. The Southeast Collaboratory for Structural Genomics: a high-throughput gene to structure factory. Accounts of chemical research. 2003; 36:191–8. [PubMed: 12641476]

189. Goh C-S, Lan N, Douglas SM, Wu B, Echols N, Smith A, Milburn D, Montelione GT, Zhao H, Gerstein M. Mining the structural genomics pipeline: identification of protein properties that affect high-throughput experimental analysis. Journal of molecular biology. 2004; 336:115–30. [PubMed: 14741208]

190. Elsliger M-A, Deacon AM, Godzik A, Lesley SA, Wooley J, Wuthrich K, Wilson IA. The JCSG high-throughput structural biology pipeline. Acta Crystallographica Section F. 2010; 66:1137–42.

191. Hui R, Edwards A. High-throughput protein crystallization. Journal of structural biology. 2003; 142:154–61. [PubMed: 12718927]

192. Zheng B, Roach LS, Ismagilov RF. Screening of Protein Crystallization Conditions on a Microfluidic Chip Using Nanoliter-Size Droplets. Journal of the American Chemical Society. 2003; 125:11170–1. [PubMed: 16220918]

193. Santarsiero BD, Yegian DT, Lee CC, Spraggon G, Gu J, Scheibe D, Uber DC, Cornell EW, Nordmeyer RA, Kolbe WF, Jin J, Jones AL, Jaklevic JM, Schultz PG, Stevens RC. An approach to rapid protein crystallization using nanodroplets. Journal of Applied Crystallography. 2002; 35:278–81.

194. Berry IM, Dym O, Esnouf RM, Harlos K, Meged R, Perrakis A, Sussman JL, Walter TS, Wilson J, Messerschmidt A. SPINE high-throughput crystallization, crystal imaging and recognition techniques: current state, performance analysis, new technologies and future aspects. Acta crystallographica Section D, Biological crystallography. 2006; 62:1137–49. Epub 2006/09/27.

195. Cohen AE, Ellis PJ, Miller MD, Deacon AM, Phizackerley RP. An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot. Journal of applied crystallography. 2002; 35:720–6. [PubMed: 24899734]

196. Snell G, Cork C, Nordmeyer R, Cornell E, Meigs G, Yegian D, Jaklevic J, Jin J, Stevens RC, Earnest T. Automated sample mounting and alignment system for biological crystallography at a synchrotron source. Structure (London, England : 1993). 2004; 12:537–45.

197. Gonzalez A, Moorhead P, McPhillips S, Song J, Sharp K, Taylor J, Adams P, Sauter N, Soltis S. Web-Ice: integrated data collection and analysis for macromolecular crystallography. Journal of Applied Crystallography. 2008; 41:176–84.

198. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. Acta crystallographica Section D, Biological crystallography. 2006; 62:859–66. Epub 2006/07/21.

199. Liu G, Shen Y, Atreya HS, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TA, Arrowsmith CH, Montelione GT, Szyperski T. NMR data collection and analysis protocol for high-throughput protein structure determination. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:10487–92. Epub 2005/07/20. [PubMed: 16027363]

200. Gabanyi MJ, Adams PD, Arnold K, Bordoli L, Carter LG, Flippen-Andersen J, Gifford L, Haas J, Kouranov A, McLaughlin WA, Micallef DI, Minor W, Shah R, Schwede T, Tao YP, Westbrook JD, Zimmerman M, Berman HM. The Structural Biology Knowledgebase: a portal to protein structures, sequences, functions, and methods. J Struct Funct Genomics. 2011; 12:45–54. Epub 2011/04/08. [PubMed: 21472436]

201. Su XD, Liang Y, Li L, Nan J, Brostromer E, Liu P, Dong Y, Xian D. A large-scale, high-efficiency and low-cost platform for structural genomics studies. Acta crystallographica Section D, Biological crystallography. 2006; 62:843–51. Epub 2006/07/21.

202. Brostromer E, Nan J, Su XD. An automated image-collection system for crystallization experiments using SBS standard microplates. Acta crystallographica Section D, Biological crystallography. 2007; 63:119–25. Epub 2007/01/24.

203. Ajdiċ D, McShan WM, McLaughlin RE, Savic G, Chang J, Carson MB, Primeaux C, Tian R, Kenton S, Jia H. Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen. Proceedings of the National Academy of Sciences. 2002; 99:14434–9.

204. Neutze R, Moffat K. Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. Current Opinion in Structural Biology. 2012; 22:651–9. [PubMed: 23021004]

205. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FR, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang M, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Homke A, Reich C, Pietschner D, Struder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuhnel KU, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jonsson O, Svenda M, Stern S, Nass K, Andritschke R, Schroter CD, Krasniqi F, Bott M, Schmidt KE, Wang X, Grotjohann I, Holton JM, Barends TR, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JC. Femtosecond X-ray protein nanocrystallography. Nature. 2011; 470:73–7. Epub 2011/02/05. [PubMed: 21293373]

206. Seibert MM, Ekeberg T, Maia FR, Svenda M, Andreasson J, Jönsson O, Odic D, Iwan B, Rocker A, Westphal D. Single mimivirus particles intercepted and imaged with an X-ray laser. Nature. 2011; 470:78–81. [PubMed: 21293374]

207. Chapman HN, Barty A, Bogan MJ, Boutet S, Frank M, Hau-Riege SP, Marchesini S, Woods BW, Bajt S, Benner WH. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. Nature Physics. 2006; 2:839–43.

208. Redecke L, Nass K, DePonte DP, White TA, Rehders D, Barty A, Stellato F, Liang M, Barends TR, Boutet S. Natively inhibited Trypanosoma brucei Cathepsin B structure determined by using an x-ray laser. Science (New York, NY). 2013; 339:227–30.

209. Barends TR, Foucar L, Botha S, Doak RB, Shoeman RL, Nass K, Koglin JE, Williams GJ, Boutet S, Messerschmidt M. De novo protein crystal structure determination from X-ray free-electron laser data. Nature. 2014; 505:244–7. [PubMed: 24270807]

210. Qu K, Zhou L, Dong Y-H. An improved integration method in serial femtosecond crystallography. Acta Crystallographica Section D. 2014; 70:1202–11.
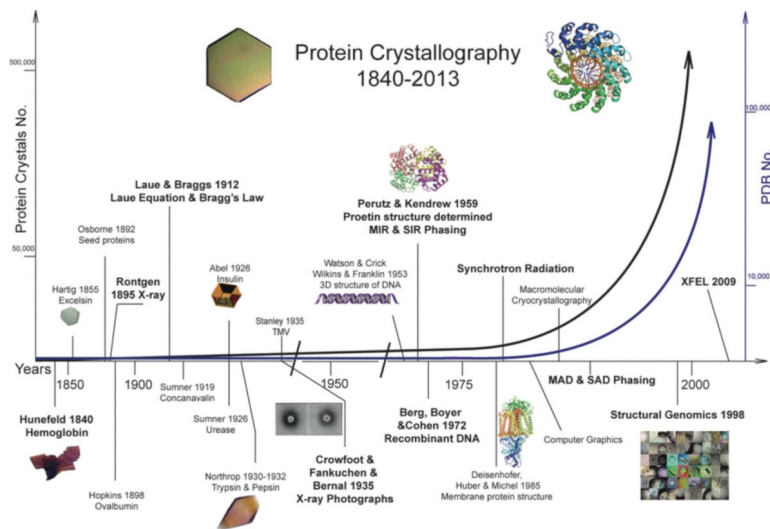
Figure 1.

A description of the development of protein crystallography with time, showing rapid exponential growth of protein crystallography since 1840s, the exponential growth of both proteins crystallized and protein crystal structures solved is well underway in the recent 20 years, block lettered events represent some of the most important milestones for protein crystallography.
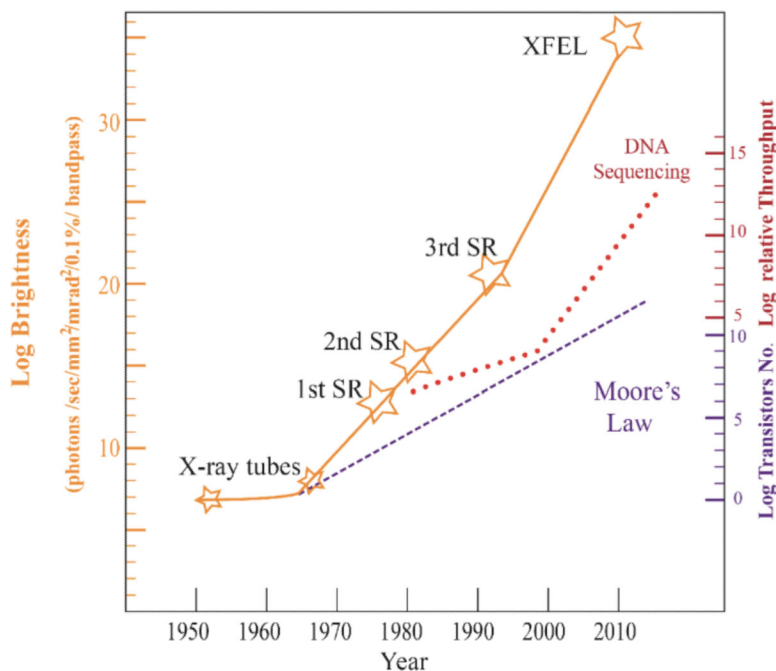
Figure 2.
The three representative technology fields with most rapidly increasing rates. All three have directly contributed to structural biology. In fact, among the three, electronics governed by the famous Moore's law has had the slowest but most stable exponential growth since 1970s. It has increased more than 10 orders of magnitude over the last 40 plus years and is responsible for the explosive growth of computers, IT and telecommunications industry. The red line represents the recent rapid growth of the so-call next generation sequencing (NGS) which has increased about 10 orders of magnitude just over the last 10 years. DNA (gene) sequences direct protein synthesis that is indispensible in recombinant DNA technology and therefore in solving protein structures. NGS makes it possible to get an overall estimation of how many proteins are in biological systems and that we might need to crystallize and visualize crystal structures. X-ray source intensities have increased more than 20 orders of magnitude since 1980s from the dedicated X-ray production source of 2nd generation SR to current XFEL sources.
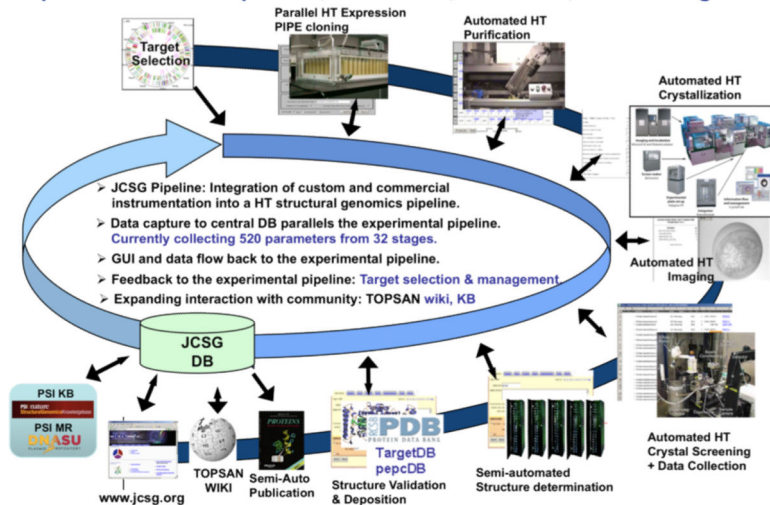
Figure 3.
Structure determination pipeline from the Joint Center for Structural Genomics (JCSG), USA. (Figure courtesy of I. A. Wilson).

Figure 4.
Over the last several years, the Su laboratory at Peking University has built up technological platforms of high-throughput (HTP) methods for structural biology studies, including target selection; HTP and semi-robotic gene cloning; protein expression; protein purification; crystallization and crystal structure determinations.

Table 1

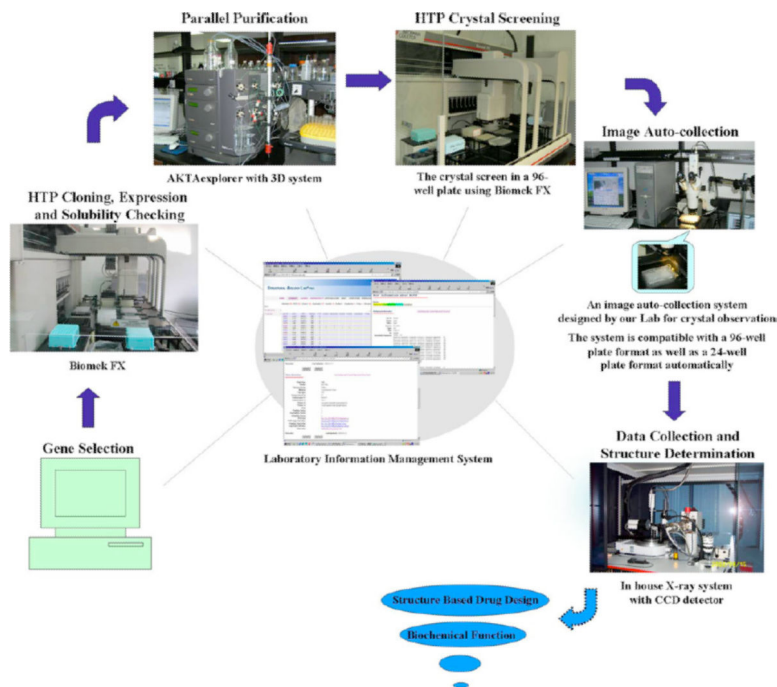Nobel prizes as milestones of discoveries and technique developments relevant to protein crystallography (including X-ray, electron and neutron crystallography), and to protein structure-function relationships. Bold letters indicate prizes for the most crucial technique developments in protein crystallography. (from: http://www.nobelprize.org/)

Year	Sub.	Nobel Laureates	Reasons for the prizes
1901	**P**	**Germany: Wilhelm Conrad Röntgen**	**Discovery of the remarkable rays (X-ray)**
1914	**P**	**Germany: Max von Laue**	**Discovery of the diffraction of X-rays by crystals**
1915	**P**	**UK: Sir William Henry Bragg** **UK: William Lawrence Bragg**	**Analysis of crystal structure by means of X-rays**
1917	P	UK: Charles Glover Barkla	Discovery of the characteristic Röntgen (X-ray) radiation of the elements
1924	P	Sweden: Karl Manne Georg Siegbahn	Discoveries and research in the field of X-ray spectroscopy
1926	C	Sweden: Theodor Svedberg	Development of the analytical ultracentrifuge
1937	P	USA: Clinton Joseph Davisson UK: George Paget Thomson	Discovery of the diffraction of electrons by crystals
1939[*]	**P**	**USA: Ernest Lawrence**	**Invention and development of the cyclotron**
1946[**]	**C**	**USA: James B. Sumner** **USA: John Howard Northrop** **USA: Wendell Meredith Stanley**	**Discovery that enzymes are proteins that can be purified and crystallized and so can viruses**
1954	**C**	**USA: Linus Carl Pauling**	**Research into the nature of the chemical bond**
1958	**C**	**UK: Frederick Sanger**	**Protein sequencing**
1962[**]	**C**	**UK: John Charles Kendrew** **UK: Max Ferdinand Perutz**	**Protein structure determination**
1962	**P&M**	**UK: Francis Harry Compton Crick** **USA: James Watson** **UK: Maurice Hugh Frederick Wilkins**	**Discovery of DNA double helix**
1964	C	UK: Dorothy Crowfoot Hodgkin	Determination of the structures of penicillin and vitamin B_{12}
1972	C	USA: Christian Borhmer Anfinsen USA: Stanford Moore USA: William Howard Stein	Principles that govern the folding of protein; Principles related to the biological activity of the enzyme
1976	C	USA: William Nunn Lipscomb Jr.	Studies on the structure of boranes
1980	**C**	**UK: Frederick Sanger** **USA: Paul Berg** **USA: Walter Gilbert**	**Fundamental studies of the biochemistry of nucleic acids; Determination of base sequences in nucleic acids**
1982	**C**	**UK: Aaron Klug**	**Development of crystallographic electron microscopy**
1985	C	USA: Herbert Aaron Hauptman USA: Jerome Karle	Direct methods for the determination of crystal structures
1986	**P**	**Germany: Ernst Ruska** **Germany: Gerd Binnig** **Switzerland: Heinrich Rohrer**	**Design of the first electron microscope; design of the scanning tunneling microscope**
1988	**C**	**Germany: Johann Deisenhofer** **Germany: Robert Huber** **Germany: Hartmut Michel**	**Determination of the three-dimensional structure of a photosynthetic reaction centre**
1992	**P**	**Switzerland: Georges Charpak**	**Invention and development of the multiwire proportional chamber detectors.**
1993[***]	**C**	**Canada: Michael Smith** **USA: Kary Banks Mullis**	**Site-directed mutagenesis: polymerase chain reaction (PCR) method**

Year	Sub.	Nobel Laureates	Reasons for the prizes
1994	**P**	**Canada: Bertram Neville Brockhouse** **USA: Clifford Glenwood Shull**	**Development of neutron spectroscopy and the neutron diffraction technique**
1997	C	USA: Paul Delos Boyer UK: John Ernest Walker Denmark: Jens Christian Skou	Enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP)
2003	C	USA: Peter Agre USA: Roderick MacKinnon	Structural and mechanistic studies of water and ion channels
2006	C	USA: Roger David Kornberg	Studies of the molecular basis of eukaryotic transcription
2009	**P**	**USA: Willard S. Boyle; George E. Smith**	**Invention of an imaging semiconductor circuit - the CCD sensor**
2009	C	USA: Venkatraman Ramakrishnan Israel: Ada Yonath, USA: Thomas Arthur Steitz	Studies of the structure and function of the ribosome
2011	C	Israel Dan Shechtman	Discovery of quasicrystals
2012	C	USA: Robert J. Lefkowitz; Brian K. Kobilka	Studies of G-protein-coupled receptors
2013	**C**	**USA Martin Karplus** **USA Michael Levitt** **USA Ariel Warshel**	**Development of multi-scale models for complex chemical systems**

Sub. Indicates the category of the prize (P, Physics; C, Chemistry; P&M, Physiology and Medicine)

[*] **Synchrotron radiation (SR)** is one of the most important techniques responsible for the rapid development of protein crystallography; the modern SR has been built on synchrotron instrument (similar to the type invented by Ernest Lawrence, although with different purposes). For recent reviews on SR instrumentation, see [7, 8].

[**] The most important prize was to recognize Max Perutz and John Kendrew's original and fundamental work to establish the protein crystallographic method. The heroic work of more than 25 years consistent and persistent efforts of Max Perutz led to the final determination of hemoglobin and other protein crystal structures [9, 10, 11].

[***] Molecular biology techniques (DNA sequencing and manipulation) such as PCR (polymerase chain reaction) and mutagenesis were crucial developments for biology including structural biology. Another key technique responsible for rapid development of protein crystallography is recombinant DNA technology and expression of extrinsic proteins in bacterial hosts as invented by Stanley Cohen and Herbert Boyer [12, 13, 14].