

The eyes of chemistry

(Dedicated to Geoffrey Roughton)

Isabel Usón^{1,2} & Claudia Millán,²

¹ICREA, ²IBMB, CSIC Barcelona (Spain).



For most humans, *vision* constitutes the main way to apprehend the world. We understand information best when we can see it in a three-dimensional frame. This explains the impact of X-ray crystallography from its onset, a century ago, as it allows seeing in atomic detail the main players in chemical and biological processes. However, deriving a three-dimensional structure from the diffraction experiment requires overcoming *the phase problem*. The evolution of crystallographic methods along the quest to retrieve the phases, which are lost in the diffraction experiment, has been linked to milestones in chemistry and biology. The field of crystallography has arguably been marked by the character of its pioneers, which possibly determined the percentage

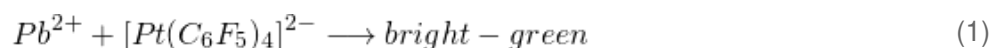
of women scientists, the collaborative spirit or the support of education and science in emergent economies.

Our group has developed structural methods to exploit the stereochemical knowledge present in small, yet very accurate, fragments. Their use to solve the central problem of crystallography, the phase problem, is implemented in our software ARCIMBOLDO. We are extending the use of fragments to map interpretation, other diffraction methods currently undergoing exciting development and structural bioinformatics. As illustrated in the painting by GIUSEPPE ARCIMBOLDO, the information content derived from a correct combination of fragments goes beyond their simple addition.

This article reflects the “Eilerslie talk” held jointly (alternated contributions) by CLAUDIA MILLÁN and ISABEL USÓN at 9 Adam Road, Cambridge, on the 28th July 2017.

An image is worth a thousand spectra

In 1991, towards the end of my Ph.D. after the usual four years of dedicated work, I was starting to panic that what I could write in the thesis would be more fitting for the XIX than for the XX century. I had accumulated a lot of information on the synthesis and properties of organometallic platinum clusters but paradoxically, I did not know what they were or how to explain their properties. For example, I knew that from the reaction:



I could isolate this luminescent cluster, which I had characterised by its spectra and determining its elemental composition. In solution, it was elusive, its instability precluding the use of solvents needed for NMR. I also had probed its reactivity but... I had no clue of what was happening at the molecular level. I knew all along that a *crystal structure* was needed. Indeed, solving the problems to obtain stable crystals, to measure the diffraction intensities and to determine the atomic structure displayed in **Fig. 1**, allowed placing all previous information in a visual framework and settled conclusively our many questions [1]. The structure confirmed the intended direct bonds between both types of metal centres and the interactions to the fluorine substituents in the ligands. I was delighted, and I thought about how

many others before me must have shared this same joy of literally seeing the answers so long pursued. But the structure also raised new questions: why was the coordination around lead, in oxidation state II, linear? Should there not be a stereochemically active lone pair of electrons? Why then was the structure not bent around the lead centre?

The Braggs shaped the character of crystallography

About one hundred years earlier, in 1895, X-rays had been discovered by RÖNTGEN [2]. Their ability to penetrate matter and yield image information (Fig. 2, left) had been patent from the onset. Diffraction by crystals, by MAX VON LAUE, WALTHER FRIEDRICH and PAUL KNIPPING, served to simultaneously establish the nature of X-rays and crystals [3]. From these findings, WILLIAM BRAGG and his son, sir LAWRENCE BRAGG realized that in the case of molecules, the natural grating provided by crystals would be required to amplify the scattering signal, by bringing a large number of equivalent atoms to add their contributions [4]. We cannot see X-rays and if we want to experiment with

light scattering on macroscopic objects and how it reflects periodicity and the underlying blocks making up the periodic object, we should choose the appropriate wavelength: monochromatic visible light. Among much excellent material available on-line to visualize diffraction on gratings, we would suggest a YouTube video [5] (Video 1) showing green light scattered on *periodic*, everyday objects, such as spans of thread, strings of beads, screws, spiral springs mimicking the double helix structure adopted by DNA or sieves illustrating lattices! As can be appreciated in the video, we are not seeing an image of the illuminated items, but the pattern is obviously related to the periodicity and to the underlying structure of the object.

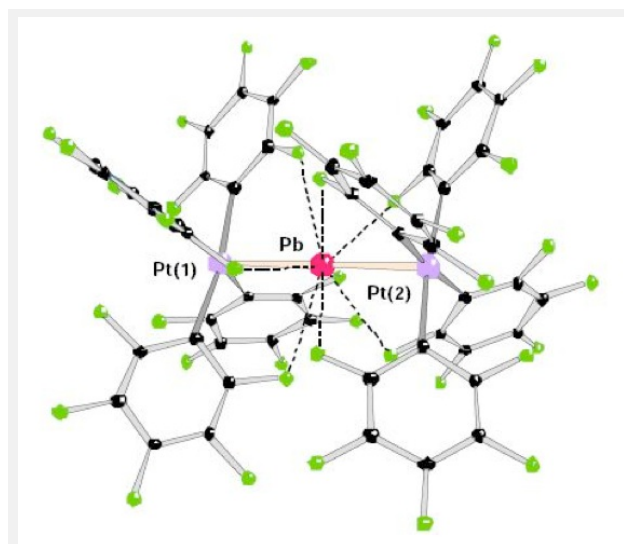


Figure 1. The three-dimensional view of a molecule in atomic detail. It provides a visual framework to all functional knowledge associated, as in the case of the anion $[\text{Pt}(\text{C}_6\text{F}_5)_4]^{2-}$ [1].

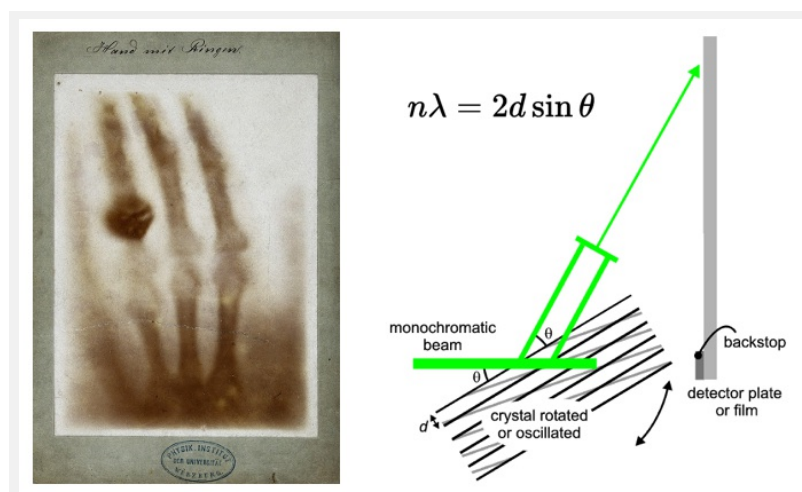


Figure 2. Röntgen's discovery and geometric illustration of

Bragg's Law. Left, image recorded by Conrad Röntgen on 22^d December 1895 showing the hand of his wife, Anna Bertha Ludwig.

Right, constructive interference geometry of X-rays incident to a crystal sample.

The fundamental law of diffraction is named *Bragg's Law*, and establishes the geometry of the diffraction pattern (Fig. 2, right). We can think of diffraction as reflection on sets of planes running through the crystal. Only at certain angles are the waves diffracted from different planes shifted by a whole number of wavelengths apart, i.e. in phase. For such angles, the intensity of the diffracted beams can be recorded on a detector. At other angles, the waves reflected from different planes are out of phase and cancel out.

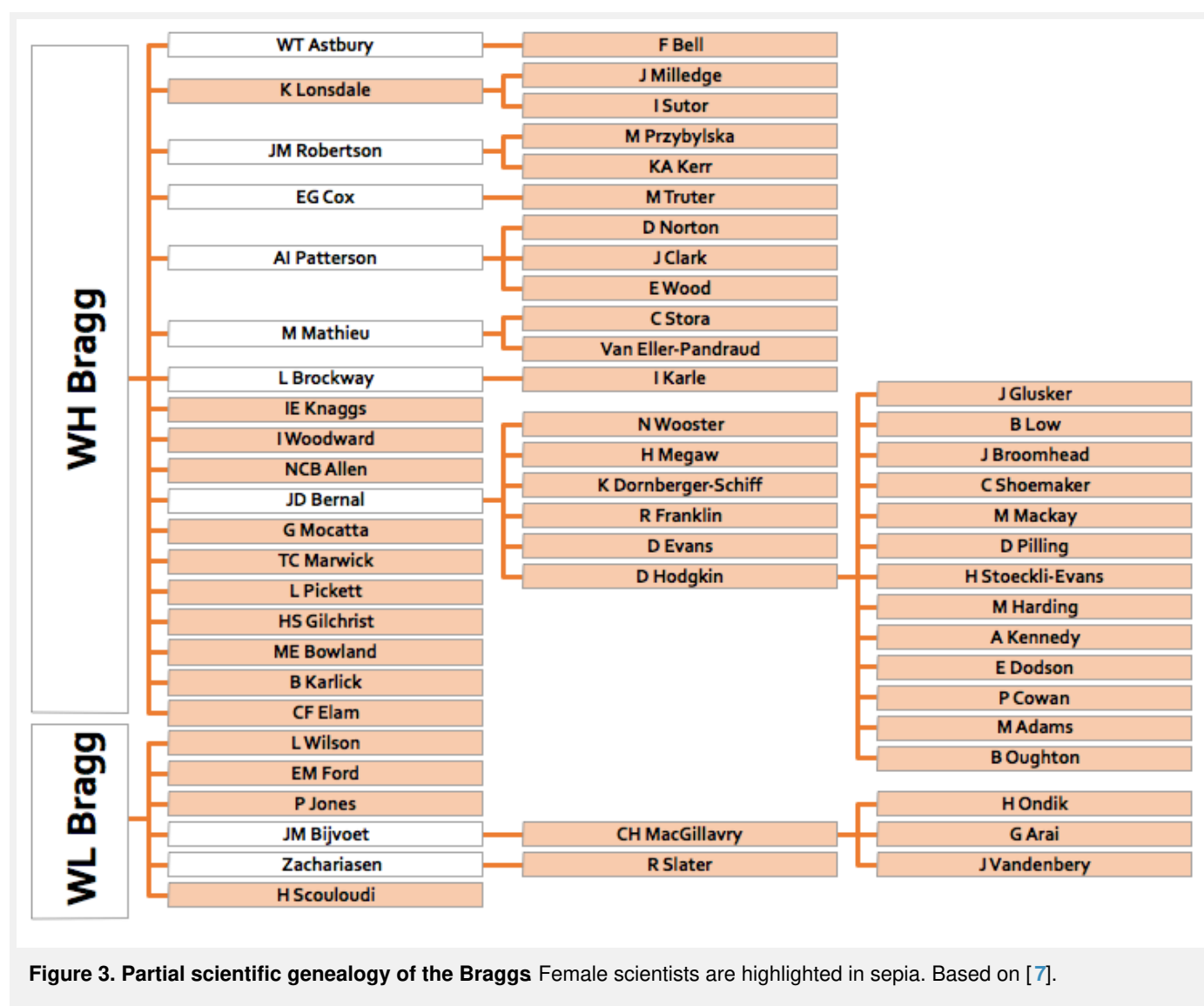
LAWRENCE BRAGG, was the first to perform structural analysis by X-ray crystallography, determining the structures of various inorganic salts, such as NaCl [6]. The BRAGGS analysed the diffraction pattern and figured out how it related to the structure in real space, placing the atoms that composed structures. Their

work established the grounds for X-ray crystallography, as we know it today and was awarded a Nobel Prize in 1915,

following MAX VON LAUE's prize from 1914.

Mentoring crystallographers regardless of their gender

The BRAGGS were not only pioneers in scientific terms. They also shaped the field by mentoring and influencing a new generation of crystallographers. In particular, support of women scientists can be appreciated in their scientific descent and in the opportunities they promoted women to, thus treating them as equals for the first time [7]. Prominent examples comprise JOAN EVANS, who was invited in 1923 by WILLIAM BRAGG to be the first woman to ever deliver a Discourse at the Royal Institution; LUCY WILSON, who was LAWRENCE BRAGG's first research student in 1923-1924; KATHLEEN LONSDALE, one of the first two female Fellows of the Royal Society elected in 1945, had been a student in the group of WILLIAM BRAGG; HELEN MEGAW, the first female staff member in Cambridge's legendary Cavendish Laboratory, was appointed in 1946 by LAWRENCE BRAGG.



In the genealogy reproduced on Fig. 3, displaying part of the scientific descent of the BRAGGS it is noticeable the unusually high numbers of distinguished women scientists. In fact, possibly due to the BRAGGS' direct intervention, crystallography has been differentiated from other scientific fields by a higher participation of women, providing role models for young scientists and also for the collaborative, open, joy-in-discovery attitude in many of its distinguished members. However, the actual percentage of women in crystallography was much lower than the one commonly perceived. While the popular feeling is that nearly half of the researchers in this field were women, in truth they amounted to less than 15% of entries in the World directory of crystallographers by the year 1981 (updated figures are available but crystallography is currently practised by many scientists who do not define themselves primarily as

crystallographers).

This biased perception must have been brought about by having such prominent figures in the field as DOROTHY HODGKIN, KATHLEEN LONSDALE, ROSALIND FRANKLIN, ISABELLA KARLE and still active in the present, ADA YONATH or ELEANOR DODSON. Even in Spain, where the incorporation of women to academia cannot be compared with the UK, SAGRARIO MARTÍNEZ-CARRERA pioneered the development of crystallographic computing from the 50s on [8]. Nobel laureate DOROTHY HODGKIN, who determined such key structures as penicillin and vitamin B12, was determinedly involved in the cause of scientists from, at the time, developing countries, in particular India and China. ELEANOR DODSON, nucleated the “collaborative computing project number 4” (CCP4), which has provided a unique model of cooperation in crystallography and done much to actively support science and mentoring worldwide.

Regarding relatives, the wife of LAWRENCE BRAGG was ALICE HOPKINSON. She and her cousin happened to share the same maiden name and both families used to live in the same street in Cambridge. The other ALICE HOPKINSON married FRANCIS JOHN WORSLEY ROUGHTON. ALICE ROUGHTON was the first woman to obtain a Ph.D. in psychiatry in Cambridge. Not only did she study: she practised taking her commitment to an extreme. She was literally living with her patients, including psychotic cases or prospective convict teenagers, which she brought into her home. After World War II, she heard of the German physicists imprisoned in Farm Hall, near Cambridge. She contrived to reach an agreement whereby HAHN, HEISENBERG and their colleagues would be fetched and brought to Cambridge every evening to participate in the academic life and dinner provided in high tables. Later on, she continued throwing her house open to refugees from every conflict, scholars, artists and stranded people from every nationality. ALICE ROUGHTON was a psychiatrist, a medical campaigner, a pioneer in the movement against nuclear weapons and a conservationist and her inspiring biography has been recently published [9]. But the story of structure and chemistry brings us back to JACK ROUGHTON, professor of Colloidal Science at Trinity College.

The physiology of respiration

The role of atomic structure in providing a visual framework to all our indirect information and prompting new questions extends to many chemical contexts: for instance, something as immediate as the physiology of respiration, which we all exercise 13 times in a minute. Much was known of this process, thanks to the works of the physiologist JACK ROUGHTON in Cambridge, before an atomic model could be envisaged. As a result of his efforts (which required constructing the apparatus to measure the reactions of haemoglobin with gases) the oxygen equilibrium curve, the thermodynamics and some of the kinetics of the reaction with oxygen seemed to be well understood. Also established were the binding of oxygen at the iron centre, whereas carbon dioxide was binding a different site, in the peptide chain, or the cooperativity among the four different haemoglobin subunits in the active species [10]. The Bohr effect, first discovered by physiologist CHRISTIAN BOHR in 1904, explained how hydrogen ions and carbon dioxide affect the affinity of oxygen in haemoglobin. If the pH was below the normal physiological pH of 7.4, haemoglobin would not bind oxygen as well. But some vital gaps and anomalies remained by the time the structure was determined by MAX PERUTZ [11].

Again, relating the previous knowledge to a three-dimensional frame provided conclusive information about the mechanism and the large structural changes involved, but opened new questions. For example, it was totally unexpected that the active sites, which were acting in cooperation, would be separated by large distances, rather than in immediate proximity. With this new structural insight, JACK ROUGHTON continued researching the biochemistry of respiration from a new perspective [12].

The phase problem

As mentioned above, crystallography does not render a direct image, and no lenses for X-rays are available to reconstruct it. In order to calculate an electron density map, both the intensities (proportional to the squares of the *amplitudes* F) and the *phases* ϕ of the scattered beams, hkl , would be required, but only the former are recorded in the diffraction experiment:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl}^{+\infty} |F(hkl)| \cdot e^{-2\pi i[hx+ky+lz-\phi(hkl)]}$$

Amplitudes
Phases?

The problem of phasing is highly non-linear and has a very poor radius of convergence, thus search, rather than minimisation methods are needed. Chemical structures can usually be solved *ab initio*, not requiring previous structural knowledge of the unknown structure or collection of additional experimental data. Directly solving the phase problem for small chemical structures, with up to 200 atoms is possible due to the excellent diffraction properties of typically well ordered crystals, allowing the measurement of many more independent reflections. This highly overdetermined problem, requiring comparatively few parameters, can be solved by computational brute force direct methods. The assumption of atomicity imposes statistical restraints on the phases. For the development of their equations to derive phases from the intensities, HERBERT HAUPTMAN and JEROME KARLE received the Nobel Prize in 1985 [13]. It has been often rued, that ISABELLA KARLE, who actually got the method to work and solved the first structures this way [14], did not get to share in the award.

Alternatively, the presence of a few heavy atoms can be used to solve the structure of chemical molecules by the Patterson function (the Fourier transform with F^2 as coefficients), which can be calculated directly from the experimental data and gives information about interatomic distances, relating atoms with significantly more electrons than the rest of the structure [15].

In the case of macromolecules, the larger number of parameters and lower proportion of independent measurements accessible makes *ab initio* methods unsuccessful. Crystals tend to be less perfect, as half of the volume is filled by disordered water, but this in turn opens the door to new phasing methods [16]. Experimental phasing is based on inducing or differentiating a small molecule within the macromolecule. For instance, by introducing heavy atoms like Hg. Because of the many electrons at these atoms, the X-ray intensities from derivatives differ sensibly from their native counterpart and these differences can be exploited to determine the structure of the heavy atoms, which in turn can be used to provide reference phases for the whole macromolecule. The first protein structures, those of myoglobin and haemoglobin, were so determined. Alternatively, a related structure can provide starting phases for the unknown one through Molecular Replacement: this requires placing the related structure in the unit cell of the target one, to best match the data.

Antibiotics constitute a class of molecules typically falling in between small molecules and macromolecules. They tend to give crystals with more than 200 independent atoms but so compactly packed that no space is left for disordered solvent. Too large for classic direct methods, these crystals could not be modified diffusing solutions with chemicals and closely enough structures would not be available. The structure of many antibiotics was only achieved in the 90's, after 50 years had passed from the determination of the structure of penicillin. These and other "large small molecules" needed recycling between the real space and reciprocal space formulation of the problem, constraining atomicity in both. For instance, vancomycin, a last resort antibiotic administered in hospitals against particularly resistant strains to common antibiotics, had not seen its structure determined until 1996, even though crystals diffracting to atomic resolution had been available for over two decades. Such direct methods-based dual-space recycling methods finally succeeded in obtaining the vancomycin structure [17].

Fragments in phasing, map and structure interpretation

It is not surprising that in the absence of atomic resolution, enforcing atomicity is of limited use. Therefore, dual space recycling methods succeeded in extending the scope of direct methods to larger structures but remained tied to the requirement of exceptionally good data and heavily overdetermined problems, i.e. data to atomic resolution. Instead, they became essential in the solution of substructures of heavy atoms and anomalous scatterers required in experimental phasing [18]. For such substructures, resolution is still "atomic" in the sense that their components are separated by longer distances and thus, resolved.

At the typical resolutions reached in macromolecular crystallography, rather than exploiting atomicity as a constraint, it was necessary to resort to the fact that macromolecular structures contain fragments of known geometry. Therefore, our group developed methods to exploit the stereochemical knowledge present in small, yet very accurate, structural units such as secondary structure fragments and their association into local folds [19]. Their use to solve the phase problem is implemented in our software *ARCIMBOLDO*, named in analogy to the portraits this artist painted in the XVIth century out of common objects such as fruits and vegetables. We assemble structural hypotheses out of common fragments of secondary structure or small local folds, such as alpha helices or small beta sheets. This is achieved with the molecular replacement methods implemented in *PHASER* based on Bayesian statistics [20], which provide a sensitive guide to decision making. If one of such substructures, comprising some 6% of the total structure, is correct and accurate to 0.5 Å rmsd to the true structure, density modification and automatic map interpretation with *SHELXE* [21] reveals the “portrait” of our protein. As most trials remain a still life, massive, parallel computing is needed in difficult cases. A video reproducing the process can be seen in our YouTube channel [22] ([Video 2](#)). As illustrated in the painting by GIUSEPPE ARCIMBOLDO, the information content derived from a correct combination of fragments goes beyond their simple addition. This has required developing our own particular toolbox for the very detailed view required in phasing [23], which can be extended to the solution of other problems. In particular, we are also extending this view to map interpretation in autotracing and general structure interpretation. The recent resolution revolution experimented in electron diffraction methods [24], has brought cryo-electron microscopy into the world of high resolution, and thus quantitative structure determination and many computing methods are being developed starting from crystallographic ones.

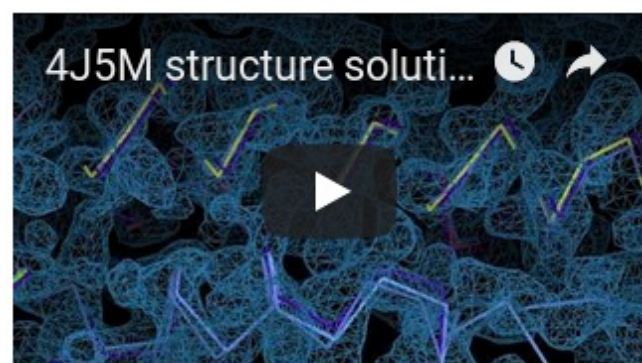
Acknowledgments

AIRLIE McCoy is gratefully acknowledged for her lecture on the BRAGG symposium, at the ECM 28 in Warwick, UK. We thank NICOLAS SOLER for expert feedback.

Movies



Video 1. Diffraction gratings. Experiments with diffraction gratings made from everyday objects. A telescope magnifies small bending angles of light so the diffraction patterns can be easily observed on a screen. Standard YouTube License 2014 [5].



Video 2. 4J5M structure solution using ARCIMBOLDO_SHREDDER spheres. In this video, we describe schematically the process of structure solution with ARCIMBOLDO_SHREDDER using the x-ray diffraction data from the cargo binding domain from human myosin Vb. Standard YouTube License 2017 [22].

References

1. USÓN R, FORNIÉS J, FALVELLO LR, USÓN MA, USÓN I. "Synthesis and molecular structure of bis(tetrabutylammonium)bis[tetrakis(perfluorophenyl)palladium]plumbate(2-), the first lead(II) compound linearly bonded to two metal atoms." *Inorg Chem*, **1992**, 31: 3697. DOI.
2. RÖNTGEN WC. "Ueber eine neue Art von Strahlen. Vorläufige Mittheilung". In: *Sonderabdruck aus den Sitzungsberichten der Würzburger Physik.-medic. Gesellschaft*. Stahel, Würzburg, **1895**. URL.
3. FRIEDRICH W, KNIPPING P, VON LAUE M. "Interferenz-Erscheinungen bei Röntgenstrahlen." *Sitz. ber. Bayer. Akad. Wiss.*, **1912**, 1912,14: 303. URL.
4. BRAGG WH, BRAGG WL. "The Reflection of X-rays by Crystals." *Proc R Soc A*, **1913**, 88: 428. DOI.
5. "Diffraction gratings." *YouTube* (channel: SuperLaser123), 00:03:19 (2014). ID: jzmqeRp_tmK.
6. BRAGG WL. "The structure of some crystals as indicated by their diffraction of X-rays." *Proc R Soc A*, **1913**, 89: 248. DOI.
7. JULIAN MM. "Women in crystallography". In: *Women of science : righting the record*, KASS-SIMON G, FARNES P, NASH D (Ed). Bloomington : Indiana University Press, **1990**, p. 335.
8. MARTÍNEZ-CARRERA S. "Commission on crystallographic apparatus." *Acta Cryst B*, **1987**, 43: 408. DOI.
9. PUIGGRÓS XM, Alice's. *La biografía de la doctora humanista Alice Roughton en el Cambridge del siglo XX. L'art de la memoria*, **2018**.
10. FERGUSON JKW, ROUGHTON FJW. "The chemical relationships and physiological importance of carbamino compounds of CO₂ with haemoglobin." *J Physiol*, **1934**, 83: 87. DOI.
11. PERUTZ MF, ROSSMANN MG, CULLIS AF, MUIRHEAD H, WILL G, NORTH ACT. "Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis." *Nature*, **1960**, 185: 416. DOI.
12. ROUGHTON FJW. "Some recent work on the interactions of oxygen, carbon dioxide and haemoglobin." *Biochem J*, **1970**, 117: 801. DOI.
13. KARLE J, HAUPTMAN H. "A theory of phase determination for the four types of non-centrosymmetric space groups 1P222, 2P22, 3P12, 3P22." *Acta Cryst*, **1956**, 9: 635. DOI.
14. KARLE IL. "Molecular formula, configuration and conformation by X-ray analysis." *Pure Appl Chem*, **1977**, 49: 1291. DOI.
15. PATTERSON AL. "A Fourier Series Method for the Determination of the Components of Interatomic Distances in Crystals." *Phys Rev*, **1934**, 46: 372. DOI.
16. HENDRICKSON WA. "Evolution of diffraction methods for solving crystal structures." *Acta Cryst A*, **2012**, 69: 51. DOI.
17. SHELDRIK GM, HAUPTMAN HA, WEEKS CM, MILLER R, USÓN I. "Ab initio phasing". In: *International Tables for Crystallography*, Vol. F: Crystallography of biological macromolecules, ROSSMANN MG, ARNOLD E (Ed). Springer, Dordrecht, **2006**, p. 333. DOI.
18. WEEKS CM, ADAMS PD, BERENDZEN J, BRUNGER AT, DODSON EJ, GROSSE-KUNSTLEVE RW, SCHNEIDER TR, SHELDRIK GM, TERWILLIGER TC, TURKENBURG MG, USÓN I. "Automatic Solution of Heavy-Atom Substructures". In: *Methods in Enzymology*, Vol. 374, CARTER CW, SWEET RM (Ed). Elsevier, **2003**, p. 37. DOI.
19. MILLÁN C, SAMMITO M, USÓN I. "Macromolecular ab initio phasing enforcing secondary and tertiary structure." *IUCrJ*, **2015**, 2: 95. DOI.
20. OEFFNER RD, AFONINE PV, MILLÁN C, SAMMITO M, USÓN I, READ RJ, MCCOY AJ. "On the application of the expected log-likelihood gain to decision making in molecular replacement." *Acta Cryst D*, **2018**, 74: 245. DOI.
21. USÓN I, SHELDRIK GM. "An introduction to experimental phasing of macromolecules illustrated by SHELXmathsemicolon new autotracing features." *Acta Cryst D*, **2018**, 74: 106. DOI.
22. "4J5M structure solution using ARCIMBOLDO_SHREDDER spheres." *YouTube* (channel: ArcimboldoTeam), 00:03:04 (2017). ID: KdmmujCit3o.
23. SAMMITO M, MILLÁN C, RODRÍGUEZ DD, DE ILARDUYA IM, MEINDL K, MARINO ID, PETRILLO G, BUEY RM, DE PEREDA JM, ZETH K, SHELDRIK GM, USÓN I. "Exploiting tertiary structure through local folds for crystallographic phasing." *Nat Methods*, **2013**, 10: 1099. DOI.
24. KUHLEBRANDT W. "The Resolution Revolution." *Science*, **2014**, 343: 1443. DOI.

ISABEL USÓN

ICREA Research Professor,
Crystallographic methods group,
Structural Biology Unit – SBU,
Institute of Molecular Biology of Barcelona – IBMB,
Barcelona (Spain).
E-mail: iufcri@ibmb.csic.es

CLAUDIA MILLÁN

Crystallographic methods group,
Structural Biology Unit – SBU,
Institute of Molecular Biology of Barcelona – IBMB,
Barcelona (Spain).
E-mail: cmncri@ibmb.csic.es

EDITORS

Jesús Salgado
Jorge Alegre-Cebollada
Xavier Daura
Teresa Giráldez

ISSN 2445-4311

CONTACT

SBE - Sociedad de Biofísica de España
Secretaría SBE, IQFR-CSIC,
C/Serrano 119, 28006 Madrid
Email: sbe_secretaria@sbe.es
WEB: <http://www.sbe.es>

SPONSORS



Biofísica: Biophysics Magazine by SBE - Sociedad de Biofísica de España.

Design based on a Theme by Alx. Powered by WordPress. PDF export using wkhtmltopdf.