

Received 11 January 2022

Accepted 16 March 2022

Edited by M. Spackman, University of Western Australia, Australia

Keywords: crystal electron density; motion in crystals; diffuse scattering.

Crystal structures

Hans-Beat Bürgi*

Department of Chemistry, Biochemistry and Pharmacy, University of Berne, Freiestrasse 12, Bern, CH-3012, Switzerland, and Department of Chemistry, University of Zürich, Winterthurerstrasse 190, Zürich, CH-8057, Switzerland.

*Correspondence e-mail: hans-beat.buergi@unibe.ch

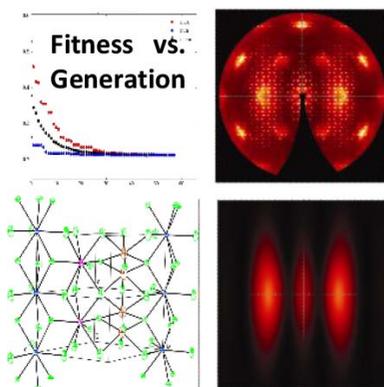
A personal view is offered on various solved and open problems related to crystal structures: the present state of reconstructing the crystal electron density from X-ray diffraction data; characterization of atomic and molecular motion from a combination of atomic displacement parameters and quantum chemical calculations; Bragg diffraction and diffuse scattering: twins, but different; models of real (as opposed to ideal) crystal structures from diffuse scattering; exploiting unexplored neighbourhoods of crystallography to mathematics, physics and chemistry.

1. Introduction

Since the Braggs determined the first crystal structures *ca* 100 years ago, single-crystal X-ray structure determination has become so well understood that both the acquisition and the interpretation of the diffraction data are now largely automated and make it the analytical method of choice for the characterization of the chemical composition and the three-dimensional (3D) structure of most newly synthesized chemical compounds.

This phenomenal achievement is due to several generations of pioneering crystallographers taking advantage of a variety of technical advances. In a nutshell: radiation sources have developed from homemade X-ray tubes to highly intense and tunable synchrotrons, X-ray free electron lasers (XFEL) and ultrafast electron diffractometers; signal observation evolved from photographic film to energy-discriminating electronic detectors. Although crystals are still rotated in the beam of probing radiation as they were 100 years ago, this is no longer done by hand, but under complete computer control. Direct methods for solving the phase problem were invented in the mid-1950s, got a Nobel in 1985 and have now been replaced almost completely by dual-space phase retrieval algorithms, such as charge flipping. Software for interpreting the experiment in combination with the staggering improvement of computational resources allow on-the-fly interpretation of the diffraction data and return results within hours instead of months or years. Another factor for the success of crystal structure analysis came from outside crystallography. Synthetic chemists decided that delegating the arduous task of structure determination by chemical means to a crystallographic technique freed their resources to invent new and better methods of chemical synthesis.

My involvement with crystal structures began in the mid-1960s during my PhD work with Jack Dunitz when the first automatic diffractometers and mainframe computers came on the scene in crystallography. Direct methods were still in the



development stage. Structure refinement was done at the rate of two or three least-squares cycles per night, often in block-diagonal approximation. Completing a crystal structure determination took several months and there was plenty of time to think about the result. My interests in crystal structures developed with some of the changes in the field. When the number of available structures started to increase exponentially, it became possible to look for structure correlations across families of more or less related compounds. The (continuing) emphasis on static crystal structures naturally kindled my interest in its dynamic aspects, either by considering structural fragments as snapshots that could trace chemical reaction coordinates if properly collated (Bürgi, 1975) or in the form of thermal motion and atomic displacement parameters (Bürgi & Capelli, 2000). Eventually the irritations associated with disorder pushed me towards its signature, diffuse scattering. In the following, I sketch my opinions on some of these topics, biased in length and detail by my areas of interest and by my own work during the last 50 years.¹

2. X-ray diffraction and the crystal electron density

X-ray diffraction is an experimental method to determine the electron-density function of a crystal. In practice, this is done at various levels of sophistication: measure Bragg diffraction intensities, find electron-density maxima with the help of spherical atomic form factors, interpret them as nuclear positions and account for the smearing of the electron density due to atomic motion with the independent atom model (IAM). This is the bread-and-butter business of crystal structure determination. Advanced methods, such as Hirshfeld Atom Refinement (HAR), work with tailor-made nonspherical Hirshfeld atom form factors (Hirshfeld, 1977). HAR and multipole modelling (MM) scrutinize the electron density for effects due to chemical bonding and analyze it for finer details, such as the polarization of atoms or molecules in the crystal field (Jayatilaka & Dittrich, 2008; Capelli *et al.*, 2014). One result from HAR is particularly significant: hydrogen and other light-atom parameters can be determined with an accuracy mostly comparable to that obtained from a much more expensive and time-consuming neutron diffraction experiment (Woińska *et al.*, 2016; Köhler *et al.*, 2019). Many variations and improvements of the HAR method are possible and some are being explored (Chodkiewicz *et al.*, 2020). In an ultimate step, information on polarization by the crystal field, as well as the influence of electron correlation and of relativistic effects on the electron density, are extracted [X-ray Constrained Wavefunction method (XCW); Jayatilaka & Grimwood, 2001; Davidson *et al.*, 2022a,b].

As ever finer details of the electron-density function are extracted from the experimental data, the question arises

whether this information is accurate, *i.e.* (largely) free of systematic errors, and precise, *i.e.* associated with credible standard uncertainties. Once the answer is in the affirmative, XCW and related methods represent the latest chapter of the fascinating and successful story of retrieving the ground-state electron density from Bragg diffraction data, a story whose outline was given in 1915 by Peter Debye (1915). The potential and limitations of XCW are still being explored [Genoni & Macchi, 2020; Davidson *et al.*, 2022a,b; see also the recent special issue on *Quantum Crystallography* in *Acta Crystallographica Section B* (https://journals.iucr.org/special_issues/2021/QCrOM2020)]. Some comments on and criticism of this method are given elsewhere (Bürgi & Genoni, 2022).

Efficiency in crystal structure analysis depends not only on experimental hardware and software. A time-consuming part of its workflow is the crystallization of micron-sized crystals. Recently, X-ray diffraction from such crystals has been complemented by the rediscovered method of single-crystal electron diffraction (ED) applied to nanocrystals. ED accelerates structure determination because nanocrystals are the size usually obtained first in synthesis. At present, the gain in time is achieved at the expense of accuracy, at least until the structure determination software based on kinematic diffraction theory is replaced by comparably efficient software based on the more complex dynamic scattering theory (Palatinus *et al.*, 2015a,b). The method is sometimes considered as part of the nanocrystal revolution because ‘it provides single-crystal data of structure solution and refinement quality, allowing the atomic structure determination of those materials that remained hitherto unknown because of their limited crystallinity’ (Gemmi *et al.*, 2019; Gruene *et al.*, 2021).

The seemingly straightforward process of retrieving the electron-density function from Bragg data has a few stumbling blocks. Some of them become more and more important as more and more information is extracted from the primary data. The most important one is to assure sufficient accuracy and precision of the data on the basis of a full physical and technical understanding of the diffraction experiment and the appropriately designed processing of the raw data. No crystal shows ideal translational symmetry throughout and, at all times, there are defects, disorder and the unavoidable thermal motion mentioned above. Such features lead to elastic and inelastic diffuse scattering between and beneath the Bragg reflections. To obtain the highly accurate Bragg intensities needed for detailed analyses, such as XCW fitting, the Bragg and diffuse signals need to be carefully distinguished and separated. X-ray absorption has to be properly and accurately accounted for. Absorption is accompanied by anomalous dispersion, which is not isotropic as generally assumed (Schiltz & Bricogne, 2008). In rare cases, corrections are needed, allowing for the breakdown of the kinematic scattering theory (extinction). While all of these effects are well known, they are often accounted for in terms of some approximate empirical procedure of unknown accuracy rather than the respective physical models.

¹ References given in this text are not exhaustive, but are meant to point to the relevant, more or less recent, literature.

3. Atomic and molecular motion: harmonic, quasi-harmonic, anharmonic and in real time

Being quantum mechanical objects, atoms in crystals vibrate down to a temperature of 0 K, an effect described by the mean-square atomic displacement parameters (ADPs). With increasingly accurate data and more comprehensive models of the electron density, the physical information content of ADPs improves as well since the ADPs tend to absorb fewer inadequacies of the model of the electron density. Atomic motion is governed by the interatomic and intermolecular potential functions. Extracting such information from ADPs is a largely unexplored field.

Interpreting ADPs in terms of models of motion in crystals faces a major limitation: ADPs pertain to individual atoms and say nothing about the correlation of motion between pairs of atoms. If there is one or a small number of strong chemical bonds between such a pair, their motion in the direction of the internuclear vector can reasonably be assumed to be highly correlated, implying that differences of mean-square displacements along this vector are close to zero (except for H–X and other Y–X bonds with order of magnitude mass differences). This assumption can be tested experimentally (Hirshfeld differences; Hirshfeld, 1976). If two atoms in a molecule are further apart, their motion is increasingly less correlated and the Hirshfeld differences may become significant. Given a sufficient number of rigid atomic connections in a molecule or molecular fragment, it can often be considered as a rigid or semi-rigid body whose librational, translational and internal motion can be extracted from the ADPs (Schomaker & Trueblood, 1968). In general, one or more parameters of such models are indeterminate (Schomaker & Trueblood, 1998).

Problems with indeterminate parameters can be circumvented by measuring the temperature dependence of the ADPs. With such data, the type and degree of correlated motion can be retrieved on the basis of normal coordinate analysis. The mathematical model extracts vibrational frequencies and the associated collective atomic displacement patterns (normal modes) from the ADPs (Bürgi & Capelli, 2000). The procedure is the mathematical inverse of extracting force constants and normal modes from vibrational spectra. Both types of analysis are based on parabolic, *i.e.* harmonic, potential energy functions and are susceptible to a shortage of experimental data. In vibrational spectroscopy, isotopic substitution can help. In the case of ADPs, their nonlinear dependence on temperature at the transition from the low-temperature quantum to the high-temperature classical regime, as well as the presence of anharmonicity, provide extra information (Bürgi *et al.*, 2000). The mathematical complexity arising from anharmonicity is circumvented by the quasi-harmonic approximation, which assumes a slow change of harmonic frequencies with temperature (quantified by the Grüneisen parameter; Grüneisen, 1926). Diffuse scattering is another source of information on the coupling of atomic motion (see below).

The normal-mode approach considers molecular motion as taking place in the mean field of the crystal lattice. Low-frequency degrees of freedom, such as molecular libration, translation and soft intramolecular deformation (*e.g.* torsion and out-of-plane), are the main contributors to the ADPs. The high-frequency intramolecular bond stretching and angle bending vibrations contribute relatively little. To get a balanced description of the inter- and intramolecular dynamics in the crystal, ADP analysis for the soft low-frequency vibrations is combined with quantum chemical calculations of the high-frequency intramolecular vibrations. Unlike the latter, the former are difficult to model with quantum chemical calculations and susceptible to effects of anharmonicity because they depend primarily on intermolecular, often anharmonic, potential functions. From a combined ADP and quantum chemical analysis, the specific heat and other thermodynamic quantities can be obtained with statistical mechanics. Given the physically more meaningful ADPs available from HAR and optimal ways of combining the crystallographic and quantum chemical data, this approach has unexplored potential towards better estimates of the relative stability of polymorphs (Aree *et al.*, 2014). Using the same approach, motion may be distinguished from static positional disorder at a resolution of a few hundredths of an Ångström (Å) as has been shown in the case of benzene. Ermer (1987) convincingly argued that the ADPs of benzene are compatible with a disordered arrangement of a threefold-symmetric localized cyclohexatriene structure. Analysis of the combined ADPs from neutron diffraction data measured at 15 and 123 K unambiguously confirm the sixfold-symmetric delocalized structure of the benzene molecule (Bürgi & Capelli, 2003). An unexplored possibility is to combine ADPs from isotopomers with sufficient mass difference, *e.g.* H and D, into a single model of motion.

In an extension of the molecular-mean-field normal-mode approach, entire unit cells rather than individual molecules have been considered. This requires that motions of molecules relative to each other must be taken into account explicitly rather than in a mean-field approximation. Here the lack of information on the correlation of motion is compensated by calculating displacement patterns (phonons) quantum chemically at the origin of the Brillouin zone (Γ -point). The calculated frequencies of these low-frequency intermolecular motions, which are governed primarily by nonbonded interactions, are difficult to represent faithfully by quantum calculations. Therefore, the calculated Γ -point displacement patterns are fixed and their low-energy frequencies optimized to best explain the observed ADPs. The approach is harmonic and based on data measured at a single temperature, usually ~ 100 K (Hoser *et al.*, 2021). The model represents an average over the Brillouin zone. Extensions of this model to additional points in the Brillouin zone or of models which include multi-temperature data and thus anharmonicity would seem possible, and may even be worthwhile, but have not yet been investigated.

Good-quality diffraction data sometimes reveal the limits of describing atomic motion in terms of a second-order ADP-

tensor which assumes (quasi-)harmonic potentials. Anharmonic models are then required to describe the atomic displacements from their mean positions. In diffraction experiments, the term ‘anharmonicity’ subsumes several phenomena. One is thermal expansion/contraction, which implies that the distances between the average positions of atoms and molecules change with temperature. For experiments done at a single temperature, this effect is not apparent. For multi-temperature experiments, this effect can be accounted for with quasi-harmonic models (with the Grüneisen constant; Grüneisen, 1926).

Another meaning of anharmonicity refers to the inadequacy of describing the probability density function (pdf) of atomic displacements with 3D Gaussian functions expressed in *linear* mean-square displacements along the three directions of space. To describe *curvilinear* motion higher order, anharmonic models of the pdf are needed, such as the Gram-Charlier series (Johnson & Levy, 1974; Herbst-Irmer *et al.*, 2013). However, this definition is tied to the use of crystal coordinates. If one were to use coordinates more adapted to the type of motion to be described, the associated pdf may well be Gaussian again. Consider libration as an example: while it displaces atoms on an arc, not in a line, the pdf expressed in terms of a mean-square libration angle may well be Gaussian, *i.e.* the motion is harmonic in librational coordinates. This may sound like sophistry, but an understanding of this point helps the description of atomic motion in terms of a minimum of parameters and thus minimal correlations between them. For the example of libration, the higher-order term deforming the ellipsoidal pdf based on the Gaussian ADP model into a banana shape is well approximated by a function of the ADPs and the atomic positions of the librating atoms, and does not require extra parameters (Willis & Pawley, 1970). This issue arises whenever the transformation between the ADP and the alternative coordinate system is nonlinear. The nature and energy of correlated motions associated with anharmonic pdfs have hardly been investigated.

Studying motion in crystals is not restricted to measuring time-averaged expectation values, such as ADPs. Synchrotron and XFEL radiation have a pulsed time structure that can be exploited for photo-induced stationary-state experiments providing information on excited-state molecular structures. Photoexcitation produces a transient concentration of excited-state molecules that can be probed with different delay times (pump-and-probe) and time resolution, hours to seconds on laboratory instruments and micro- to picoseconds with synchrotrons. Such experiments produce reaction rates and activation energies of chemical reactions (Coppens, 2017). The ultimate goal is to track chemical processes with serial femtosecond crystallography on XFELs and to illustrate them through ‘molecular movies’, as the catchword goes (Deresz *et al.*, 2021). Ultrafast electron diffraction (UED) operates in the same time range, benefits from a stronger interaction between the sample and the probing electrons and is applied to crystalline or gaseous samples (Aseyev *et al.*, 2020). Eventually motion and its driving forces at the atomic or molecular level

will have to be correlated with and actively translated to macroscopic motion, a program nicely summarized in a perspective by Naumov *et al.* (2020).

4. Diffuse scattering versus Bragg diffraction

As mentioned above, no crystal shows ideal translational symmetry throughout and, at all times, there are defects, disorder and unavoidable thermal motion. The effect of displacing electron density from its mean position is to damp the diffracted Bragg intensities, increasingly so with increasing scattering angle. This loss of intensity is essentially compensated by the appearance of diffuse scattering (through conservation of energy), either as inelastic thermal diffuse scattering (TDS) due to atomic thermal motion or as elastic disorder diffuse scattering (DDS) due to permanent positional or other kinds of disorder.

Diffuse scattering is not a rare phenomenon. Every one of the more than 1.5 million structures in the various structural databases shows TDS. About 20% of the entries in the Cambridge Structural Database (CSD; Groom *et al.*, 2016) show static disorder: occupation of atomic positions by different elements, alternative positions for a given atom, different molecular orientations or conformations, *etc.* Such disorder produces DDS. Depending on its origin, it may be found anywhere in reciprocal space. Biomolecules are especially notorious for a fast decay of Bragg intensities towards high scattering angle. In return, they show richly structured diffuse scattering all over reciprocal space (Xu *et al.*, 2021). Being smeared out over substantial parts of reciprocal volumes, diffuse scattering generally looks weak compared to the intense zero-dimensional Bragg peaks. However, its integral over reciprocal space may well be of the same order of magnitude as the sum of the Bragg intensities. Because of its continuous nature, measuring diffuse scattering is best done with an area detector; 100 years ago, this was photographic film (Lonsdale, 1942). Nowadays, reliable diffuse data are obtained from strong radiation sources, *e.g.* from synchrotrons, with fast electronic pixel detectors.

TDS is found mostly around and beneath the Bragg reflections, thus raising the question whether the presently used processing of the raw data is correctly distinguishing between the two. Studying either therefore requires the two to be reliably separated from each other. Admittedly, there are many factors affecting the shape and composition of intensity in the neighbourhood of a Bragg peak. They include properties of the crystal, the X-ray beam, the detector, the general background, *etc.* (Tsirelson & Ozerov, 1996). The way these factors are considered and documented in commercially available data-processing programs for Bragg intensity estimation is somewhat nontransparent. In the absence of well-defined algorithms and computer programs which account for these factors in terms of physical models, approximate empirical procedures are being used to obtain Bragg intensities. Niepötter *et al.* (2015) played with the integration box used to extract structure factors from the raw data and found one that led ‘to a substantial improvement in the model

quality and removes the differences in the models (of the electron densities at 15 and 100 K)'. Algorithms for obtaining the diffuse signal by removing Bragg intensities from total scattering work with even cruder guesses (*e.g.* the punch-and-fill method; Weber & Simonov, 2012). It would appear that the intricacies of separating Bragg diffraction from diffuse scattering deserve more attention by a community which tries to extract ever more detailed physical and chemical data from diffraction experiments.

TDS and DDS beneath and around Bragg peaks is due primarily to the long-wave low-energy acoustic phonons which determine the elastic tensor of the crystal and to small static atomic displacements. If the elastic tensor is known, the TDS can be calculated and introduced into the processing of the diffraction data (Willis & Pryor, 1975). A recent review critically compares experimentally determined elastic tensors of molecular crystals with quantum chemically computed ones (Spackman *et al.*, 2021). It has also been shown that the elastic tensor can be determined from a single-crystal diffraction experiment using monochromatic X-rays (Wehinger *et al.*, 2017). It seems, however, that this knowledge is hardly used to systematically separate diffuse scattering from Bragg diffraction, although such a separation would be advantageous for the interpretation of both types of data (Lucas, 1971)

5. Information in diffuse scattering

Diffuse scattering is not just a nuisance, it measures the deviations from perfect translational symmetry. If it can be interpreted, a more realistic picture of the real, as opposed to the idealized, average structure is obtained. Understanding the structural nature of disorder often holds the key to understanding and controlling materials properties.

Interpreting diffuse scattering data is orders of magnitude more complex than solving structures from Bragg data. The latter can rely on the symmetry of 230 space groups, several algorithms to solve the phase problem, sophisticated least-squares programs to optimize structural models and a highly automated workflow. Although disordered structures usually show a definite, albeit physically or chemically impossible, average structure (*e.g.* atoms with fractional occupation), the deviations from the average that are of interest cannot be enumerated as easily as can the space groups. There are only a few methods for 'structure solution', *i.e.* characterizing the differences between unit cells in the disordered crystal (Welberry & Weber, 2016). Characterizing and optimizing disorder can require substantial computational resources (Michels-Clark *et al.*, 2013, 2016).

From 'total scattering' powder data including Bragg and diffuse scattering, a one-dimensional (1D) pair distribution function (1D-PDF) can be obtained, *i.e.* a distribution of all interatomic vectors in the crystal weighted by the product of the scattering powers of the corresponding atom pair and smeared by the (static or dynamic) mean-square deviations of the interatomic distances (not the ADPs). From such distributions, a structural model with short-range coherence may be derived (Billinge, 2019). With 3D 'total-scattering' data from a

single crystal, an analogous but 3D difference pair distribution function (3D- Δ PDF) can be obtained. It is the difference between the 3D-PDF (analogous to the 1D-PDF, but also showing the direction of the interatomic vectors) and the Patterson function which is the pair distribution function of the periodic average structure derived from the Bragg intensities. A 3D- Δ PDF brings into prominence the deviations from perfect periodicity (Weber & Simonov, 2012). As in the case of 1D-PDFs, 3D- Δ PDFs serve as a basis for constructing atomistic models accounting for disorder by defining probabilistic interactions, *e.g.* 'interaction energies', between the disordered alternative parts of the structure. Atomistic modelling treats – atom-by-atom or molecule-by-molecule – a smaller or larger number of unit cells as a representative piece of a disordered crystal. An originally random distribution of structural fragments is rearranged with a Monte Carlo or reverse Monte Carlo algorithm until the deviation between the observed and calculated scattering is minimal. Even for only a few thousand cells containing a few molecules each, the required computing resources for structure modelling become very large very quickly.

Disorder models are typically defined with sets of 'interaction energies', probabilities and geometric and other parameters. Their magnitudes need to be optimized until the model of diffuse scattering best fits the experimental one. Several global optimization algorithms, such as differential evolution, genetic algorithms or particle swarm optimization, have been compared (Michels-Clark *et al.*, 2013, 2016). In the presence of highly structured diffuse scattering, direct optimization of the interaction parameters in a mean-field type of approach is a possible alternative (Schmidt *et al.*, 2021). Whatever method is used, the required computational resources are still orders of magnitude larger than those for even the most detailed analysis of Bragg diffraction. The field of disordered structure determination thus provides ample opportunities for developing new and efficient methods to obtain and interpret scattering data, to characterize disorder structurally and to optimize the corresponding models. One limitation of even the best models needs to be mentioned, however. The disordered structures resulting from the above methods are not photographs of the sample used to measure the diffuse scattering, they are merely a piece of crystal whose structural properties are statistically equivalent to the sample under investigation.

Given the costs and complications encountered in interpreting diffuse scattering, it is not surprising that the number of structures with well documented disorder is relatively small, at perhaps a few thousand structures. This number is to be compared to the more than 1.5 million single-crystal structures deposited in several databases. These relatively few data sample a nearly infinite number of possible deviations from any ordered structure. The potential richness of disordered structures raises the question whether – and if so how – disordered structures may be classified in terms of simple rules governing correlated disorder. There is hope: as Goodwin and his collaborators have put it 'We take encouragement from the observation that certain types of correlated disorder recur in

completely different fields. This recurrence of specific forms of disorder hints at the possibility of a universal language for describing correlated disordered states, much like the space groups of classical crystallography' (Keen & Goodwin, 2015). The field of determining and interpreting disordered structures and finding their 'universal language' would seem to promise many new challenges and interesting discoveries.

6. Additional remarks

Crystal structure determination by diffraction methods is often said to sit at the intersection of mathematics, physics and chemistry or molecular biology. Crystallography can continue to take advantage of this privilege. Mathematics has more to contribute to crystallography than the tools required for interpreting diffraction experiments (symmetry, geometry, Fourier transforms, *etc.*); physics is not only quantum mechanics, and confirming structure at the atomic scale is but a small part of chemistry. There are many more overlaps between these fields and crystallography.

The surge of the chemistry of Metal–Organic Frameworks (MOFs) during the last 25 years was accompanied by a resurgence of the study of crystal structures in terms of mathematical concepts, such as periodic graphs, nets, knots and weavings. These objects need to be enumerated, topologically characterized, classified and documented in databases, knowledge that proves to be a fertile basis for inventing and synthesizing new MOFs with new properties (Delgado-Friedrichs *et al.*, 2017; Liu *et al.*, 2018).

Another example is crystallization. In contrast to the abundant knowledge on crystal structures themselves, surprisingly little detailed information is available for the process of crystallization from solution. Genesis at the atomic level of liquid-like prenucleation clusters, formation of nanoparticles (amorphous, partially ordered or crystalline), their aggregation, Ostwald ripening, other crystal growth processes and phase transitions have been and will need to be studied with a wide variety of methods, such as concentration, activity and other solution methods, thermodynamics, analytical ultracentrifugation, kinetic experiments, atomic simulations, NMR, EXAFS, FT–IR, *etc.* (Gebauer *et al.*, 2014). Proponents of classical nucleation theory still argue with advocates of prenucleation clusters over the best description of the very early stages of crystal formation (Gebauer *et al.*, 2018).

The structure prediction competitions for nonbiological molecules organized by the CSD use quantum chemical, force field and other methods with increasing, but still limited, success (Reilly *et al.*, 2016). After the stunning success of protein structure prediction with Machine Learning (ML) algorithms (Tunyasuvunakool *et al.*, 2021), one wonders whether meaningful structural descriptors and corresponding ML methods could be developed to complement the conventional tools of structure prediction. The propensity of small organic molecules to crystallize has been investigated using such a method (Pereira, 2020). ML with data from the CSD has predicted the density of energetic compounds (explosives) based on the molecular formula alone (Nguyen *et al.*, 2021). To

help structure determination by solid-state NMR experiments, chemical shifts in molecular solids have been calculated and validated with a combination of ML and DFT calculations on known structures from the CSD (Paruzzo *et al.*, 2018). The interesting aspect of these applications of ML is their capability to discover patterns, use them for predictions and thereby provide useful tools for solving problems. The difficult aspect of ML for a scientist of my generation is that these methods do not directly provide an understanding of the phenomena in terms of basic mathematical, physical and chemical principles. My younger colleagues may see this differently.

There is the vast area of materials with their optical, electrical, magnetic and mechanical properties (with a strong, but not exclusive, emphasis on photovoltaics, batteries and catalysis). Materials development is now a transdisciplinary effort; the 'Materials Genomics Initiative' is an example (de Pablo *et al.*, 2019). It uses ML and cheminformatics methods that can rely on information from experiments and experimental databases, such as the Inorganic Crystal Structure Database (ICSD), as well as the open-access database of theoretically computed physical properties from the 'Materials Project' (Jain *et al.*, 2013). An in-vogue field of study is multiferroics, single-phase materials that show multiple, intimately coupled ferroic orders, *e.g.* ferroelectric, ferromagnetic and ferroelastic. The symmetry-dependent coupling gives control of one parameter by another, *e.g.* magnetic control of ferroelectricity and electric control of magnetic ordering. Because of the widespread practical applications of multiferroics, their miniaturization to nanoscopic dimension is an active field of research for both inorganic (Ramesh & Martin, 2021) and organic materials (Huang *et al.*, 2019).

Mechanical properties of conventional materials such as metals, alloys or ceramics are well studied. For molecular materials these properties are less well studied. They are likely to have a much richer phenomenology because atoms and molecules in crystals stick together through a wide range of bonding interactions: strong covalent and ionic bonding, intermediate-strength hydrogen, halogen and related bonding, and weak van der Waals interactions (Naumov *et al.*, 2020).

An example from materials science that I particularly like is the thermoelectric effect, the transformation of a temperature difference into a voltage and *vice versa*. The efficiency of a thermoelectric system depends on several parameters: it is proportional to the Seebeck coefficient, a materials specific constant, and to its electrical conductivity, but inversely proportional to its thermal conductivity, *i.e.* to its *electronic structure and its thermal motion properties*. The latter can be kept small by introducing disorder into the crystal structure of the material. Understanding these parameters thus entails all aspects of crystal structure mentioned above. For an example (chosen at random), see Zhang *et al.* (2021).

It is next to impossible to make specific predictions of future developments. Some unsolved problems and possible areas of research have been suggested above. It is uncertain whether some of these will make it into the limelight of a popular and well-funded part of crystallography and which of them will be

appreciated by a few connoisseurs only. One thing appears to be clear though: new and imaginative multiple combinations of subfields of mathematics, physics and chemistry with diffraction experiments will lead to fascinating research. Last, but not least, reliable additional driving forces for future developments are: *more, faster and better*. While this list sounds pretty unimaginative, on closer inspection it is these aspects that were some of the essential ingredients for making crystal structure determination what it has become during the last 100 years.

References

- Aree, T., Bürgi, H.-B., Chernyshov, D. & Törnroos, K. W. (2014). *J. Phys. Chem. A*, **118**, 9951–9959.
- Aseyev, S. A., Ryabov, E. A., Mironov, B. N. & Ischenko, A. A. (2020). *Crystals*, **10**, 452.
- Billinge, S. J. L. (2019). *Phil. Trans. R. Soc. A*, **377**, 20180413.
- Bürgi, H.-B. (1975). *Angew. Chem. Int. Ed. Engl.* **14**, 460–473.
- Bürgi, H.-B. & Genoni, A. (2022). *Acta Cryst.* **B78**, 298–304.
- Bürgi, H. B. & Capelli, S. C. (2000). *Acta Cryst.* **A56**, 403–412.
- Bürgi, H.-B. & Capelli, S. C. (2003). *Helv. Chim. Acta*, **86**, 1625–1640.
- Bürgi, H. B., Capelli, S. C. & Birkedal, H. (2000). *Acta Cryst.* **A56**, 425–435.
- Capelli, S. C., Bürgi, H.-B., Dittrich, B., Grabowsky, S. & Jayatilaka, D. (2014). *IUCrJ*, **1**, 361–379.
- Chodkiewicz, M. L., Wońska, M. & Woźniak, K. (2020). *IUCrJ*, **7**, 1199–1215.
- Coppens, P. (2017). *Struct. Dyn.* **4**, 032102.
- Davidson, M. L., Grabowsky, S. & Jayatilaka, D. (2022a). *Acta Cryst.* **B78**, 312–332.
- Davidson, M. L., Grabowsky, S. & Jayatilaka, D. (2022b). *Acta Cryst.* **B78**, 397–415.
- Debye, P. (1915). *Ann. Phys.* **351**, 809–823.
- Delgado-Friedrichs, O., Hyde, S. T., O’Keeffe, M. & Yaghi, O. M. (2017). *Struct. Chem.* **28**, 39–44.
- Deresz, K. A., Łaski, P., Kamiński, R. & Jarzemska, K. N. (2021). *Crystals*, **11**, 1345.
- Ermer, O. (1987). *Angew. Chem. Int. Ed. Engl.* **26**, 782–784.
- Gebauer, D., Kellermeier, M., Gale, J. D., Bergström, L. & Cölfen, H. (2014). *Chem. Soc. Rev.* **43**, 2348–2371.
- Gebauer, D., Raiteri, P., Gale, J. & Cölfen, H. (2018). *Am. J. Sci.* **318**, 969–988.
- Gemmi, M., Mugnaioli, E., Gorelik, T. E., Kolb, U., Palatinus, L., Boullay, P., Hovmöller, S. & Abrahams, J. P. (2019). *ACS Cent. Sci.* **5**, 1315–1329.
- Genoni, A. & Macchi, P. (2020). *Crystals*, **10**, 473.
- Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* **B72**, 171–179.
- Gruene, T., Holstein, J. J., Clever, G. H. & Keppler, B. (2021). *Nat. Rev. Chem.* **5**, 660–668.
- Grüneisen, E. (1926). *Handbuch der Physik*, edited by H. Geiger & K. Scheel, Vol. 10, pp. 1–59. Berlin: Springer.
- Herbst-Irmer, R., Henn, J., Holstein, J. J., Hübschle, C. B., Dittrich, B., Stern, D., Kratzert, D. & Stalke, D. (2013). *J. Phys. Chem. A*, **117**, 633–641.
- Hirshfeld, F. L. (1976). *Acta Cryst.* **A32**, 239–244.
- Hirshfeld, F. L. (1977). *Theor. Chim. Acta*, **44**, 129–138.
- Hoser, A. A., Szttylko, M., Trzybiński, D. & Madsen, A. Ø. (2021). *Chem. Commun.* **57**, 9370–9373.
- Huang, Y., Wang, Z., Chen, Z. & Zhang, Q. (2019). *Angew. Chem. Int. Ed.* **58**, 9696–9711.
- Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G. & Persson, K. A. (2013). *APL Mater.* **1**, 011002.
- Jayatilaka, D. & Dittrich, B. (2008). *Acta Cryst.* **A64**, 383–393.
- Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* **A57**, 76–86.
- Johnson, C. K. & Levy, H. A. (1974). Editors. *International Tables for X-ray Crystallography*, Vol. 4. Birmingham: Kynoch Press.
- Keen, D. A. & Goodwin, A. L. (2015). *Nature*, **521**, 303–309.
- Köhler, C., Lübber, J., Krause, L., Hoffmann, C., Herbst-Irmer, R. & Stalke, D. (2019). *Acta Cryst.* **B75**, 434–441.
- Liu, Y. Z., O’Keeffe, M., Treacy, M. M. J. & Yaghi, O. M. (2018). *Chem. Soc. Rev.* **47**, 4642–4664.
- Lonsdale, K. (1942). *Proc. Phys. Soc.* **54**, 314–353.
- Lucas, B. W. (1971). *Acta Cryst.* **A27**, 256–259.
- Michels-Clark, T. M., Lynch, V. E., Hoffmann, C. M., Hauser, J., Weber, T., Harrison, R. & Bürgi, H. B. (2013). *J. Appl. Cryst.* **46**, 1616–1625.
- Michels-Clark, T. M., Lynch, V. E., Hoffmann, C. M., Hauser, J., Weber, T., Harrison, R. & Bürgi, H. B. (2016). *J. Appl. Cryst.* **49**, 713–714.
- Naumov, P., Karothu, D. P., Ahmed, E., Catalano, L., Commins, P., Mahmoud Halabi, J., Al-Handawi, M. B. & Li, L. (2020). *J. Am. Chem. Soc.* **142**, 13256–13272.
- Nguyen, P., Loveland, D., Kim, J. T., Karande, P., Hiszpanski, A. M. & Han, T. Y. (2021). *J. Chem. Inf. Model.* **61**, 2147–2158.
- Niepötter, B., Herbst-Irmer, R. & Stalke, D. (2015). *J. Appl. Cryst.* **48**, 1485–1497.
- Pablo, J. J. de, *et al.* (2019). *NPJ Comput. Mater.* **5**, 1–23.
- Palatinus, L., Corrêa, C. A., Steciuk, G., Jacob, D., Roussel, P., Boullay, P., Klementová, M., Gemmi, M., Kopeček, J., Domeneghetti, M. C., Cámara, F. & Petříček, V. (2015a). *Acta Cryst.* **B71**, 740–751.
- Palatinus, L., Petříček, V. & Corrêa, C. A. (2015b). *Acta Cryst.* **A71**, 235–244.
- Paruzzo, F. M., Hofstetter, A., Musil, F., De, S., Ceriotti, M. & Emsley, L. (2018). *Nat. Commun.* **9**, 4501.
- Pereira, F. (2020). *CrystEngComm*, **22**, 2817–2826.
- Ramesh, R. & Martin, L. W. (2021). *Riv. Nuovo Cim.* **44**, 251–289.
- Reilly, A. M., *et al.* (2016). *Acta Cryst.* **B72**, 439–459.
- Schiltz, M. & Bricogne, G. (2008). *Acta Cryst.* **D64**, 711–729.
- Schmidt, E. M., Bulled, J. M. & Goodwin, A. L. (2021). arXiv: 2107.13336.
- Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* **B24**, 63–76.
- Schomaker, V. & Trueblood, K. N. (1998). *Acta Cryst.* **B54**, 507–514.
- Spackman, P. R., Grosjean, A., Thomas, S. P., Karothu, D. P., Naumov, P. & Spackman, M. A. (2021). *Angew. Chem. Int. Ed.* **61**, e202110716.
- Tsirelson, V. G. & Ozerov, R. P. (1996). In *Electron Density and Bonding in Crystals: Principles, Theory and X-ray Diffraction Experiments in Solid State Physics and Chemistry*. Bristol, Philadelphia: Institute of Physics Publishing.
- Tunyasuvunakool, K., *et al.* (2021). *Nature*, **596**, 590–596.
- Weber, T. & Simonov, A. (2012). *Z. Kristallogr.* **227**, 238–247.
- Wehinger, B., Mirone, A., Krisch, M. & Bosak, A. (2017). *PRL*, **118**, 035502.
- Welberry, R. T. & Weber, T. (2016). *Crystallogr. Rev.* **22**, 2–78.
- Willis, B. T. M. & Pawley, G. S. (1970). *Acta Cryst.* **A26**, 254–259.
- Willis, B. T. M. & Pryor, A. W. (1975). In *Thermal Vibrations in Crystallography*. London: Cambridge University Press.
- Wońska, M., Grabowsky, S., Dominiak, P. M., Woźniak, K. & Jayatilaka, D. (2016). *Sci. Adv.* **2**, e1600192.
- Xu, D., Meisburger, S. P. & Ando, N. (2021). *Biochemistry*, **60**, 2331–2340.
- Zhang, J., Roth, N., Tolborg, K., Takahashi, S., Song, L., Bondesgaard, M., Nishibori, E. & Iversen, B. B. (2021). *Nat. Commun.* **12**, 6709.